
JMIR Formative Research

Impact Factor (2022): 2.2

Volume 5 (2021), Issue 5 ISSN 2561-326X Editor in Chief: Gunther Eysenbach, MD, MPH, FACMI

Contents

Original Papers

- Perception of Mental Health Care Professionals in Saudi Arabia on Computerized Cognitive Behavioral Therapy: Observational Cross-sectional Study ([e26294](#))
Ahmad AlHadi, Khawla Alammari, Lojain Alsiwat, Nojood Alhaidri, Nouf Alabdulkarim, Nouf Altwaijri, Shamma Alsohaili. 5
- Impact of an Online Gastrointestinal Symptom History Taker on Physician Documentation and Charting Time: Pragmatic Controlled Trial ([e23599](#))
Natalie Hall, Sameer Berry, Jack Aguilar, Elizabeth Brier, Parth Shah, Derek Cheng, Jeremy Herman, Theodore Stein, Brennan Spiegel, Christopher Almaro. 13
- Counselor Efficiency at Providing Feedback in a Technology-Based Behavioral Weight Loss Intervention: Longitudinal Analysis ([e23974](#))
Margaret Fahey, Robert Klesges, Mehmet Kocak, Leslie Gladney, Gerald Talcott, Rebecca Krukowski. 20
- Factors Influencing Telehealth Implementation and Use in Frontier Critical Access Hospitals: Qualitative Study ([e24118](#))
Saira Haque, Sydney DeStefano, Alison Banger, Regina Rutledge, Melissa Romaine. 27
- A Resilience-Building App to Support the Mental Health of Health Care Workers in the COVID-19 Era: Design Process, Distribution, and Evaluation ([e26590](#))
Eddy Golden, Micol Zweig, Matteo Danieletto, Kyle Landell, Girish Nadkarni, Erwin Bottinger, Lindsay Katz, Ricardo Somarriba, Vansh Sharma, Craig Katz, Deborah Marin, Jonathan DePierro, Dennis Charney. 33
- A Novel Digital Patient-Reported Outcome Platform (Noona) for Clinical Use in Patients With Cancer: Pilot Study Assessing Suitability ([e16156](#))
Maria Peltola, Paula Poikonen-Saksela, Johanna Mattson, Timo Parkkari. 45
- Usability and Acceptance of Wearable Biosensors in Forensic Psychiatry: Cross-sectional Questionnaire Study ([e18096](#))
Pieter de Looff, Henk Nijman, Robert Didden, Matthijs Noordzij. 53
- Integrating User-Centered Design and Behavioral Science to Design a Mobile Intervention for Obesity and Binge Eating: Mixed Methods Analysis ([e23809](#))
Andrea Graham, Sean Munson, Madhu Reddy, Sarah Neubert, Emilie Green, Angela Chang, Bonnie Spring, David Mohr, Jennifer Wildes. 7
- A Breast Cancer Smartphone App to Navigate the Breast Cancer Journey: Mixed Methods Study ([e28668](#))
Serena Petrocchi, Chiara Filippini, Giacomo Montagna, Marta Bonollo, Olivia Pagani, Francesco Meani. 84

Video Consultation as an Adequate Alternative to Face-to-Face Consultation in Continuous Positive Airway Pressure Use for Newly Diagnosed Patients With Obstructive Sleep Apnea: Randomized Controlled Trial (e20779)	
Laura Kooij, Petra Vos, Antoon Dijkstra, Elisabeth Roovers, Wim van Harten.	95
Using Multimodal Assessments to Capture Personalized Contexts of College Student Well-being in 2020: Case Study (e26186)	
Jocelyn Lai, Amir Rahmani, Asal Yunusova, Alexander Rivera, Sina Labbaf, Sirui Hu, Nikil Dutt, Ramesh Jain, Jessica Borelli.	108
Evidence of Human-Level Bonds Established With a Digital Conversational Agent: Cross-sectional, Retrospective Observational Study (e27868)	
Alison Darcy, Jade Daniels, David Salinger, Paul Wicks, Athena Robinson.	123
Feasibility of a Voice-Enabled Medical Diary App (SpeakHealth) for Caregivers of Children With Special Health Care Needs and Health Care Providers: Mixed Methods Study (e25503)	
Emre Sezgin, Garey Noritz, Simon Lin, Yungui Huang.	130
The Adaptive GameSquad Xbox-Based Physical Activity and Health Coaching Intervention for Youth With Neurodevelopmental and Psychiatric Diagnoses: Pilot Feasibility Study (e24566)	
April Bowling, James Slavet, Chelsea Hendrick, Robbie Beyl, Phillip Nauta, Marilyn Augustyn, Mediatix Mbamalu, Carol Curtin, Linda Bandini, Aviva Must, Amanda Staiano.	165
mHealth Interventions to Support Prescription Opioid Tapering in Patients With Chronic Pain: Qualitative Study of Patients' Perspectives (e25969)	
Michael Magee, Amy McNeillage, Nicholas Avery, Paul Glare, Claire Ashton-James.	176
Optimization of Patient Flow in Urgent Care Centers Using a Digital Tool for Recording Patient Symptoms and History: Simulation Study (e26402)	
Maryam Montazeri, Jan Multmeier, Claire Novorol, Shubhanan Upadhyay, Paul Wicks, Stephen Gilbert.	189
A Provider-Facing eHealth Tool for Transitioning Youth With Special Health Care Needs From Pediatric to Adult Care: Mixed Methods, User-Engaged Usability Study (e22915)	
Darcy McMaughan, Sherry Lin, Jennifer Ozmetin, Judith Beverly, Joshua Brog, Emily Naiser.	199
Remote Rating of Atopic Dermatitis Severity Using Photo-Based Assessments: Proof-of-Concept and Reliability Evaluation (e24766)	
Zarqa Ali, Kristina Joergensen, Anders Andersen, Andrei Chiriac, Theis Bjerre-Christensen, Ionela Manole, Ana-Maria Dutei, Irina Deaconescu, Alina Suru, Adina Serban, Ari Isberg, Priyanka Dahiya, Simon Thomsen, John Zibert.	212
Changes in Perceived Stress Following a 10-Week Digital Mindfulness-Based Stress Reduction Program: Retrospective Study (e25078)	
Aarathi Venkatesan, Holly Krymis, Jenny Scharff, Art Waber.	220
Utilizing Lean Software Methods To Improve Acceptance of Global eHealth Initiatives: Results From the Implementation of the Basic Emergency Care App (e14851)	
Christian Rose, Taylor Nichols, Daniel Hackner, Julia Chang, Steven Straube, Willem Jooste, Hendry Sawe, Andrea Tenner.	232
Exploring Canadian Children's Social Media Use, Digital Literacy, and Quality of Life: Pilot Cross-sectional Survey Study (e18771)	
Lorie Donelle, Danica Facca, Shauna Burke, Bradley Hiebert, Emma Bender, Stephen Ling.	243
Neural Machine Translation-Based Automated Current Procedural Terminology Classification System Using Procedure Text: Development and Validation Study (e22461)	
Hyeon Joo, Michael Burns, Sai Kalidaikurichi Lakshmanan, Yaokun Hu, V Vydiswaran.	254

<hr/>		
Bispectral Index Alterations and Associations With Autonomic Changes During Hypnosis in Trauma Center Researchers: Formative Evaluation Study (e24044)		
C Dunham, Amanda Burger, Barbara Hileman, Elisha Chance, Amy Hutchinson.	270	
Digitalizing a Brief Intervention to Reduce Intrusive Memories of Psychological Trauma for Health Care Staff Working During COVID-19: Exploratory Pilot Study With Nurses (e27473)		
Laura Singh, Marie Kanstrup, Katherine Depa, Ann-Charlotte Falk, Veronica Lindström, Oili Dahl, Katarina Göransson, Ann Rudman, Emily Holmes.	288	
The Effect of Training on Participant Adherence With a Reporting Time Frame for Momentary Subjective Experiences in Ecological Momentary Assessment: Cognitive Interview Study (e28007)		
Cheng Wen, Doerte Junghaenel, David Newman, Stefan Schneider, Marilyn Mendez, Sarah Goldstein, Sarah Velasco, Joshua Smyth, Arthur Stone.	307	
Patient-Provider Text Messaging and Video Calling Among Case-Managed Patients Living With HIV: Formative Acceptability and Feasibility Study (e22513)		
Virginia Fonner, Samuel Kennedy, Rohan Desai, Christie Eichberg, Lisa Martin, Eric Meissner.	318	
Recorded Mental Health Recovery Narratives as a Resource for People Affected by Mental Health Problems: Development of the Narrative Experiences Online (NEON) Intervention (e24417)		
Mike Slade, Stefan Rennick-Egglestone, Joy Llewellyn-Beardsley, Caroline Yeo, James Roe, Sylvia Bailey, Roger Smith, Susie Booth, Julian Harrison, Adarsh Bhogal, Patricia Penas Morán, Ada Hui, Dania Quadri, Clare Robinson, Melanie Smuk, Marianne Farkas, Larry Davidson, Lian van der Krieke, Emily Slade, Carmel Bond, Joe Nicholson, Andrew Grundy, Ashleigh Charles, Laurie Hare-Duke, Kristian Pollock, Fiona Ng. .		
3	3	0
A Digital Platform for Facilitating Personalized Dementia Care in Nursing Homes: Formative Evaluation Study (e25705)		
Gubing Wang, Armagan Albayrak, Gerd Kortuem, Tischa van der Cammen.	354	
A Clinician-Controlled Just-in-time Adaptive Intervention System (CBT+) Designed to Promote Acquisition and Utilization of Cognitive Behavioral Therapy Skills in Bulimia Nervosa: Development and Preliminary Evaluation Study (e18261)		
Adrienne Juarascio, Paakhi Srivastava, Emily Presseller, Kelsey Clark, Stephanie Manasse, Evan Forman.	373	
Displayed Depression Symptoms on Facebook at Two Time Points: Content Analysis (e20179)		
Megan Moreno, Quintin Gaus, Megan Wilt, Alina Arseniev-Koehler, Adrienne Ton, Molly Adrian, Ann VanderStoep.	391	
Using Social Media for Qualitative Health Research in Danish Women of Reproductive Age: Online Focus Group Study on Facebook (e24108)		
Camilla Temmesen, Henriette Nielsen, Heidi Andersen, Kathrine Birch Petersen, Jane Clemensen.	404	
Use of a Self-guided Computerized Cognitive Behavioral Tool During COVID-19: Evaluation Study (e26989)		
Isadora Detweiler Guarino, Devin Cowan, Abigail Fellows, Jay Buckley.	416	
Using a Multisectoral Approach to Advance Health Equity in Rural Arizona: Community-Engaged Survey Development and Implementation Study (e25577)		
Mark Remiker, Samantha Sabo, Dulce Jiménez, Alexandra Samarron Longorio, Carmenlita Chief, Heather Williamson, Nicolette Teufel-Shone.		
4	3	7
Reporting of Differences in Taste Between Branded and Unbranded Cigarettes by Smokers Blinded to Cigarette Branding: Within-Person, Randomized Crossover Study (e24446)		
Nasser BinDhim, Nora Althumiri, Mada Basyouni, Rasha Almubark, Zaied Alkhamaali, Weam Banjar, Mohammed Zamakhshary, Khaled AlKattan.	449	
Physical Activity Patterns and Neighborhood Characteristics of First-Generation Latina Immigrants Living in Arizona: Cross-sectional Study (e25663)		
Rodney Joseph, Sonia Vega-López, SeungYong Han.	457	

Original Paper

Perception of Mental Health Care Professionals in Saudi Arabia on Computerized Cognitive Behavioral Therapy: Observational Cross-sectional Study

Ahmad N AlHadi^{1,2*}, MBBS, MD; Khawla A Alammari^{2,3*}, MBBS; Lojain J Alsiwat^{2,3*}, MBBS; Nojood E Alhaidri^{2,3*}, MBBS; Nouf H Alabdulkarim^{2,3*}, MBBS; Nouf A Altwaijri^{2,3*}, MBBS; Shamma A AlSohaili^{2,3*}, MBBS

¹Department of Psychiatry, College of Medicine, King Saud University, Riyadh, Saudi Arabia

²SABIC Psychological Health Research & Applications Chair, Department of Psychiatry, College of Medicine, King Saud University, Riyadh, Saudi Arabia

³College of Medicine, King Saud University, Riyadh, Saudi Arabia

* all authors contributed equally

Corresponding Author:

Ahmad N AlHadi, MBBS, MD

Department of Psychiatry

College of Medicine

King Saud University

PO Box 2925

Riyadh, 11461

Saudi Arabia

Phone: 966 118066346

Email: alhadi@ksu.edu.sa

Abstract

Background: Mental health disorders are common in Saudi Arabia with a 34% lifetime prevalence. Cognitive behavioral therapy (CBT), a type of psychotherapy, is an evidence-based intervention for the majority of mental disorders. Although the demand for CBT is increasing, unfortunately, there are few therapists available to meet this demand and the therapy is expensive. Computerized cognitive behavioral therapy (cCBT) is a new modality that can help fill this gap.

Objective: We aimed to measure the knowledge of cCBT among mental health care professionals in Saudi Arabia, and to evaluate their attitudes and preferences toward cCBT.

Methods: This quantitative observational cross-sectional study used a convenience sample, selecting mental health care professionals working in the tertiary hospitals of Saudi Arabia. The participants received a self-administered electronic questionnaire through data collectors measuring their demographics, knowledge, and attitudes about cCBT, and their beliefs about the efficacy of using computers in therapy.

Results: Among the 121 participating mental health care professionals, the mean age was 36.55 years and 60.3% were women. Most of the participants expressed uncertainty and demonstrated a lack of knowledge regarding cCBT. However, the majority of participants indicated a positive attitude toward using computers in therapy. Participants agreed with the principles of cCBT, believed in its efficacy, and were generally confident in using computers. Among the notable results, participants having a clinical license and with cCBT experience had more knowledge of cCBT. The overall attitude toward cCBT was not affected by demographic or work-related factors.

Conclusions: Mental health care professionals in Saudi Arabia need more education and training regarding cCBT; however, their attitude toward its use and their comfort in using computers in general show great promise. Further research is needed to assess the acceptance of cCBT by patients in Saudi Arabia, in addition to clinical trials measuring its effectiveness in the Saudi population.

(JMIR Form Res 2021;5(5):e26294) doi:[10.2196/26294](https://doi.org/10.2196/26294)

KEYWORDS

CBT; iCBT; cCBT; knowledge; attitude; mental health care professionals; computer usage; psychotherapy; therapy; cognitive behavioral therapy; health care worker; perception; Saudi Arabia; preference; mental health

Introduction

Mental health disorders are very common in Saudi Arabia, which are present in approximately 34% of the general population [1]. Cognitive behavioral therapy (CBT) is a talk-based psychotherapy that can aid in managing psychological problems by changing thinking and behavior. Research has shown that the demand for CBT is increasing [2]. Unfortunately, there are few therapists available to meet the demand and high cost [3]. To fill the gap caused by the increasing demand and low supply for CBT services, the spectrum of CBT self-help resources is growing, ranging from self-help books [4] to self-help guided programs delivered via technology-based interventions, which have been used in health care settings.

Computerized cognitive behavioral therapy (cCBT) provides a flexible health care delivery process in which patients can start their therapy with a low-intensity intervention involving only limited practitioner support. cCBT provides many advantages for the user, such as flexibility and privacy, as patients can start cCBT at any favorable setting and time [5]. Furthermore, cCBT can be delivered with or without therapist guidance [6]. In the United Kingdom, the application of conventional one-on-one CBT is constrained by the limited number of qualified therapists, with wait times for one-on-one CBT ranging between 6 months and 2 years [7]. This limitation has driven the development of self-applied CBT alternatives such as self-help books and electronic-based programs [7].

Technology-based interventions and computer-aided psychotherapy, including virtual apps and internet-based solutions, provide an attractive alternative in digitally developed countries such as Saudi Arabia, where most of the population has access to computers and mobile phones [7]. A cross-sectional study performed by the Communications and Information Technology Commission in Saudi Arabia showed that among 3000 participants aged between 12 and 65 years, 73% had access to a desktop, laptop, or tablet computer [8]. In particular, cCBT provides substantial accessibility, scalability, and flexibility benefits over conventional one-on-one CBT, and has been suggested as an effective treatment [7,9-16], especially for low-intensity interventions [7,17]. Several previous studies have displayed great diversity in the types of mental health care professionals included for evaluation, with some covering psychologists [18] or young professionals [19] exclusively, and subsequent studies covering an extended spectrum of mental health care professionals [20]. Most mental health care professionals appear to consider cCBT as inferior to one-on-one therapy, with 17%-33% declaring that cCBT can yield similar results to traditional practice [18,20,21].

Moreover, most mental health care professionals consider that therapy using computers is better utilized for protection and for mild to moderate psychological conditions [20-23]. However, a notable percentage of mental health care professionals consider the use of cCBT in their future management plans, showing a

positive attitude about the promising involvement of cCBT [18,20,21]. The knowledge of cCBT among mental health professionals is generally low [7,24], implying that knowledge of the effectiveness and availability of programs requires further development [18,19,23]. Detailed investigations revealed that some mental health care professionals were ill-informed regarding cCBT programs and the research behind them [25]. Clinician attitudes toward cCBT and other computer-based therapies were investigated in different countries such as Australia [22,23] and the United States [18,24,26,27]. Regardless of the varying results from study to study, some similarities have been found [25].

There is a rising interest in mental health care professionals' attitudes toward cCBT as a vital additional treatment modality [25]. However, no studies have been performed to date on cCBT in Saudi Arabia. In this study, we aimed to measure the knowledge of cCBT, evaluate the attitude and preference toward cCBT, and determine the usage of computers among mental health care professionals in Saudi Arabia, including psychiatrists, psychologists, and others.

Methods

Setup, Sampling, and Process

A quantitative observational cross-sectional study was performed. We translated the knowledge assessment test from Donovan et al [25]. For attitude assessment, we used the translated version the questionnaire applied in the study of Becker and Jensen-Doss [28] after obtaining approval from the authors. The translation process followed the guidelines detailed in Sousa and Rojjanasrirat [29].

This study targeted health care workers who could potentially deliver cCBT such as psychiatrists, psychologists, and social workers, among others, and was carried out between February and March of 2018 on a clustered sample spanning tertiary hospitals in Riyadh, Saudi Arabia. Due to poor responses and limited resources, an expanded data collection to span electronically collected data from mental health professionals in the Kingdom of Saudi Arabia was needed; therefore, we altered the sampling technique to convenience sampling. Similar to the Australian study results [25], the target sample size was calculated using the following equation by setting Z_{α} to 1.282, S to 1.48, and d to 0.2, to obtain 90% confidence: $n = Z_{\alpha}^2 S^2 / d^2 = (1.282^2) 1.48^2 / (0.2)^2 = 90$, where Z_{α} is the normal deviate, reflecting the type I error (calculated for 80% error); S is the standard deviation; and d is the accuracy of the estimate.

Considering a nonresponse rate of 30%, the total target sample size was 117 participants (90+27). With relaxation of the sampling technique, a total sample size of 121 participants was reached. A pilot study was performed with 10 mental health care professionals at King Khalid University Hospital in Riyadh (who were not included in the sample) to estimate the time needed to fill out the survey, assess the questionnaire's

comprehensibility, and determine any additional logistical requirements.

Ethical Considerations

All participants were informed of the purpose of the study and their right to withdraw at any time without any obligation toward the study team via a consent form. No incentives or rewards were provided for participation. The study design and purpose of the study were approved by the King Saud University Institutional Review Board before data collection commenced.

Questionnaire

Design

The questionnaire was designed as a self-administrated electronic form to maximize the response rate and cover multiple demographics.

Knowledge Scale

Knowledge was measured using six statements regarding computerized interventions. Participants chose between options of “true,” “false,” or “unsure.” This scale was adapted from Donovan et al [25].

Attitude Scale

Attitude and comfort toward computer-assisted therapy were measured using the Computer-Assisted Therapy Attitudes Scale [28]. This is an 11-item questionnaire that measures efficacy (belief in efficacy) and comfort using a 5-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). We reversed the scores for negative items; thus, in all cases, higher scores indicate a more positive attitude.

Demographics and Work-Related Questions

The third part of the questionnaire included demographics and work-related questions such as those related to CBT experience and clinical license.

Data Analysis

Data were analyzed using Statistical Package for Social Sciences version 22 (IBM Corp). Continuous variables such as age are expressed as mean (SD), whereas categorical variables are expressed as frequency and percentages. The *t* test and one-way analysis of variance were used to compare continuous variables. A *P* value <.05 was considered statistically significant.

Results

The questionnaire was sent to 132 mental health care professionals. We excluded 10 participants from the final analysis because they were also involved in our pilot study. Only one person refused to participate in the survey, and we had a near 100% completion rate from the participants who started the electronic survey, for a total of 121 participants, most of whom were women (73/121, 60.3%). The demographic characteristics of the participants, and the mean total scores for knowledge about cCBT and the feeling of using computers in therapy are summarized in Table 1. There was no statistically significant difference in knowledge with respect to demographic and work-related factors except for having a license and experience with cCBT. Knowledge was higher for clinicians who are licensed and who had experience with cCBT.

To measure the knowledge of cCBT, a test of cCBT facts was used, and the results are summarized in Table 2. The test is composed of six statements about computerized interventions, and the participants could answer with “true,” “false,” or “unsure.” The sum of answered items reflects the knowledge of cCBT; the higher the scores, the greater the knowledge of cCBT. The knowledge scale showed fair internal reliability, with Cronbach α =.601. Most of the participants showed uncertainty and lack of knowledge.

The results assessing the mental health care professionals' feelings toward using computer-assisted therapy programs are summarized in Table 3. The attitude scale showed good internal reliability with Cronbach α =.819.

Table 1. Demographic characteristics of the participants, and total scores for knowledge about computerized cognitive therapy and feelings toward using computers in therapy (N=121).

Characteristic	Value	Knowledge about computerized cognitive therapy		Feeling comfortable toward using a computer in therapy	
		Correlation coefficient or mean (SD) ^a	P value	Correlation coefficient or mean (SD) ^b	P value
Age (years), mean (SD), median	36.55 (9.11), 37.00	-0.040	.67	0.051	.58
Gender, n (%)			.31		.16
Female	73 (60.3)	0.86 (1.28)		37.07 (6.16)	
Male	48 (39.7)	0.79 (1.03)		37.71 (4.92)	
Specialty, n (%)			.69		.32
Psychiatry	29 (23.97)	1.04 (1.21)		37.43 (4.24)	
Psychology	80 (66.12)	0.76 (1.19)		37.01 (6.00)	
Sociology	8 (6.61)	1.00 (1.20)		41.25 (6.96)	
Nursing	2 (1.65)	1.00 (1.41)		35.00 (4.24)	
Other	2 (1.65)	0.00 (0.00)		35.00 (5.66)	
Area of residence, n (%)			.72		.13
Riyadh	80 (66.67)	0.80 (1.16)		36.51 (5.70)	
Makkah	19 (15.83)	0.74 (1.10)		39.58 (5.59)	
Madinah	3 (2.50)	0.33 (.58)		35.33 (7.23)	
Eastern	12 (10.00)	1.33 (1.61)		39.25 (4.65)	
Aser	2 (1.65)	1.50 (2.12)		35.00 (0.00)	
Jazan	2 (1.67)	0.50 (0.71)		43.00 (4.24)	
Al Baha	2 (1.67)	0.50 (0.71)		41.50 (0.71)	
Primary work setting, n (%)			.23		.48
Public Hospital	49 (40.50)	0.980 (1.41)		36.47 (5.80)	
Psychiatric hospital	24 (19.83)	1.08 (1.28)		38.76 (5.01)	
Primary care center	5 (4.13)	1.00 (0.70)		39.60 (4.28)	
Home care facilities	13 (10.74)	0.31 (0.48)		37.23 (7.79)	
Other	30 (24.79)	0.60 (0.89)		37.20 (5.16)	
Professional license, n (%)			.01		.07
Yes	112 (92.6)	0.88 (1.21)		37.22 (5.49)	
No	9 (7.44)	0.22 (0.44)		38.67 (8.00)	
Education level, n (%)			.12		.75
Fellowship/board	20 (16.53)	1.20 (1.32)		37.10 (4.53)	
PhD	6 (4.96)	0.33 (0.52)		39.83 (4.54)	
Master	40 (33.06)	1.03 (1.49)		37.20 (7.30)	
Bachelor	55 (45.45)	0.62 (0.85)		37.23 (4.85)	
Previous use of a computer-assisted therapy program, n (%)			.001		.32
Yes	7 (5.79)	2.57 (1.13)		39.43 (5.16)	
No	114 (94.21)	0.73 (1.11)		37.20 (5.71)	

^aBased on a total possible score of 6.^bBased on a total possible score of 55.

Table 2. Knowledge about computerized cognitive behavioral therapy (N=121).

Question (correct answer)	True, n (%)	Unsure, n (%)	False, n (%)
Computerized interventions are only available online (False)	9 (7.4)	83 (68.6)	29 (24.0)
All computerized interventions involve therapist contact (False)	32 (26.4)	70 (57.9)	19 (15.7)
Computerized interventions are less effective than face-to-face therapy (False)	36 (29.8)	76 (62.8)	9 (7.4)
Computerized interventions automatically tailor to individual needs (False)	31 (25.6)	81 (66.9)	9 (7.4)
People who receive computerized interventions are generally satisfied (True)	12 (9.9)	106 (87.6)	3 (2.5)
Computerized interventions are not interactive (False)	20 (16.5)	78 (65.4)	23 (19.0)

Table 3. Therapist attitudes and access to computer-assisted therapy (N=121).

Question	Strongly disagree, n (%)	Disagree, n (%)	Neither agree nor disagree, n (%)	Agree, n (%)	Strongly agree, n (%)
If given the opportunity and training, I would like to use computers in therapy	1 (0.8)	2 (1.6)	17 (13.9)	66 (54.1)	36 (29.5)
I feel apprehensive about using computers during therapy	28 (23.0)	42 (34.4)	39 (32.0)	11 (9.0)	2 (1.6)
I am afraid that if I begin to use computers, I will become dependent upon them and lose some of my own skills	18 (14.8)	48 (39.3)	34 (27.9)	18 (14.8)	4 (3.3)
Using computers in therapy will interfere with rapport	5 (4.1)	30 (24.6)	48 (39.3)	32 (26.2)	7 (5.7)
My clients will be more likely to drop out of treatment if I use a computer program as part of the therapy	9 (7.4)	42 (34.4)	49 (40.2)	22 (18.0)	0 (0)
My clients would find it engaging to learn new skills using a computer	0 (0)	6 (4.9)	50 (41.0)	57 (46.7)	9 (7.4)
I believe that using computer programs in therapy will lead to better outcomes for my clients	0 (0)	9 (7.4)	56 (45.9)	47 (38.5)	10 (8.2)
The challenge of learning about the use of computers in therapy seems overwhelming to me	9 (7.4)	47 (38.5)	46 (37.7)	20 (16.4)	0 (0)
I am confident that I can learn the skills to use computer-assisted therapy	1 (0.8)	2 (1.6)	14 (11.5)	80 (65.6)	25 (20.5)
My clients are not sufficiently computer savvy to use computers in therapy	0 (0)	21 (17.2)	52 (42.6)	46 (37.7)	3 (2.5)
I have sufficient access to computers to use them in sessions	19 (15.6)	42 (34.4)	33 (27.0)	22 (18.0)	6 (4.9)

Discussion

Principal Findings

Our study assessed mental health care professionals' knowledge and attitudes toward cCBT. Most of the participants were uncertain and lacked knowledge. However, the majority of participants agreed and believed in cCBT, and were confident about using it. One of the main reasons for this lack of knowledge is the lack of availability of cCBT programs in Arabic. Moreover, it appears that lack of knowledge about cCBT is not limited to Saudi Arabia, as this issue has also been demonstrated in other regions [7,24]. Nevertheless, this lack of knowledge does not greatly influence the attitude toward using computers, which is somewhat similar to the findings in Australia [25]. This suggests that lack of knowledge about cCBT

is easy to overcome and does not affect the attitude toward cCBT. However, the authors of the Australian study concluded that it is possible to change the knowledge of mental health workers about cCBT, at least to some extent, by providing them with accurate information and demonstrations as needed [25]. In another study, the attitude toward computer usage showed a positive trend, which may provide more opportunities to benefit from cCBT [26]. These results indicate a promising future for more involvement with this approach [18,20,21].

In another study, Computer-Based Training Attitudes Scale scores were higher among therapists who reported having previously used computer-based training. Negative responses toward computer-based training largely originated from those facing greater practical barriers to the use of computer-based training [26]. In our study, there was no difference in cCBT

knowledge between professionals with different backgrounds (eg, psychologists, psychiatrists, social workers, and others), which is similar to the results of the Australian study [25]. The reason could be that there is no difference in teaching and training among different specialties regarding cCBT. Most of our participants were unsure of their answers when asked about cCBT, which is similar to the results reported by Donovan et al [25] in Australia, as most of their participants had little to no knowledge of cCBT.

Clinically, cCBT can be beneficial. Reviewed advantages include flexibility in time and location, cost-effectiveness, reduction of personal stigma, time of the mental health care professional, time of waiting for treatment, the behavior of asking for help and guidance, and being satisfied by the provided treatment [30]. Despite the benefits mentioned, to accept cCBT as an alternative to the usual approach, more research should focus on how it works as an intervention [30]. There are some noted disadvantages of cCBT. Some primary concerns raised are the possible inadequacy of the therapeutic relationship and omission of the professional's contact. Another concern is that the program is not customized to each user [21]. Moreover, some professionals doubt the competence of care provided through computers [30], and the majority still consider that the usual intervention is more effective and preferable [18,20,21].

Future Directions

Future research should aim to recruit greater numbers of participants with various levels of training, skills, and different backgrounds so that effective comparisons can be made [25]. Future research should also provide clinicians with a chance to use cCBT with their clients to investigate if they would choose it. This could help generate a behavioral measure rather than assessing "intention" only [25]. The knowledge assessment in our study was very narrow, with only six questions. Future

research should include a larger number of questions assessing knowledge to ensure that any information the participants may have is reflected in their knowledge scores [25]. In addition to the suggestions for future research, replicating the study several times over the next 10 years would be interesting as changes in the country are occurring rapidly [25]. In addition, assessing patients' attitudes in Saudi Arabia toward cCBT would be very helpful.

Limitations

We used an electronic questionnaire to reach the participants in Riyadh through data collectors. We did not receive a sufficient number of responses because of the limited mental health care professionals in Riyadh, and therefore we reached out to different cities in Saudi Arabia by sending an invitation through WhatsApp groups and emails. Future research should expand this approach by including a larger participant population. Future research should also provide participants with a broader range and more items in the knowledge test.

Conclusions

The results of this study suggest that mental health care professionals in Saudi Arabia are in need of more education and training regarding cCBT; however, their attitudes toward its use, and their comfort in using computers in general, show great promise.

Lack of knowledge did not affect the participants' attitude toward cCBT, as they demonstrated a positive attitude overall. In addition, we recognize that mental health care professionals need more involvement in various up-to-date therapeutic approaches and need more resources for cCBT in Arabic. Further research should be performed to assess patients' acceptance of cCBT in Saudi Arabia along with clinical trials measuring its effectiveness in the Saudi population.

Acknowledgments

Preliminary results of this project were presented at the 27th European Congress of Psychiatry, April 6-9, 2019, in Warsaw, Poland. This research was funded by the SABIC Psychological Health Research and Applications Chair, Department of Psychiatry, College of Medicine, Deanship of Post Graduate Teaching, King Saud University. We would like to thank all of the people who participated in the research for their patience and kindness.

Conflicts of Interest

None declared.

References

1. Altwaijri YA, Al-Subaie AS, Al-Habeeb A, Bilal L, Al-Desouki M, Aradati M, et al. Lifetime prevalence and age-of-onset distributions of mental disorders in the Saudi National Mental Health Survey. *Int J Methods Psychiatr Res* 2020 Sep;29(3):e1836 [FREE Full text] [doi: [10.1002/mpr.1836](https://doi.org/10.1002/mpr.1836)] [Medline: [33245595](https://pubmed.ncbi.nlm.nih.gov/33245595/)]
2. Villemare-Krajden R, Myhr G. Evaluating the use of a computerized CBT program for outpatients on a waitlist in a university CBT unit. *J Psychiatr Pract* 2019 Jul;25(4):268-278. [doi: [10.1097/PRA.0000000000000396](https://doi.org/10.1097/PRA.0000000000000396)] [Medline: [31291207](https://pubmed.ncbi.nlm.nih.gov/31291207/)]
3. March S, Spence SH, Donovan CL. The efficacy of an internet-based cognitive-behavioral therapy intervention for child anxiety disorders. *J Pediatr Psychol* 2009 Jun;34(5):474-487. [doi: [10.1093/jpepsy/jsn099](https://doi.org/10.1093/jpepsy/jsn099)] [Medline: [18794187](https://pubmed.ncbi.nlm.nih.gov/18794187/)]
4. Martinez R, Whitfield G, Dafters R, Williams C. Can people read self-help manuals for depression? A challenge for the stepped care model and book prescription schemes. *Behav Cogn Psychother* 2007 Dec 18;36(1):89-97. [doi: [10.1017/s1352465807004067](https://doi.org/10.1017/s1352465807004067)]

5. Green KE, Iverson KM. Computerized cognitive-behavioral therapy in a stepped care model of treatment. *Prof Psychol Res Pract* 2009;40(1):96-103. [doi: [10.1037/a0012847](https://doi.org/10.1037/a0012847)]
6. Høifødt RS, Lillevoll KR, Griffiths KM, Wilsgaard T, Eisemann M, Waterloo K, et al. The clinical effectiveness of web-based cognitive behavioral therapy with face-to-face therapist support for depressed primary care patients: randomized controlled trial. *J Med Internet Res* 2013 Aug 05;15(8):e153 [FREE Full text] [doi: [10.2196/jmir.2714](https://doi.org/10.2196/jmir.2714)] [Medline: [23916965](https://pubmed.ncbi.nlm.nih.gov/23916965/)]
7. Du E, Quayle E, Macleod H. Service providers' perceptions on the uptake of computerised cognitive behavioural therapy (CCBT). *PsychNology J* 2013;11(3):213-233.
8. Individuals Report ICT Survey Results, 2015. Communications and Information Technology Commission, Kingdom of Saudi Arabia. 2015. URL: https://www.citc.gov.sa/en/reportsandstudies/studies/Documents/PublicIndividualReport2015V5_En.pdf [accessed 2020-07-13]
9. Proudfoot JG. Computer-based treatment for anxiety and depression: is it feasible? Is it effective? *Neurosci Biobehav Rev* 2004 May;28(3):353-363. [doi: [10.1016/j.neubiorev.2004.03.008](https://doi.org/10.1016/j.neubiorev.2004.03.008)] [Medline: [15225977](https://pubmed.ncbi.nlm.nih.gov/15225977/)]
10. Kaltenthaler E, Brazier J, De Nigris E, Tumor I, Ferriter M, Beverley C, et al. Computerised cognitive behaviour therapy for depression and anxiety update: a systematic review and economic evaluation. *Health Technol Assess* 2006 Sep;10(33):iii, xi-ixiv, 1. [doi: [10.3310/hta10330](https://doi.org/10.3310/hta10330)] [Medline: [16959169](https://pubmed.ncbi.nlm.nih.gov/16959169/)]
11. Kaltenthaler E, Parry G, Beverley C, Ferriter M. Computerised cognitive-behavioural therapy for depression: systematic review. *Br J Psychiatry* 2008 Sep;193(3):181-184. [doi: [10.1192/bjp.bp.106.025981](https://doi.org/10.1192/bjp.bp.106.025981)] [Medline: [18757972](https://pubmed.ncbi.nlm.nih.gov/18757972/)]
12. Spek V, Cuijpers P, Nyklíček I, Riper H, Keyzer J, Pop V. Internet-based cognitive behaviour therapy for symptoms of depression and anxiety: a meta-analysis. *Psychol Med* 2007 Mar;37(3):319-328. [doi: [10.1017/S0033291706008944](https://doi.org/10.1017/S0033291706008944)] [Medline: [17112400](https://pubmed.ncbi.nlm.nih.gov/17112400/)]
13. Andrews G, Cuijpers P, Craske MG, McEvoy P, Titov N. Computer therapy for the anxiety and depressive disorders is effective, acceptable and practical health care: a meta-analysis. *PLoS One* 2010 Oct 13;5(10):e13196 [FREE Full text] [doi: [10.1371/journal.pone.0013196](https://doi.org/10.1371/journal.pone.0013196)] [Medline: [20967242](https://pubmed.ncbi.nlm.nih.gov/20967242/)]
14. Gerhards SAH, de Graaf LE, Jacobs LE, Severens JL, Huibers MJH, Arntz A, et al. Economic evaluation of online computerised cognitive-behavioural therapy without support for depression in primary care: randomised trial. *Br J Psychiatry* 2010 Apr;196(4):310-318. [doi: [10.1192/bjp.bp.109.065748](https://doi.org/10.1192/bjp.bp.109.065748)] [Medline: [20357309](https://pubmed.ncbi.nlm.nih.gov/20357309/)]
15. Berger T, Hämmerli K, Gubser N, Andersson G, Caspar F. Internet-based treatment of depression: a randomized controlled trial comparing guided with unguided self-help. *Cogn Behav Ther* 2011;40(4):251-266. [doi: [10.1080/16506073.2011.616531](https://doi.org/10.1080/16506073.2011.616531)] [Medline: [22060248](https://pubmed.ncbi.nlm.nih.gov/22060248/)]
16. Lintvedt OK, Griffiths KM, Sørensen K, Østvik AR, Wang CEA, Eisemann M, et al. Evaluating the effectiveness and efficacy of unguided internet-based self-help intervention for the prevention of depression: a randomized controlled trial. *Clin Psychol Psychother* 2013;20(1):10-27. [doi: [10.1002/cpp.770](https://doi.org/10.1002/cpp.770)] [Medline: [21887811](https://pubmed.ncbi.nlm.nih.gov/21887811/)]
17. Rhodes H, Grant S. A review of computerised cognitive behavioural therapy (cCBT) for depression. *Cumbria Partnersh J Res Pract Learn* 2012;2(1):2-9.
18. Perle JG, Langsam LC, Randel A, Lutchman S, Levine AB, Odland AP, et al. Attitudes toward psychological telehealth: current and future clinical psychologists' opinions of internet-based interventions. *J Clin Psychol* 2013 Jan;69(1):100-113. [doi: [10.1002/jclp.21912](https://doi.org/10.1002/jclp.21912)] [Medline: [22975897](https://pubmed.ncbi.nlm.nih.gov/22975897/)]
19. Fleming T, Merry S. Youth work service providers' attitudes towards computerized CBT for adolescents. *Behav Cogn Psychother* 2013 May;41(3):265-279. [doi: [10.1017/S1352465812000306](https://doi.org/10.1017/S1352465812000306)] [Medline: [22591811](https://pubmed.ncbi.nlm.nih.gov/22591811/)]
20. Vigerland S, Ljótsson B, Bergdahl Gustafsson F, Hagert S, Thulin U, Andersson G, et al. Attitudes towards the use of computerized cognitive behavior therapy (cCBT) with children and adolescents: A survey among Swedish mental health professionals. *Internet Interv* 2014 Jul;1(3):111-117. [doi: [10.1016/j.invent.2014.06.002](https://doi.org/10.1016/j.invent.2014.06.002)]
21. Stallard P, Richardson T, Velleman S. Clinicians' attitudes towards the use of computerized cognitive behaviour therapy (cCBT) with children and adolescents. *Behav Cogn Psychother* 2010 Oct;38(5):545-560. [doi: [10.1017/S1352465810000421](https://doi.org/10.1017/S1352465810000421)] [Medline: [20615273](https://pubmed.ncbi.nlm.nih.gov/20615273/)]
22. Sinclair C, Holloway K, Riley G, Auret K. Online mental health resources in rural Australia: clinician perceptions of acceptability. *J Med Internet Res* 2013 Sep 05;15(9):e193 [FREE Full text] [doi: [10.2196/jmir.2772](https://doi.org/10.2196/jmir.2772)] [Medline: [24007949](https://pubmed.ncbi.nlm.nih.gov/24007949/)]
23. Gun SY, Titov N, Andrews G. Acceptability of Internet treatment of anxiety and depression. *Australas Psychiatry* 2011 Jun;19(3):259-264. [doi: [10.3109/10398562.2011.562295](https://doi.org/10.3109/10398562.2011.562295)] [Medline: [21682626](https://pubmed.ncbi.nlm.nih.gov/21682626/)]
24. Wells M, Mitchell KJ, Finkelhor D, Becker-Blease KA. Online mental health treatment: concerns and considerations. *Cyberpsychol Behav* 2007 Jun;10(3):453-459. [doi: [10.1089/cpb.2006.9933](https://doi.org/10.1089/cpb.2006.9933)] [Medline: [17594270](https://pubmed.ncbi.nlm.nih.gov/17594270/)]
25. Donovan CL, Poole C, Boyes N, Redgate J, March S. Australian mental health worker attitudes towards cCBT: What is the role of knowledge? Are there differences? Can we change them? *Internet Interv* 2015 Nov;2(4):372-381. [doi: [10.1016/j.invent.2015.09.001](https://doi.org/10.1016/j.invent.2015.09.001)]
26. Becker EM, Jensen-Doss A. Therapist attitudes towards computer-based trainings. *Adm Policy Ment Health* 2014 Nov;41(6):845-854. [doi: [10.1007/s10488-013-0522-4](https://doi.org/10.1007/s10488-013-0522-4)] [Medline: [24150441](https://pubmed.ncbi.nlm.nih.gov/24150441/)]
27. Cook JM, Schnurr PP, Biyanova T, Coyne JC. Apples don't fall far from the tree: influences on psychotherapists' adoption and sustained use of new therapies. *Psychiatr Serv* 2009 May;60(5):671-676 [FREE Full text] [doi: [10.1176/appi.ps.60.5.671](https://doi.org/10.1176/appi.ps.60.5.671)] [Medline: [19411356](https://pubmed.ncbi.nlm.nih.gov/19411356/)]

28. Becker EM, Jensen-Doss A. Computer-assisted therapies: examination of therapist-level barriers to their use. *Behav Ther* 2013 Dec;44(4):614-624. [doi: [10.1016/j.beth.2013.05.002](https://doi.org/10.1016/j.beth.2013.05.002)] [Medline: [24094787](https://pubmed.ncbi.nlm.nih.gov/24094787/)]
29. Sousa V, Rojjanasrirat W. Translation, adaptation and validation of instruments or scales for use in cross-cultural health care research: a clear and user-friendly guideline. *J Eval Clin Pract* 2011 Apr;17(2):268-274. [doi: [10.1111/j.1365-2753.2010.01434.x](https://doi.org/10.1111/j.1365-2753.2010.01434.x)] [Medline: [20874835](https://pubmed.ncbi.nlm.nih.gov/20874835/)]
30. Musiat P, Tarrrier N. Collateral outcomes in e-mental health: a systematic review of the evidence for added benefits of computerized cognitive behavior therapy interventions for mental health. *Psychol Med* 2014 Nov;44(15):3137-3150. [doi: [10.1017/S0033291714000245](https://doi.org/10.1017/S0033291714000245)] [Medline: [25065947](https://pubmed.ncbi.nlm.nih.gov/25065947/)]

Abbreviations

CBT: cognitive behavioral therapy

cCBT: computerized cognitive behavioral therapy

Edited by G Eysenbach; submitted 12.12.20; peer-reviewed by K Mathiasen; comments to author 22.01.21; revised version received 01.02.21; accepted 13.04.21; published 03.05.21.

Please cite as:

AlHadi AN, Alammari KA, Alsiwat LJ, Alhaidri NE, Alabdulkarim NH, Altwaijri NA, AlSohaili SA

Perception of Mental Health Care Professionals in Saudi Arabia on Computerized Cognitive Behavioral Therapy: Observational Cross-sectional Study

JMIR Form Res 2021;5(5):e26294

URL: <https://formative.jmir.org/2021/5/e26294>

doi: [10.2196/26294](https://doi.org/10.2196/26294)

PMID: [33938810](https://pubmed.ncbi.nlm.nih.gov/33938810/)

©Ahmad N AlHadi, Khawla A Alammari, Lojain J Alsiwat, Nojood E Alhaidri, Nouf H Alabdulkarim, Nouf A Altwaijri, Shamma A AlSohaili. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 03.05.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Impact of an Online Gastrointestinal Symptom History Taker on Physician Documentation and Charting Time: Pragmatic Controlled Trial

Natalie J Hall¹, MD; Sameer K Berry¹, MD, MBA; Jack Aguilar¹, MD; Elizabeth Brier¹, MD; Parth Shah¹, MD; Derek Cheng², MD; Jeremy Herman², MD; Theodore Stein², MD; Brennan M R Spiegel^{1,2,3,4,5*}, MD, MSHS; Christopher V Almario^{1,2,3,4,5*}, MD, MSHPM

¹Department of Medicine, Cedars-Sinai Medical Center, Los Angeles, CA, United States

²Karsh Division of Gastroenterology and Hepatology, Cedars-Sinai Medical Center, Los Angeles, CA, United States

³Division of Health Services Research, Cedars-Sinai Medical Center, Los Angeles, CA, United States

⁴Division of Informatics, Cedars-Sinai Medical Center, Los Angeles, CA, United States

⁵Cedars-Sinai Center for Outcomes Research and Education (CS-CORE), Los Angeles, CA, United States

*these authors contributed equally

Corresponding Author:

Christopher V Almario, MD, MSHPM
Karsh Division of Gastroenterology and Hepatology
Cedars-Sinai Medical Center
116 N Robertson Blvd
Suite 800
Los Angeles, CA, 90048
United States
Phone: 1 310 423 4462
Email: Christopher.Almario@csmc.edu

Abstract

Background: A potential benefit of electronic health records (EHRs) is that they could potentially save clinician time and improve documentation by auto-generating the history of present illness (HPI) in partnership with patients prior to the clinic visit. We developed an online patient portal called AEGIS (Automated Evaluation of Gastrointestinal [GI] Symptoms) that systematically collects patient GI symptom information and then transforms the data into a narrative HPI that is available for physicians to review in the EHR prior to seeing the patient.

Objective: This study aimed to compare whether use of an online GI symptom history taker called AEGIS improves physician-centric outcomes vs usual care.

Methods: We conducted a pragmatic controlled trial among adults aged ≥ 18 years scheduled for a new patient visit at 4 GI clinics at an academic medical center. Patients who completed AEGIS were matched with controls in the intervention period who did not complete AEGIS as well as controls who underwent usual care in the pre-intervention period. Of note, the pre-intervention control group was formed as it was not subject to contamination bias, unlike for post-intervention controls. We then compared the following outcomes among groups: (1) documentation of alarm symptoms, (2) documentation of family history of GI malignancy, (3) number of follow-up visits in a 6-month period, (4) number of tests ordered in a 6-month period, and (5) charting time (difference between appointment time and time the encounter was closed). Multivariable regression models were used to adjust for potential confounding.

Results: Of the 774 patients who were invited to complete AEGIS, 116 (15.0%) finished it prior to their visit. The 116 AEGIS patients were then matched with 343 and 102 controls in the pre- and post-intervention periods, respectively. There were no statistically significant differences among the groups for documentation of alarm symptoms and GI cancer family history, number of follow-up visits and ordered tests, or charting time (all $P > .05$).

Conclusions: Use of a validated online HPI-generation portal did not improve physician documentation or reduce workload. Given universal adoption of EHRs, further research examining how to optimally leverage patient portals for improving outcomes are needed.

KEYWORDS

patient-provider portal; computer-generated patient history; patient-reported outcome; gastrointestinal; EHR

Introduction

To facilitate communication between patients and physicians in electronic health record (EHR)-integrated environments, we developed an online patient portal (MyGiHealth) that uses a computer algorithm called Automated Evaluation of Gastrointestinal (GI) Symptoms (AEGIS) to systematically collect patients' symptom information before the clinic visit. Once collected, the data are transformed into a full narrative history of present illness (HPI) that clinicians can review prior to meeting the patient. While our prior studies noted that AEGIS creates higher quality HPIs and collects more alarm features vs physicians [1,2], we found that it did not improve patient satisfaction or shared decision making when compared to usual care in a controlled trial [3]. Therefore, the objective of this study was to investigate whether AEGIS improved physician-centric outcomes vs usual care.

Methods

We performed a pragmatic controlled study among adults aged ≥ 18 years scheduled for a new patient visit at an academic GI teaching practice and 3 community-based GI clinics at Cedars-Sinai Medical Center. This study was approved by the Cedars-Sinai Institutional Review Board, Los Angeles, CA (Pro45243).

During the intervention period (April 17, 2017-February 7, 2018), patients were invited via email to complete AEGIS via the MyGiHealth app 1 week prior to their visit. We describe AEGIS elsewhere [1-4], but in brief, the algorithm systematically assesses patients' GI symptoms and then transforms the data into a full narrative HPI as shown in [Figure 1](#). For patients who completed AEGIS, their physicians were notified 1 day before the visit that their HPIs were uploaded to the notes section of our EHR (Epic, Verona, WI) for review (see [Multimedia Appendix 1](#) for the email notification that was sent to physicians by research study staff). To identify individuals for the 2 control groups who were comparable to those in the intervention arm, each patient who completed AEGIS was matched (age ± 3 years, sex, race/ethnicity, clinic) with up to 4 patients in the pre-intervention period (October 6, 2015-April 6, 2017) and 1 patient in the intervention period who did not complete AEGIS. Of note, the pre-intervention control group was formed as it was not subject to contamination bias, unlike for post-intervention controls (ie, after physicians reviewed AEGIS reports for those in the intervention arm, they might have been more apt to take and document more thorough HPIs for their control patients). Moreover, age, sex, race/ethnicity, and clinic were selected as matching variables as they were readily available for automated extraction from the EHR and we hypothesized at the outset that they may have correlated with our outcomes.

Figure 1. Sample Automated Evaluation of Gastrointestinal Symptoms (AEGIS) history of present illness (HPI) [3], which was composed entirely by the AEGIS software and based on the patient's responses to questions about their gastrointestinal symptoms; then, the HPI is uploaded into the electronic health record where the physician can review it prior to seeing the patient as well as copy the HPI into their consult note and modify it as needed based on the subsequent clinical encounter. GERD: gastroesophageal reflux disease; NSAID: nonsteroidal anti-inflammatory drug.

HPI: [Patient] is a 34-year-old male who reports a history of Celiac disease and now presents with abdominal pain. The pain first started 8 months ago, and typically lasts for 2 hours at a time. Over the past week, the pain occurred once a day. He describes the pain as "burning" and "gnawing", says it is located in the epigastrium, and reports the pain has been "quite severe" and "quite a bit bothersome" in the past week. It does not radiate. It is associated with eating food. It typically occurs around 10-30 minutes after starting to eat. It usually comes on suddenly. It is not associated with bowel movements. The pain is somewhat relieved by reducing stress. The pain does not awaken him from sleep. He does not report early satiety. He does not report diabetes, gallstones, GERD, pancreatitis, or peptic ulcer. He does not take aspirin or NSAIDs.

He also reports diarrhea and bowel incontinence. The patient does not report dysphagia, heartburn, bloating, constipation, nausea, or vomiting.

He does not report blood in his bowel movements, black stools, vomiting blood, unintended weight loss, diminished appetite, or fevers. He has no history of abdominal surgeries. There is a family history of colorectal cancer.

Once recruitment ceased, 5 resident physicians reviewed patient charts and collected outcomes data using a REDCap data abstraction sheet [5]: (1) documentation of alarm symptom(s) in initial note (hematochezia, melena, hematemesis, unintentional weight loss, fevers); (2) documentation of family history of GI malignancy in initial note (colorectal, esophagus, gallbladder, liver, stomach, pancreas, or throat cancer); (3) charting time (time initial encounter closed minus the appointment time); (4) number of follow-up visits in a 6-month period; (5) number of lab, endoscopy, and imaging tests ordered in a 6-month period.

Statistical analyses were performed using Stata 13.1 (StataCorp LP, College Station, TX). A two-tailed $P < .05$ was considered significant in all analyses. For bivariate analyses, we used analysis of variance and chi-squared tests to compare continuous and categorical data, respectively. We also conducted multivariable regression analyses on our outcomes and included group assignment, patient age, sex, race/ethnicity, and clinic site as covariates when appropriate. Specifically, logistic regression analyses were performed on the AEGIS completion

and documentation of alarm symptoms and GI cancer family history outcomes. We used linear regression analysis to compare charting time between the intervention and control groups. Lastly, numbers of follow-up visits and tests ordered within 6 months of the initial visit were compared using zero-inflated negative binomial and negative binomial regression models, respectively.

Results

Of the 774 patients invited to complete AEGIS (Multimedia Appendix 2 shows demographics), 116 (15.0%) completed it before their clinic visit. Table 1 shows results from the regression analysis on completion of the app; older individuals and Asians were less likely to complete AEGIS. No significant associations were seen between app completion and the remaining racial/ethnic groups, sex, and clinic. Among those who completed AEGIS, the consultants' initial notes for 47 (40.5%) of the 116 patients contained at least a portion of the computer-generated report.

Table 1. Predictors of completing the Automated Evaluation of Gastrointestinal Symptoms (AEGIS) prior to the clinic visit (N=774).

Variable	Completed AEGIS (n=116)	OR ^a (95% CI) ^b
Age (years), mean (SD)	49.9 (16.1)	0.985 (0.972-0.998)
Sex, n (%)		
Male	45 (14.9)	reference
Female	71 (15.1)	0.99 (0.66-1.51)
Race/ethnicity, n (%)		
Non-Hispanic white	83 (17.9)	Reference
Non-Hispanic black	11 (11.0)	0.57 (0.29-1.13)
Latino	10 (13.7)	0.66 (0.32-1.36)
Non-Hispanic Asian	5 (7.3)	0.33 (0.13-0.85)
Other/unknown	7 (10.5)	0.48 (0.21-1.11)
Clinic, n (%)		
Resident/fellow GI ^c clinic	19 (17.4)	Reference
Physician A	37 (18.2)	1.08 (0.58-2.02)
Physician B	29 (12.7)	0.68 (0.36-1.30)
Physician C	31 (13.3)	0.74 (0.38-1.42)

^aOR: odds ratio.

^bThe logistic regression model adjusted for all covariates in the table.

^cGI: gastrointestinal.

Patients who completed AEGIS (n=116) were matched with 343 patients from the pre-intervention period and 102 from the intervention period who did not complete AEGIS. Table 2 lists

their demographics; the groups were largely similar in age, sex, race/ethnicity, clinic, and reason for consult.

Table 2. Demographics of those in the matched cohort analysis (N=561).

Variable	Control group: pre-AEGIS ^a period (n=343)	Control group: did not complete AEGIS (n=102)	Intervention group: completed AEGIS (n=116)	<i>P</i> ^b
Age (years), mean (SD)	51.4 (16.1)	53.7 (16.2)	49.9 (16.1)	.21
Sex, n (%)				.51
Male	136 (39.7)	34 (33.3)	45 (38.8)	
Female	207 (60.4)	68 (66.7)	71 (61.2)	
Race/ethnicity, n (%)				.25
Non-Hispanic white	264 (77.0)	86 (84.3)	83 (71.6)	
Non-Hispanic black	35 (10.2)	7 (6.9)	11 (9.5)	
Latino	23 (6.7)	2 (2.0)	10 (8.6)	
Non-Hispanic Asian	10 (2.9)	5 (4.9)	5 (4.3)	
Other/unknown	11 (3.2)	2 (2.0)	7 (6.0)	
Clinic, n (%)				.67
Resident/fellow GI ^c clinic	43 (12.5)	9 (8.8)	19 (16.4)	
Physician A	112 (32.7)	30 (29.4)	37 (31.9)	
Physician B	90 (26.2)	28 (27.5)	29 (25.0)	
Physician C	98 (28.6)	35 (34.3)	31 (26.7)	
Reason for consult, n (%)				
Abdominal pain	77 (22.5)	27 (26.5)	28 (24.1)	.69
Anemia evaluation	2 (0.6)	1 (1.0)	0 (0)	.60
Bloating	30 (8.8)	17 (16.7)	21 (18.1)	.008
Blood in stool	18 (5.3)	2 (2.0)	5 (4.3)	.37
Bowel incontinence	2 (0.6)	1 (1.0)	1 (0.9)	.90
Colorectal cancer screening	101 (29.5)	25 (24.5)	34 (29.3)	.61
Constipation	48 (14.0)	13 (12.8)	19 (16.4)	.73
Diarrhea	42 (12.2)	10 (9.8)	17 (14.7)	.55
Dysphagia	15 (4.4)	1 (1.0)	4 (3.5)	.27
Gastroesophageal reflux disease	46 (13.4)	15 (14.7)	33 (28.5)	.001
Inflammatory bowel disease	19 (5.5)	4 (3.9)	6 (5.2)	.81
Liver disease	2 (0.6)	1 (1.0)	1 (0.9)	.90
Nausea/vomiting	25 (7.3)	9 (8.8)	8 (6.9)	.84
Rectal pain	2 (0.6)	0 (0)	2 (1.7)	.29
Other	47 (13.7)	19 (18.6)	16 (13.8)	.45

^aAEGIS: Automated Evaluation of Gastrointestinal Symptoms.

^b*P* value from analysis of variance test (continuous data) or chi-squared test (categorical data).

^cGI: gastrointestinal.

In [Table 3](#), we present the physician-centric outcomes stratified by group. No differences were seen for documentation of alarm symptoms and GI cancer family history, EHR charting time, or numbers of follow-up visits and ordered tests.

Table 3. Physician-related outcomes according to study group (N=561).

Variable	Control group: pre-AEGIS ^a period (n=343; reference)	Control group: did not complete AEGIS (n=102)	Adjusted <i>P</i>	Intervention group: completed AEGIS (n=116)	Adjusted <i>P</i>
Documentation of an alarm symptom in initial note ^b , n (%)	61 (17.8)	22 (21.6)	.18	18 (15.5)	.39
Documentation of GI ^c cancer family history in initial note ^b , n (%)	64 (18.7)	20 (19.6)	.86	27 (23.3)	.28
Charting time, which is the time until initial EHR ^d chart encounter was closed ^e (hours), median (IQR)	3.1 (1.4-9.2)	3.3 (1.0-12.7)	.34	3.7 (1.1-10.0)	.58
Number of follow-up visits within the 6-month period ^f , median (IQR)	0 (0-1)	0 (0-0)	.22	0 (0-1)	.11
Number of tests ordered within the 6-month period ^g , median (IQR)	1 (1-3)	1 (0-2)	.21	1 (0-3)	.85

^aAEGIS: Automated Evaluation of Gastrointestinal Symptoms.

^bLogistic regression model adjusted for group assignment, patient age, sex, race/ethnicity, and clinic.

^cGI: gastrointestinal.

^dEHR: electronic health record.

^eLinear regression model adjusted for group assignment, patient age, sex, race/ethnicity, and clinic. Patients seen in the resident or fellow GI clinic (n=71) were excluded from this analysis as trainees first needed to complete their note before attendings could review or edit the note and close the encounter. Patients of Physicians A-C who were seen earlier or later than their originally scheduled appointment time (n=92) were also excluded from this analysis.

^fZero-inflated negative binomial regression model adjusted for group assignment, patient age, sex, race/ethnicity, and clinic.

^gNegative binomial regression model adjusted for group assignment, patient age, sex, race/ethnicity, and clinic.

Discussion

We discovered that uptake of AEGIS was low, as only 15% of patients accessed the online portal. Surprisingly, this rate was lower than that seen in our prior AEGIS trial (37%) focused on patient-centric outcomes [3]. This was even despite our use of email invitations (the original study used mailings), which we initially hypothesized would increase uptake as the email included a direct AEGIS hyperlink. Of note, research staff emailed invitations directly to patients; it is possible that sending invitations through the EHR patient portal may have enhanced uptake. Prior literature illustrates that patients are accepting of EHR portals [6-8] and they may be more willing to complete interventions sent through official health system platforms. Further research examining optimal methods for deploying digital interventions in EHR-integrated environments are needed.

While AEGIS was built to enhance patient-physician communication by systematically collecting salient components of the history, one-time use of the app did not increase documentation of alarm symptoms or family history of GI malignancy in the initial note. This suggests that physicians in our study may adequately screen for and document relevant red flags and family history. Alternatively, given our pragmatic design, clinicians may not have reviewed the AEGIS report or incorporated it into their note for some patients. We also noted that the app and its computer-generated HPIs did not impact

health care utilization or charting time. In short, we did not find that leveraging an online HPI-generation portal measurably improved physician-centric outcomes in this study.

Of note, we previously found that AEGIS collected more alarm features when compared to physicians [2]. The discordant results likely relate to the different study designs; in the prior observational study, AEGIS was completed by patients after their clinic visit, rather than before the visit as in our current study. It is possible that first consulting with the physician subsequently prompted patients, after further introspection, to report more alarm features through AEGIS than were discussed and documented by the physician in clinic. We also previously found that AEGIS creates higher quality HPIs versus those written by doctors [1]. However, based on our findings here and in a prior multicenter controlled trial focused on patient-centered outcomes [3], simply making the comprehensive AEGIS HPIs available for review in the EHR is insufficient for improving care. Further research is needed to determine how best to optimally implement and use these computer-generated data in clinical workflows in order to enhance outcomes.

A limitation of our study was that we could not fully assess whether and how closely physicians reviewed the AEGIS reports in the EHR. While 40.5% of physicians' notes for the intervention patients contained a portion of the AEGIS report, we do not know how rigorously they reviewed the report after

copying it into their notes. On the other hand, for the remaining patients, it is possible that clinicians thoroughly read the AEGIS report in the EHR but chose to not copy and paste it into their official consultant notes. Development of novel, effective methods for alerting and assessing how clinicians use newly uploaded, app-generated data in the EHR and that maximize its use at the point of care are urgently needed. Another limitation was that AEGIS was administered as a one-time intervention; longitudinal use of the app for tracking symptom severity could have impacted outcomes such as numbers of follow-up visits and ordered tests. Notably, longitudinal symptom monitoring via a portal decreased emergency room visits and improved survival among patients with metastatic cancer [9,10].

In summary, we found that uptake of AEGIS was low, as less than 1 in 6 patients completed it before their visit. Moreover, one-time use of a carefully developed and validated patient-provider portal did not improve documentation of key elements of the note nor reduce clinician work burden. This is disappointing as it is well known that the EHR has greatly increased physician charting time [11,12]; our goal has been to identify ways to reduce physician burden through clinically meaningful, EHR-enabled automation. Yet, even in taking care to maximize the benefits of the EHR to support physician-centric outcomes, we were unable to demonstrate a benefit. Given the near universal adoption of EHRs [13-15], further research examining how best to develop and implement digital interventions in EHR-integrated environments for improving both patient and physician outcomes is critical.

Acknowledgments

This study was funded by a Junior Faculty Development Grant from the American College of Gastroenterology awarded to CVA. CVA and BMRS are supported by a CTSI grant from the NIH/NCATS UL1TR001881-01. CVA is also supported by a loan repayment award from the NIH/NIDDK L30 DK106734. Support for the MyGiHealth portal that administers AEGIS was obtained from Ironwood Pharmaceuticals. The Cedars-Sinai Center for Outcomes Research and Education (CS-CORE) is supported by The Marc and Sheri Rapaport Fund for Digital Health Sciences & Precision Health.

Conflicts of Interest

BMRS has served on advisory boards and received grant support from Ironwood Pharmaceuticals. CVA has a stock option grant in My Total Health. The remaining authors do not have any relevant disclosures.

Multimedia Appendix 1

Supplementary Figure 1. Sample email sent to physicians the day before their clinic notifying them of the patients that completed AEGIS.

[PNG File , 171 KB - [formative_v5i5e23599_app1.png](#)]

Multimedia Appendix 2

Supplementary Table 1. Demographics of the cohort invited to complete AEGIS (N=774).

[DOCX File , 16 KB - [formative_v5i5e23599_app2.docx](#)]

References

1. Almario CV, Chey W, Kaung A, Whitman C, Fuller G, Reid M, et al. Computer-generated vs. physician-documented history of present illness (HPI): results of a blinded comparison. *Am J Gastroenterol* 2015 Jan;110(1):170-179 [FREE Full text] [doi: [10.1038/ajg.2014.356](#)] [Medline: [25461620](#)]
2. Almario CV, Chey WD, Iriana S, Dailey F, Robbins K, Patel AV, et al. Computer versus physician identification of gastrointestinal alarm features. *Int J Med Inform* 2015 Dec;84(12):1111-1117 [FREE Full text] [doi: [10.1016/j.ijmedinf.2015.07.006](#)] [Medline: [26254875](#)]
3. Almario CV, Chey WD, Khanna D, Mosadeghi S, Ahmed S, Afghani E, et al. Impact of National Institutes of Health Gastrointestinal PROMIS measures in clinical practice: results of a multicenter controlled trial. *Am J Gastroenterol* 2016 Nov;111(11):1546-1556 [FREE Full text] [doi: [10.1038/ajg.2016.305](#)] [Medline: [27481311](#)]
4. Ahmed S, Almario CV, Chey WD, Robbins LA, Chang B, Ahn J, et al. Electronic patient agenda forms: comparing agreement between the reason for specialty consultation reported by referring providers and patients. *Inform Health Soc Care* 2019 Jan;44(1):105-113 [FREE Full text] [doi: [10.1080/17538157.2018.1437041](#)] [Medline: [29509054](#)]
5. 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](#)]
6. Collins SA, Rozenblum R, Leung WY, Morrison CR, Stade DL, McNally K, et al. Acute care patient portals: a qualitative study of stakeholder perspectives on current practices. *J Am Med Inform Assoc* 2017 Apr 01;24(e1):e9-e17 [FREE Full text] [doi: [10.1093/jamia/ocw081](#)] [Medline: [27357830](#)]

7. Patel V, Barker W, Siminerio E. Individuals' access and use of their online medical record nationwide. The Office of the National Coordinator for Health Information Technology (ONC). 2014 Sep. URL: https://www.healthit.gov/sites/default/files/consumeraccessdatabrief_9_10_14.pdf [accessed 2021-04-18]
8. Goldzweig CL, Orshansky G, Paige NM, Towfigh AA, Haggstrom DA, Miake-Lye I, et al. Electronic patient portals: evidence on health outcomes, satisfaction, efficiency, and attitudes: a systematic review. *Ann Intern Med* 2013 Nov 19;159(10):677-687. [doi: [10.7326/0003-4819-159-10-201311190-00006](https://doi.org/10.7326/0003-4819-159-10-201311190-00006)] [Medline: [24247673](https://pubmed.ncbi.nlm.nih.gov/24247673/)]
9. Basch E, Deal AM, Kris MG, Scher HI, Hudis CA, Sabbatini P, et al. Symptom monitoring with patient-reported outcomes during routine cancer treatment: a randomized controlled trial. *J Clin Oncol* 2016 Feb 20;34(6):557-565 [FREE Full text] [doi: [10.1200/JCO.2015.63.0830](https://doi.org/10.1200/JCO.2015.63.0830)] [Medline: [26644527](https://pubmed.ncbi.nlm.nih.gov/26644527/)]
10. Basch E, Deal AM, Dueck AC, Scher HI, Kris MG, Hudis C, et al. Overall survival results of a trial assessing patient-reported outcomes for symptom monitoring during routine cancer treatment. *JAMA* 2017 Jul 11;318(2):197-198 [FREE Full text] [doi: [10.1001/jama.2017.7156](https://doi.org/10.1001/jama.2017.7156)] [Medline: [28586821](https://pubmed.ncbi.nlm.nih.gov/28586821/)]
11. Baumann LA, Baker J, Elshaug AG. The impact of electronic health record systems on clinical documentation times: A systematic review. *Health Policy* 2018 Aug;122(8):827-836. [doi: [10.1016/j.healthpol.2018.05.014](https://doi.org/10.1016/j.healthpol.2018.05.014)] [Medline: [29895467](https://pubmed.ncbi.nlm.nih.gov/29895467/)]
12. Sinsky C, Colligan L, Li L, Prgomet M, Reynolds S, Goeders L, et al. Allocation of physician time in ambulatory practice: a time and motion study in 4 specialties. *Ann Intern Med* 2016 Dec 06;165(11):753-760. [doi: [10.7326/M16-0961](https://doi.org/10.7326/M16-0961)] [Medline: [27595430](https://pubmed.ncbi.nlm.nih.gov/27595430/)]
13. Adler-Milstein J, Holmgren AJ, Kralovec P, Worzala C, Searcy T, Patel V. Electronic health record adoption in US hospitals: the emergence of a digital "advanced use" divide. *J Am Med Inform Assoc* 2017 Nov 01;24(6):1142-1148 [FREE Full text] [doi: [10.1093/jamia/ocx080](https://doi.org/10.1093/jamia/ocx080)] [Medline: [29016973](https://pubmed.ncbi.nlm.nih.gov/29016973/)]
14. Fraccaro P, Vigo M, Balatsoukas P, Buchan IE, Peek N, van der Veer SN. Patient portal adoption rates: a systematic literature review and meta-analysis. *Stud Health Technol Inform* 2017;245:79-83. [Medline: [29295056](https://pubmed.ncbi.nlm.nih.gov/29295056/)]
15. Myrick KL, Ogburn DF, Ward BW. Percentage of office-based physicians using any electronic health record (EHR)/electronic medical record (EMR) system and physicians that have a certified EHR/EMR system, by U.S. state: National Electronic Health Records Survey, 2017. National Center for Health Statistics. 2019. URL: https://www.cdc.gov/nchs/data/nehrs/2017_NEHRS_Web_Table_EHR_State.pdf [accessed 2021-04-18]

Abbreviations

AEGIS: Automated Evaluation of Gastrointestinal Symptoms
EHR: electronic health record
GI: gastrointestinal
HPI: history of present illness
OR: odds ratio

Edited by G Eysenbach; submitted 17.08.20; peer-reviewed by G Orshansky, J Chaplin; comments to author 02.11.20; revised version received 18.01.21; accepted 11.04.21; published 04.05.21.

Please cite as:

Hall NJ, Berry SK, Aguilar J, Brier E, Shah P, Cheng D, Herman J, Stein T, Spiegel BMR, Almario CV

Impact of an Online Gastrointestinal Symptom History Taker on Physician Documentation and Charting Time: Pragmatic Controlled Trial

JMIR Form Res 2021;5(5):e23599

URL: <https://formative.jmir.org/2021/5/e23599>

doi: [10.2196/23599](https://doi.org/10.2196/23599)

PMID: [33944789](https://pubmed.ncbi.nlm.nih.gov/33944789/)

©Natalie J Hall, Sameer K Berry, Jack Aguilar, Elizabeth Brier, Parth Shah, Derek Cheng, Jeremy Herman, Theodore Stein, Brennan M R Spiegel, Christopher V Almario. Originally published in JMIR Formative Research (<https://formative.jmir.org/>), 04.05.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Counselor Efficiency at Providing Feedback in a Technology-Based Behavioral Weight Loss Intervention: Longitudinal Analysis

Margaret C Fahey¹, MA; Robert C Klesges², PhD; Mehmet Kocak³, PhD; Leslie A Gladney², MS; Gerald W Talcott², PhD; Rebecca A Krukowski³, PhD

¹Psychology Department, The University of Memphis, Memphis, TN, United States

²School of Medicine, University of Virginia, Charlottesville, VA, United States

³Department of Preventive Medicine, College of Medicine, University of Tennessee Health Science Center, Memphis, TN, United States

Corresponding Author:

Rebecca A Krukowski, PhD

Department of Preventive Medicine

College of Medicine

University of Tennessee Health Science Center

66 N Pauline St

Memphis, TN, 38105

United States

Phone: 1 901 448 2426

Email: rkrukows@uthsc.edu

Abstract

Background: Feedback for participants' self-monitoring is a crucial and costly component of technology-based weight loss interventions. Detailed examination of interventionist time when reviewing and providing feedback for online self-monitoring data is lacking.

Objective: The aim of this study was to longitudinally examine the time counselors spent providing feedback on participant self-monitoring data (ie, diet, physical activity, weight) in a 12-month technology-based weight loss intervention. We hypothesized that counselors would compose feedback for participants more quickly over time.

Methods: The time the lay counselors (N=10) spent reviewing self-monitoring records and providing feedback to participants via email was longitudinally examined for all counselors across the three years of study implementation. Descriptives were observed for counselor feedback duration across counselors by 12 annual quarters (ie, 3-month periods). Differences in overall duration times by each consecutive annual quarter were analyzed using Wilcoxon-Mann-Whitney tests.

Results: There was a decrease in counselor feedback duration from the first to second quarter (mean 53 to 46 minutes; $P<.001$), and from the second to third (mean 46 to 30 minutes; $P<.001$). A trend suggested a decrease from the third to fourth quarter (mean 30 to 26 minutes; $P=.053$), but no changes were found in subsequent quarters. Consistent with the hypothesis, counselors may be increasing their efficiency in providing feedback; across 12 months, counselors spent less time reviewing participant self-monitoring and composing feedback (decreasing from mean 53 to 26 minutes).

Conclusions: Counselors used increasingly less time to review online self-monitoring data and compose feedback after the initial 9 months of study implementation. Results inform counselor costs for future technology-based behavioral weight loss interventions. For example, regardless of increasing counselor efficiency, 25-30 minutes per feedback message is a high cost for interventions. One possibility for reducing costs would be generating computer-automated feedback.

Trial Registration: ClinicalTrials.gov NCT02063178; <https://clinicaltrials.gov/ct2/show/NCT02063178>

(*JMIR Form Res* 2021;5(5):e23974) doi:[10.2196/23974](https://doi.org/10.2196/23974)

KEYWORDS

technology-based intervention; counselor communication; counselor feedback; counselor; weight loss; lifestyle; wellness

Introduction

Consistent weight and dietary self-monitoring are key elements for successful weight loss in both in-person [1] and technology-based programs [2], and using technology for self-monitoring (eg, apps, smart scales) can increase self-monitoring adherence [3-5]. Personalized self-monitoring feedback on the frequency of weight, dietary, and exercise monitoring; reinforcing comments about weight loss behaviors; and presentation of behavior change possibilities are core elements of behavioral weight loss interventions [6-8]. Feedback on self-monitoring data appears to be a crucial component of these interventions since it is associated with greater self-monitoring engagement as well as greater weight loss in interventions [9-12]. In recent years, technology-based communication (eg, email) has been increasingly used for counselor's feedback [9], particularly since participants are now able to self-monitor food intake, weight, and physical activity online using either researcher-developed or commercial websites/apps [13] rather than using paper and pencil diaries.

In studies that have evaluated the cost-effectiveness of behavioral weight loss interventions [14-25], interventionist compensation emerges in the available studies as one of the costliest components [21,24,25], including time for conducting the sessions and for providing feedback (ie, review of self-monitoring data, composing feedback). Interventionist costs, however, are often bundled in these analyses [21,24,25]; that is, combining time required for the sessions together with time required for providing feedback as well as other intervention tasks. The amount of time for sessions is often quite rigid (eg, 60-90 minutes for group sessions; 20-30 minutes for individual sessions), with standard outlines of material to cover, but little is known about the time associated with providing self-monitoring feedback.

One study examined costs of providing email feedback for online self-monitoring data, based on retrospective self-reported estimates of counselor time in a "typical week" and "after substantial implementation experience," but bundled all intervention costs together (ie, session preparation, conducting the group session, review of self-monitoring journals, periodic contact of participants who might miss sessions or who have questions, posting on the bulletin board, record keeping, technical work by the webmaster) [21,24]. To our knowledge, only one behavioral intervention specifically reported isolated counselor feedback time, using an average weekly estimate for the entire intervention, but it is not clear whether this information was collected contemporaneously or retrospectively [26]. Thus, previous information is limited to averages, and it is not clear that real-time data collection of each feedback message composed has been examined. Further, because counselor efficiency might increase over time, it will be important to examine the potential impact of implementation experience.

Thus, the purpose of this study was to longitudinally examine the time required to review self-monitoring data and compose feedback among newly trained counselors for participants engaged in a 12-month behavioral weight loss intervention. These data will be important since a detailed examination of interventionist time when reviewing and providing feedback for online self-monitoring data can inform the cost-effectiveness of technology-based programs (including the time for new counselors to "peak" in efficiency) and serve as a baseline for comparison if strategies are implemented for increasing efficiency. This study examined time spent on counselor self-monitoring feedback across the three years that the weight loss intervention was implemented. We hypothesize that counselors will deliver feedback on participant diet, physical activity, and weight self-monitoring more quickly over time.

Methods

Participants

Individuals receiving self-monitoring feedback in the behavioral weight loss intervention were active duty military personnel stationed at Lackland Air Force Base in San Antonio, Texas, enrolled in the Fit Blue weight loss study (2014-2017) [27,28]. Recruitment used posters and bulletins, on-base presentations, advertisements, and word of mouth. Those interested were phone screened by study staff to assess eligibility (ie, >1 year left on base, >18 years of age, BMI >25.0 kg/m², health care provider clearance, computer/email access). At baseline, 248 participants were randomized to either a counselor-initiated or self-paced 12-month intervention condition. Conditions varied in intensity and self-initiation required but were similar in intervention goals.

Self-monitoring Components

Overview

Participants were asked to self-monitor food intake, physical activity, and weight daily. To record food intake and physical activity, participants used the Lose It! app or website and permitted their counselor to access this information. To monitor weight, participants used the BodyTrace e-scale provided to them at baseline, which uploaded to a secure personalized website. In the counselor-initiated condition, counselors provided feedback on dietary, physical activity, and weight self-monitoring at the same frequency as telephone sessions (ie, weekly for 4 months, then biweekly for 4 months, then monthly for 4 months) via email (28 total). The self-paced condition was provided feedback via email when requested (up to 28), although in practice this feedback was rarely requested in the self-paced condition [28]. The time estimates presented in the analyses are based on 2670 emails (Table 1). The protocol was approved by the Institutional Review Board of the Wilford Hall Ambulatory Surgical Center in San Antonio, Texas, and acknowledged by the Institutional Review Board at the University of Tennessee Health Science Center.

Table 1. Change in feedback durations across counselors.

Quarters from first review	Feedback duration across counselors (minutes)				
	Number of feedback emails	Quartile 1	Median	Mean (SD)	Quartile 3
1	335	30	30	53 (54)	60
2	380	20	30	46 (47)	45
3	324	20	30	30 (25)	30
4	286	15	25	26 (19)	30
5	317	20	30	26 (12)	30
6	309	17	25	27 (15)	30
7	249	20	30	28 (14)	30
8	231	25	30	27 (7)	30
≥9	239	25	30	28 (12)	30

Counselor Characteristics and Training

Counselors (N=10) held bachelor's or master's degrees (ie, social work, counseling/psychology, child and family development, nursing, justice administration); however, they were considered lay interventionists since no prior counseling or research experience was required. Counselors were hired based on their interest in providing behavioral interventions and in research, and they were either retired from the military or familiar with military culture. All counselors were new to providing self-monitoring feedback in a weight loss intervention. Counselors attended a week-long training on the study protocol, behavioral weight management principles, feedback, and motivational interviewing. Counselors were taught to construct emails as "feedback sandwiches," with reinforcement of behaviors sandwiched around identification of potential areas of behavior change, consistent with guidance from other studies [29]. Counselors submitted practice emails that were discussed among counselors led by the principal investigator (RAK). Two counselors joined the team midway through the study timeline and were similarly trained. All counselors participated in ongoing biweekly 1-hour supervision and 1-hour motivational interviewing training and received roughly quarterly retraining on providing self-monitoring feedback to maintain and improve this skill.

Documentation of Self-monitoring Feedback Duration

Counselors contemporaneously logged the time it took them to review online self-monitoring data and construct each personalized email in the electronic study database.

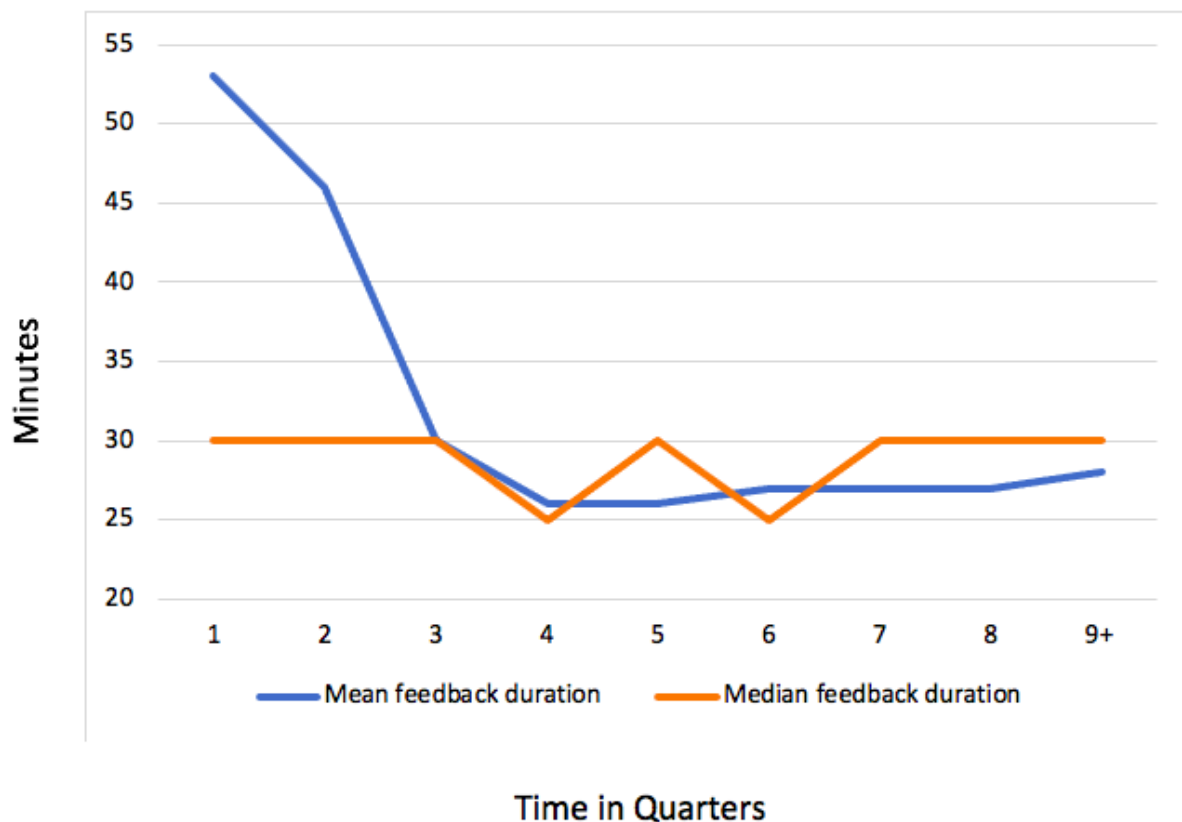
Data Analysis

All feedback duration times were included in analyses regardless of condition. Descriptives (ie, median, mean, SD, first quartile, third quartile, quartile range) were observed for feedback duration across counselors by 12 annual quarters (ie, 3-month periods). For counselors who joined the team later (N=2), the first quarter they provided feedback was compiled with the first quarter of feedback from the original counselors. Differences in overall duration times by each consecutive annual quarter were analyzed using the Wilcoxon-Mann-Whitney test.

Results

Across all counselors, there was a significant decrease in overall duration to review self-monitoring data and compose feedback messages from the first to second quarter ($P<.001$; IQR 30-60 minutes versus 20-45 minutes; mean 53, SD 56 versus mean 46, SD 47; Table 1). There was a significant decrease in overall duration from the second to third quarter ($P<.001$; IQR 20-45 minutes versus 20-30 minutes; mean 46, SD 47 versus mean 30, SD 25). A nonsignificant trend suggested a decrease in duration from the third to fourth quarter ($P=.053$; IQR 20-30 minutes versus 15-30 minutes; mean 30, SD 25 versus mean 26, SD 19). There was no significant change in duration between later quarters (Table 1). Median time ranged from 25 to 30 minutes (Table 1). Median and mean feedback durations across all quarters are presented in Figure 1.

Figure 1. Feedback duration across intervention quarters. Counselor efficiency increases over time.



Discussion

Principal Findings

Counselors needed increasingly less time to review online self-monitoring data and compose personalized feedback to participants over the first 9 months the behavioral weight loss intervention was implemented, with a nonsignificant trend suggesting increased efficiency for the next 3 months as well. Thus, after 12 months, the mean amount of time spent in reviewing self-monitoring data and composing feedback decreased from 53 to 26 minutes. When examining median times, there was less variation across all quarters (ie, 25-30 minutes); however, a decrease in IQR in the first three quarters was notable. This narrowing time range indicated that counselors composed these feedback messages more consistently near the median (ie, 30 minutes) over time.

Although the standard deviation decreased, it remained high, likely due to variability across individual counselors and participant characteristics. Some participants might have logged similar data to previous weeks, or only logged one day, requiring shorter feedback messages. In response to these self-monitoring situations, counselors might write, “You continued to meet your calorie and fat goals and continue to make regular choices of fruits, vegetables, whole grains, and low- or no-calorie beverages” or “You met your calorie and fat goals on the one day that you were able to log. What were the barriers for logging on the other days?” Additionally, perhaps some counselors were quicker at crafting feedback than others.

Mean time was 32 minutes per feedback message, which is higher than the 10-15 minutes found by Hunter et al [26], during which counselors similarly provided feedback on food intake, exercise, and weight. However, only time spent providing feedback was reported, which—unlike the current study—did not include time reviewing self-monitoring data [26]. Further, it is unclear if the counselors in this previous research logged feedback duration contemporaneously or retrospectively [26].

The decrease in time for feedback messages might be influenced by multiple factors. Perhaps counselors became more efficient at reviewing self-monitoring data and constructing feedback with experience and additional training. Another possibility is that counselors provided less detailed feedback and became “sloppier” over time. However, this is less likely given that the overall intervention results indicated participants experienced significant weight loss at 4-month and 12-month outcomes [28], and periodic retraining on self-monitoring feedback was conducted to increase the likelihood of maintaining good quality feedback. Further, participants might have needed increasingly less feedback about their behaviors over the 12 months that they participated in the intervention. However, our study compared time periods specific to counselor experience, which included the three years that the 12-month intervention was implemented. Thus, individual participants cycled throughout the three years that our results were analyzed and overlapped with participants at other stages of intervention. Nonetheless, future research should rate the quality of feedback over time alongside the time required to compose it.

Regardless of increasing counselor efficiency, 25-30 minutes per feedback message is a high cost for interventions, even cost-effective interventions such as this one [24]. In order to improve dissemination of behavioral weight management programs to all who are eligible and interested, particularly in settings (eg, primary care) without individuals who may have the time, training, and supervising experience to provide self-monitoring feedback, it may be beneficial to develop strategies for decreasing the amount of time required for crafting each message. Some possibilities for reducing the costs of individualized counselor feedback would be generating computer-automated feedback or facilitating peer-group interaction to help promote self-monitoring behaviors [30]. Although computer-automated feedback was comparable to counselor feedback in the short term in previous research, automated feedback was less effective for long-term weight loss [12]. Given limited research on computer-generated feedback [12,31,32] and the common use of counselor-generated feedback [9-12], it is clearly still important to understand in detail the time costs associated with counselor feedback. These details can inform future program budgets, especially since counselor compensation is one of the largest costs [21,24,25]. However, since computer-tailored feedback might be more regularly incorporated into future interventions, a better understanding of the efficiency of counselors in crafting

feedback may be beneficial to compare these different modalities.

There are several limitations to consider. Counselors self-reported time spent reviewing self-monitoring data and constructing feedback, which may be less accurate compared to objective measurements. However, this level of detail is much greater than previous studies. In addition, findings are based on lay counselors and may differ from other weight loss professionals (eg, dietitians). Finally, future studies might examine a consortium of behavioral weight management studies in order to have a larger sample size of interventionists in analyses that examine these questions.

Conclusion

Current findings suggest that counselors, across the initial 9 months of a behavioral weight loss intervention, become quicker at reviewing participant self-monitoring data and composing individualized feedback. Although there was individual variability, findings indicate that after 9-12 months of experience, counselors composed self-monitoring feedback more consistently in about 30 minutes. Despite indications of increased counselor efficiency, time per feedback message was only reduced to 25-30 minutes. Weight loss programs might consider testing computer-automated feedback with human tailoring to reduce counselor time.

Acknowledgments

The study was funded by the National Institute of Diabetes and Digestive and Kidney Diseases (RO1 DK097158) of the National Institutes of Health, with the title of "Dissemination of the Look Ahead Weight Management Treatment in the Military," and RCK and RAK as Principal Investigators.

The research represents a Collaborative Research and Development Agreement with the United States Air Force (CRADA #13-168-SG-C13001). The trial is registered on ClinicalTrials.gov (NCT 02063178). The opinions expressed in this document are solely those of the authors and do not represent an endorsement by or the views of the United States Air Force, the Department of Defense, or the United States Government. Finally, we would like to thank the participants and the research team for their dedication to the research.

Authors' Contributions

MCF wrote the first draft of the manuscript and conducted a literature review for the study. RCK was involved in study conceptualization, design, and implementation. MK ran statistical analyses for the study. LAG was involved in study conceptualization, design, and implementation. GWT was involved in study conceptualization, design, and implementation. RAK was involved in study conceptualization, study design, study implementation, and mentorship of student author. All authors reviewed and revised the manuscript draft.

Conflicts of Interest

None declared.

References

1. Boutelle K, Kirschenbaum D. Further support for consistent self-monitoring as a vital component of successful weight control. *Obes Res* 1998 May;6(3):219-224 [FREE Full text] [doi: [10.1002/j.1550-8528.1998.tb00340.x](https://doi.org/10.1002/j.1550-8528.1998.tb00340.x)] [Medline: [9618126](https://pubmed.ncbi.nlm.nih.gov/9618126/)]
2. Burke LE, Conroy MB, Sereika SM, Elci OU, Styn MA, Acharya SD, et al. The effect of electronic self-monitoring on weight loss and dietary intake: a randomized behavioral weight loss trial. *Obesity (Silver Spring)* 2011 Feb;19(2):338-344 [FREE Full text] [doi: [10.1038/oby.2010.208](https://doi.org/10.1038/oby.2010.208)] [Medline: [20847736](https://pubmed.ncbi.nlm.nih.gov/20847736/)]
3. Allen JK, Stephens J, Dennison Himmelfarb CR, Stewart KJ, Hauck S. Randomized controlled pilot study testing use of smartphone technology for obesity treatment. *J Obes* 2013;2013:151597 [FREE Full text] [doi: [10.1155/2013/151597](https://doi.org/10.1155/2013/151597)] [Medline: [24392223](https://pubmed.ncbi.nlm.nih.gov/24392223/)]

4. West JH, Belvedere LM, Andreasen R, Frandsen C, Hall PC, Crookston BT. Controlling Your "App"etite: How Diet and Nutrition-Related Mobile Apps Lead to Behavior Change. *JMIR Mhealth Uhealth* 2017 Jul 10;5(7):e95 [[FREE Full text](#)] [doi: [10.2196/mhealth.7410](https://doi.org/10.2196/mhealth.7410)] [Medline: [28694241](#)]
5. Laing BY, Mangione CM, Tseng C, Leng M, Vaisberg E, Mahida M, et al. Effectiveness of a smartphone application for weight loss compared with usual care in overweight primary care patients: a randomized, controlled trial. *Ann Intern Med* 2014 Nov 18;161(10 Suppl):S5-12 [[FREE Full text](#)] [doi: [10.7326/M13-3005](https://doi.org/10.7326/M13-3005)] [Medline: [25402403](#)]
6. Steinberg DM, Levine EL, Askew S, Foley P, Bennett GG. Daily text messaging for weight control among racial and ethnic minority women: randomized controlled pilot study. *J Med Internet Res* 2013 Nov 18;15(11):e244 [[FREE Full text](#)] [doi: [10.2196/jmir.2844](https://doi.org/10.2196/jmir.2844)] [Medline: [24246427](#)]
7. Diabetes Prevention Program (DPP) Research Group. The Diabetes Prevention Program (DPP): description of lifestyle intervention. *Diabetes Care* 2002 Dec 01;25(12):2165-2171 [[FREE Full text](#)] [doi: [10.2337/diacare.25.12.2165](https://doi.org/10.2337/diacare.25.12.2165)] [Medline: [12453955](#)]
8. Look AHEAD Research Group, Wadden TA, West DS, Delahanty L, Jakicic J, Rejeski J, et al. The Look AHEAD study: a description of the lifestyle intervention and the evidence supporting it. *Obesity (Silver Spring)* 2006 May;14(5):737-752 [[FREE Full text](#)] [doi: [10.1038/oby.2006.84](https://doi.org/10.1038/oby.2006.84)] [Medline: [16855180](#)]
9. Khaylis A, Yiaslas T, Bergstrom J, Gore-Felton C. A review of efficacious technology-based weight-loss interventions: five key components. *Telemed J E Health* 2010 Nov;16(9):931-938 [[FREE Full text](#)] [doi: [10.1089/tmj.2010.0065](https://doi.org/10.1089/tmj.2010.0065)] [Medline: [21091286](#)]
10. Krukowski RA, Harvey-Berino J, Ashikaga T, Thomas CS, Micco N. Internet-based weight control: the relationship between web features and weight loss. *Telemed J E Health* 2008 Oct;14(8):775-782 [[FREE Full text](#)] [doi: [10.1089/tmj.2007.0132](https://doi.org/10.1089/tmj.2007.0132)] [Medline: [18954247](#)]
11. Tate DF, Wing RR, Winnett RA. Using Internet technology to deliver a behavioral weight loss program. *JAMA* 2001 Mar 07;285(9):1172-1177. [doi: [10.1001/jama.285.9.1172](https://doi.org/10.1001/jama.285.9.1172)] [Medline: [11231746](#)]
12. Tate DF, Jackvony EH, Wing RR. A randomized trial comparing human e-mail counseling, computer-automated tailored counseling, and no counseling in an Internet weight loss program. *Arch Intern Med* 2006 Aug 14;166(15):1620-1625. [doi: [10.1001/archinte.166.15.1620](https://doi.org/10.1001/archinte.166.15.1620)] [Medline: [16908795](#)]
13. Spring B, Duncan JM, Janke EA, Kozak AT, McFadden HG, DeMott A, et al. Integrating technology into standard weight loss treatment: a randomized controlled trial. *JAMA Intern Med* 2013 Jan 28;173(2):105-111 [[FREE Full text](#)] [doi: [10.1001/jamainternmed.2013.1221](https://doi.org/10.1001/jamainternmed.2013.1221)] [Medline: [23229890](#)]
14. Smith KJ, Kuo S, Zgibor JC, McTigue KM, Hess R, Bhargava T, et al. Cost effectiveness of an internet-delivered lifestyle intervention in primary care patients with high cardiovascular risk. *Prev Med* 2016 Jun;87:103-109. [doi: [10.1016/j.ypmed.2016.02.036](https://doi.org/10.1016/j.ypmed.2016.02.036)] [Medline: [26921656](#)]
15. Smith KJ, Hsu HE, Roberts MS, Kramer MK, Orchard TJ, Piatt GA, et al. Cost-effectiveness analysis of efforts to reduce risk of type 2 diabetes and cardiovascular disease in southwestern Pennsylvania, 2005-2007. *Prev Chronic Dis* 2010 Sep;7(5):A109 [[FREE Full text](#)] [Medline: [20712936](#)]
16. Hollenbeak CS, Weinstock RS, Cibula D, Delahanty LM, Trief PM. Cost-effectiveness of SHINE: A Telephone Translation of the Diabetes Prevention Program. *Health Serv Insights* 2016;9:21-28 [[FREE Full text](#)] [doi: [10.4137/HSI.S39084](https://doi.org/10.4137/HSI.S39084)] [Medline: [27429556](#)]
17. Herman WH, Hoerger TJ, Brandle M, Hicks K, Sorensen S, Zhang P, Diabetes Prevention Program Research Group. The cost-effectiveness of lifestyle modification or metformin in preventing type 2 diabetes in adults with impaired glucose tolerance. *Ann Intern Med* 2005 Mar 01;142(5):323-332 [[FREE Full text](#)] [doi: [10.7326/0003-4819-142-5-200503010-00007](https://doi.org/10.7326/0003-4819-142-5-200503010-00007)] [Medline: [15738451](#)]
18. Diabetes Prevention Program Research Group. Within-trial cost-effectiveness of lifestyle intervention or metformin for the primary prevention of type 2 diabetes. *Diabetes Care* 2003 Sep 22;26(9):2518-2523 [[FREE Full text](#)] [doi: [10.2337/diacare.26.9.2518](https://doi.org/10.2337/diacare.26.9.2518)] [Medline: [12941712](#)]
19. Rushing J, Wing R, Wadden TA, Knowler WC, Lawlor M, Evans M, Look AHEAD Research Group. Cost of intervention delivery in a lifestyle weight loss trial in type 2 diabetes: results from the Look AHEAD clinical trial. *Obes Sci Pract* 2017 Mar 24;3(1):15-24 [[FREE Full text](#)] [doi: [10.1002/osp4.92](https://doi.org/10.1002/osp4.92)] [Medline: [28392928](#)]
20. Krukowski RA, Pope RA, Love S, Lensing S, Felix HC, Prewitt TE, et al. Examination of costs for a lay health educator-delivered translation of the Diabetes Prevention Program in senior centers. *Prev Med* 2013 Oct;57(4):400-402 [[FREE Full text](#)] [doi: [10.1016/j.ypmed.2013.06.027](https://doi.org/10.1016/j.ypmed.2013.06.027)] [Medline: [23831492](#)]
21. Krukowski RA, Tilford JM, Harvey-Berino J, West DS. Comparing behavioral weight loss modalities: incremental cost-effectiveness of an internet-based versus an in-person condition. *Obesity (Silver Spring)* 2011 Aug;19(8):1629-1635 [[FREE Full text](#)] [doi: [10.1038/oby.2010.341](https://doi.org/10.1038/oby.2010.341)] [Medline: [21253001](#)]
22. McGloin AF, Eslami S. Digital and social media opportunities for dietary behaviour change. *Proc Nutr Soc* 2014 Oct 16;74(2):139-148. [doi: [10.1017/s0029665114001505](https://doi.org/10.1017/s0029665114001505)]
23. Okorodudu DE, Bosworth HB, Corsino L. Innovative interventions to promote behavioral change in overweight or obese individuals: A review of the literature. *Ann Med* 2015 May 10;47(3):179-185 [[FREE Full text](#)] [doi: [10.3109/07853890.2014.931102](https://doi.org/10.3109/07853890.2014.931102)] [Medline: [25011006](#)]

24. Manz KC, Waters TM, Clifton HE, Kocak M, Klesges RC, Talcott GW, et al. Cost-Effectiveness of a Weight Loss Intervention: An Adaptation of the Look AHEAD Lifestyle Intervention in the US Military. *Obesity (Silver Spring)* 2020 Jan 26;28(1):89-96 [FREE Full text] [doi: [10.1002/oby.22681](https://doi.org/10.1002/oby.22681)] [Medline: [31773873](https://pubmed.ncbi.nlm.nih.gov/31773873/)]
25. Ritzwoller DP, Glasgow RE, Sukhanova AY, Bennett GG, Warner ET, Greaney ML, Be Fit Be Well study investigators. Economic analyses of the Be Fit Be Well program: a weight loss program for community health centers. *J Gen Intern Med* 2013 Dec 4;28(12):1581-1588 [FREE Full text] [doi: [10.1007/s11606-013-2492-3](https://doi.org/10.1007/s11606-013-2492-3)] [Medline: [23733374](https://pubmed.ncbi.nlm.nih.gov/23733374/)]
26. Hunter CM, Peterson AL, Alvarez LM, Poston WC, Brundige AR, Haddock CK, et al. Weight management using the internet a randomized controlled trial. *Am J Prev Med* 2008 Feb;34(2):119-126. [doi: [10.1016/j.amepre.2007.09.026](https://doi.org/10.1016/j.amepre.2007.09.026)] [Medline: [18201641](https://pubmed.ncbi.nlm.nih.gov/18201641/)]
27. Krukowski RA, Hare ME, Talcott GW, Johnson KC, Richey PA, Kocak M, et al. Dissemination of the Look AHEAD intensive lifestyle intervention in the United States Air Force: study rationale, design and methods. *Contemp Clin Trials* 2015 Jan;40:232-239 [FREE Full text] [doi: [10.1016/j.cct.2014.12.014](https://doi.org/10.1016/j.cct.2014.12.014)] [Medline: [25545025](https://pubmed.ncbi.nlm.nih.gov/25545025/)]
28. Krukowski RA, Hare ME, Talcott GW, Gladney LA, Johnson KC, Richey PA, et al. Dissemination of the Look AHEAD Intensive Lifestyle Intervention in the United States Military: A Randomized Controlled Trial. *Obesity (Silver Spring)* 2018 Oct 11;26(10):1558-1565 [FREE Full text] [doi: [10.1002/oby.22293](https://doi.org/10.1002/oby.22293)] [Medline: [30277030](https://pubmed.ncbi.nlm.nih.gov/30277030/)]
29. Semler L. Coaching tips for offering feedback on participant self-monitoring records. Diabetes Prevention Support Center. 2016 Feb 16. URL: <https://www.diabetesprevention.pitt.edu/coaching-tips-for-offering-feedback-on-participant-self-monitoring-records/> [accessed 2020-12-19]
30. Venditti EM, Wylie-Rosett J, Delahanty LM, Mele L, Hoskin MA, Edelstein SL, Diabetes Prevention Program Research Group. Short and long-term lifestyle coaching approaches used to address diverse participant barriers to weight loss and physical activity adherence. *Int J Behav Nutr Phys Act* 2014 Feb 12;11(1):16 [FREE Full text] [doi: [10.1186/1479-5868-11-16](https://doi.org/10.1186/1479-5868-11-16)] [Medline: [24521153](https://pubmed.ncbi.nlm.nih.gov/24521153/)]
31. West D, Krukowski R, Larsen C. Treatment of obesity in community settings. In: Wadden TA, Bray GA, editors. *Handbook of Obesity Treatment (Second Edition)*. New York, NY: The Guilford Press; 2018.
32. Appel LJ, Clark JM, Yeh H, Wang N, Coughlin JW, Daumit G, et al. Comparative Effectiveness of Weight-Loss Interventions in Clinical Practice. *N Engl J Med* 2011 Nov 24;365(21):1959-1968. [doi: [10.1056/nejmoa1108660](https://doi.org/10.1056/nejmoa1108660)]

Edited by G Eysenbach; submitted 05.09.20; peer-reviewed by J Hong, HL Tam; comments to author 18.11.20; revised version received 02.02.21; accepted 13.04.21; published 05.05.21.

Please cite as:

Fahey MC, Klesges RC, Kocak M, Gladney LA, Talcott GW, Krukowski RA

Counselor Efficiency at Providing Feedback in a Technology-Based Behavioral Weight Loss Intervention: Longitudinal Analysis
JMIR Form Res 2021;5(5):e23974

URL: <https://formative.jmir.org/2021/5/e23974>

doi: [10.2196/23974](https://doi.org/10.2196/23974)

PMID: [33949954](https://pubmed.ncbi.nlm.nih.gov/33949954/)

©Margaret C Fahey, Robert C Klesges, Mehmet Kocak, Leslie A Gladney, Gerald W Talcott, Rebecca A Krukowski. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 05.05.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Factors Influencing Telehealth Implementation and Use in Frontier Critical Access Hospitals: Qualitative Study

Saira Naim Haque¹, PhD, MHSA; Sydney DeStefano¹, MPH; Alison Banger¹, MPH; Regina Rutledge¹, PhD; Melissa Romaire¹, PhD

RTI International, Research Triangle Park, NC, United States

Corresponding Author:

Melissa Romaire, PhD

RTI International

3040 E Cornwallis Street

Research Triangle Park, NC, 27709

United States

Phone: 1 919 541 6894

Email: mromaire@rti.org

Abstract

Background: Telehealth has potential to help individuals in rural areas overcome geographical barriers and to improve access to care. The factors that influence the implementation and use of telehealth in critical access hospitals are in need of exploration.

Objective: The aim of this study is to understand the factors that influenced telehealth uptake and use in a set of frontier critical access hospitals in the United States.

Methods: This work was conducted as part of a larger evaluation of a Centers for Medicare & Medicaid Services–funded demonstration program to expand cost-based reimbursement for services for Medicare beneficiaries for frontier critical access hospitals. Our sample was 8 critical access hospitals in Montana, Nevada, and North Dakota that implemented the telehealth aspect of that demonstration. We reviewed applications and progress reports for the demonstration program and conducted in-person site visits. We used a semistructured discussion guide to facilitate conversations with clinical, administrative, and information technology staff. Using NVivo software (QSR International), we coded the notes from the interviews and then analyzed the themes.

Results: Several factors influenced the implementation and use of telehealth in critical access hospitals, including making changes to workflow and infrastructure as well as practitioner acceptance and availability. Participants also cited technical assistance and support for implementation as supportive factors.

Conclusions: Frontier critical access hospitals may adopt telehealth to overcome challenges such as distance from specialty practitioners and workforce challenges. Telehealth can be used for provider-to-patient and provider-to-provider interactions to improve access to care, remove barriers, and improve quality. However, the ability of telehealth to improve outcomes is limited by factors such as workflow and infrastructure changes, practitioner acceptance and availability, and financing.

(*JMIR Form Res* 2021;5(5):e24118) doi:[10.2196/24118](https://doi.org/10.2196/24118)

KEYWORDS

telehealth; rural health; health IT; telemedicine; virtual care

Introduction

Telehealth has potential to help individuals in rural areas overcome geographical barriers and to improve access to care. Barriers to access include travel time, transportation, cost, and other logistical considerations, which telehealth can help address [1]. Access to care can be enhanced through telehealth to connect practitioners and patients who are not co-located, which

enables patients to receive care that is not otherwise available locally.

Improvements in outcomes through telehealth largely occur through improved management of chronic disease and improved follow-up [2,3]. Many rural areas are not adequately served by practitioners; therefore, routine follow-up can involve extensive travel, resulting in delays in receiving care, missed appointments, and other issues [4,5]. Telehealth can promote improved adherence to care plans and can contribute to

improved outcomes and reduced costs through averted hospital stays and emergency department visits [3,6].

Although telehealth use has become more prevalent, prior to the COVID-19 pandemic, it was not widespread; thus, its benefits have not been fully realized [7-9]. For telehealth to show potential benefits in the long term beyond the public health emergency, its services must fit into changes in the health care landscape, including the advent of alternative care delivery and payment models, greater integration of physical and behavioral health services, and efforts to address workforce challenges [10,11].

To improve uptake of telehealth in communities where telehealth may be most needed and where benefits to patient access to health care may be significant, policy makers have turned to critical access hospitals (CAHs) because CAHs are charged with providing health care in rural communities that would otherwise not have basic inpatient, outpatient, and emergency care services. Recognizing the central role that CAHs play in integrating and providing access to health care in rural communities, the Centers for Medicare & Medicaid Services (CMS) implemented the 3-year Frontier Community Health Integration Project (FCHIP) Demonstration. FCHIP addressed four service areas—telehealth, skilled nursing facility care, home health services, and ambulance services—in 10 CAHs in Montana, Nevada, and North Dakota [2]. These frontier areas are characterized as having fewer than 10 people per square mile and are predominantly located in the western part of the United States.

This article reports on the qualitative findings for the first year of the telehealth component of the demonstration. We sought to identify and describe barriers and facilitators influencing telehealth adoption for CAHs. These findings can inform other telehealth efforts in rural areas.

Methods

This study was conducted as part of an evaluation of the FCHIP Demonstration project, which includes telehealth as a component. The overall study received approval from the Research Triangle Institute Institutional Review Board. Primary and secondary data on hospital landscape and FCHIP implementation activities were collected through in-person site visits and demonstration-related document review for each of the participating FCHIP CAHs. Site visit data collection included key informant interviews with a variety of stakeholders at each CAH.

Site Visits

Researchers with expertise in rural health and health informatics conducted on-site interviews with each CAH during June and July of 2017. We identified key stakeholders at each of the sites and interviewed this convenience sample, which included hospital leadership (eg, hospital administrators, medical directors), practitioners, and administrative support staff. We developed discussion guides for each category of informant. Discussion topics included any changes made (eg, staffing, infrastructure), implementation progress, perceived impacts, and any facilitators and barriers to implementation. We

conducted interviews with 36 individuals at 8 CAHs participating in the telehealth component of FCHIP. Two FCHIP-participating CAHs focused demonstration activities on service areas other than telehealth. We were not able to secure interviews with patients or caregivers.

Document Review

The evaluation team collected information about each CAH from documents that the CAH completed as part of the evaluation, including applications, progress plans, and facility overview/quarterly reports. These reports were submitted to CMS, which shared them with the evaluation team. These documents were used to triangulate with the interview findings.

Data Analysis

This work is part of a larger evaluation to understand the impact of the FCHIP Demonstration on participating CAHs. The evaluation focused on how the FCHIP Demonstration can improve access to care, integration, and delivery. We synthesized the qualitative information from site visits and the document review. We developed a codebook based on the literature and the overall goals of the evaluation. Items in the codebook included workflow, practitioner acceptance, infrastructure, and financing. We then coded data using NVivo (QSR International), a qualitative software analysis package. Next, we developed key themes using discussion topics from the interview protocols, the document reviews, and overall evaluation goals. Finally, we generated reports by each theme for review and analysis.

Results

CAHs expanded their capabilities to provide different types of services via telehealth with the hope that these new services would allow more community members to remain in the local community for care. Factors that influenced telehealth implementation and use include workflow, practitioner availability and acceptance, infrastructure and cost, and sustainability.

Workflow

Workflow changes included adding telehealth to staff responsibilities, education, training, and outreach. To implement or expand telehealth services, CAHs added new responsibilities to existing staff members' workloads; none of the CAHs reported hiring new staff for telehealth. New responsibilities included setting up equipment, managing referrals, and coordinating between CAHs and distant providers. Staff who assumed these new duties were nurses, administrative support staff, ward clerks, and medical assistants.

The addition of responsibilities was predicated on training. For example, staff received training on how to use equipment, document telehealth encounters for reimbursement, and incorporate telehealth encounters into the workflow. Providing outreach to surrounding communities was an additional responsibility for CAH staff. Outreach included explaining how telehealth worked, demonstrating the value of telehealth, and overcoming resistance to change in referring and using telehealth

services. Outreach to communities included different audiences, such as practitioners and the general community.

Practitioners in the frontier community who provide referrals to telehealth services, distant providers who deliver services, and staff who support them also received training and outreach to improve telehealth acceptance so they could change their workflows to accommodate telehealth. Interviewees had mixed feelings on training, indicating that training on using telehealth technologies and working with partners providing distant services ranged from straightforward to onerous. As one CAH administrator indicated:

Being able to help patients is fabulous. As far as our staff go, it is a learning curve. We are all new to what we need to be doing. ... Now we have streamlined ... Now we have a system that works.

Staff at one site described conducting outreach to encourage telehealth referrals. Rather than waiting for practitioners to refer patients for telehealth services, the site directed staff to take a more active role in telehealth decisions. The staff started asking patients if they would be interested in telehealth services, then called referring practitioners to see whether telehealth would be appropriate. Staff required some training on making calls, referring practitioners, and engaging in patient outreach.

Practitioner Availability and Acceptance

Sites participating in telehealth noted that administrative barriers presented the greatest challenges during the demonstration. Several sites noted that they were not able to provide services due to impediments to securing properly credentialed practitioners with availability. These sites faced challenges in obtaining the proper state credentials that give distant providers the necessary privileges to provide services for CAH patients. Furthermore, practitioners and administrators described efforts to use telehealth to provide specialty care, particularly dermatology and behavioral health services, but were told that distant providers were at capacity and were not accepting any more patients through telehealth.

Additionally, multiple sites noted that telehealth service provision was outside the usual scope of work for many CAH practitioners and staff and was often met with reluctance to change. Using telehealth requires a referral and additional paperwork; thus, some medical directors preferred their historical approach of calling a specialist colleague (by telephone) to obtain immediate answers. Participants indicated that having a practitioner champion was key to increasing volume because practitioners make telehealth referrals. As one participant indicated, "Providers are the big issue of adopting telehealth because they have to give the referrals." This challenge highlights the importance of practitioner outreach as a facilitator so that telehealth services are available to patients.

Infrastructure

Telehealth implementation required a few infrastructure and operational changes, which could be barriers to implementation and use. Infrastructure modification included technical and physical changes. Some sites modified their electronic health records (EHRs) to provide a consistent mechanism to document telehealth services. Other sites required equipment and physical

changes. For example, some sites purchased mobile telehealth equipment so that they did not have to share equipment between the CAH-owned clinic and the CAH itself.

Cost and Sustainability

Hospital staff reported that upfront costs for equipment combined with very low use would likely lead to a negative impact on financial performance. CAH staff members noted the high cost to procure the necessary telehealth equipment using their regular operating budget and expressed a desire for alternative sources of funding for this upfront cost. At hospitals that invested in infrastructure and equipment, it was noted that an increase in volume was needed to obtain sufficient reimbursements to offset the investments and improve financial performance.

Hospital staff reported varied experiences in billing for telehealth services. One hospital administrator mentioned that she was given a list of Current Procedural Terminology codes for Medicare telehealth claims and was able to use them to bill for these services easily. However, a staff member at another site noted that their EHR system was not equipped to properly document the service dates and subsequently bill for telehealth services. Although one hospital administrator felt that Medicare billing guidance was straightforward, some felt that guidance on Medicaid billing was lacking or complex, which led to confusion when billing for services.

Discussion

Principal Findings

CAH staff identified several obstacles to telehealth: difficulty establishing the needed relationship with appropriate distant providers due to credentialing issues, capacity limits of distant providers, potential distant providers in different health care networks, and distant providers unwilling to use telehealth. Another barrier was unwillingness to integrate telehealth into care delivery by referring providers in the patients' communities. Contextual factors such as accommodating process flows of local practitioners, patients, and staff was another barrier. Limited resources to support telehealth was another environmental factor that influenced telehealth uptake. Although the clinical staff we interviewed recognized the benefits of telehealth, they also indicated that workflow changes were needed so that practitioners would refer to telehealth services rather than face-to-face encounters and incorporate telehealth into their clinical and administrative practices.

Factors Influencing Telehealth Uptake and Use

Staff Acceptance

Telehealth represents a change in delivery mechanism and thus requires acceptance by practitioners, staff, and patients alike. CAH staff indicated that because practitioners are the gatekeepers who provide referrals to telehealth services, their engagement was critical to telehealth uptake, which was consistent with other findings [12]. In addition, if staff were engaged with telehealth, they were more likely to serve as champions internally and share their excitement with patients, which also resonates with findings from the literature [13-16].

Administrators across CAHs indicated the importance of getting practitioners and staff excited about and engaged with telehealth. Engagement was a key success factor.

Operational Considerations

Telehealth implementation reflects an organizational shift that requires some operational changes including clinical and administrative workflows, relationships with distant providers, and technical considerations [17]. Frontier CAHs generally seek specialty services and support from distant providers. Thus, operationalizing the relationship with distant providers requires structure that must be in place before implementation [4,18]. Staff at frontier CAHs indicated that ensuring that their providers could easily refer patients to distant specialty providers was a key success factor for telehealth uptake. Operational considerations also include items for hospital leadership, such as credentialing distant providers so that they can provide services at the facility [19]. The CAH staff indicated that credentialing of providers was a significant barrier.

In addition to relationships between providers, workflow considerations include ensuring that services appear seamless to the patient [20]. CAHs must identify processes (that may change ahead of time) so that policies are not developed on an ad hoc basis [21]. CAH staff indicated that they need implementation support to help identify what items needed to be changed and materials such as implementation guides to support telehealth implementation. Processes to consider include identifying and scheduling patients [18], making referrals and sharing information, and consenting. CAH staff reported that spending time on these processes helped facilitate care coordination.

Financial Considerations

Implementing and using telehealth involves financial considerations, primarily expenses due to infrastructure, staffing, and training, and offsetting those expenses with reimbursement. In addition, costs must be allocated for education, outreach, and workflow integration efforts. Although reimbursement for telehealth services is increasing, it still varies widely and is inconsistent [12,13,22,23]. CAH staff reported needing assistance with revenue cycles and with billing for services under the demonstration, particularly when they had not previously done so. Other financial considerations included structuring arrangements with distant providers. Telehealth involves sharing patients with practitioners who are in different locations, and CAHs needed assistance with billing processes to ensure payment for services [4,5,24].

Limitations

Some limitations may have affected this work. First, our major data source for this report was the responses that the authors collected during key informant interviews. The number of interviewees was small, which limits generalizability. Our secondary data source was reports, although many of these were developed by the same people we interviewed. Although the goal of the interviews was to obtain feedback (including viewpoints) from a variety of stakeholders, there is no guarantee that the individuals who participated in the interviews are representative of the entire staff at the CAHs. In addition, some of the interviewees were responsible for creating some of the reports reviewed in the document review. Therefore, the analysis results from the qualitative data may not represent all perspectives.

Furthermore, this work focuses on a small number of participating CAHs (n=8) in three states. Each of these CAHs was in a designated frontier area; therefore, the population was less dense than is generally found in rural areas. Each CAH has different internal and community resources available that may affect the successful implementation of telehealth. In addition, all of the CAHs were in the western part of the country. These factors may impact generalizability to other rural areas.

Conclusion

Frontier CAHs face challenges, including distance and provider availability, that telehealth implementation and use could address. Telehealth can be used for provider-to-patient and provider-to-provider interactions. To realize the promise of telehealth, frontier CAHs need support to improve factors that influence implementation and use.

During the COVID-19 pandemic, many organizations moved to telehealth very rapidly [25]. Although many of the initial increases in telehealth uptake were not sustained, telehealth encounters are above prepandemic levels [26,27]. Thus, understanding barriers and facilitators to telehealth generally can be useful for postpandemic planning. Across CAHs, support needs include change management, service identification for the community, credentialing, and reimbursement. Factors such as necessary workflow and infrastructure changes, practitioner acceptance and availability, and financing must be addressed to improve telehealth uptake. Supporting telehealth includes organizational and policy-level changes that can increase access and improve outcomes in rural communities.

Acknowledgments

The authors wish to acknowledge Fred Thomas for his support of this work. The authors also are grateful to the participants for sharing their time and experiences.

Conflicts of Interest

None declared.

References

1. Bullock DR, Vehe RK, Zhang L, Correll CK. Telemedicine and other care models in pediatric rheumatology: an exploratory study of parents' perceptions of barriers to care and care preferences. *Pediatr Rheumatol Online Journal* 2017 Jul 11;15(1):55-57 [FREE Full text] [doi: [10.1186/s12969-017-0184-y](https://doi.org/10.1186/s12969-017-0184-y)] [Medline: [28693580](https://pubmed.ncbi.nlm.nih.gov/28693580/)]
2. Frontier Community Health Integration Project Demonstration. Centers for Medicare & Medicaid Services. 2020. URL: <https://innovation.cms.gov/innovation-models/frontier-community-health-integration-project-demonstration> [accessed 2020-08-31]
3. Report to Congress: E-health and Telemedicine. U.S. Department of Health and Human Services. 2016 Aug 12. URL: <https://aspe.hhs.gov/system/files/pdf/206751/TelemedicineE-HealthReport.pdf> [accessed 2020-08-31] [WebCite Cache ID [76ccHjshS](https://www.webcitation.org/76ccHjshS)]
4. Jennett P, Jackson A, Healy T, Ho K, Kazanjian A, Woollard R, et al. A study of a rural community's readiness for telehealth. *J Telemed Telecare* 2003;9(5):259-263. [doi: [10.1258/135763303769211265](https://doi.org/10.1258/135763303769211265)] [Medline: [14599328](https://pubmed.ncbi.nlm.nih.gov/14599328/)]
5. Jennett PA, Gagnon MP, Brandstadt HK. Preparing for success: readiness models for rural telehealth. *J Postgrad Med* 2005;51(4):279-285 [FREE Full text] [Medline: [16388170](https://pubmed.ncbi.nlm.nih.gov/16388170/)]
6. Mandated report: Telehealth services and the Medicare program. MedPAC. 2016 Dec 01. URL: http://www.medpac.gov/docs/default-source/reports/mar18_medpac_ch16_sec.pdf [accessed 2020-08-30]
7. Implementation of telehealth in community clinics. California Health Care Foundation. 2016. URL: <http://www.chcf.org/publications/2010/11/implementation-telehealth-community-clinics> [accessed 2020-09-01]
8. Telehealth services: rural health series. Centers for Medicare & Medicaid Services. 2016. URL: <https://www.cms.gov/Outreach-and-Education/Medicare-Learning-Network-MLN/MLNProducts/downloads/TelehealthSrvcsfctsht.pdf> [accessed 2020-08-31]
9. Jones L, Grech C. The patient experience of remote telemonitoring for heart failure in the rural setting: a literature review. *Contemp Nurse* 2016;52(2-3):230-243. [doi: [10.1080/10376178.2015.1111154](https://doi.org/10.1080/10376178.2015.1111154)] [Medline: [26525077](https://pubmed.ncbi.nlm.nih.gov/26525077/)]
10. Brooks E, Turvey C, Augusterfer EF. Provider barriers to telemental health: obstacles overcome, obstacles remaining. *Telemed J E Health* 2013 Jun;19(6):433-437. [doi: [10.1089/tmj.2013.0068](https://doi.org/10.1089/tmj.2013.0068)] [Medline: [23590176](https://pubmed.ncbi.nlm.nih.gov/23590176/)]
11. Lowery CL, Bronstein JM, Benton TL, Fletcher DA. Distributing medical expertise: the evolution and impact of telemedicine in Arkansas. *Health Aff (Millwood)* 2014 Feb;33(2):235-243. [doi: [10.1377/hlthaff.2013.1001](https://doi.org/10.1377/hlthaff.2013.1001)] [Medline: [24493766](https://pubmed.ncbi.nlm.nih.gov/24493766/)]
12. Vimarlund V, Le RC. Barriers and opportunities to the widespread adoption of telemedicine: a bi-country evaluation. *Stud Health Technol Inform* 2013;192:933. [Medline: [23920707](https://pubmed.ncbi.nlm.nih.gov/23920707/)]
13. Teichgräber UK, Benter T, Klihs L, Schröder RJ, Hidajat N, Dörken B, et al. Project graph technique for time management in abdominal ultrasound evaluations. Article in German. *Ultraschall Med* 1999 Dec;20(6):236-241. [doi: [10.1055/s-1999-8916](https://doi.org/10.1055/s-1999-8916)] [Medline: [10670068](https://pubmed.ncbi.nlm.nih.gov/10670068/)]
14. Yusif S, Hafeez-Baig A, Soar J. e-Health readiness assessment factors and measuring tools: A systematic review. *Int J Med Inform* 2017 Nov;107:56-64. [doi: [10.1016/j.ijmedinf.2017.08.006](https://doi.org/10.1016/j.ijmedinf.2017.08.006)] [Medline: [29029692](https://pubmed.ncbi.nlm.nih.gov/29029692/)]
15. Nayar P, McFarland KK, Chandak A, Gupta N. Readiness for teledentistry: validation of a tool for oral health professionals. *J Med Syst* 2017 Jan;41(1):4. [doi: [10.1007/s10916-016-0654-7](https://doi.org/10.1007/s10916-016-0654-7)] [Medline: [27822871](https://pubmed.ncbi.nlm.nih.gov/27822871/)]
16. Baird MB, Whitney L, Caedo CE. Experiences and attitudes among psychiatric mental health advanced practice nurses in the use of telemental health: results of an online survey. *J Am Psychiatr Nurses Assoc* 2018;24(3):235-240. [doi: [10.1177/1078390317717330](https://doi.org/10.1177/1078390317717330)] [Medline: [28748728](https://pubmed.ncbi.nlm.nih.gov/28748728/)]
17. Maunder K, Walton K, Williams P, Ferguson M, Beck E. A framework for eHealth readiness of dietitians. *Int J Med Inform* 2018 Jul;115:43-52. [doi: [10.1016/j.ijmedinf.2018.04.002](https://doi.org/10.1016/j.ijmedinf.2018.04.002)] [Medline: [29779719](https://pubmed.ncbi.nlm.nih.gov/29779719/)]
18. James HE. Pediatric neurosurgery telemedicine clinics: a model to provide care to geographically underserved areas of the United States and its territories. *J Neurosurg Pediatr* 2016 Dec;25(6):753-757. [doi: [10.3171/2016.6.PEDS16202](https://doi.org/10.3171/2016.6.PEDS16202)] [Medline: [27589599](https://pubmed.ncbi.nlm.nih.gov/27589599/)]
19. Crane M. Exploring telehealth models. *Med Econ* 2014 Jul 25;91(14):17-20. [Medline: [25233550](https://pubmed.ncbi.nlm.nih.gov/25233550/)]
20. Becevic M, Boren S, Mutrux R, Shah Z, Banerjee S. User satisfaction with telehealth: study of patients, providers, and coordinators. *Health Care Manag (Frederick)* 2015;34(4):337-349. [doi: [10.1097/HCM.0000000000000081](https://doi.org/10.1097/HCM.0000000000000081)] [Medline: [26506296](https://pubmed.ncbi.nlm.nih.gov/26506296/)]
21. Kreofsky BLH, Blegen RN, Lokken TG, Kapraun SM, Bushman MS, Demaerschalk BM. Sustainable telemedicine: designing and building infrastructure to support a comprehensive telemedicine practice. *Telemed J E Health* 2018 Dec;24(12):1021-1025. [doi: [10.1089/tmj.2017.0291](https://doi.org/10.1089/tmj.2017.0291)] [Medline: [29658828](https://pubmed.ncbi.nlm.nih.gov/29658828/)]
22. Campagna M, Naka F, Lu J. Teledermatology: an updated overview of clinical applications and reimbursement policies. *Int J Womens Dermatol* 2017 Sep;3(3):176-179 [FREE Full text] [doi: [10.1016/j.ijwd.2017.04.002](https://doi.org/10.1016/j.ijwd.2017.04.002)] [Medline: [28831431](https://pubmed.ncbi.nlm.nih.gov/28831431/)]
23. Giambrone D, Rao BK, Esfahani A, Rao S. Obstacles hindering the mainstream practice of teledermatopathology. *J Am Acad Dermatol* 2014 Oct;71(4):772-780. [doi: [10.1016/j.jaad.2014.04.043](https://doi.org/10.1016/j.jaad.2014.04.043)] [Medline: [24906611](https://pubmed.ncbi.nlm.nih.gov/24906611/)]
24. Paul DL, McDaniel RR. Facilitating telemedicine project sustainability in medically underserved areas: a healthcare provider participant perspective. *BMC Health Serv Res* 2016 Apr 26;16:148 [FREE Full text] [doi: [10.1186/s12913-016-1401-y](https://doi.org/10.1186/s12913-016-1401-y)] [Medline: [27112268](https://pubmed.ncbi.nlm.nih.gov/27112268/)]
25. Expansion of telehealth during COVID-19 pandemic. Epic Health Research Network. 2020 May 05. URL: <https://ehrn.org/expansion-of-telehealth-during-covid-19-pandemic/> [accessed 2020-10-01]

26. As office visits fall, telehealth takes hold. Epic Health Research Network. 2020 Jun 12. URL: <https://ehrn.org/as-office-visits-fall-telehealth-takes-hold/> [accessed 2020-10-01]
27. Monthly telehealth regional tracker, March 2021. FAIR Health. 2021. URL: <https://www.fairhealth.org/states-by-the-numbers/telehealth> [accessed 2021-02-01]

Abbreviations

CAH: critical access hospital

CMS: Centers for Medicare & Medicaid Services

EHR: electronic health record

FCHIP: Frontier Community Health Integration Project

Edited by G Eysenbach; submitted 04.09.20; peer-reviewed by J Soar, P Yellowlees, C Hawke; comments to author 23.12.20; revised version received 23.12.20; accepted 13.04.21; published 05.05.21.

Please cite as:

Haque SN, DeStefano S, Banger A, Rutledge R, Romaine M

Factors Influencing Telehealth Implementation and Use in Frontier Critical Access Hospitals: Qualitative Study

JMIR Form Res 2021;5(5):e24118

URL: <https://formative.jmir.org/2021/5/e24118>

doi: [10.2196/24118](https://doi.org/10.2196/24118)

PMID: [33949958](https://pubmed.ncbi.nlm.nih.gov/33949958/)

©Saira Naim Haque, Sydney DeStefano, Alison Banger, Regina Rutledge, Melissa Romaine. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 05.05.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

A Resilience-Building App to Support the Mental Health of Health Care Workers in the COVID-19 Era: Design Process, Distribution, and Evaluation

Eddy A Golden¹, MPH; Micol Zweig¹, MPH; Matteo Danieleto¹, PhD; Kyle Landell¹, BA; Girish Nadkarni¹, MD, MPH; Erwin Bottinger^{1,2}, MD; Lindsay Katz^{3,4}, BA; Ricardo Somarriba⁵, BA; Vansh Sharma^{3,4}, MD; Craig L Katz^{3,4}, MD; Deborah B Marin^{3,4}, MD; Jonathan DePierro^{3,4}, PhD; Dennis S Charney^{4,6,7}, MD

¹Hasso Plattner Institute for Digital Health at Mount Sinai, Icahn School of Medicine at Mount Sinai, New York, NY, United States

²Digital Health Center, Hasso Plattner Institute, University of Potsdam, Potsdam, Germany

³Center for Stress, Resilience, and Personal Growth, Icahn School of Medicine at Mount Sinai, New York, NY, United States

⁴Department of Psychiatry, Icahn School of Medicine at Mount Sinai, New York, NY, United States

⁵Academic Informatics & Technology, Icahn School of Medicine at Mount Sinai, New York, NY, United States

⁶Department of Pharmacology, Icahn School of Medicine at Mount Sinai, New York, NY, United States

⁷Department of Neuroscience, Icahn School of Medicine at Mount Sinai, New York, NY, United States

Corresponding Author:

Eddy A Golden, MPH

Hasso Plattner Institute for Digital Health at Mount Sinai

Icahn School of Medicine at Mount Sinai

770 Lexington Avenue, 14th Fl

New York, NY, 10065

United States

Phone: 1 4014398798

Email: eddye.golden@mssm.edu

Abstract

Background: The COVID-19 pandemic has resulted in increased strain on health care systems and negative psychological effects on health care workers (HCWs). This is anticipated to result in long-term negative mental health effects on the population, with HCWs representing a particularly vulnerable group. The scope of the COVID-19 pandemic necessitates the development of a scalable mental health platform to provide services to large numbers of at-risk or affected individuals. The Mount Sinai Health System in New York City was at the epicenter of the pandemic in the United States.

Objective: The Center for Stress, Resilience, and Personal Growth (CSRPG) was created to address the current and anticipated psychological impact of the pandemic on the HCWs in the health system. The mission of the Center is to support the resilience and mental health of employees through educational offerings, outreach, and clinical care. Our aim was to build a mobile app to support the newly founded Center in its mission.

Methods: We built the app as a standalone digital platform that hosts a suite of tools that users can interact with on a daily basis. With consideration for the Center's aims, we determined the overall vision, initiatives, and goals for the Wellness Hub app, followed by specific milestone tasks and deliverables for development. We defined the app's primary features based on the mental health assessment and needs of HCWs. Feature definition was informed by the results of a resilience survey widely distributed to Mount Sinai HCWs and by the resources offered at CSRPG, including workshop content.

Results: We launched our app over the course of two phases, the first phase being a "soft" launch and the second being a broader launch to all of Mount Sinai. Of the 231 HCWs who downloaded the app, 173 (74.9%) completed our baseline assessment of all mental health screeners in the app. Results from the baseline assessment show that more than half of the users demonstrate a need for support in at least one psychological area. As of 3 months after the Phase 2 launch, approximately 55% of users re-entered the app after their first opening to explore additional features, with an average of 4 app openings per person.

Conclusions: To address the mental health needs of HCWs during the COVID-19 pandemic, the Wellness Hub app was built and deployed throughout the Mount Sinai Health System. To our knowledge, this is the first resilience app of its kind. The Wellness Hub app is a promising proof of concept, with room to grow, for those who wish to build a secure mobile health app

to support their employees, communities, or others in managing and improving mental and physical well-being. It is a novel tool offering mental health support broadly.

(*JMIR Form Res* 2021;5(5):e26590) doi:[10.2196/26590](https://doi.org/10.2196/26590)

KEYWORDS

mHealth; resilience; mental health; COVID-19; HCWs; digital health; health app; mental health platform; mobile phone

Introduction

COVID-19 has resulted in over 41 million infections with a worldwide case fatality ratio of approximately 2.6% as of early November 2020 [1,2]. Disease spread has been facilitated by a prolonged incubation period and variable symptomatology and severity. As a result, there is a tremendous burden on the health care system, with health care workers (HCWs) experiencing increased stress and demand and increased risk of COVID-19 infection, particularly for patient-facing staff [3]. With consideration for the existing issues of burnout and stress within the health care population, the additional impact of COVID-19 on HCWs, who represent a vulnerable group during this pandemic, is a growing concern [4-6]. Results from a survey of 2579 frontline HCWs at Mount Sinai Hospital in New York City during spring 2020 showed that 39% of respondents screened positive for COVID-19–related posttraumatic stress disorder (PTSD), generalized anxiety disorder, or depression. This finding displays the growing need for additional mental health support among HCWs as a result of the pandemic [7].

In recent years, mobile health (mHealth) apps on smartphones have become ubiquitous tools for personal health management and behavior tracking [8,9]. mHealth apps can provide individuals with continuous feedback on health status and progress, push notification reminders, and other useful engagement features. These remote, digital capabilities pose a major advantage for smartphones because they provide users with a mechanism to engage with, study, and interpret their health behaviors more regularly and comprehensively than in an average sequence of in-person doctors' appointments or study visits [10,11]. Informed by the literature that has considered the techniques that drive behavior change through mHealth apps as well as their effectiveness [12-15], implementation of mHealth apps for personal health and behavior management can be fine-tuned and leveraged to become a powerful tool in behavior modification and overall lifestyle improvement.

mHealth apps have been rigorously implemented during the COVID-19 pandemic in order to support contact tracing, symptom tracking, and behavioral management for individuals across the globe. These apps have been deployed to support health care systems, governments, and communities to reduce the spread of COVID-19 and its negative long-term effects on the stress and well-being of individuals [16-18].

The growing knowledge and popularity around mHealth apps, in combination with the imperative to support frontline workers [7,19], allowed us to launch a mental health–focused smartphone app for HCWs in New York's Mount Sinai Health System. It offers tools and support for users to bolster emotional well-being [20]. In its current state, the app is anonymous and the data

collected in this phase will offer insight into how to optimize the features and content.

Methods

Overview

The Wellness Hub app is designed as a tool for general mental health maintenance and resilience and connects users to the mental health services available through the Center for Stress, Resilience, and Personal Growth (CSRPG) and Mount Sinai at large [21]. A team of psychiatrists and psychologists affiliated with CSRPG provided all the content for the app and partnered with the Digital Discovery Program at the Hasso Plattner Institute for Digital Health at Mount Sinai to lead the software development. With consideration for the Center's aims, we determined the overall vision, initiatives, and goals for the Wellness Hub app, followed by specific milestone tasks and deliverables for development. Content and design framework were shared between the development team and CSRPG. The first version of the app was released after 2.5 months of iterative development.

The app is a standalone platform that hosts a suite of tools that users can interact with daily. While initially only available within Mount Sinai, we envisage the app being applicable and deployable in the future to other outside settings where there is a need to offer resilience-building tools to HCWs. Content (eg, lists of local supportive resources) could be tailored to specific health care systems.

Resilience Content

Resilience-focused content and user tasks for the app were adapted from resilience training workshops provided to Mount Sinai Health System staff starting in July 2020, in direct response to the impact of the pandemic on staff well-being. The Center's resilience workshops draw upon the work of Southwick and Charney, who outlined 10 factors contributing to personal resilience from laboratory research and interviews with trauma survivors, including prisoners of war during the Vietnam War [22].

The Center's staff developed a peer co-led resilience training curriculum and first piloted a series of eleven 45-minute workshop meetings [21]. Virtual training of over 70 peer leaders was conducted over the summer of 2020, with an initial focus on nurses and physician assistants. As of November 2020, 46 unique staff members have participated in the workshops, with time constraints often cited as reasons for nonattendance. Based on these practical considerations, we condensed the intervention to five meetings in the fall of 2020 focusing on the following: (1) realistic optimism, (2) facing fears and active coping, (3) social support and utilizing resilient role models, (4) self-care,

and (5) faith, meaning, and spirituality. Video content has been developed around these five topics hosted in the app—on Mount Sinai's YouTube channel and on the Center's website—to facilitate engagement of staff who do not have the time to participate in person [23].

App Framework

Overview

We defined the app's primary features based on the mental health assessment and needs of HCWs. Feature definition was informed by the results of a resilience survey widely distributed to Mount Sinai HCWs [7] and by the resources offered at CSRPG, including workshop content.

Most prominently, app users answer standardized mental health surveys with subsequent score-based feedback that incorporates various exercises for building resilience. Progress is visualized over time based on individual survey scores, showing green (good) or yellow (cautionary, screen-in) feedback depending on the score, allowing users to track their mental health in a regular, quantitative, and accessible way. A user will receive randomized feedback text based on their score (ie, good vs cautionary). If a user scores in a way that suggests a need for mental health care (ie, cautionary), a feedback screen appears suggesting that they seek additional mental health and wellness resources and lists available resources and tools directly on the screen for easy access. App users can write private digital journal entries and access videos for relaxation, including some created specifically for Mount Sinai employees. Users also have the ability to tailor their resources and videos to suit their mindset and interests by favoriting specific ones for ease of access. The app also offers quick links to Mount Sinai wellness-related resources (eg, spiritual care and behavioral health clinics). The app is intended to be used in conjunction with CSRPG's resilience workshops and, therefore, includes key information and activities related to resilience development. We selected features based on their usability as it relates to mental health management and their compatibility with behavior change theory [12,24]. Our primary aim was to build features that support feedback and self-monitoring in a simple and straightforward manner through surveys and journaling and, in this way, support behavior change. Resources in the app—videos and referral resources—serve as reinforcement and added support if desired.

Mental Health Surveys

The app contains mental health surveys presented on the main screen. Users are first introduced to the surveys in a *baseline* fashion, where all questions are offered at one time, in one baseline survey. After completion, users will see each survey individualized on the home screen and can take each survey as many times as they would like. The surveys offered to users are validated measures commonly used in research and clinical practice, though they are available to answer consistently in the app as opposed to on a validated schedule. This is because the surveys are meant to serve as a feedback mechanism for users, where they can regularly assess their well-being and mental health as well as receive consistent feedback based on their results. The surveys were chosen based on prior knowledge

about disorders that occur in the context of trauma as well as data derived from a previous survey distributed among Mount Sinai Hospital HCWs [7]. All surveys are presented to users in plain language, as opposed to their clinical name, with scores calculated as indicated in the original validated scoring instructions. The following surveys are included in the app:

1. Ecological mood assessment ("My Daily Feelings"). We included a daily feelings meter with emoticons indicating sadness to happiness on a scale from 1 to 7 for participants to rate their daily mood. This measure is similar to that utilized in prior research with smartphone apps [25].
2. PHQ-8 (8-item Patient Health Questionnaire) and GAD-2 (2-item Generalized Anxiety Disorder scale) ("My Mood"). The PHQ-8 and GAD-2 screen for depression and overall anxiety, respectively, over the past 2 weeks [26,27]. The PHQ-8 is identical to the PHQ-9, except that the item pertaining to suicidality is omitted [28]. PHQ-8 and GAD-2 items are presented on a scale from 0 to 3, with higher scores indicating more severe depression and anxiety, respectively. A PHQ-8 score of 10 or higher was used to indicate probable depression, while a score of 3 or higher on the GAD-2 suggested clinically significant anxiety problems.
3. PCL-5 (PTSD Checklist for DSM-5 [Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition]) ("My Stress Reactions"). We utilized the PCL-5, a newly published brief 4-item version of the longer 20-item PCL-5, anchored to the COVID-19 experience, to evaluate posttraumatic stress reactions [29,30]. Items are presented on a scale from 0 (not at all) to 4 (extremely). Based on initial work with this brief measure, a score of 8 or higher was used as a cut point for probable PTSD [29].
4. CD-RISC2 (2-item Connor-Davidson Resilience Scale) ("My Coping Resources"). We utilized the CD-RISC2, a previously validated short form of the CD-RISC [31,32]. The scale ranges from 0 (not true at all) to 4 (true nearly all the time), with higher scores indicating greater resilience. Based on a review of the available normative data, a score of 5 or lower was used to indicate current problems with resilience for purposes of providing feedback [31].
5. WHO-5 (5-item World Health Organization Well-Being Index) ("My Well-being"). The WHO-5 is a 5-item globally utilized measure of overall well-being [33]. Items are presented on a scale from 0 (at no time) to 5 (all the time). Higher scores indicate greater well-being. Scores are summed and multiplied by 4 to yield an index ranging from 0 to 100; scores at or below 50 are suggestive of depression.
6. AUDIT-C (Alcohol Use Disorders Identification Test-Concise) ("My Alcohol Use"). The AUDIT-C is a widely utilized 3-item screening measure for alcohol misuse [34]. Items are presented on a scale from 0 to 4; scores of 3 or higher were taken to be suggestive of problems with alcohol misuse.
7. Spiritual struggle items ("My Spirituality"). Based on prior research, two items were included in the app to measure problems with spirituality based on prior research: (1) "Do you struggle with loss of meaning and joy in your life?" and (2) "Do you currently have what you would describe as spiritual or religious struggles?" [35]. These items were

presented on a 4-point Likert scale ranging from 0 (not at all) to 3 (a great deal). Endorsement of 1 or higher on either item triggered feedback regarding addressing loss of meaning or spiritual problems.

Figure 1 shows the available surveys on the app, as well as the weekly progress screen and the survey-specific progress screen. Figure 2 shows the randomized feedback sent to users based on their survey scores.

Figure 1. Available surveys (left), weekly progress screen (center), and survey-specific progress screen (right).

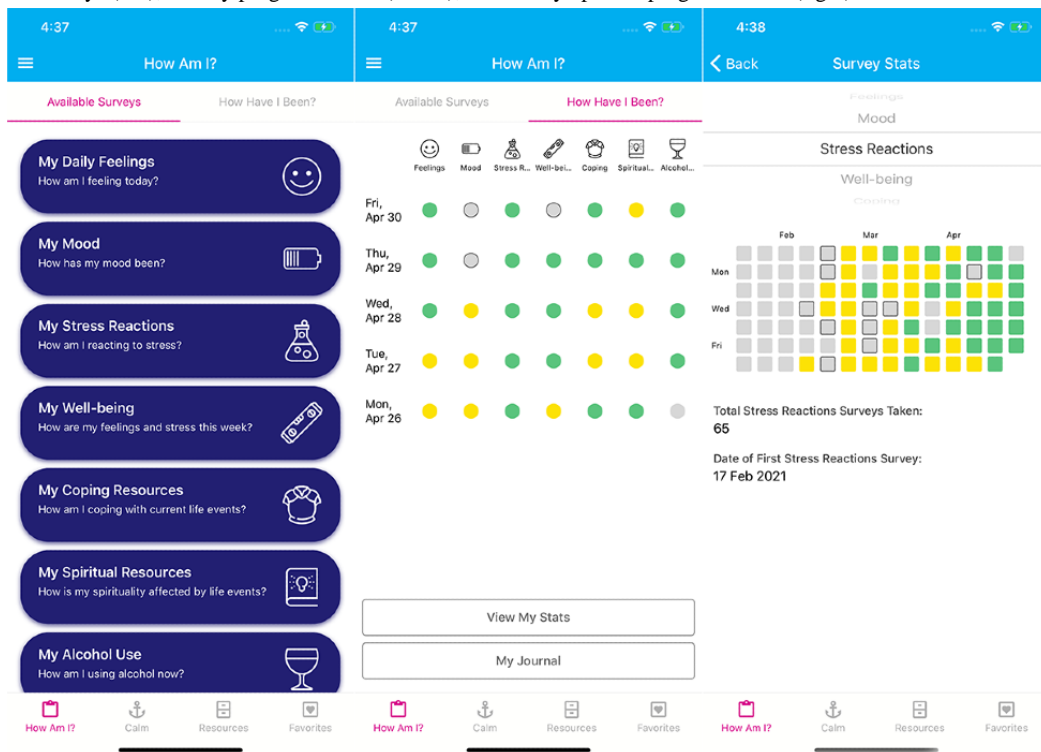
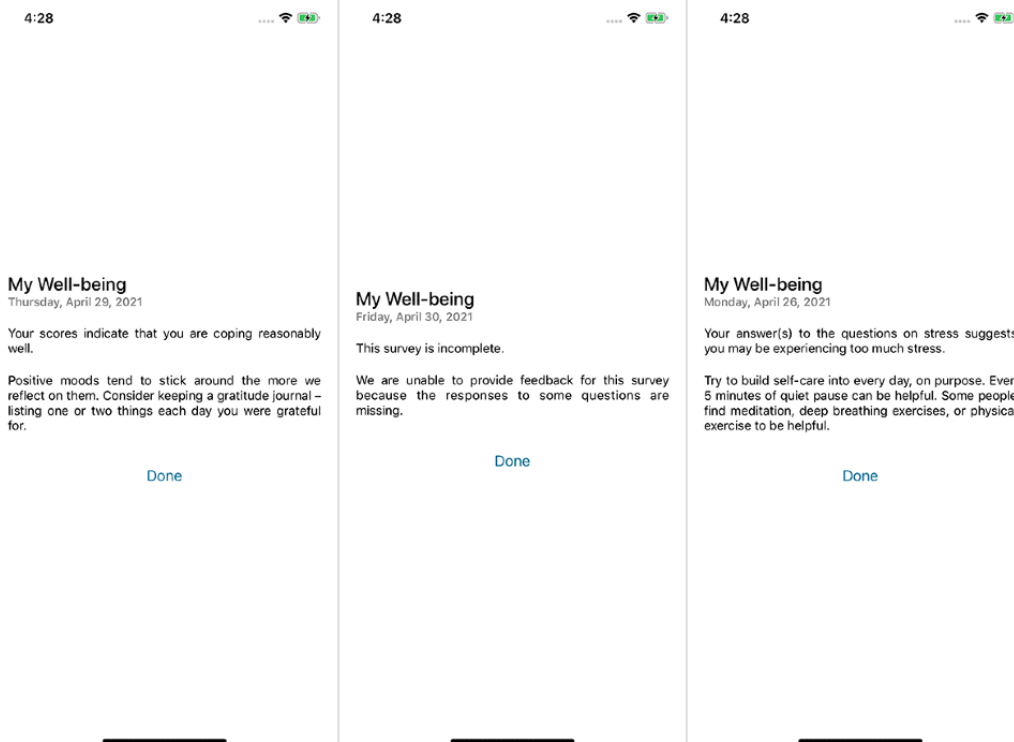


Figure 2. Randomized feedback based on survey score: good (left), incomplete (middle), or cautionary (right).



My Resources and My Journal

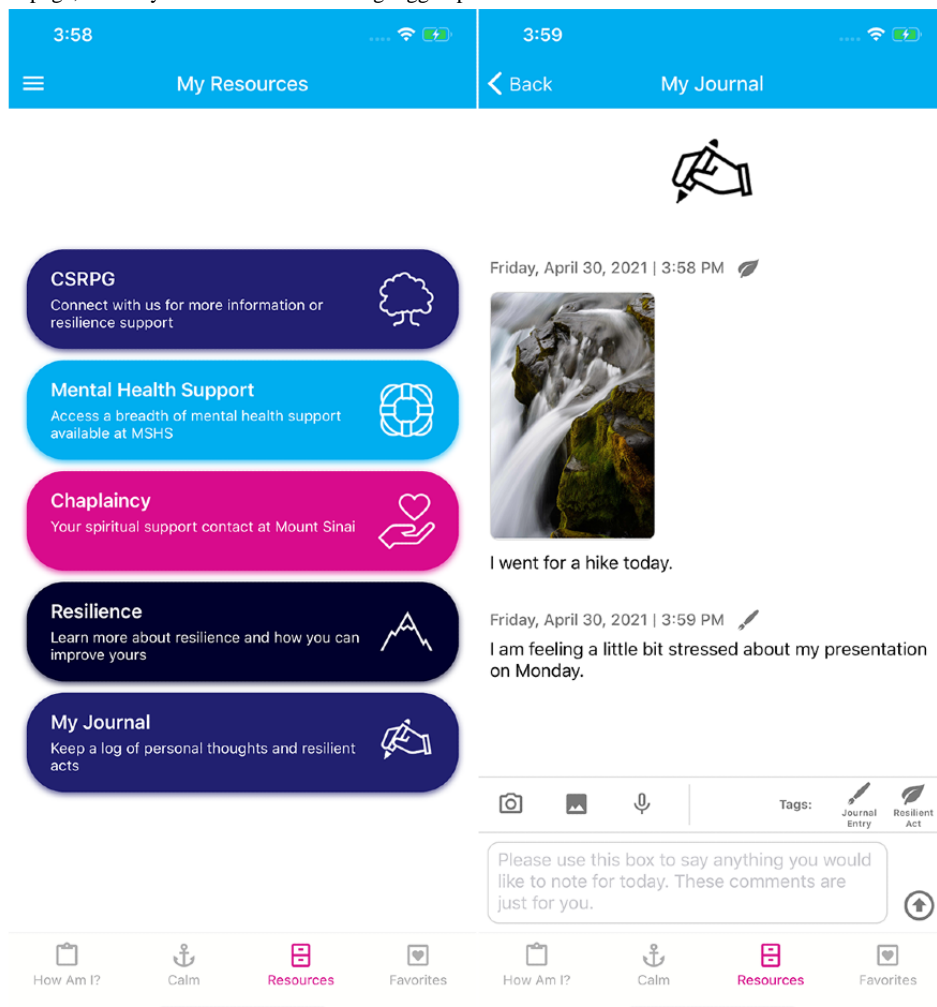
We offer a library of resources to our Wellness Hub users. Whether based on survey feedback, general curiosity, or additional mental health needs, the app provides access to the Mount Sinai mental health support network and offers resources and tools external to Mount Sinai.

For individuals looking for direct interaction with a mental health expert or support network, the app provides contact information for CSRPG, chaplaincy, and other behavioral practices and offices at Mount Sinai. We offer a page focused on promoting resilience and building a resilience plan. This

page also offers additional CSRPG workshop content and videos.

The “My Resources” tool includes a journal function, which allows users to write down any relevant thoughts or experiences that they feel are important to log. These entries serve as a way to self-monitor their feelings and thoughts and can be in the form of text, image, or voice recording. They can also be tagged as a regular journal entry or a resilience act. This tagging feature allows users to apply what they learn in the CSRPG workshops, along with the *Resilience* page in “My Resources” in order to discern for themselves what experiences are helping them to build their resilience (Figure 3).

Figure 3. My Resources page, with My Journal feature showing tagged photo and text entries.



Another feature provides HCWs with easily accessible relaxation content. The “My Calm” tool offers users a library of relaxation videos, some from the public web and some specifically designed for Mount Sinai employees [36]. These videos range in topic and duration so a user can select something that works for them in the moment. For example, an app user may choose a short, 2-minute video if they are on break during the workday, but may prefer to engage with a longer, more involved yoga video in the evening or to start the day.

General App Features

Overview

Wellness Hub also includes general app and health app–related features. Our data management page allows users to download their data and/or delete all data from the app. Because the app may hold sensitive or personal information related to a user’s feelings or mental health, we maintain an anonymous back-end server, meaning we do not collect identifiable data from any of our users. Our only way of identifying a specific user is if they provide us with their universally unique identifier that lives in the “My Profile” section of the app.

A notification-setting tool allows users to set the time of day they want to receive notifications about checking in with the app to engage with the available features. Users can also turn off notifications completely.

Additionally, and in accordance with standard practice, we also include information about the app, its developers, and additional copyright or reference information for content hosted in the app.

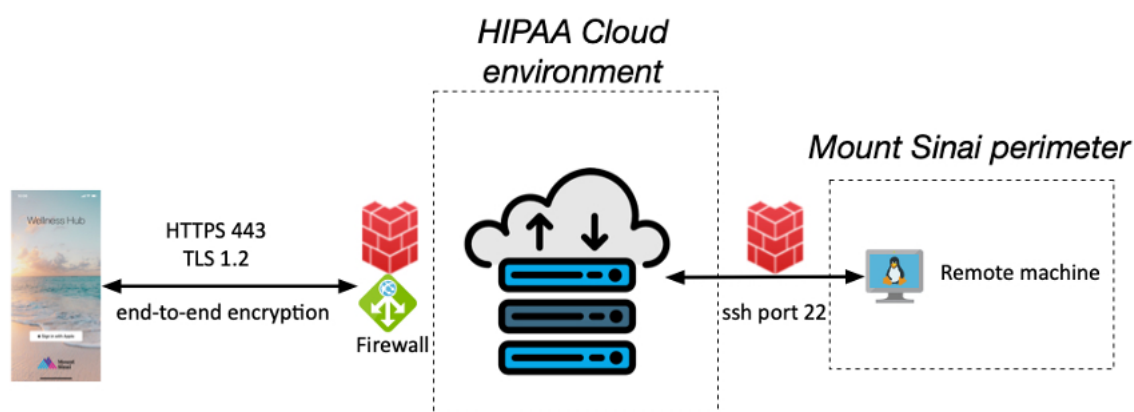
Data Privacy and Security

Data privacy and security are important considering the content of the app and its potentially stigmatizing and/or personal content, and we ensure that all anonymous data are stored with limited access within the Mount Sinai secure network. Further, all questions in the app are optional, including demographics.

Wellness Hub app data are hosted in a Health Insurance Portability and Accountability Act (HIPAA)-compliant ecosystem. For quality control, robust data transfer, and data persistence, the Wellness Hub app stores the data in an encrypted environment. The environment is fully encrypted, and the metadata are stored in secure non-Structured Query Language (NoSQL) instances for quality control.

The data transferred from the app to the cloud are sent through an https protocol with end-to-end encryption. Authentication protection at the beginning of the data ingestion is performed, and all data requests external to the server are blocked. No personally identifiable information is ever transferred from the app to the back end. The server manager cannot identify the data's owner, effectively making the ingested data anonymous (Figure 4).

Figure 4. Data ingestion diagram. HIPAA: Health Insurance Portability and Accountability Act; ssh: secure shell; TLS: Transport Layer Security.



The cloud system is only accessible through a secure shell with a public-private key, and IP address restrictions are in place to prevent cyberattacks. The database is restricted only to the app data and does not host any data from outside the app.

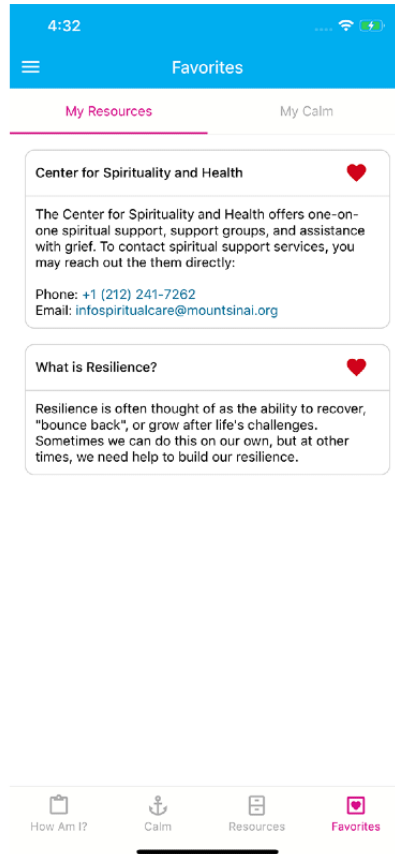
The app processes the survey scores locally and initializes the data transfer. Scores are given either a valid value or an invalid value, depending on whether the user completes a full survey or not. The app also implements single sign-on authentication, either Apple or Google, depending on which operating system the app is running. Consistent with our approach to privacy, the

app has a nonutilization timer that expires if the app is not used for 2 minutes, and the timer forces a log-out when it expires.

Personalization

The “Favorites” tool in Wellness Hub allows users to tailor the content of the app based on what they find most helpful to them. All resources and videos from the “My Resources” and “My Calm” pages can be selected as favorites and will subsequently appear as a personalized library in this view. Here, preferred resources can be easily accessible (Figure 5).

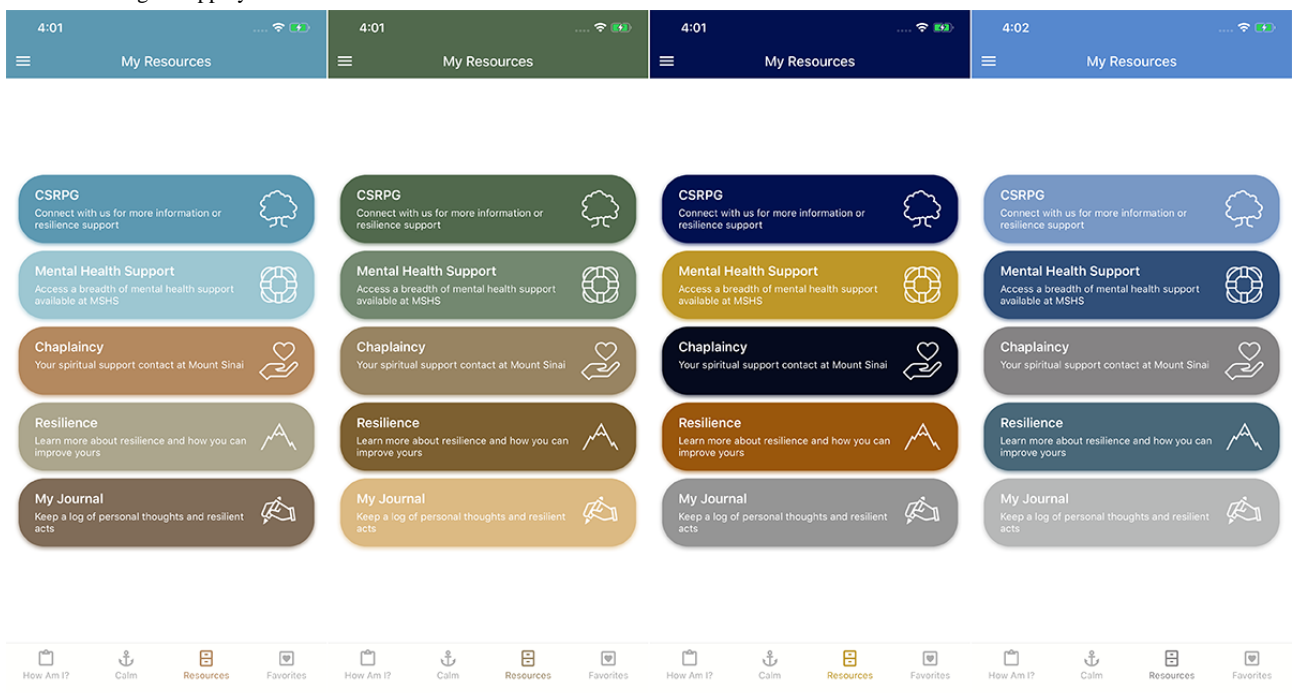
Figure 5. Example of the My Favorites page with resources and contact pages “hearted” and easily accessible.



Beyond tailoring resources, we also allow users to further personalize their experience through design. In addition to the standard Mount Sinai colors, we include a feature allowing users to choose the color scheme they prefer for their app, with numerous options meant to offer a calming effect to the user

(Figure 6). This serves as a simple way to allow personalization of the app, in the same way that they can personalize their favorite content and track their journal entries and survey responses.

Figure 6. Tailoring the app by color scheme.



Demographics

For future use, we have included a basic and optional “About Me” anonymous demographics survey in the standard profile section of the app in order to improve the personalization of resources based on age, role at Mount Sinai Health System, and other demographic or COVID-19–related features. This section can also help us better understand the breakdown of our user population and develop new or amend existing features to better address and target outreach to an underserved user population.

Distribution

This app is available for both iOS and Android devices and is currently only available to employees of the Mount Sinai Health System. Due to this employment restriction in distribution, we developed this as a custom app that is only available for download using a unique redemption code or link. The code or link can only be used one time, thus limiting the possibility of users distributing the app outside of the health system. In order to verify the current employment status of the user, we created a page on the Mount Sinai intranet in which users must first log in, thus validating their employment, and subsequently select whether they want to download the app for iOS or Android. Next, they enter the email address to which they want to receive installation instructions. The intranet page automatically pulls a unique redemption code from a list on the back end and sends download instructions, including the unique code, to the interested employee. The code is then marked as “used” and unable to be supplied to another person.

The Wellness Hub app was launched between June and September 2020, as a Phase 1 pilot in which we manually distributed redemption codes to interested Mount Sinai Health System employees while further developing the app in anticipation of the Phase 2 launch in October 2020, to correspond with Mental Health Awareness Month. At the time of the Phase 2 launch in early October, CSRPG introduced the app to the Mount Sinai Health System community through regular presentations, in-hospital messaging, and email notifications.

Key Performance Indicators

Prior to app development, we established key performance indicators (KPIs) to measure the success of the product and level of engagement of our app users. Our KPIs included total number of HCWs downloading the app, total number of baseline surveys submitted, number of users submitting additional surveys postbaseline, and general app engagement metrics (eg, number of app openings).

Results

Overview

This paper outlines the design and launch process for the Wellness Hub app. Development time for the platform was variable between Phase 1 and Phase 2. From initial ideation to the current version, we estimate a total of approximately 2.5 months for development and launch. This time frame is reflective of one software engineer, one product manager, and one designer working on the app. Since the app was developed in-house at Mount Sinai, the platform’s primary cost is for database management, which is a regular ongoing service fee. As of March 2021, there were new downloads each month and no reports of bugs or software crashes, suggesting a successful technical launch of the Wellness Hub app.

We analyzed several KPIs to determine the success of the app early after the launch. These early results serve to showcase the initial launch of the app. Between July 27 and October 13, 2020, during our Phase 1 pilot launch, which included manual app distribution and word-of-mouth recruitment, 74 users downloaded the app. We received informal feedback from many of these early users, which allowed us to refine the app and include additional features and adapt existing ones to support our user population. This led to the inclusion of relaxation videos and the improvement of our survey feedback content and progress visualizations.

Our broader, Phase 2 launch was on October 14, 2020, and included greater recruitment and easily accessible app distribution methods. Within 3 months of Phase 2, we received 157 additional app downloads, totaling our reach and app use to 231 users over the course of 4.5 months.

Of the 231 HCWs who downloaded the app, 173 (74.9%) completed our baseline assessment of all mental health screeners available in the app. As of 3 months after the Phase 2 launch, approximately 55% of users re-entered the app after their first opening to explore additional features, with an average of 4 app openings per person.

While the platform is meant primarily to provide individualized feedback to users, and not analyze a cohort of users together, we have compiled a brief overview of baseline results.

Slightly more than half of the baseline survey responses showed *screen-in* cautionary scores on the WHO-5 (97/173, 56.1%), AUDIT-C (79/159, 49.7%), and spiritual struggle items (113/159, 71.1%). This demonstrates a clear need for additional HCW support related to overall well-being and spiritual struggles (Table 1).

Table 1. Survey-specific completion at baseline and screen-in distribution.

Survey type	Total completed, n	Yellow score (screen-in), n (%)	Green score (doing okay), n (%)	Total incomplete, n (%)
Daily feelings	173	105 (60.7)	65 (37.6)	3 (1.7)
WHO-5 ^a	173	97 (56.1)	71 (41.0)	5 (2.9)
AUDIT-C ^b	159	79 (49.7)	75 (47.2)	5 (3.1)
GAD-2 ^c	102	38 (37.3)	64 (62.7)	0 (0)
PHQ-8 ^d	160	45 (28.1)	105 (65.6)	10 (6.3)
PCL-5 ^e	159	53 (33.3)	98 (61.6)	8 (5.0)
CD-RISC-2 ^f	159	41 (25.8)	113 (71.1)	5 (3.1)
Spiritual struggle	159	113 (71.1)	40 (25.2)	6 (3.8)

^aWHO-5: 5-item World Health Organization Well-Being Index.

^bAUDIT-C: Alcohol Use Disorders Identification Test-Concise.

^cGAD-2: 2-item Generalized Anxiety Disorder scale.

^dPHQ-8: 8-item Patient Health Questionnaire.

^ePCL-5: Posttraumatic Stress Disorder Checklist for DSM-5 (Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition).

^fCD-RISC-2: 2-item Connor-Davidson Resilience Scale.

On average, users took 2.06 (SD 6.5) surveys (screeners) after completing the baseline assessment. Of all the surveys available on the app, the daily feelings (“My Daily Feelings”), PHQ-8 (“My Mood”), and PCL-5 (“My Stress Reactions”) surveys had been taken the most, with 410, 368, and 318 total completions, respectively. Considering the everyday availability of these surveys for users on the platform, they were not taken on the validated schedule for which the surveys are primarily intended.

As our study of app usage proceeds, we will review resilience-related indicators beyond the baseline assessments taken by users. We intend to review additional KPIs that align with longer-term use of the platform.

Strengths and Limitations of the App

The strength of the app is that the Center was able to quickly deploy a tool based on existing resilience literature in order to address a mental health need within the frontline HCW community amidst the initial wave of the COVID-19 pandemic.

The main limitation of the app is that, as it is not a research study in its current version, there is no strict protocol given to users for how they must interact with the app. As a result, there is significant variability in app usage and our ability to track user experiences relative to their demographics. Future versions of the app may include a research intervention that will empirically test the use and effectiveness of the app as a supplement to CSRPG workshops.

Additionally, we consider selection bias; users may already have a bias or inclination toward resilience building, or recognition of their need for wellness tools, thus leading them to search for an app-based tool. We may be missing an important HCW population that is both lacking in resources and would not know how to find or use this tool effectively. We have to consider an outreach strategy that will promote greater access or we must consider making these tools available in multiple formats and venues. Lastly, the app is available to a specific cohort with specific access: HCWs in the Mount Sinai Health

System. In its current state, it is meant to serve partly as a supplement to CSRPG workshops held at Mount Sinai. This may impact external validation of the app’s efficacy in building resilience outside of this specific cohort.

Discussion

mHealth apps are an increasingly used tool for people to maintain their physical and mental wellness [8]. mHealth apps focused on engaging features, linked to behavior change theory, are crucial in helping to empower people to take control of their own health [37].

There are many publicly available resilience-building platforms similar to Wellness Hub, such as COVID Coach and Calm. COVID Coach, developed by the US Department of Veterans Affairs, offers education about coping, self-care tools, progress visualization, and national mental health resources [38]. Calm is a meditation app that offers video and audio programs related to meditation and mindfulness [39]. The Wellness Hub app stands out due to its targeted nature, in that it provides specific and accessible health system resources to its users and also offers validated mental health surveys along with progress visualization. With HCWs feeling the brunt of the pandemic stress, it makes sense to offer them resources that are close to home and relevant to their work.

To address the mental health needs of HCWs during the COVID-19 pandemic, the Wellness Hub app was built and deployed throughout the Mount Sinai Health System. To our knowledge, this is the first resilience app targeted to HCWs and applied within the context of a global trauma. It includes validated surveys and questionnaires based on prior knowledge of disorders that occur in the context of trauma, a journal tool to allow users to log their feelings and activities related to resilience, and numerous Mount Sinai-specific resources and videos targeted to those seeking additional support and tools to build resilience. The app’s initial results, showing that more

than half of the baseline assessments resulted in *screen-in* scores, are consistent with previous findings at Mount Sinai in the same cohort via a hospital-wide early pandemic survey [7].

The limited level of app use and engagement thus far may be related to the constrained recruitment strategies to date, considering the lack of in-person advertising, enrollment, and distribution due to the pandemic. Additionally, the slower uptake is consistent with our expectations of an initial app release, especially considering the involved download process and our active and occupied target audience of frontline HCWs and others in the health system.

We intend to continue developing this app by improving features based on feedback from our users, including the translation of the platform into Spanish and the integration of Mount Sinai employee and health care services, such as appointment

scheduling. We will move beyond assessment by adding specific resilience-promoting interventions, action items, and tools. Potential future improvements include *badges* that users can obtain in order to encourage them through their resilience journey as well as inclusion of materials from the CSRPG workshop curriculum in specific *learning* sections and as thought prompts in an expanded version of our journal tool. As the CSRPG continues to grow and COVID-19's effects continue to be felt across the health care system, we intend to use our KPIs, user feedback, and guidance from the Center workshops to determine the future direction of the app and its content.

The Wellness Hub app is a promising proof of concept for those who wish to build a secure mHealth app to support their employees, communities, or others in managing and improving mental and physical well-being. It offers a novel tool to offer mental health support broadly.

Acknowledgments

The authors would like to thank the faculty, staff, trainees, and students within the Mount Sinai Health System who responded to the COVID-19 pandemic. We also thank the Sinai Central team for their consistent support in this effort.

Conflicts of Interest

DSC is a co-inventor of a pending provisional patent application filed by the Icahn School of Medicine at Mount Sinai covering this technology. The technology is not yet licensed.

References

1. Estimating mortality from COVID-19. World Health Organization. 2020 Aug 04. URL: <https://www.who.int/publications/i/item/WHO-2019-nCoV-Sci-Brief-Mortality-2020.1> [accessed 2021-04-28]
2. WHO Coronavirus (COVID-19) Dashboard. World Health Organization. URL: <https://covid19.who.int/> [accessed 2021-04-28]
3. Misra-Hebert AD, Jehi L, Ji X, Nowacki AS, Gordon S, Terpeluk P, et al. Impact of the COVID-19 pandemic on healthcare workers' risk of infection and outcomes in a large, integrated health system. *J Gen Intern Med* 2020 Nov;35(11):3293-3301 [FREE Full text] [doi: [10.1007/s11606-020-06171-9](https://doi.org/10.1007/s11606-020-06171-9)] [Medline: [32875500](https://pubmed.ncbi.nlm.nih.gov/32875500/)]
4. Willard-Grace R, Knox M, Huang B, Hammer H, Kivlahan C, Grumbach K. Burnout and health care workforce turnover. *Ann Fam Med* 2019 Jan;17(1):36-41 [FREE Full text] [doi: [10.1370/afm.2338](https://doi.org/10.1370/afm.2338)] [Medline: [30670393](https://pubmed.ncbi.nlm.nih.gov/30670393/)]
5. Shreffler J, Petrey J, Huecker M. The impact of COVID-19 on healthcare worker wellness: A scoping review. *West J Emerg Med* 2020 Aug 17;21(5):1059-1066 [FREE Full text] [doi: [10.5811/westjem.2020.7.48684](https://doi.org/10.5811/westjem.2020.7.48684)] [Medline: [32970555](https://pubmed.ncbi.nlm.nih.gov/32970555/)]
6. Shaikat N, Ali DM, Razzak J. Physical and mental health impacts of COVID-19 on healthcare workers: A scoping review. *Int J Emerg Med* 2020 Jul 20;13(1):40 [FREE Full text] [doi: [10.1186/s12245-020-00299-5](https://doi.org/10.1186/s12245-020-00299-5)] [Medline: [32689925](https://pubmed.ncbi.nlm.nih.gov/32689925/)]
7. Feingold JH, Peccoraro L, Chan CC, Kaplan CA, Kaye-Kauderer H, Charney D, et al. Psychological impact of the COVID-19 pandemic on frontline health care workers during the pandemic surge in New York City. *Chronic Stress (Thousand Oaks)* 2021;5:1-13 [FREE Full text] [doi: [10.1177/2470547020977891](https://doi.org/10.1177/2470547020977891)] [Medline: [33598592](https://pubmed.ncbi.nlm.nih.gov/33598592/)]
8. McCarthy J. One in five US adults use health apps, wearable trackers. Gallup. 2019 Dec 11. URL: <https://news.gallup.com/poll/269096/one-five-adults-health-apps-wearable-trackers.aspx> [accessed 2021-04-28]
9. Sama PR, Eapen ZJ, Weinfurt KP, Shah BR, Schulman KA. An evaluation of mobile health application tools. *JMIR Mhealth Uhealth* 2014 May 01;2(2):e19 [FREE Full text] [doi: [10.2196/mhealth.3088](https://doi.org/10.2196/mhealth.3088)] [Medline: [25099179](https://pubmed.ncbi.nlm.nih.gov/25099179/)]
10. Ventola CL. Mobile devices and apps for health care professionals: Uses and benefits. *P T* 2014 May;39(5):356-364 [FREE Full text] [Medline: [24883008](https://pubmed.ncbi.nlm.nih.gov/24883008/)]
11. Sawesi S, Rashrash M, Phalakornkule K, Carpenter JS, Jones JF. The impact of information technology on patient engagement and health behavior change: A systematic review of the literature. *JMIR Med Inform* 2016 Jan 21;4(1):e1 [FREE Full text] [doi: [10.2196/medinform.4514](https://doi.org/10.2196/medinform.4514)] [Medline: [26795082](https://pubmed.ncbi.nlm.nih.gov/26795082/)]
12. Milne-Ives M, Lam C, Van Velthoven MH, Meinert E. Mobile apps for health behavior change: Protocol for a systematic review. *JMIR Res Protoc* 2020 Jan 30;9(1):e16931 [FREE Full text] [doi: [10.2196/16931](https://doi.org/10.2196/16931)] [Medline: [32012109](https://pubmed.ncbi.nlm.nih.gov/32012109/)]
13. Michie S, Yardley L, West R, Patrick K, Greaves F. Developing and evaluating digital interventions to promote behavior change in health and health care: Recommendations resulting from an international workshop. *J Med Internet Res* 2017 Jun 29;19(6):e232 [FREE Full text] [doi: [10.2196/jmir.7126](https://doi.org/10.2196/jmir.7126)] [Medline: [28663162](https://pubmed.ncbi.nlm.nih.gov/28663162/)]

14. Han M, Lee E. Effectiveness of mobile health application use to improve health behavior changes: A systematic review of randomized controlled trials. *Healthc Inform Res* 2018 Jul;24(3):207-226 [FREE Full text] [doi: [10.4258/hir.2018.24.3.207](https://doi.org/10.4258/hir.2018.24.3.207)] [Medline: [30109154](https://pubmed.ncbi.nlm.nih.gov/30109154/)]
15. Owen JE, Jaworski BK, Kuhn E, Makin-Byrd KN, Ramsey KM, Hoffman JE. mHealth in the wild: Using novel data to examine the reach, use, and impact of PTSD Coach. *JMIR Ment Health* 2015;2(1):e7 [FREE Full text] [doi: [10.2196/mental.3935](https://doi.org/10.2196/mental.3935)] [Medline: [26543913](https://pubmed.ncbi.nlm.nih.gov/26543913/)]
16. Whitelaw S, Mamas MA, Topol E, Van Spall HGC. Applications of digital technology in COVID-19 pandemic planning and response. *Lancet Digit Health* 2020 Aug;2(8):e435-e440 [FREE Full text] [doi: [10.1016/S2589-7500\(20\)30142-4](https://doi.org/10.1016/S2589-7500(20)30142-4)] [Medline: [32835201](https://pubmed.ncbi.nlm.nih.gov/32835201/)]
17. Balcombe L, De Leo D. An integrated blueprint for digital mental health services amidst COVID-19. *JMIR Ment Health* 2020 Jul 22;7(7):e21718 [FREE Full text] [doi: [10.2196/21718](https://doi.org/10.2196/21718)] [Medline: [32668402](https://pubmed.ncbi.nlm.nih.gov/32668402/)]
18. Figueroa CA, Aguilera A. The need for a mental health technology revolution in the COVID-19 pandemic. *Front Psychiatry* 2020;11:523 [FREE Full text] [doi: [10.3389/fpsy.2020.00523](https://doi.org/10.3389/fpsy.2020.00523)] [Medline: [32581891](https://pubmed.ncbi.nlm.nih.gov/32581891/)]
19. Ripp J, Peccoralo L, Charney D. Attending to the emotional well-being of the health care workforce in a New York City health system during the COVID-19 pandemic. *Acad Med* 2020 Aug;95(8):1136-1139 [FREE Full text] [doi: [10.1097/ACM.0000000000003414](https://doi.org/10.1097/ACM.0000000000003414)] [Medline: [32282344](https://pubmed.ncbi.nlm.nih.gov/32282344/)]
20. Montag C, Dagum P, Elhai JD. On the need for digital phenotyping to obtain insights into mental states in the COVID-19 pandemic. *Digit Psychol* 2020 Oct 27;1(2):40-42. [doi: [10.24989/dp.v1i2.1857](https://doi.org/10.24989/dp.v1i2.1857)]
21. DePiero J, Katz CL, Marin D, Feder A, Bevilacqua L, Sharma V, et al. Mount Sinai's Center for Stress, Resilience and Personal Growth as a model for responding to the impact of COVID-19 on health care workers. *Psychiatry Res* 2020 Nov;293:113426. [doi: [10.1016/j.psychres.2020.113426](https://doi.org/10.1016/j.psychres.2020.113426)] [Medline: [32861094](https://pubmed.ncbi.nlm.nih.gov/32861094/)]
22. Southwick SM, Charney DS. *Resilience: The Science of Mastering Life's Greatest Challenges*. 2nd edition. Cambridge, UK: Cambridge University Press; 2018.
23. Icahn School of Medicine at Mount Sinai videos. Icahn School of Medicine at Mount Sinai. URL: https://icahn.mssm.edu/about/video?videoid=KyDUPpp9GDI#playlistid=PLqLDR0CTP9_q1uFCk42yl61iOuBrnmQet [accessed 2021-04-28]
24. 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](https://doi.org/10.1007/s12160-013-9486-6)] [Medline: [23512568](https://pubmed.ncbi.nlm.nih.gov/23512568/)]
25. Kim J, Lim S, Min YH, Shin YW, Lee B, Sohn G, et al. Depression screening using daily mental-health ratings from a smartphone application for breast cancer patients. *J Med Internet Res* 2016 Aug 04;18(8):e216 [FREE Full text] [doi: [10.2196/jmir.5598](https://doi.org/10.2196/jmir.5598)] [Medline: [27492880](https://pubmed.ncbi.nlm.nih.gov/27492880/)]
26. 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](https://doi.org/10.1016/j.jad.2008.06.026)] [Medline: [18752852](https://pubmed.ncbi.nlm.nih.gov/18752852/)]
27. Plummer F, Manea L, Trepel D, McMillan D. Screening for anxiety disorders with the GAD-7 and GAD-2: A systematic review and diagnostic metaanalysis. *Gen Hosp Psychiatry* 2016;39:24-31. [doi: [10.1016/j.genhosppsych.2015.11.005](https://doi.org/10.1016/j.genhosppsych.2015.11.005)] [Medline: [26719105](https://pubmed.ncbi.nlm.nih.gov/26719105/)]
28. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: Validity of a brief depression severity measure. *J Gen Intern Med* 2001 Sep;16(9):606-613 [FREE Full text] [doi: [10.1046/j.1525-1497.2001.016009606.x](https://doi.org/10.1046/j.1525-1497.2001.016009606.x)] [Medline: [11556941](https://pubmed.ncbi.nlm.nih.gov/11556941/)]
29. Geier TJ, Hunt JC, Hanson JL, Heyrman K, Larsen SE, Brasel KJ, et al. Validation of abbreviated four- and eight-item versions of the PTSD Checklist for DSM-5 in a traumatically injured sample. *J Trauma Stress* 2020 Jun;33(3):218-226. [doi: [10.1002/jts.22478](https://doi.org/10.1002/jts.22478)] [Medline: [32277772](https://pubmed.ncbi.nlm.nih.gov/32277772/)]
30. Blevins CA, Weathers FW, Davis MT, Witte TK, Domino JL. The Posttraumatic Stress Disorder Checklist for DSM-5 (PCL-5): Development and initial psychometric evaluation. *J Trauma Stress* 2015 Dec;28(6):489-498. [doi: [10.1002/jts.22059](https://doi.org/10.1002/jts.22059)] [Medline: [26606250](https://pubmed.ncbi.nlm.nih.gov/26606250/)]
31. Vaishnavi S, Connor K, Davidson JRT. An abbreviated version of the Connor-Davidson Resilience Scale (CD-RISC), the CD-RISC2: Psychometric properties and applications in psychopharmacological trials. *Psychiatry Res* 2007 Aug 30;152(2-3):293-297 [FREE Full text] [doi: [10.1016/j.psychres.2007.01.006](https://doi.org/10.1016/j.psychres.2007.01.006)] [Medline: [17459488](https://pubmed.ncbi.nlm.nih.gov/17459488/)]
32. Connor KM, Davidson JRT. Development of a new resilience scale: The Connor-Davidson Resilience Scale (CD-RISC). *Depress Anxiety* 2003;18(2):76-82. [doi: [10.1002/da.10113](https://doi.org/10.1002/da.10113)] [Medline: [12964174](https://pubmed.ncbi.nlm.nih.gov/12964174/)]
33. Topp CW, Østergaard SD, Søndergaard S, Bech P. The WHO-5 Well-Being Index: A systematic review of the literature. *Psychother Psychosom* 2015;84(3):167-176 [FREE Full text] [doi: [10.1159/000376585](https://doi.org/10.1159/000376585)] [Medline: [25831962](https://pubmed.ncbi.nlm.nih.gov/25831962/)]
34. Bush K, Kivlahan DR, McDonnell MB, Fihn SD, Bradley KA. The AUDIT alcohol consumption questions (AUDIT-C): An effective brief screening test for problem drinking. Ambulatory Care Quality Improvement Project (ACQUIP). Alcohol Use Disorders Identification Test. *Arch Intern Med* 1998 Sep 14;158(16):1789-1795. [doi: [10.1001/archinte.158.16.1789](https://doi.org/10.1001/archinte.158.16.1789)] [Medline: [9738608](https://pubmed.ncbi.nlm.nih.gov/9738608/)]
35. King SDW, Fitchett G, Murphy PE, Pargament KI, Harrison DA, Loggers ET. Determining best methods to screen for religious/spiritual distress. *Support Care Cancer* 2017 Feb;25(2):471-479. [doi: [10.1007/s00520-016-3425-6](https://doi.org/10.1007/s00520-016-3425-6)] [Medline: [27714532](https://pubmed.ncbi.nlm.nih.gov/27714532/)]

36. Mount Sinai Calm: Virtual Classes. YouTube. URL: https://www.youtube.com/playlist?list=PLCT7BA-HcHlh1Zv0eO_3_jUV_Gzqy7gJM [accessed 2021-04-28]
37. McKay FH, Wright A, Shill J, Stephens H, Uccellini M. Using health and well-being apps for behavior change: A systematic search and rating of apps. JMIR Mhealth Uhealth 2019 Jul 04;7(7):e11926 [FREE Full text] [doi: [10.2196/11926](https://doi.org/10.2196/11926)] [Medline: [31274112](https://pubmed.ncbi.nlm.nih.gov/31274112/)]
38. Mobile apps: COVID Coach. National Center for PTSD, US Department of Veterans Affairs. URL: https://www.ptsd.va.gov/appvid/mobile/COVID_coach_app.asp [accessed 2021-04-28]
39. Calm. URL: <https://www.calm.com/> [accessed 2021-04-28]

Abbreviations

AUDIT-C: Alcohol Use Disorders Identification Test-Concise
CD-RISC2: 2-item Connor-Davidson Resilience Scale
CSRPG: Center for Stress, Resilience, and Personal Growth
DSM-5: Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition
GAD-2: 2-item Generalized Anxiety Disorder scale
HCW: health care worker
HIPAA: Health Insurance Portability and Accountability Act
ISMMS: Icahn School of Medicine at Mount Sinai
KPI: key performance indicator
mHealth: mobile health
NoSQL: non-Structured Query Language
PCL-5: Posttraumatic Stress Disorder Checklist for DSM-5
PHQ-8: 8-item Patient Health Questionnaire
PTSD: posttraumatic stress disorder
WHO-5: 5-item World Health Organization Well-Being Index

Edited by G Eysenbach; submitted 26.01.21; peer-reviewed by K Schroeder, B Nievas Soriano; comments to author 09.03.21; revised version received 24.03.21; accepted 13.04.21; published 05.05.21.

Please cite as:

Golden EA, Zweig M, Danieleto M, Landell K, Nadkarni G, Bottinger E, Katz L, Somarriba R, Sharma V, Katz CL, Marin DB, DePierro J, Charney DS

A Resilience-Building App to Support the Mental Health of Health Care Workers in the COVID-19 Era: Design Process, Distribution, and Evaluation

JMIR Form Res 2021;5(5):e26590

URL: <https://formative.jmir.org/2021/5/e26590>

doi:[10.2196/26590](https://doi.org/10.2196/26590)

PMID:[33872189](https://pubmed.ncbi.nlm.nih.gov/33872189/)

©Eddy A Golden, Micol Zweig, Matteo Danieleto, Kyle Landell, Girish Nadkarni, Erwin Bottinger, Lindsay Katz, Ricardo Somarriba, Vansh Sharma, Craig L Katz, Deborah B Marin, Jonathan DePierro, Dennis S Charney. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 05.05.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

A Novel Digital Patient-Reported Outcome Platform (Noona) for Clinical Use in Patients With Cancer: Pilot Study Assessing Suitability

Maria Kristiina Peltola¹, MD; Paula Poikonen-Saksela¹, MD, PhD; Johanna Mattson¹, MD, PhD; Timo Parkkari²

¹Department of Oncology, Helsinki University Hospital Comprehensive Cancer Center, Helsinki University, Helsinki, Finland

²Clinius Ltd, Helsinki, Finland

Corresponding Author:

Maria Kristiina Peltola, MD

Department of Oncology

Helsinki University Hospital Comprehensive Cancer Center

Helsinki University

Paciuksenkatu 3

Helsinki, 00029

Finland

Phone: 358 504900591

Fax: 358 94711

Email: maria.k.peltola@hus.fi

Abstract

Background: As the incidence of cancer is on the rise, there is a need to develop modern communication tools between patients and the medical personnel. Electronic patient-reported outcome (ePRO) measures increase the safety of cancer treatments and may have an impact on treatment outcome as well. ePRO may also provide a cost-efficient way to organize follow-up for patients with cancer. Noona is an internet-based system for patients to self-report symptoms and adverse events of cancer treatments from home via a computer or a smart device (eg, smartphone, tablet).

Objective: In this pilot study, we assessed the suitability of a novel ePRO application (Noona) for patients with cancer, nurses, and doctors at the Helsinki University Hospital, Finland.

Methods: The study included 44 patients with cancer (different solid tumor types) and 17 health care professionals (nurses or medical doctors). Patients were either operated or received systemic treatment or radiotherapy. Patients reported their symptoms to the medical staff via Noona. In addition, patients and clinicians answered a questionnaire, based on which Noona's suitability for clinical use was evaluated in terms of usability (ease of use, operability, and learnability), reliability (subjective opinion of the participant), and incidence of harmful events reported by the participants.

Results: A total of 41/44 (93%) patients and 15/17 (88%) professionals reported that the program was easy or quite easy to use; 38/44 (86%) patients and 11/17 (65%) professionals found Noona reliable, and 38/44 (86%) patients and 10/17 (59%) professionals would recommend Noona to other patients or their colleagues. No harmful incidences caused by the use of Noona were reported by the patients; however, 1 harmful incidence was reported by one of the professionals.

Conclusions: The majority of the participants felt that Noona appeared reliable and it was easy to use. Noona seems to be a useful tool for monitoring patient's symptoms during cancer therapy. Future studies will determine the impact of this ePRO platform in routine clinical practice.

(*JMIR Form Res* 2021;5(5):e16156) doi:[10.2196/16156](https://doi.org/10.2196/16156)

KEYWORDS

electronic patient-reported outcome; adverse events; patients with cancer

Introduction

According to the Finnish Cancer Registry, approximately 30,000 people receive cancer diagnoses in Finland every year, and the number of patients with cancer continues to rise.

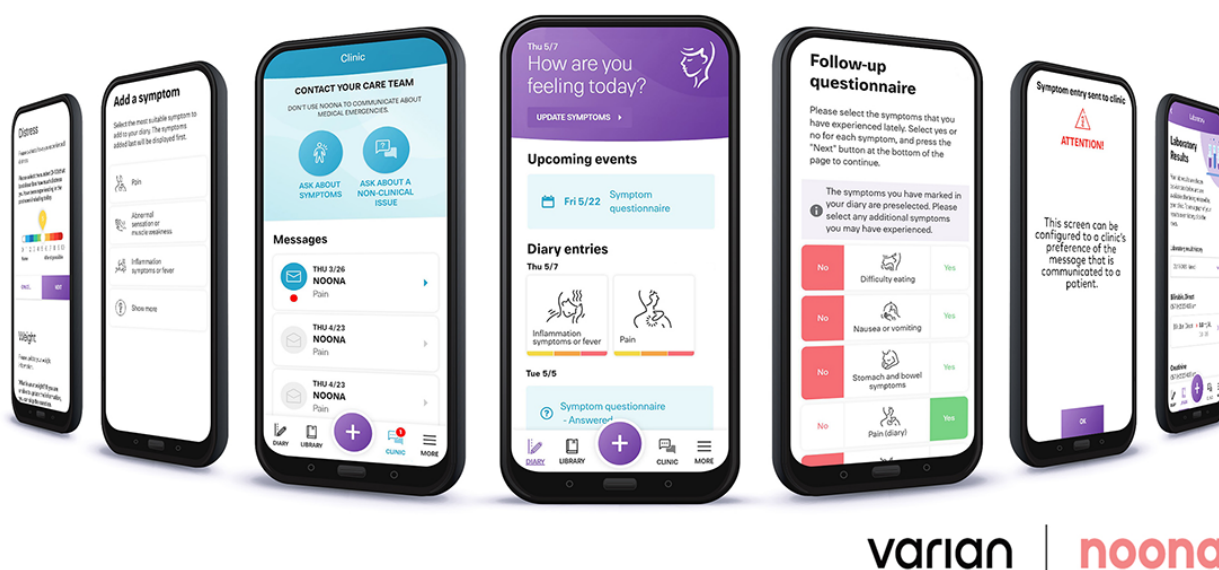
Cancer treatments cause adverse events and long-term consequences. Patients receiving radiotherapy may experience irritation, redness of skin, pain, and fatigue. If the acute reactions remain improperly treated, patients may discontinue radiotherapy, leading to a loss of local control [1].

During chemotherapy patients commonly experience side effects such as nausea, fatigue, neutropenic infections, mucositis, peripheral neuropathy, and pain. These side effects impair patients' quality of life and may require emergency room visits, hospital stays, reductions in the following chemotherapy doses, or lead to treatment interruption [2]. Recognizing adverse events of chemotherapy early remains important to ensure proper medical interventions [3].

Digital communication between patients and cancer clinics via electronic patient-reported outcomes (ePROs) enables early detection of adverse events during chemotherapy, decreases emergency room visits and hospitalization, increases quality of life, and may even improve survival [3]. In the study by Denis et al [4], after primary treatment of lung cancer, disease relapse was detected earlier and the patients lived longer if they reported their symptoms via electronic software, compared with traditional follow-up visits. The overall survival was 19 months (95% CI 12.5 to noncalculable) in the study arm, compared with 12 months (95% CI 8.6-16.4, $P=.001$) in the control arm [4]. Because of these advantages, ePROs will likely be implemented in routine cancer care in the near future [5].

The aim of this study is to describe the usability of the first version of Noona, a web-mediated PRO application.

Figure 1. Noona interface for the patients.

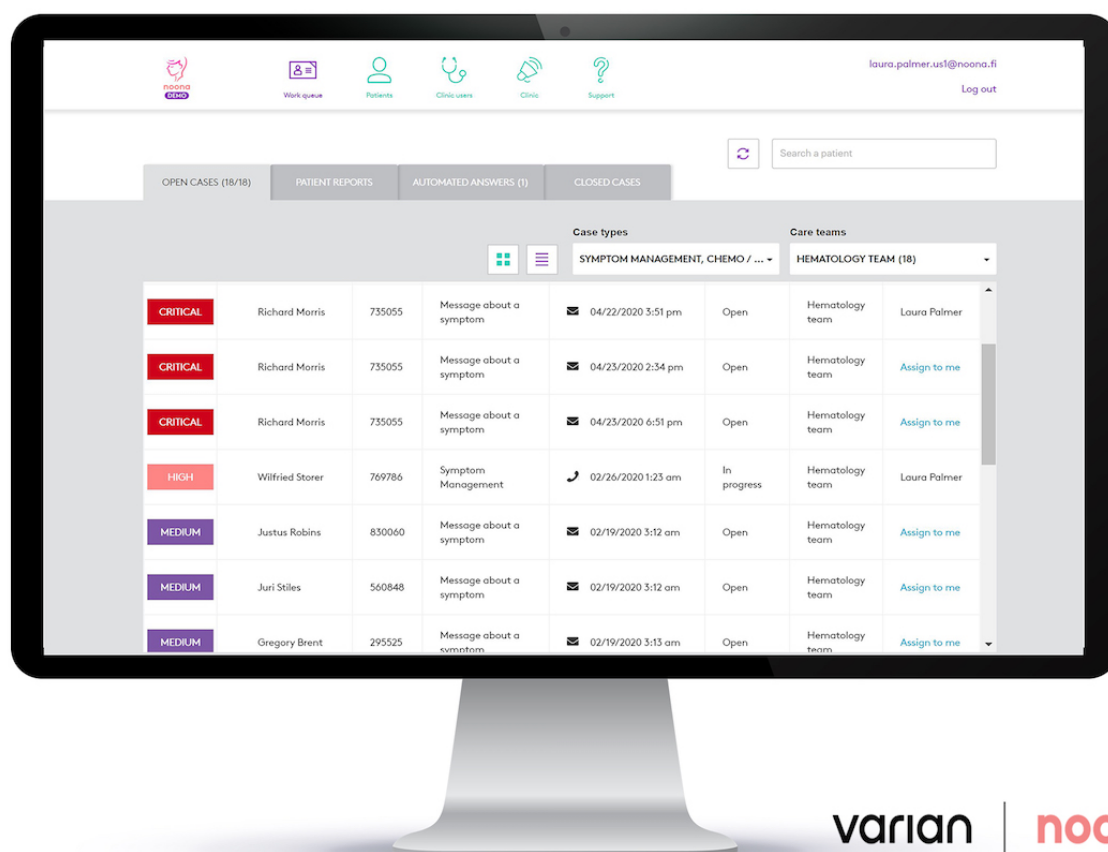


Methods

The investigated ePRO tool Noona is a web-mediated application developed by Noona Healthcare Oy in collaboration with Helsinki University Hospital Comprehensive Cancer Center. The Noona mobile service is developed for remote monitoring of patients with cancer, and to be used as a support tool for communication between patients with cancer and health care professionals. Noona is an online application, which can be used with a web browser and a suitable device, such as desktop, laptop, tablet, and smartphone. Noona has 2 clear user groups: patients with cancer and cancer care professionals, specifically nurses and doctors working in cancer hospitals. At the beginning of their treatment, patients with cancer are registered to Noona, and they will continue to use Noona during the follow-up and rehabilitation periods. During the treatment phase, Noona is designed to evaluate symptoms and recovery progress based on patient-reported outcome data. During the follow-up and rehabilitation periods, the intended use is to enable fluent and accessible communication between the patient and the professionals, and to monitor patient recovery from cancer and related symptoms.

Noona has 2 user interfaces: one for the patients (Figure 1) and the other for the professionals (Figure 2). The main functionalities of the patient interface are symptom reporting and a diary. Patients can report on cancer- and treatment-related symptoms using question wizards that cover the clinically relevant questions and evaluate the most common and relevant symptoms, such as pain, fatigue, nausea, vomiting, and bowel symptoms. In addition to the symptoms, patients may contact their clinic regarding other topics using an open question form. Patients can also receive messages from their care team via Noona.

Figure 2. Noona interface for the professionals.



varian | noona

The main functionalities for the professionals are (1) a work queue to monitor new patients who have asked assistance or responded to a scheduled questionnaire, (2) a view of patient information as well as symptom history, and (3) a possibility to send messages directly to patients. When a patient starts to use Noona, the application provides a tutorial for the patient. As part of its implementation in the hospital, nurses are trained to give patients end user support. In the first version of Noona no alarm or reminder function was included and professionals were asked to check Noona regularly as part of their clinical work.

As part of the clinical evaluation of Noona, a prospective pilot trial was designed to study its usability and reliability by recruiting both patients with cancer and medical personnel, as both groups are expected to use the software during cancer treatments (ie, chemotherapy and radiotherapy or after cancer surgery). This clinical evaluation was part of the mandatory process before introducing Noona as a medical device to the markets in Finland and countries in Europe.

Patients were eligible if they were at least 18 years old; could read Finnish; received either radiotherapy, surgery, or medical therapy as a treatment for their cancer; and had a computer, phone, or tablet with an internet access available to be able to use Noona. Patients were asked to participate in the study by their physician during their routine visit at the hospital; otherwise their treatment continued according to local guidelines. All the doctors and nurses treating patients with cancer in the Helsinki University Comprehensive Cancer Center

and willing to participate were eligible. Professionals (nurses, surgeons, medical oncologists, and radiotherapists) were recruited by the research doctor. A written informed consent was obtained from all participants. At the beginning, professionals had a training session about the use of Noona.

Patients were trained individually on the use of Noona by the research nurses and were introduced to the online tutorial. The information collected via Noona corresponded to that normally obtained in the clinical practice, and no additional information was requested. In case of a technical problem or if participants did not get a response from Noona within 24 hours, they were advised to contact their care team (nurse) by phone. Emergency patients were advised to contact their local emergency room.

A total of 45 patients with cancer and 18 health care professionals (nurses or medical doctors) were initially recruited at the Helsinki University Hospital Comprehensive Cancer Center in Finland in 2016. None of the participants declined to participate in the study; 1 patient and 1 professional were later excluded as they failed to fill in the questionnaire. Thus, the final number of patients and professionals recruited was 44 and 17, respectively. The study period was from June 26 to October 27, 2016, for the patients and from September 6 to November 1, 2016, for the professionals.

During the study period, the patients reported their symptoms and adverse events after surgery or during radiotherapy on Noona instead of reporting these via phone or at the doctor's appointment. During chemotherapy the participating nurses sent

their patients structured questionnaires via Noona a couple of days before chemotherapy infusion instead of contacting them over the phone. In addition, all patients could report using Noona how they are doing at any time. At the end of the study period, all participants answered questionnaires on the usability and reliability of Noona ([Multimedia Appendix 1](#)). Outcome measures of usability were ease of use, operability, and learnability. Participants were also asked about reliability (subjective opinion about Noona), whether they would recommend Noona to others, and if there was any harmful event related to the use of Noona. The questionnaires are described in detail in [Multimedia Appendix 1](#).

The study was approved by the Ethics Committee of the Helsinki University Hospital.

Results

Patients

The mean age of the patients (n=44) was 55 (range 28-79); 37/44 (84%) patients were female and 7/44 (16%) were male. The cancer types were breast 32/44 (73%), gynecological 1/44 (2%), melanoma 3/44 (7%), colorectal 3/44 (7%), and urological 3/44 (7%). The information on cancer type was missing from

2/44 (5%) patients. Patients were either operated (13/44, 30%), received chemotherapy (16/44, 36%), or radiotherapy (14/44, 32%). The information on treatment type was missing from one patient (2%).

The patients were asked about their overall activity in using internet, mobile services, and their experiences of using Noona. About 70% (30/44, 68%) of the patients reported using mobile or web services every day (eg, when banking, shopping, or using social media). During the study Noona was used 1-3 times per week by 15/44 patients (34%) and less than that by 29/44 (66%). The most popular devices for the use of Noona were computer (19/44 patients, 43%), smartphone (10/44 patients, 23%), and tablet (7/44 patients, 16%), whereas others used more than 1 of these mentioned devices.

The detailed information about patients' feedback on the usability of Noona in general, sending messages, and reporting of symptoms functions; their preference; and recommendation is presented in [Table 1](#). Nearly 93% (41/44) of patients reported that the program was easy or quite easy to use, 38/44 (86%) patients found Noona reliable, and 38/44 (86%) patients would recommend Noona to other patients or their colleagues. No harmful incidences caused by the use of Noona were reported by the patients.

Table 1. Assessment of Noona's usability by patients and professionals.

Question and parameter	Patients (n=44), n (%)	Professionals (n=17), n (%)
Logging was		
Easy	30 (68)	14 (82)
Not easy or difficult	12 (27)	3 (18)
Difficult	2 (5)	0 (0)
Not answered	0 (0)	0 (0)
Enough instructions/education for use		
Yes	37 (84)	16 (94)
No	6 (14)	1 (6)
Not answered	1 (2)	0 (0)
Use of Noona was		
Easy	29 (66)	5 (29)
Quite easy	12 (27)	10 (59)
Quite difficult	1 (2)	1 (6)
Difficult	0 (0)	0 (0)
Not answered	2 (5)	1 (6)
Found message function		
Easy	22 (50)	12 (71)
Difficult	0 (0)	2 (12)
Not used the function	22 (50)	2 (12)
Not answered	0 (0)	1 (6)
Found the side effect reporting function		
Easy	20 (45)	Not asked in this group
Not easy or difficult	3 (7)	
Difficult	0 (0)	
Not used the function	21 (48)	
Preference for communication		
Noona	7 (16)	Not asked in this group
Telephone	3 (7)	
No preference	12 (27)	
Not answered	22 (50)	
Noona was reliable		
Yes	38 (86)	11 (65)
No	2 (5)	4 (24)
Not answered	4 (9)	2 (12)
Would recommend Noona		
Yes	38 (86)	10 (59)
No	2 (5)	2 (12)
Not answered or no comment	4 (9)	5 (29)

Health Care Professionals

A total of 7 physicians and 10 nurses participated in the study. The mean age of the health care professionals (n=17) was 43

(range 28-58), and 14/17 (82%) of them were female, 1/17 (6%) was male, and 2/17 (12%) did not report their gender. The specialties of the doctors were surgery (2/7, 29%), medical oncology (3/7, 43%), and radiotherapy (2/7, 29%); 3/10 (30%)

nurses worked in the department of surgery, 5/10 (50%) in the department of chemotherapy, and 2/10 (20%) in the department of radiotherapy. All health care professionals reported, in general, using mobile or web services every day (several times a day). During the study, 8/17 (47%) professionals used Noona daily, 5/17 (29%) 1-3 times per week, and 4/17 (24%) less than that. A total of 3/7 (43%) doctors used Noona once a week while 4/7 (57%) used less than that; 8/10 nurses (80%) used Noona daily and 2/10 (20%) 1-3 times per week. The detailed information on the professionals' feedback related to the use of Noona in general, message functions, their preference, and recommendation of Noona is presented in [Table 1](#). As much as 15/17 (88%) professionals reported that the program was easy or quite easy to use and 11/17 (65%) professionals found Noona reliable; 10/17 (59%) professionals would recommend Noona to other patients or colleagues. One professional reported a harmful event with the use of Noona.

Participants spontaneously reported via the ePRO system that Noona "was easy to use" and "system was safe." Some participants (13 patients and 12 professionals) gave suggestions for improvements. Recommendations included adding an alarm function when a new message has arrived, adding capability for users to send pictures through the program, possibility for logging in automatically with a saved password, and having more options for the questions.

Discussion

Principal Findings

This pilot study suggests that an ePRO application called Noona is feasible and acceptable in clinical practice. Most patients and professionals found Noona easy to log in and easy or quite easy to use. None of the patients reported difficulty with using Noona. Most of both patients and professionals would recommend Noona to other patients or colleagues.

None of the patients and only 1 professional reported a harmful event related to the use of Noona. In that case, the patient (female) did not understand that her symptoms were related to the treatment of cancer and thus did not report them via Noona; however, as per the professional, a direct face-to-face contact would have probably clarified the nature of her symptoms. None of the participants declined to participate in the study and only 2 participants were excluded from the study as they did not fill

in the questionnaire. The high compliance rate suggests that participants found Noona easy enough to use. Only 2 patients had initial technical problems logging in, but these were subsequently resolved. One explanation for these promising results is that Noona has been designed in collaboration with the Helsinki University Hospital Comprehensive Cancer Center and the needs and requirements of both patients and professionals were recognized while developing Noona. A possible limitation of the study is that most participants had good digital literacy, which can limit the expansion of the results to different countries and populations.

As much as 8/10 (80%) of the nurses used Noona daily while none of the doctors did. This finding also reflects the current clinical practice, where most of the communication happens between patients and nurses.

Our results are in line with previous studies utilizing electronic data-collection tools [1,6-11], but currently no data comparing Noona with the other tools in a randomized setting are available. In their study of patients with breast cancer, Abernethy et al [6] demonstrated a high level of patient compliance and satisfaction using a tablet-based data-collection system. According to the study authors, the ePRO helped patients to identify symptoms that deserve reporting to their cancer care provider.

Advances in information technology have enabled the utilization of many ePRO systems in cancer clinics [12]. ePRO collection provides a unique opportunity to monitor symptoms in real-time and provide clinical management during cancer care [9]. Incorporation of the ePRO tool in clinical practice thus generates the opportunity to collect patient data via a comprehensive system [6]. Some studies have shown that ePROs have a positive impact on patients' satisfaction [4,13], whereas others have found that ePROs are both feasible and acceptable [5,14-16] in clinical practice because published data show that their use may improve patient's quality of life and even prognosis [4]. Because of these advantages, there is a growing interest in the use of internet-based follow-up systems.

Conclusion

Noona seems to be an easy-to-use and suitable tool to monitor patient-reported outcomes during cancer treatments. However, larger studies are needed to compare ePROs with traditional methods of contact with regard to patient preference, quality of life, resource utility, and costs.

Acknowledgments

We thank the study nurses, physicians, and patients. Noona Healthcare Oy (currently Varian) had no influence over the content of this article.

Conflicts of Interest

JM has received research funding from Noona Healthcare Oy. The remaining authors have nothing to declare.

Multimedia Appendix 1

Questionnaire on the usability and reliability of Noona.

[[DOCX File, 16 KB - formative_v5i5e16156_app1.docx](#)]

References

1. Peltola MK, Lehtikainen JS, Sippola LT, Saarilahti K, Mäkitie AA. A Novel Digital Patient-Reported Outcome Platform for Head and Neck Oncology Patients—A Pilot Study. *Clin Med Insights Ear Nose Throat* 2016 Sep 27;9:CMEN.S40219. [doi: [10.4137/cment.s40219](https://doi.org/10.4137/cment.s40219)]
2. Knoerl R, Gray E, Stricker C, Mitchell SA, Kippe K, Smith G, et al. Electronic versus paper-pencil methods for assessing chemotherapy-induced peripheral neuropathy. *Support Care Cancer* 2017 Jun 2;25(11):3437-3446. [doi: [10.1007/s00520-017-3764-y](https://doi.org/10.1007/s00520-017-3764-y)]
3. 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. *JCO* 2016 Feb 20;34(6):557-565. [doi: [10.1200/jco.2015.63.0830](https://doi.org/10.1200/jco.2015.63.0830)]
4. Denis F, Lethrosne C, Pourel N, Molinier O, Pointreau Y, Domont J, et al. Randomized Trial Comparing a Web-Mediated Follow-up With Routine Surveillance in Lung Cancer Patients. *J Natl Cancer Inst* 2017 Sep 01;109(9):1029. [doi: [10.1093/jnci/djx029](https://doi.org/10.1093/jnci/djx029)] [Medline: [28423407](https://pubmed.ncbi.nlm.nih.gov/28423407/)]
5. LeBlanc TW, Abernethy AP. Patient-reported outcomes in cancer care - hearing the patient voice at greater volume. *Nat Rev Clin Oncol* 2017 Dec 4;14(12):763-772. [doi: [10.1038/nrclinonc.2017.153](https://doi.org/10.1038/nrclinonc.2017.153)] [Medline: [28975931](https://pubmed.ncbi.nlm.nih.gov/28975931/)]
6. Abernethy AP, Herndon JE, Wheeler JL, Day JM, Hood L, Patwardhan M, et al. Feasibility and acceptability to patients of a longitudinal system for evaluating cancer-related symptoms and quality of life: pilot study of an e/ Tablet data-collection system in academic oncology. *J Pain Symptom Manage* 2009 Jun;37(6):1027-1038 [FREE Full text] [doi: [10.1016/j.jpainsymman.2008.07.011](https://doi.org/10.1016/j.jpainsymman.2008.07.011)] [Medline: [19394793](https://pubmed.ncbi.nlm.nih.gov/19394793/)]
7. Andikyan V, Rezk Y, Einstein M, Zivanovic O, Leitao M, Sonoda Y, et al. A prospective study of the feasibility and acceptability of a web-based, electronic patient-reported outcomes system in assessing patient recovery after major gynecologic cancer surgery. *Gynecologic Oncology* 2012 Mar;125:S118. [doi: [10.1016/j.ygyno.2011.12.285](https://doi.org/10.1016/j.ygyno.2011.12.285)]
8. Holzner B, Giesinger JM, Pinggera J, Zugals S, Schöpf F, Oberguggenberger AS, et al. The Computer-based Health Evaluation Software (CHES): a software for electronic patient-reported outcome monitoring. *BMC Med Inform Decis Mak* 2012 Nov 09;12(1):126 [FREE Full text] [doi: [10.1186/1472-6947-12-126](https://doi.org/10.1186/1472-6947-12-126)] [Medline: [23140270](https://pubmed.ncbi.nlm.nih.gov/23140270/)]
9. Bennett AV, Jensen RE, Basch E. Electronic patient-reported outcome systems in oncology clinical practice. *CA Cancer J Clin* 2012 Jul 18;62(5):337-347 [FREE Full text] [doi: [10.3322/caac.21150](https://doi.org/10.3322/caac.21150)] [Medline: [22811342](https://pubmed.ncbi.nlm.nih.gov/22811342/)]
10. Basch E, Barbera L, Kerrigan CL, Velikova G. Implementation of Patient-Reported Outcomes in Routine Medical Care. *American Society of Clinical Oncology Educational Book* 2018 May(38):122-134. [doi: [10.1200/edbk_200383](https://doi.org/10.1200/edbk_200383)]
11. Basch E, Artz D, Dulko D, Scher K, Sabbatini P, Hensley M, et al. Patient Online Self-Reporting of Toxicity Symptoms During Chemotherapy. *JCO* 2005 May 20;23(15):3552-3561. [doi: [10.1200/jco.2005.04.275](https://doi.org/10.1200/jco.2005.04.275)]
12. Jensen RE, Snyder CF, Abernethy AP, Basch E, Potosky AL, Roberts AC, et al. Review of Electronic Patient-Reported Outcomes Systems Used in Cancer Clinical Care. *JOP* 2014 Jul;10(4):e215-e222. [doi: [10.1200/jop.2013.001067](https://doi.org/10.1200/jop.2013.001067)]
13. 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 Jun 11;13(1):211 [FREE Full text] [doi: [10.1186/1472-6963-13-211](https://doi.org/10.1186/1472-6963-13-211)] [Medline: [23758898](https://pubmed.ncbi.nlm.nih.gov/23758898/)]
14. Abernethy AP, Zafar SY, Uronis H, Wheeler JL, Coan A, Rowe K, et al. Validation of the Patient Care Monitor (Version 2.0): a review of system assessment instrument for cancer patients. *J Pain Symptom Manage* 2010 Oct;40(4):545-558 [FREE Full text] [doi: [10.1016/j.jpainsymman.2010.01.017](https://doi.org/10.1016/j.jpainsymman.2010.01.017)] [Medline: [20579839](https://pubmed.ncbi.nlm.nih.gov/20579839/)]
15. Snyder CF, Jensen R, Courtin SO, Wu AW, Website for Outpatient QOL Assessment Research Network. PatientViewpoint: a website for patient-reported outcomes assessment. *Qual Life Res* 2009 Sep 21;18(7):793-800 [FREE Full text] [doi: [10.1007/s11136-009-9497-8](https://doi.org/10.1007/s11136-009-9497-8)] [Medline: [19544089](https://pubmed.ncbi.nlm.nih.gov/19544089/)]
16. Sztankay M, Nepl L, Wintner LM, Loth FL, Willenbacher W, Weger R, et al. Complementing clinical cancer registry data with patient reported outcomes: A feasibility study on routine electronic patient-reported outcome assessment for the Austrian Myelome Registry. *Eur J Cancer Care (Engl)* 2019 Nov;28(6):e13154 [FREE Full text] [doi: [10.1111/ecc.13154](https://doi.org/10.1111/ecc.13154)] [Medline: [31465136](https://pubmed.ncbi.nlm.nih.gov/31465136/)]

Abbreviations

ePRO: electronic patient-reported outcome

Edited by G Eysenbach; submitted 06.09.19; peer-reviewed by MS Duh, A Sharma, E Khoong; comments to author 18.05.20; revised version received 07.09.20; accepted 11.03.21; published 06.05.21.

Please cite as:

Peltola MK, Poikonen-Saksela P, Mattson J, Parkkari T

A Novel Digital Patient-Reported Outcome Platform (Noona) for Clinical Use in Patients With Cancer: Pilot Study Assessing Suitability
JMIR Form Res 2021;5(5):e16156

URL: <https://formative.jmir.org/2021/5/e16156>

doi: [10.2196/16156](https://doi.org/10.2196/16156)

PMID: [33955841](https://pubmed.ncbi.nlm.nih.gov/33955841/)

©Maria Kristiina Peltola, Paula Poikonen-Saksela, Johanna Mattson, Timo Parkkari. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 06.05.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Usability and Acceptance of Wearable Biosensors in Forensic Psychiatry: Cross-sectional Questionnaire Study

Pieter Christiaan de Looff^{1,2,3}, PhD; Henk Nijman^{1,3}, PhD; Robert Didden^{1,4}, PhD; Matthijs L Noordzij⁵, PhD

¹Behavioural Science Institute, Radboud University, Nijmegen, Netherlands

²De Borg, Den Dolder, Netherlands

³Fivoor Science and Treatment Innovation, Den Dolder, Netherlands

⁴Trajectum, Zwolle, Netherlands

⁵Department of Psychology, Health and Technology, University of Twente, Enschede, Netherlands

Corresponding Author:

Pieter Christiaan de Looff, PhD

Behavioural Science Institute

Radboud University

Postbus 9104

Nijmegen

Netherlands

Phone: 31 030 2256405

Email: peterdelooff@gmail.com

Abstract

Background: The use of wearable biosensor devices for monitoring and coaching in forensic psychiatric settings yields high expectations for improved self-regulation of emotions and behavior in clients and staff members. More so, if clients have mild intellectual disabilities (IQ 50-85), they might benefit from these biosensors as they are easy to use in everyday life, which ensures that clients can practice with the devices in multiple stress and arousal-inducing situations. However, research on (continuous) use and acceptance of biosensors in forensic psychiatry for clients with mild intellectual disabilities and their caretakers is scarce. Although wearable biosensors show promise for health care, recent research showed that the acceptance and continuous use of wearable devices in consumers is not as was anticipated, probably due to low expectations.

Objective: The main goal of this study was to investigate the associations between and determinants of the expectation of usability, the actual experienced usability, and the intention for continuous use of biosensors.

Methods: A total of 77 participants (31 forensic clients with mild intellectual disabilities and 46 forensic staff members) participated in a 1-week trial. Preceding the study, we selected 4 devices thought to benefit the participants in domains of self-regulation, physical health, or sleep. Qualitative and quantitative questionnaires were used that explored the determinants of usability, acceptance, and continuous use of biosensors. Questionnaires consisted of the System Usability Scale, the Technology Acceptance Model questionnaire, and the extended expectation confirmation model questionnaire.

Results: Only the experienced usability of the devices was associated with intended continuous use. Forensic clients scored higher on acceptance and intention for continuous use than staff members. Moderate associations were found between usability with acceptance and continuous use. Staff members showed stronger associations between usability and acceptance ($r=.80$, $P<.001$) and usability and continuous use ($r=.79$, $P<.001$) than clients, who showed more moderate correlations between usability and acceptance ($r=.46$, $P=.01$) and usability and continuous use ($r=.52$, $P=.003$). The qualitative questionnaires in general indicated that the devices were easy to use and gave clear information.

Conclusions: Contrary to expectations, it was the actual perceived usability of wearing a biosensor that was associated with continuous use and to a much lesser extent the expectancy of usability. Clients scored higher on acceptance and intention for continuous use, but associations between usability and both acceptance and continuous use were markedly stronger in staff members. This study provides clear directions on how to further investigate these associations. For example, whether this is a true effect or due to a social desirability bias in the client group must be investigated. Clients with mild intellectual disabilities might benefit from the ease of use of these devices and their continuing monitoring and coaching apps. For these clients, it is especially important to develop easy-to-use biosensors with a minimum requirement on cognitive capacity to increase usability, acceptance, and continuous use.

KEYWORDS

forensic psychiatry; wearable biosensors; intellectual disabilities; usability; acceptance; continuous use; emotion regulation; behavior regulation

Introduction

The use of wearable biosensor devices for monitoring and coaching in forensic psychiatric settings for people with intellectual disabilities and their caretakers yields high expectations for improved self-regulation of emotions and behavior. This is based on the expectation that wearable biosensor devices can be used to detect changing levels of emotional states [1] and behavior in both clients and staff members [2-5]. The devices typically collect body measurements such as heart rate, blood pressure, movement, and breathing [6,7]. It is assumed that this information can be used to quantify health, physiological stress, and well-being [8]. Users can, for instance, visualize their data in smartphone apps to track their health and stress levels or receive real-time information on their heart rate coupled to personalized, prespecified interventions [2]. This information is thought to enhance health and well-being and alleviate stress [2,6,8]. Besides the detection and prediction of emotional states and behavior, wearable biosensors may be used to promote a healthy lifestyle, especially because of their real-time data monitoring capabilities [3,9]. For example, they can be used to monitor sleep [10], nutrition, movement, cardiac disease, blood sugar [3], or epilepsy [8]. In these cases, wearable technology has the potential to lower medical costs, improve health-related behavior of users, and reduce physician time [3]. In addition, it may result in opportunities to gather accurate real-time data that will allow for the diagnosis, prevention, and treatment of various chronic diseases in a more economical manner [11-13]. This makes the potential economic benefits of these biosensors immense; health and fitness biosensors are targeted at improving healthy behavior, which will likely have a significant impact on health care costs [11]. Although some results from recent research in forensic psychiatry are promising [4,14], the (continuous) use of biosensors in everyday life, particularly in forensic psychiatry, is still in nascent stages [15].

A potentially complicating factor in the use of these biosensors are the mild intellectual disabilities and borderline intellectual functioning (MID-BIF; IQ 50-85) of the user. Clients with MID-BIF might not benefit from cognitive behavioral therapies (eg, anger management) like people with average intelligence [16] due to the complexity of such therapies. The use of biosensors might be an easier method by which to teach people about arousal-inducing events and possible self-regulation strategies. Moreover, the ambulatory nature of the devices ensures that clients can practice with the devices in multiple stress-inducing situations occurring in real life.

Although wearable biosensors show promise for health care, research into the use and acceptance of wearable biosensors is almost absent in forensic psychiatry, let alone in clients with MID-BIF. Recent consumer research, however, showed that the acceptance of wearable devices in consumers is not as was anticipated [3], and, more importantly, the long-term continuous

use (presumably following acceptance) of these devices seems low [17]. In addition, there is a need for more longitudinal research to systematically study the trends in wearable use over time. Pal et al [12] recently identified factors that might contribute to the low frequency of continuous use of wearable devices. According to these authors, there is a gap between the expectations of usability that people have before using, for example, a smartwatch and the factors that would lead to continuous use by experiences with the device (expectation–experience–continuous use). In this study, we will investigate whether it is the expectation of using the biosensor or the actual experience itself that contributes to continuous use. Examples of factors associated with usability, acceptance, and continuous use of wearable biosensor devices include the accuracy of the information from the devices, reliability and validity of the information, comfort of the devices, feedback provided, and how it is provided [3]. For instance, if a wearable biosensor device signals that a client is stressed while the client is clearly at ease, this will likely reduce the willingness to use the biosensor, as the accuracy, reliability, and validity of the information is clearly erroneous in this example.

To increase the usability, acceptance, and continuous use of biosensors, user preferences, needs, and wishes, especially for people with MID-BIF, must be known. In addition, it is necessary to determine the goals of the user and the tasks for which the biosensors will be used. Finally, the functions of user interfaces and biosensor devices should be evaluated to make them more attractive, desirable, and efficient for the users by integrating the outcomes of the evaluation [17,18]. Kim and Shin [15] argued that additional antecedents of biosensor adoption and the role of control variables should be further investigated to increase the (continuous) use and acceptance in diverse international samples. In addition, Kalantari [11] argued for more diverse samples and heterogeneous user groups. Kalantari [11] also notes that qualitative research is lacking in the field of acceptance and adoption research.

Research on the use of biosensors for clients with MID-BIF is scarce while the potential benefits might be significant. Therefore, we investigated the use of everyday wearable biosensors to establish what would lead to their (continuous) use and acceptance. Biosensor information could potentially benefit not only clients but staff members as well. To this end, qualitative and quantitative questionnaires were used to explore the psychological (preferences, needs, wishes, and goals) and functional (tasks and functions) determinants of usability, acceptance, and continuous use of biosensors in both staff members and clients. The main goal of this study was to investigate the expectation–experience–continuous use connection to see whether there is a gap between expectations of usability and the actual experience that will preferably lead to continuous use and whether there are differences between clients and staff members. In addition, we investigated the key

determinants involved in the usability, acceptance, and continuous use of biosensors using validated questionnaires. The following research questions were formulated:

- Are there differences between clients and staff members in expectations of usability and the actual experienced usability that will lead to continuous use of biosensors (expectation–experience–continuous use)?
- Which key determinants contribute to the usability, acceptance, and continuous use of biosensors in forensic psychiatry for clients with MID-BIF and staff members?

Methods

Participants and Setting

The participants for this small-scale study consisted of two groups of users: clients with MID-BIF who are residents of forensic psychiatric living units and staff members who work as nurses or sociotherapists on these forensic psychiatric living units. Clients are often referred to the units as a result of aggressive and violent behavior and are at an increased risk for severe behavior problems, offending behavior, and recidivism [19]. During their admission, these clients are encouraged to participate in treatments aimed at decreasing the risk for recidivism. The staff members who work with these clients often work irregular shifts and are at an increased risk for work-related stress and burnout symptoms [4].

Materials

After multiple sessions with a user group consisting of staff members and clients, we selected 4 devices that were thought to benefit the participants in domains of self-regulation, physical health, or sleep. All 4 devices used in this study are US Food and Drug Administration and CE approved and can be bought in regular stores (commercially available).

The Spire Stone (Spire Health) is a wearable device in the form of a stone that can be attached to a belt (men) or bra (woman). It measures the contraction of the torso to indicate the rate of breathing. The device comes with an app and classifies the respiration rate as calm, focused, or tense (see Holt et al [20], for instance). Moreover, it measures the amount of activity and sedentary behavior. The app provides users with daily overviews or direct feedback on breathing rate.

The Charge 3 (Fitbit Inc) is a physical activity tracker with a heart rate monitor that provides users with real-time feedback on heart rate and physical activity. In addition, it can provide users with information on sleep and exercise. The app provides users with detailed information on stress, sleep, and activity (see Schragger et al [21], for instance).

The vivosmart 4 (Garmin Ltd) is a physical activity tracker with a heart rate monitor that provides users with real-time feedback on energy expenditure, stress indications based on heart rate, and sleep quality assessment. The app provides users with detailed information on sleep, stress, and energy expenditure [22].

The TicWatch E (Mobvoi Inc) is a smartwatch running WearOS with a heart rate sensor. It can be used as a biofeedback device when running the Sense-It app [2], an ambulatory e-coaching

app that provides users with information on deviations in their regular heart rate. It is aimed at supporting users to better understand and recognize changes in their arousal levels.

Questionnaires

To assess determinants of usability, we evaluated user satisfaction with the wearable biosensors [3] with the System Usability Scale (SUS), a short 10-item questionnaire. To assess determinants of acceptance, we used the Technology Acceptance Model (TAM) questionnaire, one of the most frequently used theoretical frameworks for the acceptance of new technology [15]. To assess determinants of continuous use, we used a questionnaire devised by Pal et al [12] based on an extended expectation confirmation model (EECM) that consists of 10 prime factors associated with the continuous use of smartwatches.

As the questionnaires were not available in Dutch, they were translated by 3 researchers and 8 staff members who work in forensic psychiatric settings with MID-BIF clients. The questionnaires were then back-translated by native English speakers. As the formulation of the questions was deemed too complex for the MID-BIF clients, an easier version of all 3 questionnaires was constructed for MID-BIF clients consisting of fewer, more easily formulated questions. The choice of which questions to select for the short version for clients was made by the researchers and staff members based on two key principles: the question should easily be understood by the MID-BIF client and represent the implied construct to be measured.

Usability

The SUS is a 10-item questionnaire with good reliability (.85) [23,24] that can be used to assess determinants of subjective usability. It is widely used to assess the usability of different types of technology such as medical devices, software, and websites. It has a well-established standard reference [3] and is quick to administer. The SUS is scored on a 5-point Likert scale ranging from strongly disagree to strongly agree. The total score is obtained by adding the positively worded items minus 1. For negatively worded items, the score is subtracted from 5. The scores are then summed and multiplied by 2.5 to obtain a value between 0 and 100. Missing values can be assigned a neutral value of 3 in accordance with recommendations [25].

To answer the first research question on the expectation–experience–continuous use connection, the questionnaire was administered twice. The SUS questionnaire was administered preceding the study to measure the expectation of the participants. Following 1 week, the SUS was administered to measure the actual experience with the biosensors.

Acceptance

The TAM questionnaire [26] is one of the most frequently used questionnaires for the acceptance of new technology [27]. The model uses a generalized theoretical framework of technology. We used a recently developed version of this questionnaire [15] specifically aimed at smartwatches that distinguishes between 10 key determinants of smartwatches: perceptions of and attitudes toward technology, affective quality, relative

advantage, mobility, availability, perceived ease of use, intention to use, perceived usefulness, subcultural appeal, and cost. The TAM is scored on a 7-point Likert scale ranging from strongly disagree to strongly agree. The questionnaire consists of 36 questions, and all scales have reliabilities well over .70.

Continuous Use

Pal et al [12] proposed an EECM consisting of 10 prime factors associated with the continuous use of smartwatches: hedonic motivation, self-socio motivation, perceived privacy, perceived comfort, battery-life concern, perceived accuracy with functional limitations, perceived usefulness, confirmation, satisfaction, and continuous use. The EECM is scored on a 7-point Likert scale ranging from strongly disagree to strongly agree. The calculation of the total EECM score follows a similar logic as the SUS calculation. The positively phrased EECM questions were summed with the value of the score minus 1. The negatively phrased EECM questions were scored with a value of 7 minus the score (note that we did not multiply the sum as is common with the SUS). This was thought to give an adequate indication of continuous use intention. The EECM questionnaire consists of 32 questions, and all scales have reliabilities well over .70.

Short Questionnaires

The full TAM and EECM questionnaires would be too much of a burden for the clients with MID-BIF. Therefore, we selected one question from each factor on the TAM and EECM for the clients to answer. The staff members completed the full version of the questionnaire. For ease of reporting and interpretation for both clients and staff members, the results reported in this paper consist of the SUS, the short version of the TAM, and the short version of the EECM ([Multimedia Appendix 1](#)). All short questionnaires had an overall Cronbach $>.80$ in this study.

Qualitative Questionnaires

The qualitative questionnaires consisted of an individually administered semistructured interview based on the quantitative questionnaires. Participants were asked to elaborate on thoughts they had on the aspects of usability, acceptance, and continuous use. Additional determinants of usability, acceptance, and continuous use were derived from the semistructured interview in order to further explore the second research question.

Procedure

The research was conducted from May to August 2019. Recruitment of the participants was done at the sites of the living units. Participants were invited to participate and informed about the aim of the study through posters, flyers, and email. After a participant agreed to participate in the study, wearable biosensors were given to the participants with instructions on how to use them. If a user did not own a phone to connect to the app, a P smart (Huawei Device Co Ltd) was provided to the participant (although some participants were not allowed to use a phone due to the nature of their sentence and only used the app in the presence of their caretakers). One of four commercially available devices was randomly assigned to the participant. Before wearing the device, they completed the SUS questionnaire to assess their expectations of usability. The research coordinator completed the SUS questionnaire with the

participant if necessary. Sheehan and Hassiotis [28] identify several reasons why people with intellectual disabilities might experience barriers to use technology including cognitive limitations, physical and sensory impairment, and a lack of training and support. For this reason, the research coordinators assisted the clients in explaining how the devices work and completing the questionnaires.

The participants were given time to get familiar with the biosensors. The research coordinator functioned as a contact person in case the participant had any technical problems [2]. After the participant wore the device for 1 week, they completed the questionnaires on usability, acceptance, and continuous use of the wearable biosensors (which thus also resulted in a measure of experience with the device) again. In-depth qualitative interviews were conducted with 29% (22/77) of participants, 14 of whom were staff members.

Statistical Analysis

Descriptive statistics were used to describe the devices that were worn and the age, education, and gender of participants. A 2-way mixed analysis of variance (within: pre-post and between: client-staff) was used to test the main outcome on the expectation of usability with the actual experience. The SUS scores were calculated both preassessment and postassessment for each participant to determine their correlation with continuous use to answer the first research question. To test whether there was an association between the expected and experienced usability with (the intended) continuous use of the EECM questionnaire, we used Spearman correlations to answer the second research question, as the scores on the SUS were not normally distributed. To determine the key determinants that contribute to the usability, acceptance, and continuous use of biosensors, the proportion and number of responses for all questionnaires was computed for clients and staff members.

Further exploratory statistical analyses consisted of a 2-way analysis of covariance (ANCOVA) to test which demographic factors were associated with the judgments of the participants concerning the usability, acceptance, and continuous use of biosensors. Last, an analysis was performed on the qualitative questions of usability, acceptance, and continuous use regarding word frequency. All analyses were done in R (version 3.6.1, R Foundation for Statistical Computing) software [29].

Results

Participants and Setting

To investigate whether there were differences between clients and staff members in expectations of usability and the actual experienced usability that will lead to continuous use, 77 participants were included (31 clients and 46 staff members), with an age range varying from 18 to 63 (mean 34.9 [SD 10.8]) years. Participants were included from 4 mental health institutions in the Netherlands that provide forensic care for clients with MID-BIF.

Materials

Participants wore a Charge 3 (31/77), vivosmart 4 (21/77), Spire Stone (14/77), or TicWatch E (11/77; [Table 1](#)). The difference

between groups in terms of number of worn devices was mainly due to difficulty in including participants at different locations as the sites were randomly assigned two devices each. More staff members than clients were included. Some clients and staff

members did not want to answer questions regarding their gender (1/77) or education level (3/77) and were therefore set to missing.

Table 1. Descriptive statistics.

Participant	Client (n=31), n (%)	Staff (n=46), n (%)
Device		
Charge 3	16 (52)	15 (33)
vivosmart 4	7 (23)	14 (30)
Spire Stone	3 (10)	11 (24)
TicWatch E	5 (16)	6 (13)
Gender^a		
Male	20 (65)	21 (46)
Female	10 (32)	25 (54)
Education^a		
Primary	16 (52)	0 (0)
Secondary	13 (42)	17 (37)
Higher	0 (0)	28 (61)

^aSome participants were reluctant to answer questions on gender and education.

A small proportion of clients with MID-BIF were not allowed to use a mobile phone (6/77). They were therefore unable to answer questions regarding the use of the biosensor in combination with the app. The missing values were therefore imputed with the “don’t know or neutral” categories of the

questionnaires. [Table 2](#) shows the total SUS scores at the start and end of the study period for each device. The SUS scores for most devices increased comparing prescores and postscores except for the Spire Stone (both clients and staff) and TicWatch E (only staff decreased).

Table 2. Descriptive statistics of System Usability Scale scores.

Group	Product Start	n	Start	SD	End	SD end	Min start	Max start	Min end	Max end
Client	Charge 3	16	56.88	20.01	60.31	18.66	15.00	90.00	15.00	82.50
Client	vivosmart 4	7	66.43	16.00	67.86	17.82	45.00	87.50	47.50	95.00
Client	Spire Stone	3	64.17	10.10	63.33	14.65	55.00	75.00	52.50	80.00
Client	TicWatch E	5	51.50	8.59	57.50	27.33	37.50	60.00	20.00	82.50
Staff	Charge 3	15	69.17	10.42	75.50	12.00	55.00	87.50	55.00	92.50
Staff	vivosmart 4	14	74.11	5.24	76.25	8.31	65.00	82.50	65.00	92.50
Staff	Spire Stone	11	64.77	6.56	59.77	20.05	55.00	77.50	12.50	75.00
Staff	TicWatch E	6	70.83	16.93	44.58	13.73	50.00	97.50	30.00	70.00

The mean for the clients increased over time while the mean for the staff decreased ([Table 3](#)). Paired samples *t* tests showed that these differences were nonsignificant for both clients ($t=-0.647$, $P=.52$) and staff members ($t=0.645$, $P=.52$). The discrepancy between the experienced (posttest) and expected (pretest) usability seems to have been caused by the Spire Stone

and TicWatch E (see [Table 2](#)). There was no interaction effect of expected versus experienced for group ($F_{1,75}=1.82$, $P=.18$).

The correlation between the expected usability and the EECM was .18 ($P=.12$). The correlation between the experienced usability and the EECM was .54 ($P<.001$), which indicates a higher relevance for continuous use after participants wore the device.

Table 3. Descriptive statistics of System Usability Scale scores per group.

Group	Time	Variable	n	Mean (SD)
Client	Start	Score	31	58.87 (17.19)
Staff	Start	Score	46	69.84 (9.77)
Client	End	Score	31	61.85 (19.09)
Staff	End	Score	46	67.94 (17.45)

Usability

As shown in [Figure 1](#) and [Figure 2](#), at the start and end of the study, more than 75% of clients (start 26/31; end 25/31) and staff members (start 41/46; end 36/46) indicated that they were confident to use the biosensors. More than 50% of clients (start 24/31; end 20/31) and staff members (start 37/46; end 31/46) also indicated that they could imagine that most people would learn to use the biosensor very quickly, and both clients (start 16/31; end 20/31) and staff members (start 28/46; end 24/46) would like to use it frequently. At the start of the study, 61% (19/31) of clients indicated they needed help from a technical

person, in comparison to 29% (9/31) at the end of the study. More than 50% of staff members (start 29/46; end 34/46) thought the biosensor was easy to use, and by the end of the study, more than 50% of clients also agreed with that (20/31). More than 50% of staff members (start 29/46; end 36/46) did not agree that the biosensors were unnecessarily complex, did not agree that they would need the support of a technical person (start 33/46; end 36/46), did not agree that the biosensors were very cumbersome to use (start 38/46; end 38/46), and did not agree that they needed to learn many things before they could get going with the biosensors (start 35/46; end 33/46).

Figure 1. Proportion and number of responses for clients on the System Usability Scale questionnaire.

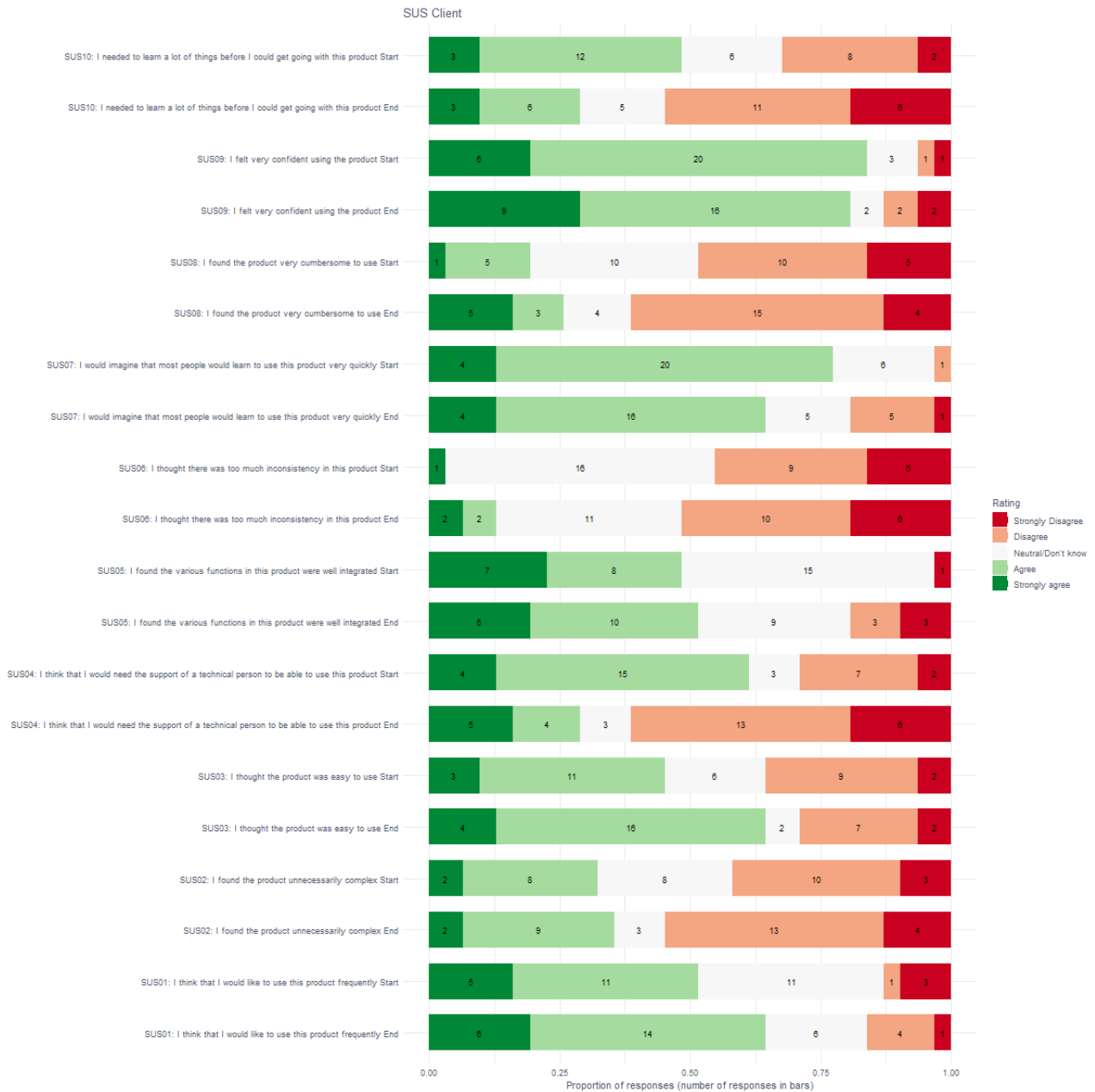
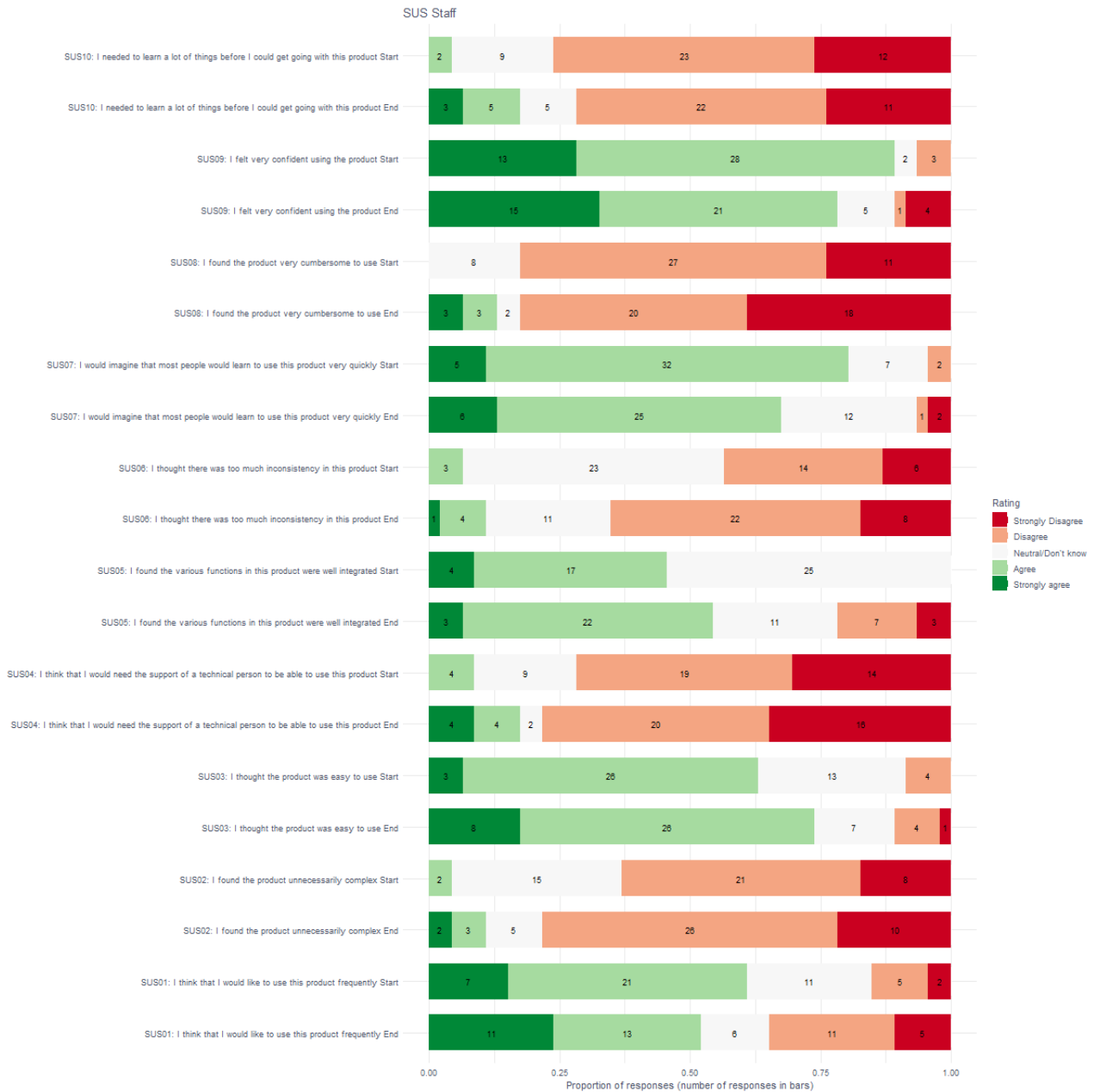


Figure 2. Proportion and number of responses for staff on the System Usability Scale questionnaire.



Acceptance

Figure 3 shows that more than 75% of clients thought that the smart watch was attractive and pleasing (26/31), they could use the device to get the desired information and service (24/31), and the advantages outweighed the disadvantages (26/31). More than 50% of clients liked the idea of using the smartwatch

(23/31), thought it was expensive (20/31), felt they could use the smartwatch anywhere (23/31), was easy to use (23/31), and was useful for doing their job (26/31). Figure 4 shows that staff members had similar patterns except that a small number of staff members thought that the smartwatch was expensive (8/46) or that it was useful for doing their job (18/46).

Figure 3. Proportion and number of responses for clients on the Technology Acceptance Model questionnaire.

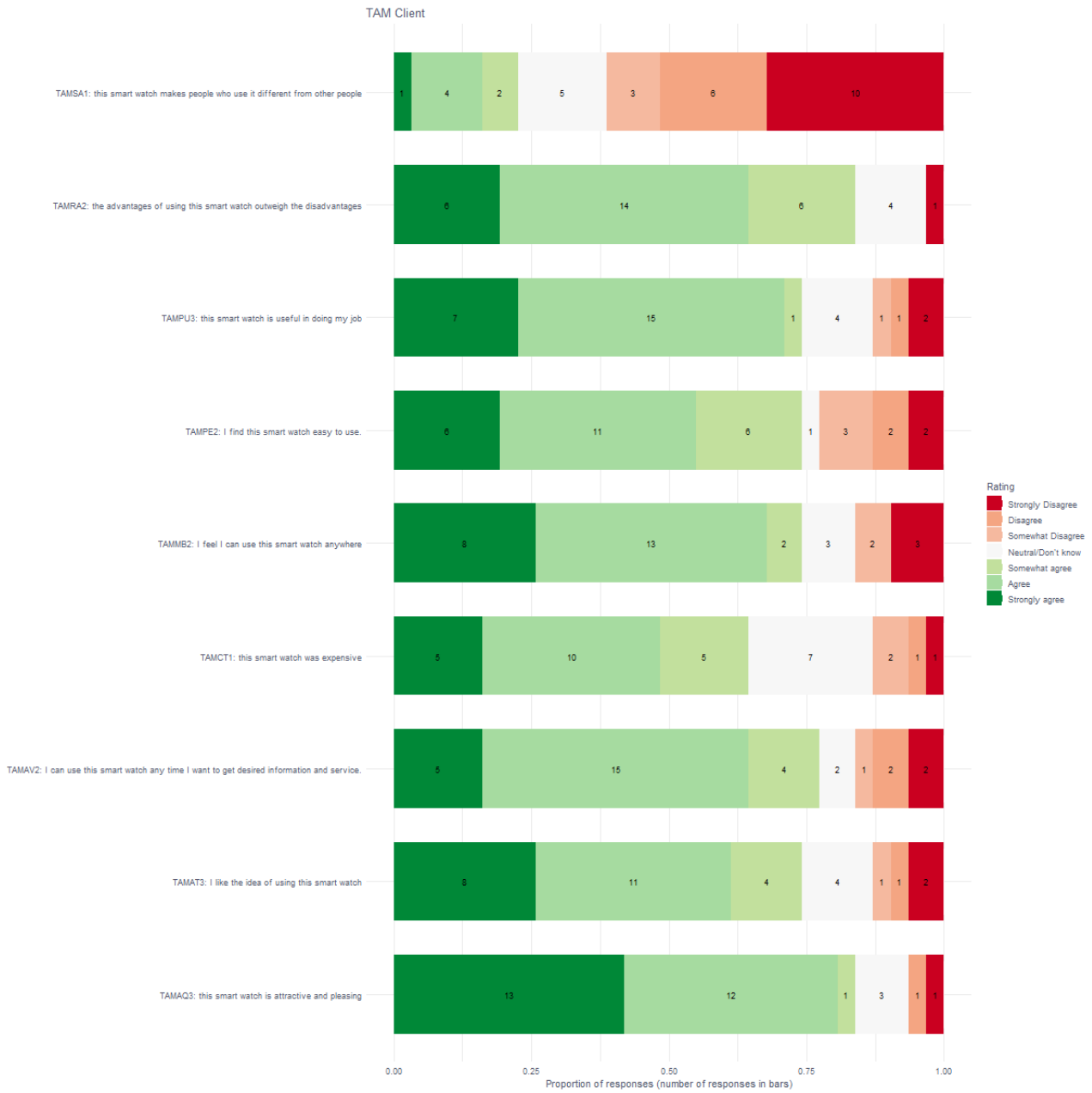
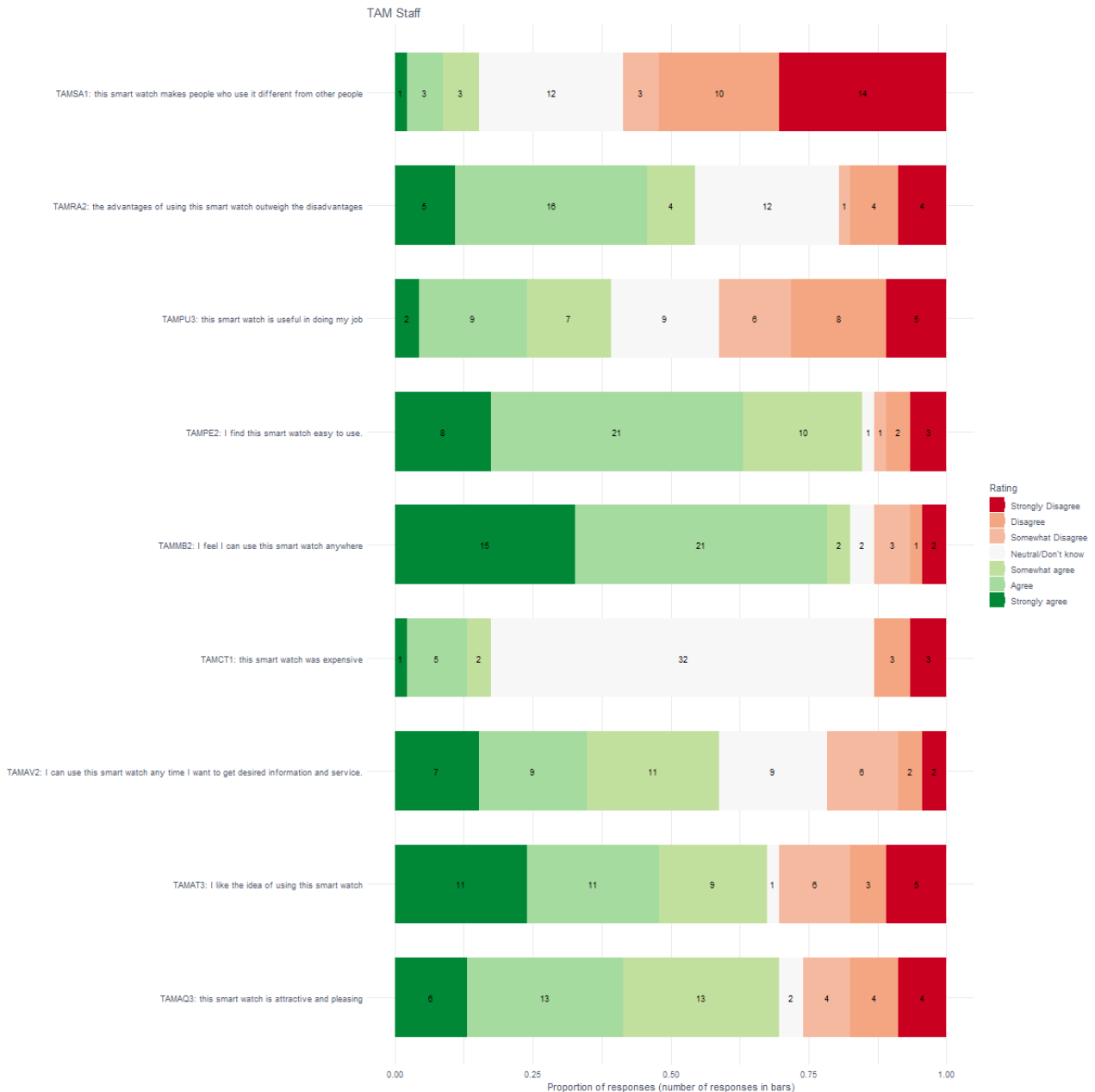


Figure 4. Proportion and number of responses for staff on the Technology Acceptance Model questionnaire.



Continuous Use

Figure 5 shows that more than 75% (26/31) of clients agreed to some extent that it was a pleasant experience, that it was entertaining (24/31), that the biosensors met demands in excess of what they expected (25/31), and that they doubted whether the fitness data were accurate (25/31). In addition, more than 75% (26/31) of clients disagreed that the biosensors were heavy and large, and that the devices needed a larger battery capacity (25/31). More than 50% of clients agreed that the experience was better than expected (20/31), that they intend to use the device in the future (19/31), that the device helps them perform

many things more conveniently (20/31), and that they have better control over their health (17/31) and did not agree that using the smart watch makes them feel uncomfortable (20/31). Figure 6 shows that the outcome for staff members was slightly different, although the patterns look similar at first glance. More than 75% (38/46) of staff members disagreed that the biosensors were heavy and large. More than 50% of staff members found the use entertaining (27/46), doubted that the fitness data were accurate (25/46), had better control over their overall health (27/46), and had an overall pleasant experience (30/46). More than 50% (29/46) did not agree that using the smartwatch made them feel uncomfortable.

Figure 5. Proportion of responses and number of responses for clients on the extended expectation confirmation model.

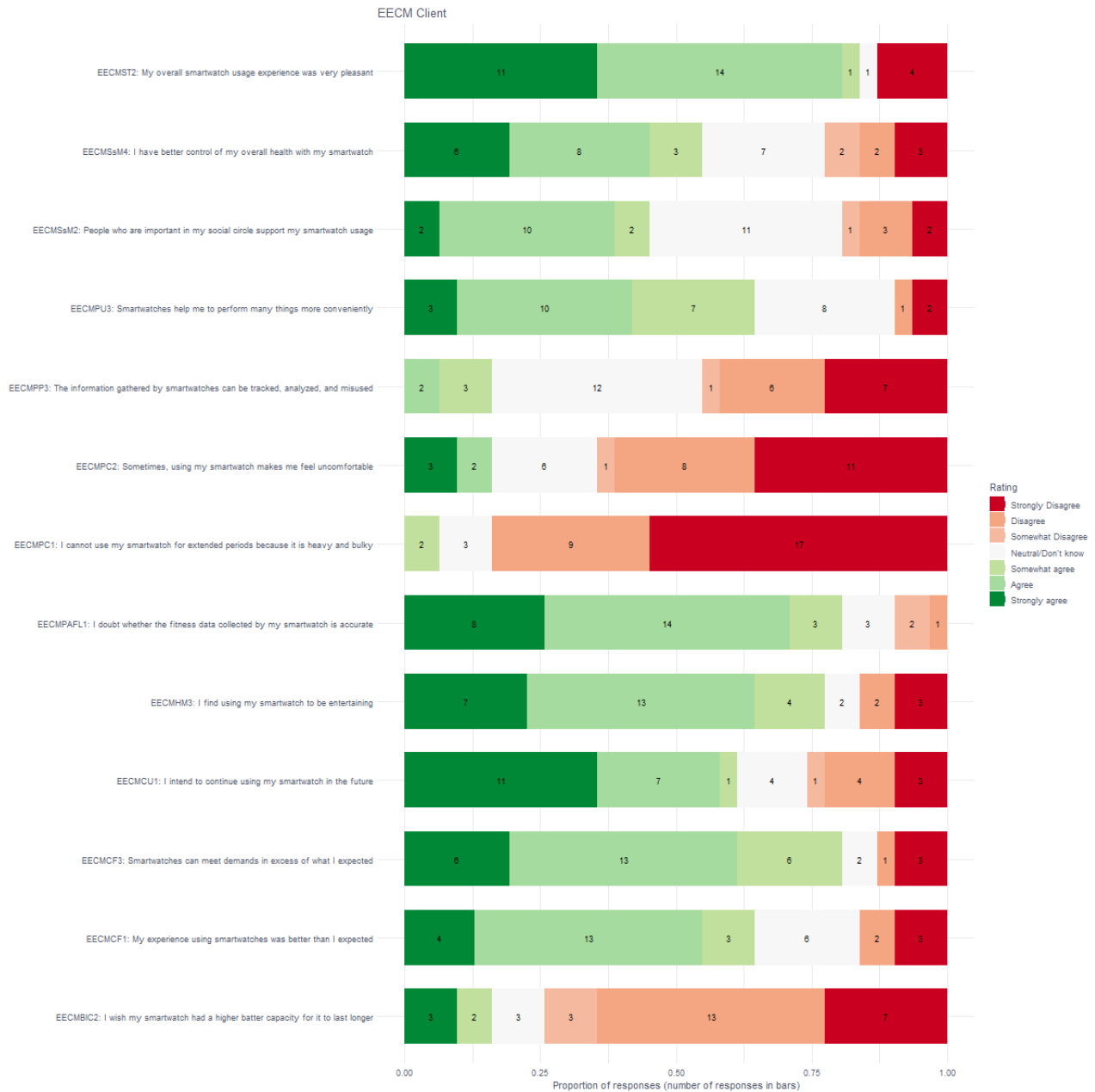
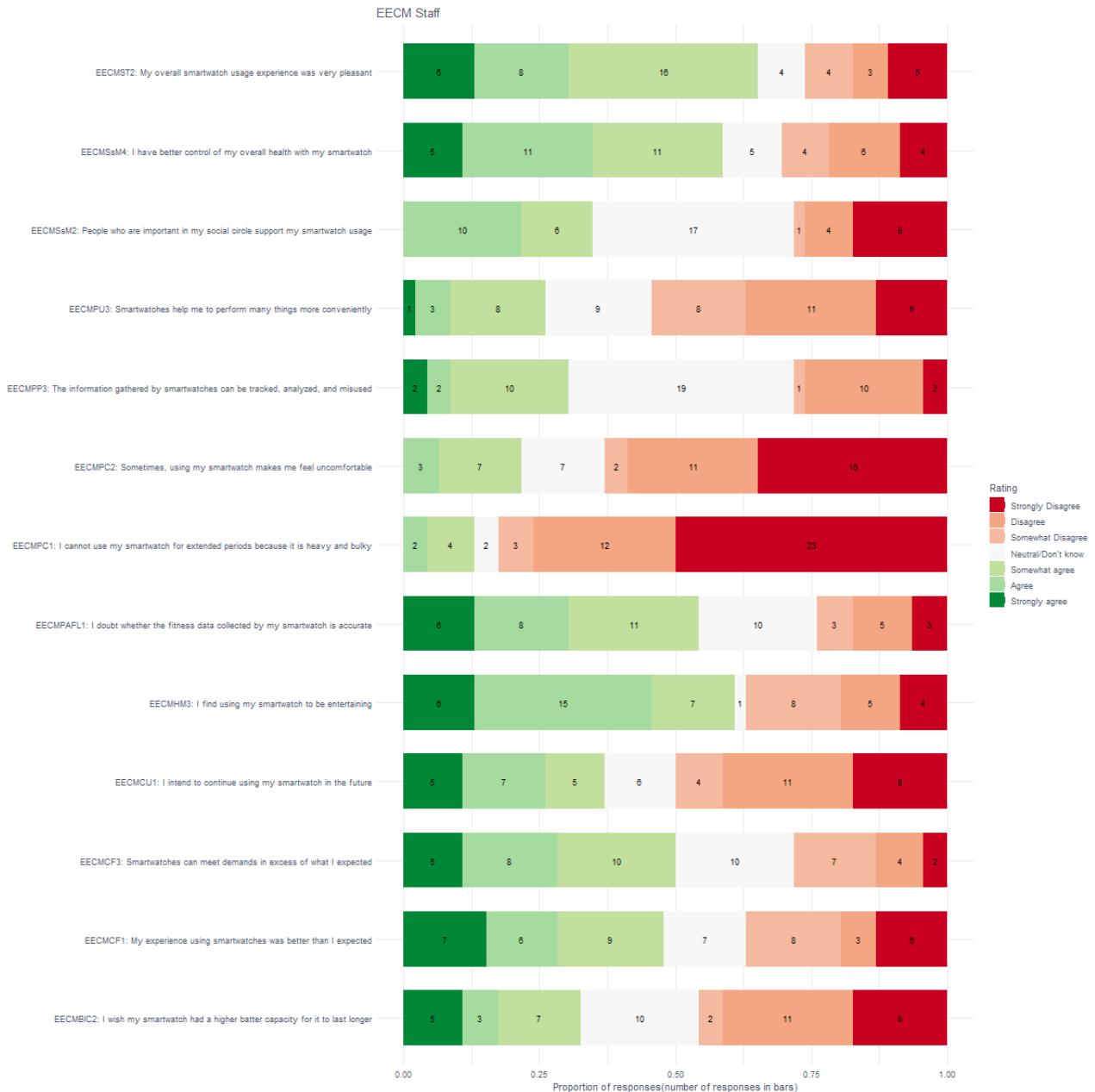


Figure 6. Proportion of responses and number of responses for staff on the extended expectation confirmation model.



Demographic Variables

We explored which demographic variables contributed to the usability (SUS), acceptance (TAM), and continuous use (EECM) of biosensors. We included age, level of education, and gender in separate 2-way ANCOVAs as the dependent variables were correlated, which prohibited a multifactorial multivariate analysis of covariance. In addition, we performed a Box-Cox transformation on the SUS and TAM as these were not normally distributed; the ANCOVAs were adjusted for age. There was a significant difference for both acceptance ($F_{1,69}=9.214, P=.003$) and continuous use ($F_{1,69}=16.607, P<.001$) for group; in both cases the clients scored higher on acceptance and intention for continuous use. There was no significant interaction between gender and education on the usability score ($F_{2,64}=0.549, P=.58$), acceptance score ($F_{2,64}=0.545, P=.58$), or continuous use score

($F_{2,64}=1.475, P=.24$). This indicates that the effect of gender on usability, acceptance, and continuous use does not depend on the level of education or vice versa. As can be seen in Table 4, a strong correlation between the acceptance (TAM) of wearable devices and the intention of continuous use (EECM) was found. In addition, moderate correlations were found between usability with acceptance and continuous use. Last, there was a weak negative correlation between usability and age. Further analysis of the correlations showed that the staff members obtained strong correlations between usability and acceptance ($r=.80, P<.001$), usability and continuous use ($r=.79, P<.001$), and acceptance with continuous use ($r=.89, P<.001$). Clients obtained moderate correlations between usability and acceptance ($r=.46, P=.01$), usability and continuous use ($r=.52, P=.003$), and strong correlations between acceptance and continuous use ($r=.75, P<.001$).

Table 4. Spearman correlations between variables.

Questionnaire	SUS ^a	EECM ^b	TAM ^c
EECM	0.54	— ^d	—
TAM	0.58	0.86	—
Age	-0.24	-0.15	-0.03

^aSUS: System Usability Scale.

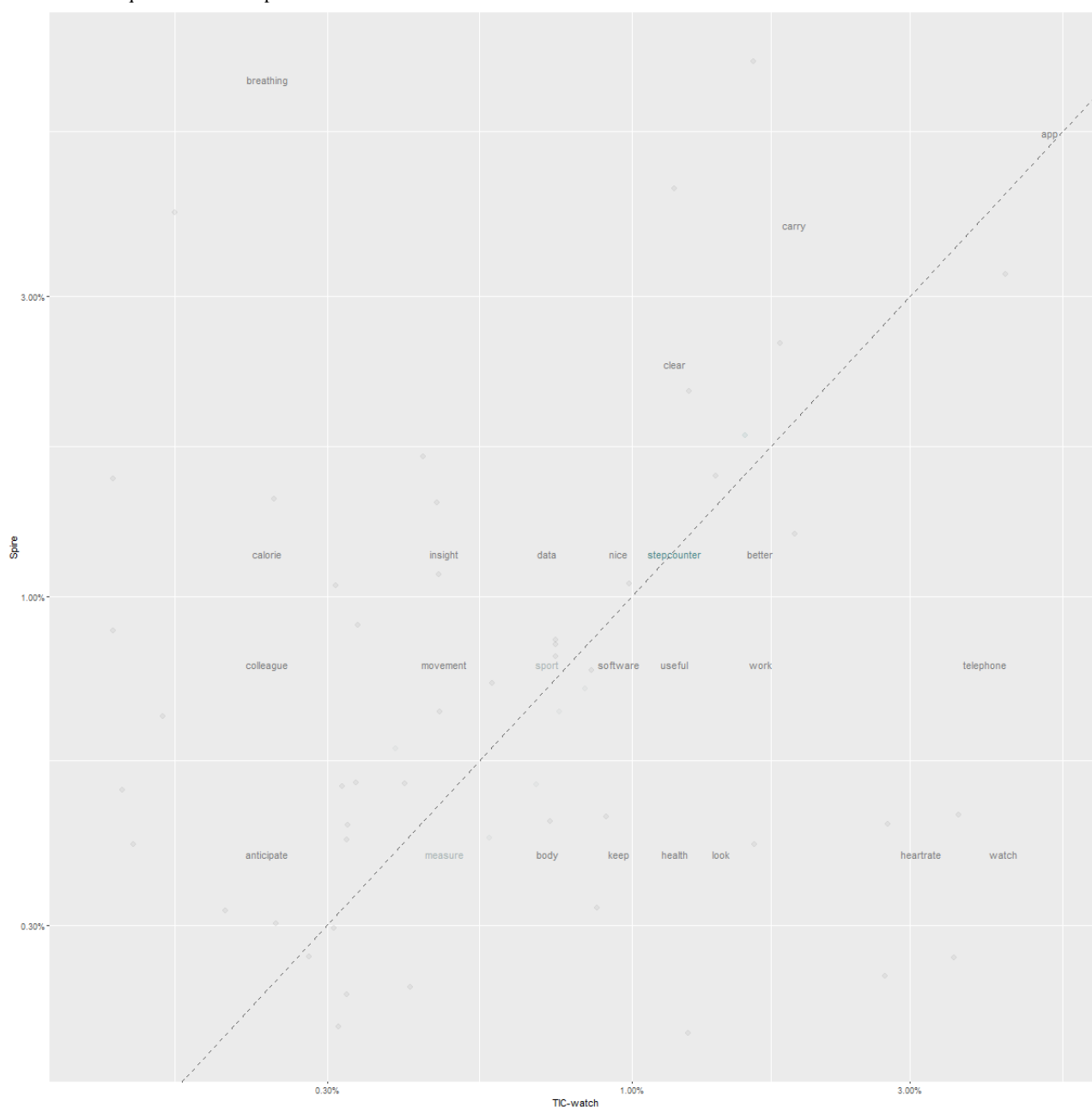
^bEECM: extended expectation confirmation model.

^cTAM: Technology Acceptance Model questionnaire.

^dNot applicable.

For analysis of qualitative questionnaires, a sample size between 5 and 50 is required [30]. In our sample, only the qualitative interviews for the Spire Stone (n=6) and TicWatch E (n=10) met this criterion and are reported here. The word frequencies for both devices are reported in Figure 7 as we wanted to compare the frequencies for the different devices. Words that are close to the dotted line share a frequency similarity for both

devices. Words that are farther away from the line have nonsimilar frequencies for both devices. As expected, breathing has a higher frequency for the Spire Stone and heart rate is higher for the TicWatch E as this is the main function for both devices. The Spire Stone and TicWatch E are generally considered easy to use and give clear information.

Figure 7. Word frequencies for the Spire Stone and TicWatch E.

Discussion

Principal Findings

In this study, we investigated whether the expectancy or the actual experience was most important for an intended continuous use of biosensor devices for monitoring and coaching in forensic psychiatry. In addition, we investigated what contributes to the usability, acceptance, and intended continuous use. The main result of the study is that it was the actual experience of wearing a biosensor that was associated with intended continuous use, and to a much lesser extent, the expectancy. This is contrary to the hypotheses of Pal et al [12] that expectancy would be important for continuous use and that a gap exists between the expectations of usability and the factors that would lead to continuous use by experiences with the device. In our study, the expected usability had only a weak positive association with

continuous use, while the actual experienced usability had a moderate positive association. These associations were markedly different between staff members and clients. The associations between usability and both acceptance and continuous use were stronger in staff members and more moderate in clients. It may well be that factors outside usability are responsible for associations with acceptance and continuous use in clients. Perhaps the questionnaires on acceptance and continuous use do not cover the full range of aspects associated with usability for clients. It may also be that clients had different expectations and experiences with the device. Further longitudinal research (ie, longer than 1 month) on this association is warranted, and it would be of interest to investigate any mediation effects of technology experience on the expectancy–continuous use association to further investigate the hypothesis. This would, however, require larger sample sizes.

In addition, a strong association between the acceptance (TAM) of wearable devices and the intention of continuous use (EECM) was found. This seems to indicate that these two questionnaires measure overlapping constructs, and the question arises whether both need to be administered. Especially when the load on participants should be kept to a minimum as in our sample with forensic MID-BIF clients. These results must be interpreted with care as the design of our study, without proper counterbalancing, and the use of short questionnaires limit the conclusions that can be drawn from the study.

As far as the determinants for usability, acceptance, and continuous use are concerned, answers from the usability scale indicated that most of the clients and staff members felt confident using the biosensors and after they wore the devices and thought that most people would learn to use the product very quickly and want to use it frequently. The acceptance scale indicated that the majority have positive attitudes toward technology, their affective quality, relative advantage, mobility, availability, and perceived ease of use. The continuous use scale showed that the majority of staff members and clients gave positive answers on satisfaction, self-socio motivation, perceived comfort, and hedonic motivation. However, the majority had doubts on the perceived accuracy and functional limitations. It is also interesting to note that a minority of staff members and clients were not positive about the usability, acceptance, and continuous use of the devices. These people might not want to use the devices or think that they need help in using the devices. For instance, a minority of people think that they would need help from a technical person to use the device. It might well be that providing them with proper support might increase their intention to use the device. Also, some find that wearing the device is uncomfortable and the accuracy of the fitness data could be improved. These devices might thus benefit from developments in accuracy and form factor [2,3,15,31]. Also, it is unclear if questions on cost of the device can be properly answered as the participants in this study did not have to pay for the devices. Participants can only guess if the device was expensive. In addition, there were no differences in gender, education, or age for usability, acceptance, and continuous use. However, it must be pointed out that our sample size was limited. Clients showed a higher score on acceptance and intention for continuous use. It must be further investigated whether this is a true effect or could be due to a social desirability bias in the forensic clients. It is interesting that clients scored higher on acceptance and continuous use as they might benefit from the ease of use of these devices and their continuing monitoring and coaching apps.

Strengths and Limitations

Two particular strengths of this study are the use of simply worded questionnaires adapted for clients with MID-BIF as this was not available in the literature. In addition, we used qualitative questionnaires and a diverse and heterogeneous sample as Kalantari [11] pointed out was missing in the literature. Three major limitations of this study are related to the validity of the questionnaires, design of the study, and duration of wearing the devices. First, although the reliabilities of the short TAM and EECM are above .80, it remains to be established if a short version of these questionnaires can validly

capture the construct intended by the original questionnaires. Furthermore, it is unknown whether these short questionnaires measure similar constructs over time and for different groups. Second, the design of the study does not warrant any comparison between devices, as the participants did not wear all 4 devices. This seriously limits the comparability of results to a within-subject comparison on the SUS. Third, participants wore the devices for only 1 week, which will have a significant impact on the measure of continuous use. The measure only applies to expectations and intentions of continuous use. It remains to be established if this intention for continuous use is associated with actual continuous use. Further limitations of this study were the limited sample size, uneven number of administered questionnaires per device, and use of a modified version of the TAM. The original theoretical framework of the TAM [27] was altered in an extended smartwatch-oriented TAM [15] and it is unclear how this may have affected the results. Furthermore, one of the devices was not a smartwatch in the strict sense, and this limits the conclusions that can be drawn for this device.

Future Research

Future research should focus on longitudinal research investigating usability, acceptance, and continuous use, should include a counterbalanced design in which all devices are worn at least once, and should investigate measurement invariance for the short questionnaires [32,33]. Biosensors in forensic psychiatry might prove to be a very useful addition in the treatment for vulnerable MID-BIF clients due to their health-related functionality. These clients often suffer from obesity, sleep disorders, and trauma-related disorders and might substantially benefit from the different functionality that biosensors potentially offer. The use of these devices yields high expectations, and it is plausible that wearable biosensor devices can be used to detect changing levels of emotional states, assist in self-regulation, and even signal imminent problem behavior, such as aggression in clients [14]. The types of biosensors that a client or staff member uses will depend, at least in part, on the problem (ie, health problems, sleep problems, fitness tracking) or use case (ie, emotion regulation, behavior modification) they have. Information on breathing and tension or focus during activities and at certain locations requires a different sensor than monitoring real-time heart rate changes, and the conditions in which the devices are used and prerequisites for using the biosensors should be as clear as possible to increase the use and maximize the potential health benefits of these sensors.

Another important topic for future research is the reliability and validity of the sensors, especially in comparison with gold standard equipment used in laboratories. Peake et al [11] reported that only 5% of wearable devices with integrated biosensors are well validated and most validation studies lack clear conclusions [34]. Not providing accurate and timely information might seriously affect the willingness to wear a device.

Last, for people with MID-BIF, it is especially important to develop easy-to-use biosensors with a minimum requirement on cognitive capacity to increase usability, acceptance, and continuous use in the future. It must be noted, however, that

clients scored similar to staff members on ease of use of available devices and higher on acceptance and (intended) continuous use. Whether clients indeed grasped the information provided by the sensors must be investigated further.

Conclusion

Actual perceived usability of wearing a biosensor and to a much lesser extent the expectancy of usability were associated with continuous use. Clients with mild intellectual disabilities might

benefit from the ease of use of wearables devices and their continuing monitoring and coaching apps. Clients scored higher on acceptance and intention for continuous use, but associations between usability and both acceptance and continuous use were markedly stronger in staff members. For clients, it is especially important to develop easy-to-use biosensors with a minimum requirement on cognitive capacity to increase usability, acceptance, and continuous use.

Acknowledgments

We would like to thank the project partners of the Sense-IT project for making the Sense-IT available for use in this study: University of Twente, Scelta/GGNet, VUmc, Arkin, and Pluryn.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Short questionnaires used in the study.

[[PDF File \(Adobe PDF File\), 475 KB - formative_v5i5e18096_app1.pdf](#)]

References

1. Kreibitz SD. Autonomic nervous system activity in emotion: a review. *Biol Psychol* 2010 Jul;84(3):394-421. [doi: [10.1016/j.biopsycho.2010.03.010](https://doi.org/10.1016/j.biopsycho.2010.03.010)] [Medline: [20371374](https://pubmed.ncbi.nlm.nih.gov/20371374/)]
2. Derks Y, Klaassen R, Westerhof G, Bohlmeijer E, Noordzij M. Development of an ambulatory biofeedback app to enhance emotional awareness in patients with borderline personality disorder: multicycle usability testing study. *JMIR Mhealth Uhealth* 2019 Oct 15;7(10):e13479 [FREE Full text] [doi: [10.2196/13479](https://doi.org/10.2196/13479)] [Medline: [31617851](https://pubmed.ncbi.nlm.nih.gov/31617851/)]
3. Liang J, Xian D, Liu X, Fu J, Zhang X, Tang B, et al. Usability study of mainstream wearable fitness devices: feature analysis and system usability scale evaluation. *JMIR Mhealth Uhealth* 2018 Nov 08;6(11):e11066 [FREE Full text] [doi: [10.2196/11066](https://doi.org/10.2196/11066)] [Medline: [30409767](https://pubmed.ncbi.nlm.nih.gov/30409767/)]
4. de Looff P, Didden R, Embregts P, Nijman H. Burnout symptoms in forensic mental health nurses: results from a longitudinal study. *Int J Ment Health Nurs* 2019 Feb;28(1):306-317. [doi: [10.1111/inm.12536](https://doi.org/10.1111/inm.12536)] [Medline: [30156026](https://pubmed.ncbi.nlm.nih.gov/30156026/)]
5. Poh MZ, Swenson NC, Picard RW. A wearable sensor for unobtrusive, long-term assessment of electrodermal activity. *IEEE Trans Biomed Eng* 2010 May;57(5):1243-1252. [doi: [10.1109/tbme.2009.2038487](https://doi.org/10.1109/tbme.2009.2038487)]
6. Cosoli G, Spinsante S, Scalise L. Wrist-worn and chest-strap wearable devices: systematic review on accuracy and metrological characteristics. *Measurement* 2020 Jul;159:107789. [doi: [10.1016/j.measurement.2020.107789](https://doi.org/10.1016/j.measurement.2020.107789)]
7. Schuurmans AAT, de Looff P, Nijhof KS, Rosada C, Scholte RHJ, Popma A, et al. Validity of the Empatica E4 wristband to measure heart rate variability (HRV) parameters: a comparison to electrocardiography (ECG). *J Med Syst* 2020 Sep 23;44(11):190 [FREE Full text] [doi: [10.1007/s10916-020-01648-w](https://doi.org/10.1007/s10916-020-01648-w)] [Medline: [32965570](https://pubmed.ncbi.nlm.nih.gov/32965570/)]
8. Peake JM, Kerr G, Sullivan JP. A critical review of consumer wearables, mobile applications, and equipment for providing biofeedback, monitoring stress, and sleep in physically active populations. *Front Physiol* 2018;9:743 [FREE Full text] [doi: [10.3389/fphys.2018.00743](https://doi.org/10.3389/fphys.2018.00743)] [Medline: [30002629](https://pubmed.ncbi.nlm.nih.gov/30002629/)]
9. Kim J, Ryu B, Cho S, Heo E, Kim Y, Lee J, et al. Impact of personal health records and wearables on health outcomes and patient response: three-arm randomized controlled trial. *JMIR Mhealth Uhealth* 2019 Jan 04;7(1):e12070 [FREE Full text] [doi: [10.2196/12070](https://doi.org/10.2196/12070)] [Medline: [30609978](https://pubmed.ncbi.nlm.nih.gov/30609978/)]
10. Sano A, Picard RW, Stickgold R. Quantitative analysis of wrist electrodermal activity during sleep. *Int J Psychophysiol* 2014 Dec;94(3):382-389 [FREE Full text] [doi: [10.1016/j.ijpsycho.2014.09.011](https://doi.org/10.1016/j.ijpsycho.2014.09.011)] [Medline: [25286449](https://pubmed.ncbi.nlm.nih.gov/25286449/)]
11. Kalantari M. Consumers' adoption of wearable technologies: literature review, synthesis, and future research agenda. *Int J Technol Mark* 2017;12(3):274-307. [doi: [10.1504/ijtmkt.2017.089665](https://doi.org/10.1504/ijtmkt.2017.089665)]
12. Pal D, Funilkul S, Vanijja V. The future of smartwatches: assessing the end-users' continuous usage using an extended expectation-confirmation model. *Univ Access Inf Soc* 2018 Oct 31;19(2):261-281. [doi: [10.1007/s10209-018-0639-z](https://doi.org/10.1007/s10209-018-0639-z)]
13. Phillips SM, Cadmus-Bertram L, Rosenberg D, Buman MP, Lynch BM. Wearable technology and physical activity in chronic disease: opportunities and challenges. *Am J Prev Med* 2018 Jan;54(1):144-150 [FREE Full text] [doi: [10.1016/j.amepre.2017.08.015](https://doi.org/10.1016/j.amepre.2017.08.015)] [Medline: [29122356](https://pubmed.ncbi.nlm.nih.gov/29122356/)]
14. de Looff P, Noordzij ML, Moerbeek M, Nijman H, Didden R, Embregts P. Changes in heart rate and skin conductance in the 30 min preceding aggressive behavior. *Psychophysiology* 2019 Oct;56(10):e13420. [doi: [10.1111/psyp.13420](https://doi.org/10.1111/psyp.13420)] [Medline: [31184379](https://pubmed.ncbi.nlm.nih.gov/31184379/)]

15. Kim KJ, Shin D. An acceptance model for smart watches. *Internet Research* 2015 Aug 03;25(4):527-541. [doi: [10.1108/IntR-05-2014-0126](https://doi.org/10.1108/IntR-05-2014-0126)]
16. Cooney P, Tunney C, O'Reilly G. A systematic review of the evidence regarding cognitive therapy skills that assist cognitive behavioural therapy in adults who have an intellectual disability. *J Appl Res Intellect Disabil* 2018 Jan;31(1):23-42. [doi: [10.1111/jar.12365](https://doi.org/10.1111/jar.12365)] [Medline: [28544303](https://pubmed.ncbi.nlm.nih.gov/28544303/)]
17. Sullivan AN, Lachman ME. Behavior change with fitness technology in sedentary adults: a review of the evidence for increasing physical activity. *Front Public Health* 2016;4:289 [FREE Full text] [doi: [10.3389/fpubh.2016.00289](https://doi.org/10.3389/fpubh.2016.00289)] [Medline: [28123997](https://pubmed.ncbi.nlm.nih.gov/28123997/)]
18. Spath D, Hermann F, Peissner M, Sproll S. User requirements collection and analysis. In: *Handbook of Human Factors and Ergonomics*. Hoboken: John Wiley & Sons, Ltd; 2012:1313-1322.
19. Hounscome J, Whittington R, Brown A, Greenhill B, McGuire J. The structured assessment of violence risk in adults with intellectual disability: a systematic review. *J Appl Res Intellect Disabil* 2018 Jan;31(1):e1-e17. [doi: [10.1111/jar.12295](https://doi.org/10.1111/jar.12295)] [Medline: [27891723](https://pubmed.ncbi.nlm.nih.gov/27891723/)]
20. Holt M, Yule B, Jackson D, Zhu M, Moraveji N. Ambulatory monitoring of respiratory effort using a clothing-adhered biosensor. 2018 Presented at: IEEE Int Symp Med Meas Appl MeMeA; 2018; Rome p. 1-6. [doi: [10.1109/memea.2018.8438678](https://doi.org/10.1109/memea.2018.8438678)]
21. Schrage JD, Shayne P, Wolf S, Das S, Patzer RE, White M, et al. Assessing the influence of a Fitbit physical activity monitor on the exercise practices of emergency medicine residents: a pilot study. *JMIR Mhealth Uhealth* 2017 Jan 31;5(1):e2 [FREE Full text] [doi: [10.2196/mhealth.6239](https://doi.org/10.2196/mhealth.6239)] [Medline: [28143805](https://pubmed.ncbi.nlm.nih.gov/28143805/)]
22. Reddy RK, Pooni R, Zaharieva DP, Senf B, El Youssef J, Dassau E, et al. Accuracy of wrist-worn activity monitors during common daily physical activities and types of structured exercise: evaluation study. *JMIR Mhealth Uhealth* 2018 Dec 10;6(12):e10338 [FREE Full text] [doi: [10.2196/10338](https://doi.org/10.2196/10338)] [Medline: [30530451](https://pubmed.ncbi.nlm.nih.gov/30530451/)]
23. Dumas J. User-based evaluations. In: Jacko JA, Sears A, editors. *The Human-Computer Interaction Handbook*. Mahwah: Lawrence Erlbaum Associates, Inc; 2003:1093-1117.
24. Brooke J. SUS: a "quick and dirty" usability scale. In: Jordan P, Thomas B, Weerdmeester B, editors. *Usability Evaluation in Industry*. Philadelphia: Taylor & Francis; 1996:189-194.
25. Lewis J. Usability testing. In: Salvendy G, editor. *Handbook of Human Factors and Ergonomics*. 4th edition. Hoboken: John Wiley & Sons; 2012:1267-1312.
26. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly* 1989 Sep;13(3):319. [doi: [10.2307/249008](https://doi.org/10.2307/249008)]
27. Venkatesh V, Bala H. Technology acceptance model 3 and a research agenda on interventions. *Decis Sci* 2008 May;39(2):273-315. [doi: [10.1111/j.1540-5915.2008.00192.x](https://doi.org/10.1111/j.1540-5915.2008.00192.x)]
28. Sheehan R, Hassiotis A. Digital mental health and intellectual disabilities: state of the evidence and future directions. *Evid Based Ment Health* 2017 Nov;20(4):107-111. [doi: [10.1136/eb-2017-102759](https://doi.org/10.1136/eb-2017-102759)] [Medline: [28947677](https://pubmed.ncbi.nlm.nih.gov/28947677/)]
29. R: a language and environment for statistical computing. R Foundation for Statistical Computing. 2014. URL: <http://www.R-project.org/> [accessed 2021-04-28]
30. Dworkin SL. Sample size policy for qualitative studies using in-depth interviews. *Arch Sex Behav* 2012 Dec;41(6):1319-1320. [doi: [10.1007/s10508-012-0016-6](https://doi.org/10.1007/s10508-012-0016-6)] [Medline: [22968493](https://pubmed.ncbi.nlm.nih.gov/22968493/)]
31. De Visser T, Bohlmeijer ET, Noordzij ML, Derks YP. mHealth in Mental Health: how to efficiently and scientifically create an ambulatory biofeedback e-coaching app for patients with borderline personality disorder. *Int J Human Factors Ergonom* 2017;5(1):61-92. [doi: [10.1504/ijhfe.2017.10009438](https://doi.org/10.1504/ijhfe.2017.10009438)]
32. de Beurs DP, Fokkema M, de Groot MH, de Keijser J, Kerkhof AJ. Longitudinal measurement invariance of the Beck Scale for Suicide Ideation. *Psychiatry Res* 2015 Feb 28;225(3):368-373 [FREE Full text] [doi: [10.1016/j.psychres.2014.11.075](https://doi.org/10.1016/j.psychres.2014.11.075)] [Medline: [25571773](https://pubmed.ncbi.nlm.nih.gov/25571773/)]
33. van de Schoot R, Lugtig P, Hox J. A checklist for testing measurement invariance. *Eur J Devel Psychol* 2012 Jul;9(4):486-492. [doi: [10.1080/17405629.2012.686740](https://doi.org/10.1080/17405629.2012.686740)]
34. van Lier HG, Pieterse ME, Garde A, Postel MG, de Haan HA, Vollenbroek-Hutten MMR, et al. A standardized validity assessment protocol for physiological signals from wearable technology: Methodological underpinnings and an application to the E4 biosensor. *Behav Res Methods* 2020 Apr;52(2):607-629 [FREE Full text] [doi: [10.3758/s13428-019-01263-9](https://doi.org/10.3758/s13428-019-01263-9)] [Medline: [31290128](https://pubmed.ncbi.nlm.nih.gov/31290128/)]

Abbreviations

ANCOVA: analysis of covariance

EECM: extended expectation confirmation model

MID-BIF: mild intellectual disabilities and borderline intellectual functioning

SUS: System Usability Scale

TAM: Technology Acceptance Model questionnaire

Edited by G Eysenbach; submitted 17.02.20; peer-reviewed by N Shen, K Liu, M Sjölander, J Brooke; comments to author 31.08.20; revised version received 14.10.20; accepted 13.04.21; published 10.05.21.

Please cite as:

de Looff PC, Nijman H, Didden R, Noordzij ML

Usability and Acceptance of Wearable Biosensors in Forensic Psychiatry: Cross-sectional Questionnaire Study

JMIR Form Res 2021;5(5):e18096

URL: <https://formative.jmir.org/2021/5/e18096>

doi: [10.2196/18096](https://doi.org/10.2196/18096)

PMID: [33970115](https://pubmed.ncbi.nlm.nih.gov/33970115/)

©Pieter Christiaan de Looff, Henk Nijman, Robert Didden, Matthijs L Noordzij. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 10.05.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Integrating User-Centered Design and Behavioral Science to Design a Mobile Intervention for Obesity and Binge Eating: Mixed Methods Analysis

Andrea K Graham^{1,2}, PhD; Sean A Munson³, PhD; Madhu Reddy^{1,4}, PhD; Sarah W Neubert^{1,5}, BA; Emilie A Green^{1,5}, BA; Angela Chang^{1,2}, BA; Bonnie Spring⁶, PhD; David C Mohr^{1,6}, PhD; Jennifer E Wildes⁷, PhD

¹Center for Behavioral Intervention Technologies, Northwestern University Feinberg School of Medicine, Chicago, IL, United States

²Department of Medical Social Sciences, Northwestern University Feinberg School of Medicine, Chicago, IL, United States

³Department of Human Centered Design & Engineering, University of Washington, Seattle, WA, United States

⁴Department of Communication Studies, Northwestern University, Chicago, IL, United States

⁵Weinberg College of Arts and Sciences, Northwestern University, Evanston, IL, United States

⁶Department of Preventive Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL, United States

⁷Department of Psychiatry & Behavioral Neuroscience, University of Chicago, Chicago, IL, United States

Corresponding Author:

Andrea K Graham, PhD

Center for Behavioral Intervention Technologies

Northwestern University Feinberg School of Medicine

750 N. Lake Shore Drive

10th Floor

Chicago, IL, 60611

United States

Phone: 1 312 503 5266

Email: andrea.graham@northwestern.edu

Abstract

Background: Accounting for how end users engage with technologies is imperative for designing an efficacious mobile behavioral intervention.

Objective: This mixed methods analysis examined the translational potential of user-centered design and basic behavioral science to inform the design of a new mobile intervention for obesity and binge eating.

Methods: A total of 22 adults (7/22, 32% non-Hispanic White; 8/22, 36% male) with self-reported obesity and recurrent binge eating (≥ 12 episodes in 3 months) who were interested in losing weight and reducing binge eating completed a prototyping design activity over 1 week. Leveraging evidence from behavioral economics on choice architecture, participants chose treatment strategies from 20 options (aligned with treatment targets composing a theoretical model of the relation between binge eating and weight) to demonstrate which strategies and treatment targets are relevant to end users. The process by which participants selected and implemented strategies and their change in outcomes were analyzed.

Results: Although prompted to select one strategy, participants selected between 1 and 3 strategies, citing perceived achievability, helpfulness, or relevance as selection reasons. Over the week, all practiced a strategy at least once; 82% (18/22) struggled with implementation, and 23% (5/22) added a new strategy. Several themes emerged on successes and challenges with implementation, yielding design implications for supporting users in behavior change. In postexperiment reflections, 82% (18/22) indicated the strategy was helpful, and 86% (19/22) planned to continue use. One-week average within-subject changes in weight (-2.2 [SD -5.0] pounds) and binge eating (-1.6 [SD -1.8] episodes) indicated small clinical improvement.

Conclusions: Applying user-centered design and basic behavioral science yielded design insights to incorporate personalization through user choice with guidance, which may enhance engagement with and potential efficacy of digital health interventions.

(*JMIR Form Res* 2021;5(5):e23809) doi:[10.2196/23809](https://doi.org/10.2196/23809)

KEYWORDS

obesity; binge eating; user-centered design; mobile intervention; engagement; experimental therapeutics

Introduction

Experimental therapeutics and the Science of Behavior Change program at the National Institutes of Health focus on measuring whether experimentally manipulated, hypothesized targets of an intervention lead to behavior change and improved clinical outcomes [1,2]. More specifically, experimental therapeutics first evaluates an intervention effect on a hypothesized mechanism (ie, target); an intervention that engages the target mechanism is then tested to determine whether changes in the target lead to changes in clinical symptoms [3]. To date, the experimental therapeutics Research Domain Criteria framework, defined by the National Institute of Mental Health, has focused on individual-level constructs (eg, cognitive systems, positive and negative valence systems) [4]. For digital (eg, online, mobile) interventions, we have suggested that experimental therapeutics also must account for user engagement as a mediator of clinical outcomes [5] because even a clinically potent intervention will fail to improve symptoms if users do not engage with it. However, engagement is a common problem for digital interventions [6], and digital behavioral interventions have been criticized for using designs that tell users what to do, which can limit considerations for user preferences that impact engagement [7,8].

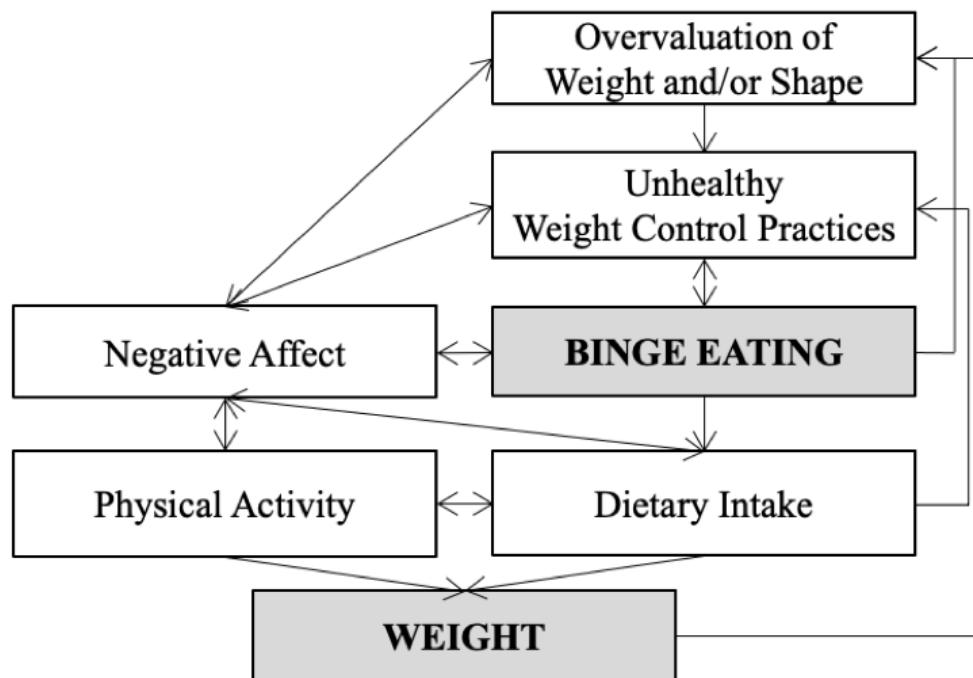
User-centered design provides a methodology for engaging deeply with end users about their needs, goals, and preferences to yield discoveries about the user experience and generate evidence for designing interventions [5,9,10]. User-centered design aims to make technologies and services engaging (eg, useful, usable, satisfying) by working collaboratively and iteratively with end users to ascertain their needs and the ways in which they interact with devices that deliver interventions [9]. As a result, digital tools achieve greater acceptability, understanding, adoption, and engagement [11-14], as well as potentially improved clinical outcomes [5,11], yet clinical scientists in health care have greatly underused design methods [15]. One reason for this underutilization may be that design methods appear to threaten the goal of maintaining fidelity to an evidence-based intervention (ie, delivering the intervention as it is intended). More specifically, conducting design activities

to understand how to deliver a digital behavioral intervention could indeed result in design decisions to modify how the intervention is delivered.

However, the approach to date of simply translating an evidence-based face-to-face intervention to a digital format has not worked; the process relies on what clinicians think users need, the way in which in-person services are delivered does not align with how people engage with their phones [16,17], and it fails to take advantage of the new affordances and opportunities offered through mobile interventions [18]. Instead, new methods are needed to help our field understand how to increase engagement with digital interventions while preserving the core psychological and behavioral principles that can achieve changes in treatment targets.

This paper aims to demonstrate the application of user-centered design and basic behavioral science to inform the design of a new mobile behavioral intervention that addresses both obesity and recurrent binge eating, an eating disorder behavior characterized by eating a large amount of food while experiencing a sense of loss of control over eating [19]. Binge eating affects up to 30% of treatment-seeking adults with obesity [20-22], and more than 75% of people with recurrent binge eating have overweight or obesity [23]. The association between binge eating and weight gain over time [24,25] makes tackling these health outcomes simultaneously important. In line with an experimental therapeutics approach [1-3], the mobile intervention design focuses on addressing putative intervention targets hypothesized as mechanisms that contribute to the cycle of binge eating and changes in weight. The theoretical model, depicted in [Figure 1](#), integrates treatment targets in evidence-based treatments for obesity or binge eating [26-28], with the goal of delivering behavioral and cognitive strategies associated with these targets within a standalone mobile intervention. Yet intervening on these targets could be achieved through several diverse behavioral and cognitive strategies. Because of this, a design lens is needed to learn which strategies are relevant to end users and identify ways to support end users in engaging with these strategies via the planned mobile intervention.

Figure 1. Theoretical model depicting the relation between binge eating and weight gain. The model integrates treatment targets (white boxes) in evidence-based treatments for obesity or binge eating [26-28]. Clinical outcomes are depicted in the gray boxes.



To this end, we applied user-centered design methods to understand how end users select strategies that could address the treatment targets and how they implement these strategies over 1 week. Although this brief period may be insufficient for an end user to determine if a strategy would work for them longer term, it allowed users to answer design questions about how to make strategies within an intervention relevant and engaging without devoting substantial resources to intervention development and deployment. To avoid the shortcomings of prior, overly prescriptive digital interventions, we had users choose their strategy from multiple options. Behavioral economics suggests that leveraging choice architecture, such as using active choice (ie, forced selection among relevant options), may improve engagement and facilitate behavior change [29,30]. Accordingly, using the active choice paradigm enabled assessing if such a feature would be relevant for the mobile intervention.

Taken together, this study employed a mixed methods approach to understand how end users would select and implement a strategy and where they would struggle in the implementation process. Findings inform design implications to increase intervention engagement and, in turn, clinical impact. This work is an example of how user-centered design and basic behavioral science can be leveraged to inform the design of a mobile intervention within an overarching program of research to establish an evidence-based intervention for obesity and binge eating. Such efforts are imperative because few publications have documented the use of design methods to evaluate digital interventions for eating disorders [31], and to our knowledge, we are the first to publish on the use of design methods to create a new digital intervention for people with eating disorders [32].

Methods

Participants

Participants were recruited using dscout (dscout Inc), an online/mobile qualitative and market research platform. Although small for behavioral science research, this sample size is consistent with research in the field of human-computer interaction [33] and was assumed to be sufficient for achieving saturation [34-36]. This size also enabled enrolling a representative sample of target intervention users with diverse perspectives while adhering to sample size constraints in dscout.

Interested individuals were invited to complete an online screening questionnaire in dscout titled “Struggles with eating and weight” that advertised a \$100 reward and 25 openings for participation and gave a brief study overview. The 15 screening questions were developed for this study to confirm eligibility; demographic data (ie, age, gender, race, city of residence) were already captured in the profile of each dscout user. Eligible participants screened positive for obesity (BMI ≥ 30 , based on self-reported height and weight) and self-reported recurrent binge eating (≥ 12 episodes in the past 3 months). For reporting weight, instructions stated, “What is your current weight? Please tell us this number based on when you measure your weight wearing light indoor clothing and without shoes.” For reporting binge eating, instructions stated: “Binge eating is when someone eats an unusually large amount of food and feels a sense of loss of control while eating.” These instructions were written to align with the definition of binge eating in the *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition* [19]. Inclusion criteria required that participants were English-speaking, nonpregnant adults (aged 18 years and older), felt they weighed more than they ought to weigh (yes/no question), struggled with their weight or were interested in

losing weight (yes/no question), were interested in reducing binge eating (yes/no question), and were willing to use an app to address these problems (yes/no question). Among respondents who met the criteria, the final cohort was selected to ensure diversity based on race/ethnicity, gender, and age.

Procedure

Enrollment

This study was approved by the Northwestern University Institutional Review Board. All enrolled participants provided online informed consent. Of those eligible, all participants who were invited (n=25) began the study. Participation was ended early for 3 individuals who stopped completing study procedures. No reasons for discontinuation were provided. Only completers (n=22) were compensated, all of whom received the \$100 compensation.

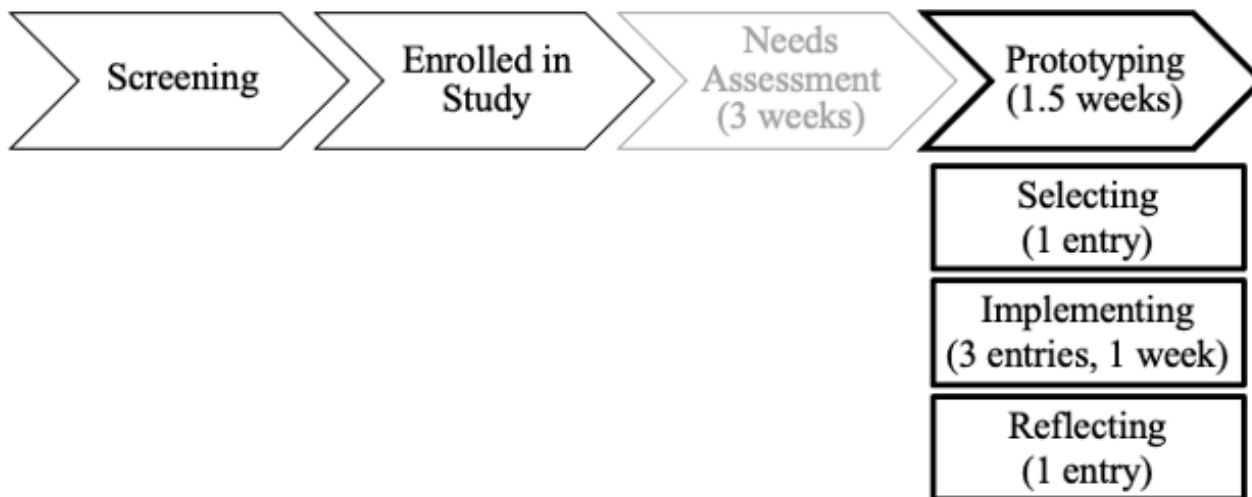
Dscout

All study procedures occurred online via the diary study feature of dscout [37]. Dscout has over 100,000 members who can respond to advertisements and complete screeners to determine eligibility for research opportunities. Users primarily engage with dscout via their smartphones, which facilitates capturing in-the-moment, in-context experiences over time. Multiple research prompts can be included in each diary study. Further, users can submit multiple entries to each research prompt to

assess experiences across contexts. Dscout has several response formats (ie, users can upload videos that are automatically transcribed, upload pictures, submit open-ended responses, and respond to multiple-choice prompts) and has an easy-to-use interface for the researcher to interact with users as needed (eg, to send reminders).

Figure 2 presents a schematic of study activities. Participants completed design research activities over 1 month using dscout. The first 3 weeks comprised a needs assessment to learn about participants' experiences with obesity and binge eating, strategies they have used to address these problems, and ideas for managing weight and eating. This paper focuses on the prototyping design activity that occurred in the final 1.5 weeks when participants were asked to "try making one change" to help with weight and binge eating by selecting and implementing a strategy for 1 week. Prototyping is used to iteratively evaluate design options conveyed through versions of a product (prototypes) [38]. Prototypes may or may not closely resemble the intended product, referred to as high- or low-fidelity prototypes. In this study, dscout was leveraged as a low-fidelity prototype to gain rapid insights for designing the delivery of strategies in the mobile intervention. As shown in Figure 2, the prototyping activity was administered via 3 research prompts. All 22 participants who began the prototyping activity completed it.

Figure 2. Schematic of participant flow through the study activities, including prototyping, the focus of this analysis, and the three research prompts it comprises. Of the 25 participants who enrolled, 22 began and completed prototyping.



Prototyping Activity

Assessment Guide

An assessment guide was created that specified the research prompts that would be administered to evaluate user experiences. Questions were drafted by AKG with input from SAM, MR, DCM, and JEW. These individuals are researchers with expertise in the treatment of eating disorders and obesity, digital interventions, and/or user-centered design. AKG and SWN (an undergraduate student) then practiced answering the questions by submitting mock entries in dscout for internal testing prior to launch with participants. AKG oversaw study administration and data collection with participants. Study procedures and

assessment items were consistent for all participants. After participants' entries were submitted, SWN edited the transcribed videorecordings for accuracy and deidentification.

Selecting a Strategy

At the start of the week, participants were prompted to submit their first entry. In this entry, participants were asked to report their weight and number of binge eating episodes in the previous week and were then prompted to select one strategy from 20 options to practice for the week. The instructions did not indicate a limit to how many strategies participants could select. The strategies aligned with the putative intervention targets of the intervention's theoretical model (Figure 1) and were based on

evidence-based behavioral and cognitive strategies in interventions for obesity and binge eating [26-28]. Using the “think aloud” design technique [9,39], participants recorded a video while picking their strategy, talking aloud about their thought process as they made their choice. Participants then provided an open-ended response indicating their guess for how the strategy would help them.

Implementing the Strategy

Over the next week, participants submitted 3 entries showing moments in which they practiced or were struggling to practice their strategy. For those struggling, participants were asked to share what was getting in the way. For each entry, participants submitted a video and an open-ended response describing the experience.

Reflecting on Implementation

At the end of the week, participants were prompted to submit their final entry. They recorded a video reflecting on how the experience went, if it matched what they guessed would happen, and the evidence they collected on whether the strategy was helpful. Responses were coded as “planned,” “somewhat as planned,” and “not as planned.” Participants answered a yes/no question on whether they would continue using the strategy in the future and reported their weight and number of binge eating episodes over the past week.

Analyses

Analyses focused on the process by which participants selected and implemented strategies and their change in outcomes. Qualitative analyses were conducted using Dedoose (SocioCultural Research Consultants), a qualitative data analysis software. Qualitative data from the baseline strategy selection

process were coded separately from the weekly entries on implementation. Qualitative data were analyzed using thematic analysis based on the methodology of Braun and Clark [40], which involved reviewing transcripts to become familiar with the data, generating codes through open coding, iteratively applying the codes to the transcripts, and organizing the codes systematically into broader themes. AKG oversaw these analyses with input and review by SWN, SAM, and MR. Quantitative data were aggregated. The difference in weight and binge eating between the start and end of the week was calculated for each participant, the average of which was then calculated to explore average within-subject 1-week changes in weight and binge eating. Given the small sample size and exploratory nature of the quantitative analyses, a significance test was not conducted.

Results

Sample Characteristics

A total of 22 participants completed all study procedures and were included in the analyses. Table 1 shows demographic information on study completers. Average age was 37.0 (SD 10.2) years; 64% (14/22) identified as female. Participants identified as White (7/22, 32%), African American/Black (6/22, 27%), Hispanic/Latino (6/22, 27%), and Asian or Pacific Islander (2/22, 9%); one participant (5%) did not report their race/ethnicity. Participants reported living in 12 US states.

At screening, average BMI was 37.1 (SD 5.4, range 30.3 to 49.4), and average number of binge eating episodes over the previous 3 months was 20.5 (SD 7.3, range 12 to 35). All participants endorsed previous attempts to lose weight, and 91% (20/22) endorsed previous attempts to stop binge eating.

Table 1. Study participant demographics.

ID	Sex	Age	Race/ethnicity	State of residence
1	Male	42	Asian or Pacific Islander	California
3	Female	45	African American/Black	Arizona
4	Female	27	Hispanic/Latino	California
5	Female	35	Hispanic/Latino	California
6	Male	42	Hispanic/Latino	Pennsylvania
7	Male	43	African American/Black	New York
8	Female	43	Prefer not to respond	South Carolina
9	Female	45	Hispanic/Latino	Illinois
10	Female	36	White	South Carolina
11	Male	47	White	Texas
12	Female	62	White	California
13	Male	30	White	Texas
14	Female	39	Asian or Pacific Islander	Illinois
15	Female	20	White	Illinois
16	Female	43	African American/Black	California
19	Male	30	Hispanic/Latino	California
20	Female	30	White	Virginia
21	Female	22	African American/Black	Illinois
22	Male	39	Hispanic/Latino	Florida
23	Female	26	White	Ohio
24	Female	22	African American/Black	North Carolina
25	Male	45	African American/Black	New Jersey

Selecting a Strategy

Participants selected 15 unique strategies, shown in [Table 2](#). Although prompted to pick 1 strategy, participants selected between 1 and 3 strategies. Most participants (15/22, 68%) selected 1 strategy, 23% (5/22) of participants selected 2 strategies, and 9% (2/22) selected 3 strategies. The most commonly selected strategy was to “plan for the meals you’ll

eat this week,” selected by 6 participants. The majority (25/31, 81%) of selected strategies were associated with the intervention targets of dietary intake and physical activity, whereas only 6 selections were associated with overvaluation of weight and/or shape, unhealthy weight control practices, and negative affect. The 5 strategies no one selected were associated with these latter 3 intervention targets.

Table 2. Selection of strategies.

Putative intervention target and strategy	Times selected	Went as planned?			Helpful ^a %
		Yes	Some	No	
Dietary intake					
Eat meals and snacks at the same time each day.	2	✓	✓	__ ^b	100
Avoid eating snacks that you didn't plan to eat.	2	✓	—	✓	50
Plan for the meals you'll eat this week.	6	✓	✓✓✓✓	✓	83
Find a buddy who will help you eat more healthfully.	2	✓✓	—	—	100
Eat smaller portions.	3	✓✓✓	—	—	100
Eat more fruits and vegetables.	2	—	—	✓✓	50
Eat less fast food.	2	✓	✓	—	100
Total selections of this target	19	—	—	—	—
Physical activity					
Regularly (approximately 3 times per week) do physical activity like walking, riding a bike, or going to the gym (unless a doctor has said it is not appropriate/healthy for you to exercise right now).	4	✓✓	✓✓	—	100
Have less screen time: watch less television and spend less time on your computer, tablet, or phone.	1	✓	—	—	100
Find a buddy who will help you be more physically active.	1	—	✓	—	100
Total selections of this target	6	—	—	—	—
Overvaluation of weight and/or shape					
When you notice yourself criticizing something about your body, stop yourself. Ask yourself: What is the evidence that the criticism is true or not true? Then think of a more balanced conclusion you can draw about your body.	0	—	—	—	—
List things you like and value about yourself as a person. Remind yourself of things that are more important to you than how your body looks or how much you weigh.	1	✓	—	—	100
Avoid spending time in front of the mirror pointing out what you think of as your "flaws."	2	✓	✓	—	100
Stop yourself when you dwell on "feeling fat." Tell yourself that "fat" is not a feeling and instead say something to yourself that is not self-blaming or self-shaming.	0	—	—	—	—
Total selections of this target	3	—	—	—	—
Unhealthy weight control practices					
Avoid skipping meals or going for long stretches of time without eating.	1	—	✓	—	100
Avoid "dieting" and cutting out certain types of foods.	0	—	—	—	—
Try eating one serving of a food that you've been avoiding because you consider it a "trigger" food for binge eating.	1	—	✓	—	0
Total selections of this target	2	—	—	—	—
Negative affect					
Do activities that make you happy and do not involve food.	0	—	—	—	—
Notice times when you're feeling down and find something that makes you feel a bit better about the situation.	0	—	—	—	—
Ask a friend or loved one to do something enjoyable together or repair a relationship in which you had a disagreement or falling out.	1	✓	—	—	100
Total selections of this target	1	—	—	—	—

^aIndicates the percentage of participants who rated their selected strategy as helpful to them.

^bNot applicable.

Participants gave a variety of reasons for selecting versus not selecting strategies. Some participants indicated they selected a strategy because it was something they already were pursuing, whereas for others, this was the reason they did not select that

strategy. Participants also indicated they selected strategies they perceived to be attainable and easy to complete or adjust to doing, that they perceived would be most helpful to them in managing their eating or weight, or that were new to them and therefore were perceived to be worth trying. Participants did not select strategies that they described as having been unhelpful in the past or that they believed would be unhelpful for achieving their goals around weight and binge eating. Some said certain strategies did not apply to them and thus would be challenging to achieve. Finally, participants did not select strategies because they thought they would fail in the implementation.

Participants who selected multiple strategies indicated their selections were motivated by a desire to capitalize on recent successes with those strategies or make progress with recently planned goals or because the strategies were perceived as serving a similar purpose. Two participants did not provide a rationale for why they selected multiple strategies.

While selecting strategies, all but 2 participants set a plan and/or identified ways to implement the strategy. Most participants (19/22, 86%) listed the benefits of maintaining a plan. Participants guessed their strategy would help them manage binge eating or weight (21/22, 95%), address eating-related triggers (17/22, 77%; eg, cravings, unplanned snacking, negative self-talk), or improve a related area (10/22, 45%; eg, increase self-esteem, happiness, or productivity; be more present with children; or improve their immune system).

Implementing the Strategies

Over the week, all participants reported they practiced implementing a strategy at least once; 82% (18/22) described moments in which they struggled with implementation. A total of 23% (5/22) of participants indicated they focused on a new or additional strategy. Participants changed strategies for 2 reasons. One reason was because they realized they would be unable to work on their originally selected choice (ie, scheduling difficulties prevented exercising with a buddy). The other reason was to address a more pressing and/or relevant problem area that had presented (eg, reduce binge eating by avoiding eating unplanned snacks, increase physical activity to avoid overeating, plan for meals and snacks to reduce overeating when very hungry, or eat less fast food after a recent increase in this behavior).

Participants experienced successes and challenges with implementation; overarching themes from these results are described here, and specific details for each strategy are detailed in [Multimedia Appendix 1](#).

The ways in which participants were successful ranged from using individually focused techniques (eg, changing a routine) to collaborating (eg, using resources and others for support); participants used techniques that best fit their needs. Participants were successful when they planned in advance, including when to eat (eg, setting a schedule), what to eat (eg, meal planning, packing healthy snacks), or how much to eat (eg, ordering smaller portions). Participants described ways they avoided triggers and unhealthy behaviors, such as avoiding eating unplanned snacks, reducing stress, or doing alternative activities to avoid overeating or triggers for overeating (eg, engaging in

physical activity, referencing a list of alternative responses in the face of triggers, or using flash cards with positive statements to combat negative thoughts about their body). A total of 18% (4/22) of participants described changing their routine to engage in healthy behaviors; 18% (4/22) practiced moderation with their strategy (eg, using smaller dishes, eating smaller amounts more frequently, or eating unhealthy snacks in moderation); 23% (5/22) leveraged resources (eg, social media, commercial entities like Weight Watchers) and other people for support; 9% (2/22) found new outlets for physical activity, like being active with pets or doing chores; and 23% (5/22) described ways they challenged negative thinking, engaged in positive self-talk, and practiced motivation-enhancing techniques to support progress. During the week, 9% of participants (2/22) reflected on the results of implementing their strategy.

Participants also described challenges that affected implementation. Participants faced challenges with integrating healthy behaviors into their schedules. They had difficulties practicing their strategy in unforeseen situations (eg, when away from home or in unplanned circumstances) and challenges with sufficient planning. A total of 32% (7/22) of participants reported challenges associated with their home environment (eg, others in the home do not support healthy eating, ordering fast food is more convenient than cooking, or being at home triggered binge eating); 27% (6/22) shared how preferences for other behaviors (eg, low desire to eat healthy foods despite planning to do so or using birthdays to justify unhealthy eating) and changes in motivation affected implementation; and 14% (3/22) described the impact of stress, low mood or energy, and financial difficulties on implementation. Finally, 23% (5/22) conveyed they misunderstood the strategy and how it should be applied.

Reflecting on Implementation

In postimplementation reflections, 41% (9/22) said implementation went as planned, 41% (9/22) said it went somewhat as planned, and 18% (4/22) said it did not go as planned. A total of 82% (18/22) endorsed their strategy as helpful. [Table 2](#) shows these ratings relative to each strategy. Of the participants who endorsed their strategy as not helpful, 75% (3/4) said implementation did not go as planned. Finally, 86% (19/22) of participants endorsed plans to continue using their strategy.

One-Week Changes in Weight and Binge Eating

At the start of the week, participants reported an average weight of 225.9 (SD 34.6, range 162 to 307) pounds and engaged in an average of 3.4 (SD 2.1, range 0 to 8) binge eating episodes over the prior week. After implementation, participants reported an average weight of 223.7 (SD 36.8, range 159 to 320) pounds and average of 1.7 (SD 1.2, range 1 to 5) binge eating episodes. Average within-subject changes in weight and binge eating were -2.2 (SD -5.0 , range -11 to 13) pounds and -1.6 (SD -1.8 , range -6 to 1) episodes, respectively.

Discussion

Principal Findings

User-centered design has the potential to improve engagement with and efficacy of behavioral interventions [5,9,10]. This study aimed to inform the design of a mobile intervention for obesity and binge eating by applying user-centered design and basic behavioral science to understand how end users would select and implement strategies associated with putative intervention targets for changing weight and binge eating. Results offered useful implications for intervention design and future research.

Design Implications

In this prototyping activity, offering a choice in selecting strategies seemed to be successful given the variation of strategies participants selected. Although we did not compare active choice to no choice and therefore cannot conclude that active choice is better than assigning strategies to participants, evidence from behavioral economics shows that prompting people to make choices from several options (ie, active choice [29,30]) can spur behavior change. This may be because such an approach capitalizes on user motivation to make a change; on the whole in this study, participants selected strategies to experience success. Helping users achieve early success could make them more likely to sustain engagement; indeed, most participants endorsed plans to continue using their strategy after the week. Sustained engagement is especially important for people with obesity as failed weight loss attempts predict reduced success in future weight loss efforts [41,42]. Thus, the first design implication is to incorporate ways to offer users choices in selecting strategies and, more broadly, the areas they want to address in treatment. Offering choices also overcomes critiques of digital interventions for using designs that limit autonomy; digital interventions often have preset curricula and prescribed behavior change goals [7,8,43].

With that said, a second, complementary design implication is to reduce or scaffold the number of strategies presented. Research on decision making shows that offering too many choices leads to choice overload [44,45]. For this study, 20 options were offered. However, several entries indicated that some strategies could be collapsed, and 5 strategies were not selected. For example, no one selected “avoid ‘dieting’ and cutting out certain types of foods,” perhaps due to misunderstanding what this strategy means and why it has clinical relevance or because users may not want to follow this recommendation. A challenge for behavior change interventions is balancing what users want with what is clinically indicated when these areas may not align. It is important to incorporate designs that make less appealing but clinically necessary strategies more enticing and relevant rather than have those strategies disregarded. Designs promoting choice could be achieved through guided customization, which facilitates user choice within a defined array of options or via credible suggestions [46]. For example, it may be beneficial to deliver a distilled set of strategies [47], particularly early in the intervention, that are appealing to end users while also aligned with best clinical practices and design features that guide users

to appropriate strategies and provide a rationale for their potential benefit to the user [48]. This distilled list or guided recommendations could be based on users’ identified problem areas or past progress, which could strengthen its perceived relevance to the user. Strategy selection around commonly avoided strategies also could be an area where guidance from a coach may be useful. Determining the optimal number of strategies presented at any one time remains to be tested, as does testing whether allowing users to select multiple strategies has benefits over selecting only single strategies at a time [49,50].

The third design implication is to define strategies when they are presented. Participants were not given definitions for the 20 options (to learn how users interpret the strategies) nor did they receive feedback as they shared their experiences throughout the week. As a result, seemingly straightforward strategies were interpreted in multiple ways. The variation in how participants interpreted strategies was particularly notable given that most selected strategies focused on the main intervention targets of behavioral weight loss treatment, changing dietary intake and physical activity, and 100% of participants endorsed prior weight loss attempts. This suggests that digital intervention designers cannot make assumptions about what and how much users know about eating and weight management; they need to educate users about strategies so they are positioned for success. Further, for nearly 60% of participants, implementation did not go as planned, likely because the strategies lacked specificity in how they should be implemented. Consequently, some participants abandoned their strategy or reported feelings of failure, disheartenment, and decreased willingness to practice that strategy again in the future—opposite the intention of offering choice to increase engagement. Based on these findings, an intervention architecture may need to include descriptions of what the strategy is, why the strategy is relevant to managing eating and weight, and how the strategy could be implemented. Presenting these details could help avoid misinterpretations, make unfamiliar strategies seem less daunting, and offer structure and scaffolding for their implementation.

However, delivering only instructional content on how to implement a strategy is likely insufficient. A challenge for technology-mediated services is moving users from qualitative, often distant goals to something concrete and actionable [51]. Thus, the fourth design implication is to provide support for implementing the strategy over time, something that was missing from the prototyping activity. Although results showed that participants already had some tacit understanding of evidence-based behavior change techniques (eg, planning for when, what, or how much to eat reflects action planning; avoiding eating unplanned snacks because doing so triggers overeating reflects information about antecedents) [52], many participants still struggled. Accordingly, findings suggest there would be utility in incorporating guidance and support as users implement their strategies.

Guidance and support could be delivered through coaching and content or app designs that model how to implement strategies. Throughout implementation, timely feedback on progress would be helpful, too, as this is an important component of health-related behavior change strategies [53,54] and

measurement-based care more broadly [55]. Because allowing participants to implement their strategy for 1 week without feedback was problematic for some, delivering feedback soon after users begin implementation may enable users to course-correct more quickly. To balance against overly prescriptive intervention designs, corrective feedback could focus on problematic implementation (eg, when the user misunderstands the strategy). Another challenge was planning how to implement strategies, including across contexts, and executing those plans. Thus, when a strategy is presented, guidance should include designs that help users plan for implementation. At the same time, participants used existing resources to support implementation, like finding recipes and physical activity videos online. Such insights suggest it can be helpful to direct users to existing resources or help users creatively harness resources in their everyday environment, which could also save the time and costs of building app-specific versions of these resources in the intervention.

Application of User-Centered Design to Behavioral Science

Much can be learned from this study in terms of applying user-centered design to drive progress for health-related behavioral interventions. This low-fidelity prototyping activity used qualitative and quantitative data to understand why and how users engage with aspects of an intervention—in this case, selecting and implementing strategies. The data collection platform and design methods enabled gathering in-the-moment perspectives from diverse participants who were matched to intended intervention users. End user perspectives were rapidly gathered with low participant burden, given that each entry required only a few minutes to complete and could be submitted from participants' smartphones. Researcher burden was also minimized through the use of remote recruitment, remote data collection with multiple response types, and automatic video transcription. The design methods generated insights for intervention design without spending time or money developing a mobile intervention or creating high-fidelity prototypes. Further, these insights were gleaned from relatively few participants.

Limitations

However, limitations should be noted. First, because procedures occurred remotely using an existing platform (ie, dscout), the

research team was unable to ask clarifying or follow-up questions about participant entries, which may have limited the number and depth of insights generated. Second, the lack of definitions for each of the strategies may have influenced strategy selection and adherence and therefore generalizability and clinical relevance of the findings. Third, the study design makes it difficult to disentangle how participant improvements in the implementation process were affected by having to submit multiple entries about their progress, as longitudinal design research itself can affect behavior [56]. Also, despite asking participants to submit 3 entries showing their progress over the week, we did not ask participants to report the total number of times they implemented their strategy over the week. Fourth, although a 1-week observation period was used to gain rapid insights into strategy selection and implementation, this timeframe may have been too short for users to confidently assess the strategy's efficacy. Fifth, the study cannot inform how users will iterate on their experiences implementing strategies or whether they will sustain engagement with implementing self-selected strategies over a longer duration. Since behavior change must occur over the long term, future design research could explore how to support users' iterative learning over time [57]. Going forward, design recommendations should be evaluated for their impact on longer term engagement and clinical improvement. Last, although participants had small average improvements in weight and binge eating, these findings should be interpreted with caution given the brief observation period and small sample size, as achieving clinical change was not the objective of the prototyping activity. Moreover, use of self-report to assess weight and binge eating can be flawed and subject to recall biases.

Conclusions

Results of this study highlight the translational potential of applying user-centered design and basic behavioral science to inform the design of a mobile behavioral intervention for obesity and binge eating. Discovering ways to make digital technologies relevant to end users is imperative to ensure these tools fit into the fabric of users' lives and therefore are used in the moments and contexts when they are needed most. Such efforts can substantially improve engagement with and potential efficacy of digital health-related behavioral interventions.

Acknowledgments

This work was supported by grants K01 DK116925 and P50 MH115837 from the National Institutes of Health.

Conflicts of Interest

AKG reports a grant from the National Institute of Diabetes and Digestive and Kidney Diseases during the study and personal fees from Actualize Therapy outside the submitted work. SAM reports a grant from the National Institute of Mental Health during the study and personal fees from Apple Inc outside the submitted work. BS reports personal fees from Apple and Actigraph outside the submitted work. DCM reports personal fees from Otsuka Pharmaceuticals, Apple Inc, Pear Therapeutics, One Mind Foundation, and Adaptive Health Inc outside the submitted work. MR, SWN, EAG, AC, and JEW have no disclosures.

Multimedia Appendix 1

Participant successes and challenges implementing each selected strategy.

[DOCX File , 42 KB - [formative_v5i5e23809_app1.docx](#)]

References

1. Riddle M, Science of Behavior Change Working Group. News from the NIH: using an experimental medicine approach to facilitate translational research. *Transl Behav Med* 2015 Dec;5(4):486-488 [FREE Full text] [doi: [10.1007/s13142-015-0333-0](#)] [Medline: [26622921](#)]
2. Nielsen L, Riddle M, King JW, NIH Science of Behavior Change Implementation Team, Aklin WM, Chen W, et al. The NIH Science of Behavior Change Program: transforming the science through a focus on mechanisms of change. *Behav Res Ther* 2018 Feb;101:3-11 [FREE Full text] [doi: [10.1016/j.brat.2017.07.002](#)] [Medline: [29110885](#)]
3. Insel TR. The NIMH experimental medicine initiative. *World Psychiatry* 2015 Jun;14(2):151-153 [FREE Full text] [doi: [10.1002/wps.20227](#)] [Medline: [26043323](#)]
4. Strategic plan for research. Bethesda: National Institute of Mental Health; 2015. URL: <https://www.nih.gov/sites/default/files/about-nih/strategic-plan-fy2016-2020-508.pdf> [accessed 2020-12-01]
5. Graham AK, Lattie EG, Mohr DC. Experimental therapeutics for digital mental health. *JAMA Psychiatry* 2019 Aug 21:1. [doi: [10.1001/jamapsychiatry.2019.2075](#)] [Medline: [31433448](#)]
6. Yardley L, Spring BJ, Riper H, Morrison LG, Crane DH, Curtis K, et al. Understanding and promoting effective engagement with digital behavior change interventions. *Am J Prev Med* 2016 Nov;51(5):833-842. [doi: [10.1016/j.amepre.2016.06.015](#)] [Medline: [27745683](#)]
7. Baumer E, Katz S, Freeman J, Adams P, Gonzales A, Pollak J. Prescriptive persuasion and open-ended social awareness: expanding the design space of mobile health. 2012 Presented at: ACM 2012 conference on Computer Supported Cooperative Work; 2012; Seattle p. 475-484. [doi: [10.1145/2145204.2145279](#)]
8. Cordeiro F, Epstein DA, Thomaz E, Bales E, Jagannathan AK, Abowd GD, et al. Barriers and negative nudges: exploring challenges in food journaling. *Proc SIGCHI Conf Hum Factor Comput Syst* 2015 Apr;2015:1159-1162 [FREE Full text] [doi: [10.1145/2702123.2702155](#)] [Medline: [26894233](#)]
9. Graham AK, Wildes JE, Reddy M, Munson SA, Barr Taylor C, Mohr DC. User-centered design for technology-enabled services for eating disorders. *Int J Eat Disord* 2019 Oct;52(10):1095-1107 [FREE Full text] [doi: [10.1002/eat.23130](#)] [Medline: [31313370](#)]
10. Klasnja P, Hekler E, Korinek E, Harlow J, Mishra S. Toward usable evidence: optimizing knowledge accumulation in HCI research on health behavior change. *Proc SIGCHI Conf Hum Factor Comput Syst* 2017 May;2017:3071-3082 [FREE Full text] [doi: [10.1145/3025453.3026013](#)] [Medline: [30272059](#)]
11. McCurdie T, Taneva S, Casselman M, Yeung M, McDaniel C, Ho W, et al. mHealth consumer apps: the case for user-centered design. *Biomed Instrum Technol* 2012;Suppl:49-56. [doi: [10.2345/0899-8205-46.s2.49](#)] [Medline: [23039777](#)]
12. Maguire M. Methods to support human-centred design. *Int J Human Comput Studies* 2001 Oct;55(4):587-634. [doi: [10.1006/ijhc.2001.0503](#)]
13. Kujala S. User involvement: A review of the benefits and challenges. *Behaviour & Information Technology* 2003 Jan;22(1):1-16. [doi: [10.1080/01449290301782](#)]
14. Kushniruk A, Nørh C. Participatory design, user involvement and health IT evaluation. *Stud Health Technol Inform* 2016;222:139-151. [Medline: [27198099](#)]
15. Altman M, Huang TTK, Breland JY. Design thinking in health care. *Prev Chronic Dis* 2018 Sep 27;15:E117 [FREE Full text] [doi: [10.5888/pcd15.180128](#)] [Medline: [30264690](#)]
16. Mohr DC, Weingardt KR, Reddy M, Schueller SM. Three problems with current digital mental health research . . . and three things we can do about them. *Psychiatr Serv* 2017 May 01;68(5):427-429 [FREE Full text] [doi: [10.1176/appi.ps.201600541](#)] [Medline: [28412890](#)]
17. Mohr DC, Riper H, Schueller SM. A solution-focused research approach to achieve an implementable revolution in digital mental health. *JAMA Psychiatry* 2018 Feb 01;75(2):113-114. [doi: [10.1001/jamapsychiatry.2017.3838](#)] [Medline: [29238805](#)]
18. Riley WT, Rivera DE, Atienza AA, Nilsen W, Allison SM, Mermelstein R. Health behavior models in the age of mobile interventions: are our theories up to the task? *Transl Behav Med* 2011 Mar;1(1):53-71 [FREE Full text] [doi: [10.1007/s13142-011-0021-7](#)] [Medline: [21796270](#)]
19. American Psychiatric Association. *Diagnostic and Statistical Manual for Mental Disorders*. 5th Edition. Washington: American Psychiatric Publishers; 2013.
20. Mitchell JE, King WC, Courcoulas A, Dakin G, Elder K, Engel S, et al. Eating behavior and eating disorders in adults before bariatric surgery. *Int J Eat Disord* 2015 Mar;48(2):215-222 [FREE Full text] [doi: [10.1002/eat.22275](#)] [Medline: [24719222](#)]
21. Wilfley D, Cítrome L, Herman B. Characteristics of binge eating disorder in relation to diagnostic criteria. *Neuropsychiatr Dis Treat* 2016;12:2213-2223 [FREE Full text] [doi: [10.2147/NDT.S107777](#)] [Medline: [27621631](#)]
22. Hudson JI, Hiripi E, Pope HG, Kessler RC. The prevalence and correlates of eating disorders in the National Comorbidity Survey Replication. *Biol Psychiatry* 2007 Feb 01;61(3):348-358 [FREE Full text] [doi: [10.1016/j.biopsych.2006.03.040](#)] [Medline: [16815322](#)]

23. Brownley KA, Berkman ND, Peat CM, Lohr KN, Cullen KE, Bann CM, et al. Binge-eating disorder in adults: a systematic review and meta-analysis. *Ann Intern Med* 2016 Sep 20;165(6):409-420 [FREE Full text] [doi: [10.7326/M15-2455](https://doi.org/10.7326/M15-2455)] [Medline: [27367316](https://pubmed.ncbi.nlm.nih.gov/27367316/)]
24. Sonnevile KR, Horton NJ, Micali N, Crosby RD, Swanson SA, Solmi F, et al. Longitudinal associations between binge eating and overeating and adverse outcomes among adolescents and young adults: does loss of control matter? *JAMA Pediatr* 2013 Feb;167(2):149-155 [FREE Full text] [doi: [10.1001/2013.jamapediatrics.12](https://doi.org/10.1001/2013.jamapediatrics.12)] [Medline: [23229786](https://pubmed.ncbi.nlm.nih.gov/23229786/)]
25. Meany G, Conceição E, Mitchell JE. Binge eating, binge eating disorder and loss of control eating: effects on weight outcomes after bariatric surgery. *Eur Eat Disord Rev* 2014 Mar;22(2):87-91 [FREE Full text] [doi: [10.1002/erv.2273](https://doi.org/10.1002/erv.2273)] [Medline: [24347539](https://pubmed.ncbi.nlm.nih.gov/24347539/)]
26. Fairburn CG, Cooper Z, Shafran R. Cognitive behaviour therapy for eating disorders: a "transdiagnostic" theory and treatment. *Behav Res Ther* 2003 May;41(5):509-528. [Medline: [12711261](https://pubmed.ncbi.nlm.nih.gov/12711261/)]
27. Wilfley DE, Hayes JF, Balantekin KN, Van Buren DJ, Epstein LH. Behavioral interventions for obesity in children and adults: evidence base, novel approaches, and translation into practice. *Am Psychol* 2018 Nov;73(8):981-993 [FREE Full text] [doi: [10.1037/amp0000293](https://doi.org/10.1037/amp0000293)] [Medline: [30394777](https://pubmed.ncbi.nlm.nih.gov/30394777/)]
28. Fairburn C. *Cognitive Behavior Therapy and Eating Disorders*. New York: Guilford Press; 2008.
29. Nease RF, Glave Frazee S, Zarin L, Miller SB. Choice architecture is a better strategy than engaging patients to spur behavior change. *Health Aff (Millwood)* 2013 Feb;32(2):242-249. [doi: [10.1377/hlthaff.2012.1075](https://doi.org/10.1377/hlthaff.2012.1075)] [Medline: [23381516](https://pubmed.ncbi.nlm.nih.gov/23381516/)]
30. Keller PA, Harlam B, Loewenstein G, Volpp KG. Enhanced active choice: a new method to motivate behavior change. *J Consumer Psychol* 2011 Oct;21(4):376-383. [doi: [10.1016/j.jcps.2011.06.003](https://doi.org/10.1016/j.jcps.2011.06.003)]
31. Nitsch M, Dimopoulos CN, Flaschberger E, Saffran K, Kruger JF, Garlock L, et al. A guided online and mobile self-help program for individuals with eating disorders: an iterative engagement and usability study. *J Med Internet Res* 2016 Jan 11;18(1):e7 [FREE Full text] [doi: [10.2196/jmir.4972](https://doi.org/10.2196/jmir.4972)] [Medline: [26753539](https://pubmed.ncbi.nlm.nih.gov/26753539/)]
32. Weinheimer EA, Chang A, Neubert SW, Wildes JE, Graham AK. Past, current, and future willingness to engage with treatment targets: applying user-centered design to inform the design of a mobile behavioral intervention. *Int J Eat Disord* 2020 Apr;53(4):611-617 [FREE Full text] [doi: [10.1002/eat.23252](https://doi.org/10.1002/eat.23252)] [Medline: [32112592](https://pubmed.ncbi.nlm.nih.gov/32112592/)]
33. Caine K. Local standards for sample size at CHI. 2016 Presented at: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems; 2016; San Jose p. 981-992. [doi: [10.1145/2858036.2858498](https://doi.org/10.1145/2858036.2858498)]
34. Guest G. How many interviews are enough? An experiment with data saturation and variability. *Field Methods* 2006 Feb 01;18(1):59-82. [doi: [10.1177/1525822X05279903](https://doi.org/10.1177/1525822X05279903)]
35. Fusch P, Ness L. Are we there yet? Data saturation in qualitative research. *Qual Rep* 2015 Sep 8;20:1408-1416. [doi: [10.46743/2160-3715/2015.2281](https://doi.org/10.46743/2160-3715/2015.2281)]
36. Stavros C, Westberg K. Using triangulation and multiple case studies to advance relationship marketing theory. *Qualitative Mrkt Res Int J* 2009 Jun 12;12(3):307-320. [doi: [10.1108/13522750910963827](https://doi.org/10.1108/13522750910963827)]
37. dscout Diary. URL: <https://dscout.com/diary> [accessed 2020-12-01]
38. Houde S, Hill C. What do prototypes prototype? In: Helander M, Landauer T, Prabhu P, editors. *Handbook of Human-Computer Interaction*. 2nd Edition. Amsterdam: Elsevier Science BV; 1997:367-381.
39. Martin B, Hanington B. *Universal Methods of Design*. Beverly: Rockport Publishers; 2012.
40. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol* 2006 Jan;3(2):77-101. [doi: [10.1191/1478088706qp063oa](https://doi.org/10.1191/1478088706qp063oa)]
41. Teixeira PJ, Going SB, Sardinha LB, Lohman TG. A review of psychosocial pre-treatment predictors of weight control. *Obes Rev* 2005 Feb;6(1):43-65. [doi: [10.1111/j.1467-789X.2005.00166.x](https://doi.org/10.1111/j.1467-789X.2005.00166.x)] [Medline: [15655038](https://pubmed.ncbi.nlm.nih.gov/15655038/)]
42. Myers VH, McVay MA, Champagne CM, Hollis JF, Coughlin JW, Funk KL, et al. Weight loss history as a predictor of weight loss: results from Phase I of the weight loss maintenance trial. *J Behav Med* 2013 Dec;36(6):574-582 [FREE Full text] [doi: [10.1007/s10865-012-9450-0](https://doi.org/10.1007/s10865-012-9450-0)] [Medline: [22907176](https://pubmed.ncbi.nlm.nih.gov/22907176/)]
43. Purpura S, Schwanda V, Williams K, Stubler W, Sengers P. Fit4life: the design of a persuasive technology promoting healthy behavior and ideal weight. 2011 Presented at: SIGCHI Conference on Human Factors in Computing Systems; 2011; Vancouver p. 423-432. [doi: [10.1145/1978942.1979003](https://doi.org/10.1145/1978942.1979003)]
44. Iyengar SS, Lepper MR. When choice is demotivating: can one desire too much of a good thing? *J Pers Soc Psychol* 2000 Dec;79(6):995-1006. [doi: [10.1037//0022-3514.79.6.995](https://doi.org/10.1037//0022-3514.79.6.995)] [Medline: [11138768](https://pubmed.ncbi.nlm.nih.gov/11138768/)]
45. Hadar L, Sood S. When knowledge is demotivating: subjective knowledge and choice overload. *Psychol Sci* 2014 Sep;25(9):1739-1747. [doi: [10.1177/0956797614539165](https://doi.org/10.1177/0956797614539165)] [Medline: [25037963](https://pubmed.ncbi.nlm.nih.gov/25037963/)]
46. Zhang R, Ringland KE, Paan M, Mohr D, Reddy M. Designing for emotional well-being: integrating persuasion and customization into mental health technologies [accepted for presentation]. 2021 Presented at: 2021 ACM CHI Conference on Human Factors in Computing Systems; 2021; Yokohama.
47. Johnson EJ, Shu SB, Dellaert BGC, Fox C, Goldstein DG, Häubl G, et al. Beyond nudges: tools of a choice architecture. *Mark Lett* 2012 May 25;23(2):487-504. [doi: [10.1007/s11002-012-9186-1](https://doi.org/10.1007/s11002-012-9186-1)]
48. Schroeder J, Karkar R, Murinova N, Fogarty J, Munson SA. Examining opportunities for goal-directed self-tracking to support chronic condition management. *Proc ACM Interact Mob Wearable Ubiquitous Technol* 2019 Dec;3(4):1 [FREE Full text] [doi: [10.1145/3369809](https://doi.org/10.1145/3369809)] [Medline: [32656490](https://pubmed.ncbi.nlm.nih.gov/32656490/)]

49. Chung C, Munson SA. Should you always have a backup? One vs. two goals in games. *Int J Hum-Comput Stud* 2017 Dec;108:32-40. [doi: [10.1016/j.ijhcs.2017.07.001](https://doi.org/10.1016/j.ijhcs.2017.07.001)]
50. Munson S, Consolvo S. Exploring goal-setting, rewards, self-monitoring, and sharing to motivate physical activity. 2012 Presented at: 2012 6th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth) and Workshops; 2012; San Diego. [doi: [10.4108/icst.pervasivehealth.2012.248691](https://doi.org/10.4108/icst.pervasivehealth.2012.248691)]
51. Niess J, Wozniak P. Supporting meaningful personal fitness: the tracker goal evolution model. 2018 Presented at: 2018 CHI Conference on Human Factors in Computing Systems; 2018; Montreal p. 1-12. [doi: [10.1145/3173574.3173745](https://doi.org/10.1145/3173574.3173745)]
52. 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](https://doi.org/10.1007/s12160-013-9486-6)] [Medline: [23512568](https://pubmed.ncbi.nlm.nih.gov/23512568/)]
53. 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. [doi: [10.1016/j.pec.2011.07.018](https://doi.org/10.1016/j.pec.2011.07.018)] [Medline: [21852063](https://pubmed.ncbi.nlm.nih.gov/21852063/)]
54. De Leon E, Fuentes LW, Cohen JE. Characterizing periodic messaging interventions across health behaviors and media: systematic review. *J Med Internet Res* 2014;16(3):e93 [FREE Full text] [doi: [10.2196/jmir.2837](https://doi.org/10.2196/jmir.2837)] [Medline: [24667840](https://pubmed.ncbi.nlm.nih.gov/24667840/)]
55. Lewis CC, Boyd M, Puspitasari A, Navarro E, Howard J, Kassab H, et al. Implementing measurement-based care in behavioral health: a review. *JAMA Psychiatry* 2019 Mar 01;76(3):324-335 [FREE Full text] [doi: [10.1001/jamapsychiatry.2018.3329](https://doi.org/10.1001/jamapsychiatry.2018.3329)] [Medline: [30566197](https://pubmed.ncbi.nlm.nih.gov/30566197/)]
56. Bhattacharya A, Liang C, Zeng E, Shukla K, Wong M, Munson S, et al. Engaging teenagers in asynchronous online groups to design for stress management. 2019 Jun Presented at: 18th ACM International Conference on Interaction Design and Children; 2019; Boise p. 26-37. [doi: [10.1145/3311927.3323140](https://doi.org/10.1145/3311927.3323140)]
57. Lee J, Walker E, Bursleson W, Kay M, Buman M, Hekler E. Self-experimentation for behavior change: design and formative evaluation of two approaches. 2017 Presented at: 2017 CHI Conference on Human Factors in Computing Systems; 2017; Denver p. 6837-6849. [doi: [10.1145/3025453.3026038](https://doi.org/10.1145/3025453.3026038)]

Edited by G Eysenbach; submitted 24.08.20; peer-reviewed by S Goldstein, A Venkatesan; comments to author 06.10.20; revised version received 01.12.20; accepted 24.12.20; published 10.05.21.

Please cite as:

Graham AK, Munson SA, Reddy M, Neubert SW, Green EA, Chang A, Spring B, Mohr DC, Wildes JE

Integrating User-Centered Design and Behavioral Science to Design a Mobile Intervention for Obesity and Binge Eating: Mixed Methods Analysis

JMIR Form Res 2021;5(5):e23809

URL: <https://formative.jmir.org/2021/5/e23809>

doi:[10.2196/23809](https://doi.org/10.2196/23809)

PMID:[33970114](https://pubmed.ncbi.nlm.nih.gov/33970114/)

©Andrea K Graham, Sean A Munson, Madhu Reddy, Sarah W Neubert, Emilie A Green, Angela Chang, Bonnie Spring, David C Mohr, Jennifer E Wildes. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 10.05.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

A Breast Cancer Smartphone App to Navigate the Breast Cancer Journey: Mixed Methods Study

Serena Petrocchi¹, PhD; Chiara Filippini², MSc; Giacomo Montagna³, MD; Marta Bonollo⁴, MD; Olivia Pagani⁵, Prof Dr, MD; Francesco Meani⁶, MD, MS

¹Institute of Communication and Health, Università della Svizzera Italiana, Lugano, Switzerland

²Applied Research Division for Cognitive and Psychological Science, European Institute of Oncology IRCCS, Milan, Italy

³Breast Surgery Service, Memorial Sloan Kettering Cancer Center, New York, NY, United States

⁴Gynecology and Obstetrics Service, Centromedico, Mendrisio, Switzerland

⁵Swiss Group for Clinical Cancer Research, Geneva University Hospitals, Geneva, Switzerland

⁶Department of Obstetrics and Gynecology, Centro di Senologia della Svizzera Italiana, Ente Ospedaliero Cantonale, Viganello, Switzerland

Corresponding Author:

Francesco Meani, MD, MS

Department of Obstetrics and Gynecology

Centro di Senologia della Svizzera Italiana

Ente Ospedaliero Cantonale

Via Capelli 1

Viganello, 6962

Switzerland

Phone: 41 091 811 63 08

Fax: 41 091 811 63 93

Email: francesco.meani@eoc.ch

Abstract

Background: Several mobile apps have been designed for patients with a diagnosis of cancer. Unfortunately, despite the promising potential and impressive spread, their effectiveness often remains unclear. Most mobile apps are developed without any medical professional involvement and quality evidence-based assessment. Furthermore, they are often implemented in clinical care before any research is performed to confirm usability, appreciation, and clinical benefits for patients.

Objective: We aimed to develop a new smartphone app (Centro di Senologia della Svizzera Italiana [CSSI]) specifically designed by breast care specialists and patients together to help breast cancer patients better understand and organize their journey through the diagnosis and treatment of cancer. We describe the development of the app and present assessments to evaluate its feasibility, usefulness, and capability to improve patient empowerment.

Methods: A mixed method study with brief longitudinal quantitative data collection and subsequent qualitative semistructured interviews was designed. Twenty breast cancer patients participated in the study (mean age 51 years, SD 10 years). The usability of the app, the user experience, and empowerment were measured after 1 month. The semistructured interviews measured the utility of the app and the necessary improvements.

Results: The app received good responses from the patients in terms of positive perception of the purpose of the app (7/20, 35%), organizing the cure path and being aware of the steps in cancer management (5/20, 25%), facilitating doctor-patient communication (4/20, 20%), and having detailed information about the resources offered by the hospital (2/20, 10%). Correlation and regression analyses showed that user experience increased the level of empowerment of patients ($B=0.31$, 95% CI 0.22-0.69; $P=.009$). The interviews suggested the need to constantly keep the app updated and to synchronize it with the hospital's electronic agenda, and carefully selecting the best time to offer the tool to final users was considered crucial.

Conclusions: Despite the very small number of participants in this study, the findings demonstrate the potential of the app and support a fully powered trial to evaluate the empowering effect of the mobile health app. More data will be gathered with an improved version of the app in the second phase involving a larger study sample.

(JMIR Form Res 2021;5(5):e28668) doi:[10.2196/28668](https://doi.org/10.2196/28668)

KEYWORDS

breast cancer; decision-making process; breast cancer patient; smartphone app; empowerment; breast cancer journey; mobile app

Introduction

Background

The time of cancer diagnosis is uncertain and worrisome. Several studies on cancer patients have shown a strong desire for information and guidance, and these needs are often unmet [1-3]. A patient binder has been introduced in many breast units to provide patients with clear and easy to understand information, as well as to organize and manage appointments, treatment schedules, medical documents, and contact details. It has been proven that such a tool could help to reduce psychological distress and improve adaptation to cancer [4,5]. In the current technological era, the old concept of the patient binder has been improved and modernized. Mobile phone apps have been designed to operate as an electronic version of the patient binder. This modern way of conveying information is believed to be even more efficacious in enhancing communication between the health care institution and patients [6,7]. Electronic applications are often programmed to include the same sections of the old binders, with the advantage of unique portability and adaptability [8].

The use of mobile apps has grown exponentially in the last few years, particularly for cancer patients [9,10], as they carry the potential to provide health-pertinent information, services, and even health care interventions in a cost-effective way. Current estimates report more than 40,000 health-related applications available on the market, often freely delivered [11]. In particular, in the field of oncology, health apps have been employed to promote prevention, promote early detection, manage cancer care, and support survivorship [12,13]. Unfortunately, despite the promising potential and impressive spread, their effectiveness often remains unclear. Mobile apps are, most of the time, developed without any quality assessment procedure (evidence based) and medical professional involvement. Furthermore, they are often implemented in clinical care before any research is done to confirm usability, appreciation, and clinical benefits for patients [11].

Two recent systematic reviews on research-tested apps for breast cancer [14,15] showed promising results from studies on the inclusion of health apps in breast cancer care and called for caution before implementing these apps in clinical practice, as the final effects on users may be unpredictable [16,17] or not disease focused [15]. Therefore, they need to be extensively and scientifically tested before making them available to the public [14,15,18].

From this standpoint, the Centro di Senologia della Svizzera Italiana (CSSI) endeavored in designing and developing (proof of concept) a new smartphone app dedicated to breast cancer patients and based on direct observation of patients' needs in everyday practice. Patients actively participated in meetings with medical staff and software developers during the whole process to develop the prototype and fine tune the tool according to their needs before this pilot test run.

The aim of the app is to help patients better understand and organize their journey through the diagnosis and treatment of cancer. In other words, it aims to develop self-management skills and to empower them in the decision-making process for treatment plans. Through psychological empowerment, individuals gain control over their lives [19] and develop cognitive abilities to deal with difficult situations and behavioral tendency to take actions [20]. Psychological empowerment is one of the most important constructs for an individual's well-being. Therefore, the app was developed to give women the ability to promote active participation in their health care, improve their sense of control, and improve the relationship and communication with their attending physicians. The app was designed to allow patients to be able to quickly access reliable information related to breast cancer and the treatment process. It offers the possibility to carry important selected medical files for second opinions or for sharing with caregivers/partners, to have a private electronic agenda available for appointments and medical checks, which can also be remotely updated by the breast unit staff, and to allow the patient to write down questions and notes on an electronic scratchpad to recall at the time of the visit. The app also provides a telephone directory with useful contact details to reach the treating team and delivers news regarding activities, conferences organized by the treating institute, and other relevant information for the patients.

Access to the content of the app is made secure by encryption and by login through a two-factor identification system. Before implementation on a large scale in clinical practice, we designed a pilot test for usability and qualitative evaluation of the app prototype by a small group of breast cancer patients to build an evidence-based foundation for its use. The test run involved patients who received a breast cancer diagnosis and focused on the following three specific areas of investigation: (1) the perceived usefulness of the app, (2) the possible effect that the use of the app may have on women empowerment, and (3) the improvements or implementations proposed by the patients themselves. The goal was to show a development model for an app dedicated to breast cancer patients, but potentially extendable to other cancer patients.

App Layout and Features

The CSSI app is organized to provide convenient access to information and selected relevant documents according to the different phases of the cancer journey. The home page offers a central large "news banner" through which the breast unit can keep patients and users updated on the activities and services that are offered. Besides, several different widgets help navigate through other sections of the app. The links section provides links to reliable breast cancer websites to access several quality information sources (prevention, detection, diagnosis, treatment, survivorship, support groups, services, etc) carefully selected and periodically updated by our team. The contacts section provides suitable information to reach the CSSI health care team members and facilities. The patient's calendar helps to keep all patient appointments organized and secured in one place and

also allows patients to put reminders and keep track of what has been arranged for them. The patient's documents section allows patients to store relevant clinical reports (radiology reports, blood test reports, histology data, and descriptions of surgical interventions). For patients, it is essential to have easy access to their relevant clinical information in order to be able to share them, for example, with other doctors in case of second opinions. The note pad appears as a loose-leaf paper and is suitable for notetaking and jotting down questions patients want to ask their health care team, as well as recording the answers, and for easily finding notes when needed.

Aims

The first aim of this study was to test the app usability for the very first time and collect information about the patients' perceived usefulness of the app. The second aim concerned the consideration of the possible positive effects that the use of the app may have on patient empowerment. Finally, the third area of interest in this study was the analysis of improvements or implementations as proposed by the patients.

Methods

Procedure

Between February 2017 and August 2018, encouraged and supported by an active group of our young patients (Anna dai Capelli Corti) [21], we developed at the Breast Unit of Italian Switzerland (CSSI) a new smartphone app specifically dedicated to breast cancer patients. The development of the app was based on direct observation of our patients' needs in everyday practice. To help us design it, patients actively participated in meetings with medical staff and software developers and provided feedback on the evaluation of the prototype in order to increase its readability, ease of use, and adherence to patient needs.

Following the prototype set up, between August 2018 and February 2020, a total of 20 patients, treated at the CSSI, were invited to take part in a test run of the new information

technology (IT) tool explicitly designed for aiding breast cancer patients with their journey through cancer treatments. The project protocol was drawn up following the indications of the Declaration of Helsinki on research involving human subjects, and the Ethical Committee of the Università della Svizzera Italiana approved the research. The participants did not receive any compensation for taking part in the project.

The patients were invited by their attending physician to participate in this research project. Inclusion criteria were age over 18 years, diagnosis of nonmetastatic operable breast cancer, and ownership of a smartphone. Adherence to this pilot project was proposed shortly after surgery at the time of the first postoperative check (T0). The breast specialist was in charge of explaining the project and obtaining informed consent with a signature from the patient. The secretarial staff briefly illustrated the structure and functions of the app and helped the patients to download it and obtain the login credentials.

Patients were asked to use the CSSI app for a minimum of 1 month. Afterward, while under adjuvant treatments or follow-up, the patients were reconvened to complete a paper questionnaire and reply to a short interview (T1). The questionnaire collected sociodemographic information (age, degree of education, and marital status), information relating to the use of the app, a subjective assessment of the characteristics of the app, and judgment regarding the personal perception of the support derived from the use of the app in managing one's care path. The short semistructured interview (lasting about 5-10 minutes) aimed at investigating the usefulness of the app, as perceived by patients, and collecting suggestions for improvements to be made to the tool or its management.

Participants

Twenty patients participated in this pilot project (mean age 51 years, SD 10 years), and all of them had an education level exceeding mandatory schooling and a partner. Table 1 presents detailed information about the sample's characteristics.

Table 1. Characteristics of the participants (N=20).

Characteristic	Value, mean (SD) or n (%)
Age (years), mean (SD)	51 (10)
Education, n (%)	
Apprenticeship/high school diploma	12 (60%)
University bachelor's degree	5 (25%)
University master's degree	3 (15%)
Marital status, n (%)	
Single	3 (15%)
Married or civil partnership	13 (65%)
Separated or divorced	3 (15%)

Measurement Tools

Questionnaires

Usage Habits

A scale consisting of six items taken from the existing Mobile App Rating Scale (MARS) [22] was applied to investigate the familiarity of the participants to the use of electronic devices (in general, and smartphones in particular) in everyday life. Specifically, questions asked the frequency of smartphone use (from 1 [several times a day] to 4 [never]), the use of other apps in addition to that provided by the hospital (yes/no), the types of apps used (gaming and entertainment apps, information apps, social apps, or any other), and the frequency of use (from 1 [several times a day] to 3 [a few times during the week]).

Usability of the App

Thirteen items were selected from the MARS to investigate the patients' subjective assessment of the app in terms of its goal, interest of usage, appropriateness of the content, easiness of functioning, reliability of the information provided, design, learning process needed to use the app, trustfulness of the source, intention of using the app in the future, and overall app evaluation. The first item (ie, assessment of the app in terms of its goal) presented a multiple-choice response with six possible options (ie, increase well-being, decrease negative emotions, organize the care process, inform about the services offered by the hospital, increase awareness, and control over the care process and simplify the relationship with the doctor). All the other items had response options ranging from 1 to 5 (with labels according to the item). Items 7, 8, and 9, which all evaluated the external graphic aspect of the app, were averaged, and the final score showed acceptable reliability and a moderate internal consistency ($\alpha=.58$, $r_s>0.29$). All other items were considered individually.

User Experience

The scale is made up of the following seven items developed ad-hoc for this study: "I think I would like to use this app frequently," "I found it complicated to use," "I think I need the support of a person who is already able to use it," "I think the features of the app are well integrated," "I found inconsistencies between the various app features," "I think most people could learn to use the app easily," and "I used the app with confidence." Response options ranged from 1 (completely disagree) to 5 (completely agree) relating to the statements. A general indicator of user experience was created by averaging the individual items. The internal consistency calculated on the seven items was low ($\alpha=.57$, $r_s>0.18$), and Cronbach alpha suggested eliminating three items ("I found it complicated to

use," "I think most people could learn to use the app easily," and "I used the app with confidence"). On recalculating the internal consistency for the remaining four items, alpha reached acceptable levels ($\alpha=.64$, $r_s>0.33$).

Empowerment

The scale consists of nine items selected from the Empowerment Scale [23]. An assessment was performed on a scale from 1 (in full disagreement) to 7 (completely agree) relating to a list of statements concerning the breast cancer treatment experience after app use. The total measure of empowerment was calculated as the average of the responses to the individual items and showed good internal consistency ($\alpha=.97$, $r_s>0.81$).

Interview

The semistructured interview was focused on the following two specific areas of investigation: the perceived utility of the app and the possible areas of improvement, with very general and open questions to reduce possible bias and unreliability of the answers.

Results

Smartphone Usage Habits

Most participants declared using the app on their smartphone several times per week (18/20, 90%). Our patient population was composed of women accustomed to mobile apps in general. The majority of them (18/20, 90%) declared keeping and using other apps on their smartphone, for example, information apps (eg, broadcast), entertainment apps (eg, games), and social apps to keep in contact with others (eg, WhatsApp Messenger). They also declared that they used those apps several times in a day (16/20, 80%).

App Usability

According to the results of the questionnaire, in terms of perception of the purpose of the app, 35% (7/20) of women declared that the tool helped them to organize the cure path and to be more aware of what is the next step in management (5/20, 25%). Another 20% (4/20) declared that the app facilitated communication with the doctor, and 10% (2/20) said that they have a clear idea about the health offered by the hospital owing to the app. A small percentage (1/20, 5%) declared that the app helped to manage negative emotions.

Table 2 shows descriptive statistics regarding the other MARS items. The maximum possible value was 5. As can be seen in the table, the mean values are quite high (>3.46), except for the values of the questions regarding trustfulness and the possibility to use the app even in the future.

Table 2. Descriptive statistics of the Mobile App Rating Scale.

Item	Score		
	Range	Mean	SD
Interest of use	2-5	4.00	0.81
Appropriateness of the contents	3-4	3.61	0.50
Easiness of functioning	2-4	3.46	0.66
Accuracy of the information	4-5	4.33	0.49
Learning process	3-5	4.46	0.66
App design	3-5	3.77	0.43
App recommendation	1-5	3.75	1.20
Trustfulness of the source	1-3	2.16	1.02
Use of the app in the future	2-3	2.41	0.51
General evaluation	2-5	3.58	0.90

Relations Among Variables

Table 3 shows the correlations among the variables. Sociodemographic variables were negatively correlated with the MARS. Specifically, older patients put higher efforts in understanding how to use the app. Moreover, a higher level of education was associated with a lower positive general evaluation of the app. Women who perceived a higher interest in using the app declared a higher general positive evaluation, in addition to those who would recommend the app and who evaluated the content as appropriate and the design as high quality. Patients who evaluated the content as appropriate

revealed a higher probability of using the app in the future and higher empowerment scores. The level of patient empowerment after using the app for a month correlated with the evaluation of the appropriateness of the content, the design of the app, the recommendations found in the app, the possibility of using the app in the future, and the general evaluation of user experience.

The regression analysis findings in **Table 4** demonstrate that empowerment increased when women perceived the content of the app as appropriate, appreciated the quality of the design, and were satisfied with the general user experience (step 2). Age was also a significant and positive predictor of empowerment (step 2).

Table 3. Correlations among the variables.

Variable	Age	Educational level	Interest of use	Appropriateness of the contents	Easiness of functioning	Accuracy of the information	Learning process	App design	App recommendation	Trustfulness of the source	Use of the app in the future	General user evaluation	User experience	Empowerment
Age														
<i>r</i>	1	0.29	0.10	-0.10	-0.07	-0.03	-0.68	0.23	0.11	-0.32	0.15	-0.09	0.17	0.14
<i>P</i> value	— ^a	.37	.75	.75	.82	.93	.01	.47	.72	.31	.65	.78	.60	.65
Educational level														
<i>r</i>	0.29	1	0.09	-0.14	-0.42	0.28	-0.18	0.39	-0.34	0.14	0.00	-0.72	-0.16	-0.11
<i>P</i> value	.37	—	.77	.65	.18	.40	.58	.21	.28	.67	.99	.009	.61	.74
Interest of use														
<i>r</i>	0.10	0.09	1	0.27	-0.25	0.12	0.29	0.56	0.58	0.03	0.58	0.33	0.19	0.56
<i>P</i> value	.75	.77	—	.39	.42	.72	.37	.06	.049	.93	.034	.29	.55	.06
Appropriateness of the contents														
<i>r</i>	-0.10	-0.14	0.27	1	-0.06	-0.39	0.20	0.21	0.45	-0.24	0.60	0.29	0.16	0.62
<i>P</i> value	.75	.65	.39	—	.86	.24	.52	.51	.14	.45	.04	.37	.62	.03
Easiness of functioning														
<i>r</i>	-0.07	-0.42	-0.25	-0.06	1	-0.07	-0.14	-0.16	-0.07	-0.52	0.22	0.27	-0.27	-0.16
<i>P</i> value	.82	.18	.42	.86	—	.84	.67	.61	.82	.09	.50	.39	.39	.63
Accuracy of the information														
<i>r</i>	-0.03	0.28	0.12	-0.39	-0.07	1	-0.36	0.17	-0.07	0.46	-0.04	-0.39	0.50	0.13
<i>P</i> value	.93	.40	.72	.24	.84	—	.27	.62	.84	.15	.91	.24	.12	.71
Learning process														
<i>r</i>	-0.68	-0.18	0.29	0.20	-0.14	-0.36	1	0.24	0.06	0.22	0.08	0.46	-0.16	0.22
<i>P</i> value	.01	.58	.37	.52	.67	.27	—	.45	.85	.49	.80	.13	.62	.49
App design														
<i>r</i>	0.23	0.39	0.56	0.21	-0.16	0.17	0.24	1	0.48	0.00	0.63	0.11	0.21	0.74
<i>P</i> value	.47	.21	.06	.51	.61	.62	.45	—	.11	.99	.03	.74	.52	.006
App recommendation														
<i>r</i>	0.11	-0.34	0.58	0.45	-0.07	-0.07	0.06	0.48	1	-0.18	0.66	0.43	0.48	0.75
<i>P</i> value	.72	.28	.049	.14	.82	.84	.85	.11	—	.58	.02	.16	.12	.005
Trustfulness of the source														
<i>r</i>	-0.32	0.14	0.03	-0.24	-0.52	0.46	0.22	0.00	-0.18	1	-0.31	-0.03	0.36	0.02
<i>P</i> value	.31	.67	.93	.45	.09	.15	.49	.99	.58	—	.32	.93	.26	.94
Use of the app in the future														
<i>r</i>	0.15	0.00	0.58	0.60	0.22	-0.04	0.08	0.63	0.66	-0.31	1	0.38	0.36	0.75
<i>P</i> value	.65	.99	<.05	.04	.50	.91	.80	.03	.02	.32	—	.22	.26	.006
General user evaluation														
<i>r</i>	-0.09	-0.72	0.33	0.29	0.27	-0.39	0.46	0.11	0.43	-0.03	0.38	1	0.25	0.47
<i>P</i> value	.78	.009	.29	.37	.39	.24	.13	.74	.16	.93	.22	—	.44	.13
User experience														
<i>r</i>	0.17	-0.16	0.19	0.16	-0.27	0.50	-0.16	0.21	0.48	0.36	0.36	0.25	1	0.59

Variable	Age	Educational level	Interest of use	Appropriateness of the contents	Easiness of functioning	Accuracy of the information	Learning process	App design	App recommendation	Trustfulness of the source	Use of the app in the future	General user evaluation	User experience	Empowerment
<i>P</i> value	.60	.61	.55	.62	.39	.12	.62	.52	.12	.26	.26	.44	—	.043
Empowerment														
<i>r</i>	0.14	−0.11	0.56	0.62	−0.16	0.13	0.22	0.74	0.75	0.02	0.75	0.47	0.59	1
<i>P</i> value	.65	.74	.06	.03	.63	.71	.49	.006	.005	.94	.006	.13	.043	—

^aNot applicable.

Table 4. Regression analysis.

Variable	Empowerment	
	B (95% CI) or %	<i>P</i> value
Step 1: Sociodemographic		
Civil status (0=no partner)	0.09 (−1.37 to 1.73)	.79
Age	0.34 (−0.04 to 0.12)	.33
Education (0=low)	−0.12 (−1.35 to 1.01)	.74
R ² (%)	13%	N/A ^a
Step 2: Experience		
Civil status (0=no partner)	−0.16 (−0.64 to 0.03)	.06
Age	0.13 (0.001 to 0.03)	.04
Education (0=low)	−0.15 (−0.55 to 0.12)	.13
Appropriateness of the contents	0.44 (0.63 to 1.39)	.003
App design	0.61 (0.81 to 2.34)	.007
App recommendation	0.18 (−0.09 to 0.43)	.12
Use of the app in the future	−0.02 (−0.56 to 0.44)	.71
User experience	0.31 (0.22 to 0.69)	.009
R ² (%)	79%	N/A

^aN/A: not applicable.

Perceived Usefulness and Suggested Improvements

At the end of the test period and after filling up the questionnaire, the patients were asked to give their opinions on the app through a semistructured interview.

The first area investigated the perceived usefulness and the reasons for their perception. Overall, 55% (11/20) of the patients declared that the app was in general useful. Some examples of the patient responses are as follows:

Yes, very useful for having all the appointments in the same application and for accessing information links. [age 55 years]

Yes, to keep track of appointments. [age 58 years]

It was particularly useful for finding reliable and verified information. [age 54 years]

The other patients (9/20, 45%) declared that the app was partially useful. Some examples of responses are as follows:

Partially useful, I find the part of the contact details very useful especially when you are worried, but I find the part relating to the reports lacking because they are loaded only at the request of the patient. [age 60 years]

There are aspects that can help, but at the same time some limits: I would increase communication between the various departments/health workers involved in the App management. [age 55 years]

No, since the agenda was not updated, I preferred to use my own (agenda). [age 55 years]

When focusing on communication, the majority of patients (15/20, 75%) did not believe that the app had influenced their relationship with the referring physician. Only a minority (4/20, 20%) of patients reported that using the app helped them to create a more direct channel of communication with the doctor, while the remaining 5% (1/20) of patients replied that they did not have a clear opinion on this aspect.

As for cancer management, the majority of patients (13/20, 65%) declared not considering the app useful as a self-management tool. Some of the reasons reported by the patients to explain this point were that the app reminded them of being sick or that they felt capable of managing cancer and the related treatment on their own. Interestingly, however, more than half of the patients (11/20, 55%) affirmed that the app increased the knowledge they had about their clinical condition and thus the awareness of their own path of clinical care.

The second area of investigation focused on the proposed improvements. All patients provided several suggestions on different aspects.

First, regarding treatment modalities, patients suggested increasing the number of links to reliable websites, with a focus on surgical treatments, pre- and postsurgical images, and news on newly developed technologies. Information on remedies to manage the side effects of cancer treatments (eg, use of laser during menopause) was frequently requested. Second, patients suggested features for self-management, with inclusion of an alarm as a reminder for medical appointments or a notification when the hospital cancels an appointment. Third, patients suggested including information regarding activities and/or resources (conferences within the hospital or other health institutions, or the existence of patient advocacy groups).

A special area of improvement focused on the doctor-patient relationship with the inclusion of a section dedicated to bidirectional communication between health care providers and patients. Interestingly, another area of improvement concerned the inclusion of a section dedicated to the costs of medical treatment. Finally, 12 patients spontaneously suggested delivering the app at a different time, preferentially before treatment, to increase its usefulness as a source of reliable information.

Discussion

General Opinions on the App's Usefulness

This work describes a structured process involving end-users to test the efficacy and perceived usefulness of a new mobile app dedicated to breast cancer patients. We also investigated possible areas of improvement for further development as proposed by the patients themselves. The aim of the app was to help patients better understand and organize their journey through the diagnosis and treatment of cancer, that is, to develop their self-management skills, empowerment, and sense of control. Overall, our patients perceived the app as easy to learn and use, accurate, and appropriate. They stated that they would use the app in the future and rated it almost 4 out of 5 points. Moreover, the app increased patient empowerment.

Association of the App's Perceived Usefulness, Sociodemographic Characteristics, and Empowerment

We found significant relationships between variables. As expected, older patients put higher efforts in understanding how to use the app compared with younger patients. This is an expected result given the use of technologies is more common among young people. Future development of the app should

take into account the age influence, mainly because breast cancer risk increases with age, peaking above 50 years [24,25].

The educational level was negatively correlated with general users' evaluation, that is, a higher level of education was associated with a lower positive general evaluation of the app. These results might be linked to the fact that the educational level is associated with literacy and health literacy [26,27]. It may be that participants with high educational levels have high literacy and therefore high knowledge and competence with the use of new technologies. As expected, there were correlations between the items of the MARS, indicating a general and overall consistency of the app rating. Women who perceived a higher interest in using the app declared a higher general positive evaluation, in addition to those who recommended the app and who evaluated the content as appropriate and the design as high quality.

However, the most interesting result was regarding the level of patient empowerment, which increased when women perceived the content of the app as appropriate, perceived the high quality of the design, and were satisfied with the general user experience of the app. This result represents the first demonstration of the perceived usefulness of the app by breast cancer patients, which was a fundamental aim of the research. Therefore, apart from a generally positive evaluation of the app, it seems that its use enhanced the sense of control over cancer and the general empowerment of women owing to the potential of technologies to switch from a paternalistic to a collaborative relationship between patients and physicians [28,29].

Suggested Improvements

From the interviews, it was found that our app, as with other mobile health (mHealth) tools, will require constant revision and updates. Nowadays, the information flow on the internet and social media is constantly in dynamic change. Future development of the app should take into account that a static app functioning as a sort of agenda without being connected to the hospital agenda, for example, or without timely updates is not useful.

There appeared to be inconsistency in the results concerning the perceived usefulness of the app. While 65% of patients declared a certain level of disappointment with the app as a self-management tool, more than half (55%) of the patients declared that it helped to increase knowledge about their clinical condition and awareness of their own path of clinical care. This might be a demonstration of the discrepancy between the effect delivered by the use of the tool and the awareness of real advantages that are not fully perceived as such. In this sense, our preliminary results seem to suggest that use of the app holds the potential to improve the sense of control over cancer and the general empowerment of women [28], which are both related to positive health outcomes [30-32].

Another very important aspect arising from the interviews is that the vast majority of patients perceived receiving the app after surgery as utterly untimely. They stressed that the app was not available as a source of reliable information when they needed it the most (before treatment). Such feedback is practice changing and will be taken into account in the next validation

phase and finally in routine clinical practice. This is, in our opinion, a clear example of the utility of patient involvement in the development and implementation process of mHealth tools.

Limitations

This study has several limitations. First, the sample showed selection bias, and its size was quite small. Our sample included a small number of breast cancer patients selected via word of mouth. These patients used the smartphone quite frequently and were accustomed to using other apps. In this sense, the sample involved women with an adequate level of literacy regarding the use of electronic tools such as the smartphone and apps. Women with a lower literacy level were underrepresented. Future research should test the app with older patients and with a larger sample size. It might also be that women who agreed to participate were the most involved in their care path, and this may represent a selection bias. Second, this study did not consider a control group of women with no access to the app for comparison. Future research should include two groups of women and measure the baseline empowerment at the time of enrollment in the research. Third, in this pilot study, the semistructured interview was conducted by the treating physician. This might have consciously or unconsciously altered the way answers were given or interpreted (observation and interviewer bias). Finally, the sample considered here included patients with breast cancer. Future research should test the app with heterogeneous groups of patients having different diagnoses and types of treatments.

Conclusions

The values of this work and the CSSI app lie in the involvement of health and communication professionals in the design and implementation processes. Of equal value is the assessment of

the quality and usability of the contents performed through the involvement of patients in a feedback process guided by scientifically validated questionnaires. In this way, our mHealth tool differs from the vast majority of other health apps on the market, which are often produced without the involvement of health care professionals and patients, devoid of any scientific basis, and not subjected to any quality assessment. We learned that the IT tool has maximum utility and obtains maximum consent when managed correctly by staff and when its features and use are clearly explained to final users. It requires constant application by managers to guarantee effective functioning and the continuous updating of content. Patient feedback also underlined the importance of the timing and delivery methods of the app. It takes time to explain how it works, even if the tool is simple. Our preliminary data seem to suggest that the best time to offer it is before surgery, but not when communicating the diagnosis. This aspect will be further investigated in the second phase of our work. Among the various features, unlike what was expected at the time of the app design, those that attracted the most attention of the patients were the electronic appointment calendar and the storage area for clinical reports.

Findings from this pilot study demonstrate the potential of the app and its validation protocol, and support a fully powered trial to evaluate the empowering effect of the mHealth app. More data will be gathered with an improved version of the app in a second phase involving a larger study population. The next step will be to extend the use to a greater number of patients and follow our patients' suggestions. We will ensure a more proactive attitude by the team responsible for the management of the IT tool and the interaction with users. We will implement changes to the software suggested through feedback, and at that point, a new assessment will be performed for quality, appreciation, and usefulness in terms of patient empowerment of the sense of control and self-management.

Acknowledgments

We specially thank the association of young patients "Anna dai Capelli Corti" (www.annadaicapellicorti.ch) for having made a fundamental contribution to this work. This very active group of young patients, who experienced the turmoil of a cancer diagnosis, the accompanying need for guidance, and the value of effective communication with health care professionals, strongly believed in this project. They not only financed it entirely, but also participated in the entire process, from initial design to the first test run. This research did not receive any specific grant from funding agencies in the public or commercial sectors. The not-for-profit patient association "Anna dai Capelli Corti" funded the app design and development of the software, with a private donation of 30,000 euros.

Authors' Contributions

SP: conceptualization, writing the original draft, data curation, review, and editing; CF: writing the original draft and data curation; GM: conceptualization; MB: review and editing; OP: conceptualization and funding acquisition; FM: conceptualization, funding acquisition, writing the original draft, data curation, review, and editing. All authors approved the final version and are accountable for all aspects of the work.

Conflicts of Interest

None declared.

References

1. The Canadian Cancer Society. Final report on the needs of people living with cancer across Canada. Toronto, Canada: The Canadian Cancer Society; 1992.

2. The Canadian Cancer Society. Breaking down the barriers: study of cancer patient and caregiver needs in Ontario. Toronto, Canada: The Canadian Cancer Society; 2003.
3. Chelf JH, Agre P, Axelrod A, Cheney L, Cole DD, Conrad K, et al. Cancer-related patient education: an overview of the last decade of evaluation and research. *Oncol Nurs Forum* 2001 Aug;28(7):1139-1147. [Medline: [11517847](#)]
4. Doak LG, Doak CC, Meade CD. Strategies to improve cancer education materials. *Oncol Nurs Forum* 1996 Sep;23(8):1305-1312. [Medline: [8883075](#)]
5. Gauthier-Frohlick D, Boyko S, Conlon M, Damore-Petingola S, Lightfoot N, Mackenzie T, et al. Evaluation of cancer patient education and services. *J Cancer Educ* 2010 Mar;25(1):43-48. [doi: [10.1007/s13187-009-0008-y](#)] [Medline: [20094826](#)]
6. Williams SA, Schreier AM. The effect of education in managing side effects in women receiving chemotherapy for treatment of breast cancer. *Oncol Nurs Forum* 2004 Feb;31(1):E16-E23. [doi: [10.1188/04.ONF.E16-E23](#)] [Medline: [14722602](#)]
7. Attai DJ, Cowher MS, Al-Hamadani M, Schoger JM, Staley AC, Landercasper J. Twitter Social Media is an Effective Tool for Breast Cancer Patient Education and Support: Patient-Reported Outcomes by Survey. *J Med Internet Res* 2015 Jul 30;17(7):e188 [FREE Full text] [doi: [10.2196/jmir.4721](#)] [Medline: [26228234](#)]
8. Young-Afat DA, van Gils CH, Bruinvels DJ, van der Pol CC, Witkamp AJ, Sijtsma S, et al. Patients' and Health Care Providers' Opinions on a Supportive Health App During Breast Cancer Treatment: A Qualitative Evaluation. *JMIR Cancer* 2016 Jun 07;2(1):e8 [FREE Full text] [doi: [10.2196/cancer.5334](#)] [Medline: [28410170](#)]
9. Whittaker R. Issues in mHealth: findings from key informant interviews. *J Med Internet Res* 2012 Oct 02;14(5):e129 [FREE Full text] [doi: [10.2196/jmir.1989](#)] [Medline: [23032424](#)]
10. Aitken M, Gauntlett C. Patient apps for improved healthcare: from novelty to mainstream. Parsippany, NJ: IMS Institute for Healthcare Informatics; 2013.
11. Mobasheri MH, Johnston M, King D, Leff D, Thiruchelvam P, Darzi A. Smartphone breast applications - what's the evidence? *Breast* 2014 Oct;23(5):683-689 [FREE Full text] [doi: [10.1016/j.breast.2014.07.006](#)] [Medline: [25153432](#)]
12. Davis SW, Oakley-Girvan I. mHealth Education Applications Along the Cancer Continuum. *J Cancer Educ* 2015 Jun;30(2):388-394. [doi: [10.1007/s13187-014-0761-4](#)] [Medline: [25482319](#)]
13. Bender JL, Yue RYK, To MJ, Deacken L, Jadad AR. A lot of action, but not in the right direction: systematic review and content analysis of smartphone applications for the prevention, detection, and management of cancer. *J Med Internet Res* 2013 Dec 23;15(12):e287 [FREE Full text] [doi: [10.2196/jmir.2661](#)] [Medline: [24366061](#)]
14. Jongerius C, Russo S, Mazzocco K, Pravettoni G. Research-Tested Mobile Apps for Breast Cancer Care: Systematic Review. *JMIR Mhealth Uhealth* 2019 Feb 11;7(2):e10930 [FREE Full text] [doi: [10.2196/10930](#)] [Medline: [30741644](#)]
15. Triberti S, Savioni L, Sebrì V, Pravettoni G. eHealth for improving quality of life in breast cancer patients: A systematic review. *Cancer Treat Rev* 2019 Mar;74:1-14. [doi: [10.1016/j.ctrv.2019.01.003](#)] [Medline: [30658289](#)]
16. Foley NM, O'Connell EP, Lehane EA, Livingstone V, Maher B, Kaimkhani S, et al. PATI: Patient accessed tailored information: A pilot study to evaluate the effect on preoperative breast cancer patients of information delivered via a mobile application. *Breast* 2016 Dec;30:54-58. [doi: [10.1016/j.breast.2016.08.012](#)] [Medline: [27611236](#)]
17. Heo J, Chun M, Lee KY, Oh Y, Noh OK, Park RW. Effects of a smartphone application on breast self-examination: a feasibility study. *Healthc Inform Res* 2013 Dec;19(4):250-260 [FREE Full text] [doi: [10.4258/hir.2013.19.4.250](#)] [Medline: [24523989](#)]
18. Collado-Borrell R, Escudero-Vilaplana V, Ribed-Sánchez A, Ibáñez-García S, Herranz-Alonso A, Sanjurjo-Sáez M. Smartphone applications for cancer patients; what we know about them? *Fam Hosp* 2016 Jan 01;40(1):25-35 [FREE Full text] [doi: [10.7399/fh.2016.40.1.8993](#)] [Medline: [26882831](#)]
19. Rappaport J. In praise of paradox: a social policy of empowerment over prevention. *Am J Community Psychol* 1981 Feb;9(1):1-25. [doi: [10.1007/BF00896357](#)] [Medline: [7223726](#)]
20. Zimmerman MA. Taking aim on empowerment research: On the distinction between individual and psychological conceptions. *American Journal of Community Psychology* 1990;18(1):169-177. [doi: [10.1007/bf00922695](#)]
21. Anna dai Capelli Corti. URL: <http://www.annadaicapellicorti.ch/> [accessed 2021-05-04]
22. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. *JMIR Mhealth Uhealth* 2015 Mar 11;3(1):e27 [FREE Full text] [doi: [10.2196/mhealth.3422](#)] [Medline: [25760773](#)]
23. Spreitzer GM. An empirical test of a comprehensive model of intrapersonal empowerment in the workplace. *Am J Community Psychol* 1995 Oct;23(5):601-629. [doi: [10.1007/BF02506984](#)] [Medline: [8851342](#)]
24. Petrocchi S, Ludolph R, Labrie NHM, Schulz P. Application of the theory of regulatory fit to promote adherence to evidence-based breast cancer screening recommendations: experimental versus longitudinal evidence. *BMJ Open* 2020 Nov 12;10(11):e037748 [FREE Full text] [doi: [10.1136/bmjopen-2020-037748](#)] [Medline: [33184078](#)]
25. Moser K, Patnick J, Beral V. Do women know that the risk of breast cancer increases with age? *Br J Gen Pract* 2007 May;57(538):404-406 [FREE Full text] [Medline: [17504593](#)]
26. van der Heide I, Wang J, Droomers M, Spreeuwenberg P, Rademakers J, Uiters E. The relationship between health, education, and health literacy: results from the Dutch Adult Literacy and Life Skills Survey. *J Health Commun* 2013;18 Suppl 1:172-184 [FREE Full text] [doi: [10.1080/10810730.2013.825668](#)] [Medline: [24093354](#)]

27. Schulz PJ, Pessina A, Hartung U, Petrocchi S. Effects of Objective and Subjective Health Literacy on Patients' Accurate Judgment of Health Information and Decision-Making Ability: Survey Study. *J Med Internet Res* 2021 Jan 21;23(1):e20457 [FREE Full text] [doi: [10.2196/20457](https://doi.org/10.2196/20457)] [Medline: [33475519](https://pubmed.ncbi.nlm.nih.gov/33475519/)]
28. Pravettoni G, Triberti S. *P5 eHealth: An Agenda for the Health Technologies of the Future*. New York, NY: Springer International Publishing; 2020.
29. Petrocchi S, Iannello P, Lecciso F, Levante A, Antonietti A, Schulz PJ. Interpersonal trust in doctor-patient relation: Evidence from dyadic analysis and association with quality of dyadic communication. *Soc Sci Med* 2019 Aug;235:112391. [doi: [10.1016/j.socscimed.2019.112391](https://doi.org/10.1016/j.socscimed.2019.112391)] [Medline: [31301438](https://pubmed.ncbi.nlm.nih.gov/31301438/)]
30. Laine C, Davidoff F. Patient-centered medicine. A professional evolution. *JAMA* 1996 Jan 10;275(2):152-156. [Medline: [8531314](https://pubmed.ncbi.nlm.nih.gov/8531314/)]
31. Báñez M, Blasco T, Fernández-Castro J, Viladrich C. A structural model of the relationships between perceived control and adaptation to illness in women with breast cancer. *J Psychosoc Oncol* 2007;25(1):21-43. [doi: [10.1300/J077v25n01_02](https://doi.org/10.1300/J077v25n01_02)] [Medline: [17360314](https://pubmed.ncbi.nlm.nih.gov/17360314/)]
32. Watson M, Homewood J, Haviland J, Bliss JM. Influence of psychological response on breast cancer survival: 10-year follow-up of a population-based cohort. *Eur J Cancer* 2005 Aug;41(12):1710-1714. [doi: [10.1016/j.ejca.2005.01.012](https://doi.org/10.1016/j.ejca.2005.01.012)] [Medline: [16098457](https://pubmed.ncbi.nlm.nih.gov/16098457/)]

Abbreviations

CSSI: Centro di Senologia della Svizzera Italiana

IT: information technology

MARS: Mobile App Rating Scale

mHealth: mobile health

Edited by G Eysenbach; submitted 10.03.21; peer-reviewed by T Spanic; comments to author 31.03.21; revised version received 12.04.21; accepted 12.04.21; published 10.05.21.

Please cite as:

Petrocchi S, Filippini C, Montagna G, Bonollo M, Pagani O, Meani F

A Breast Cancer Smartphone App to Navigate the Breast Cancer Journey: Mixed Methods Study

JMIR Form Res 2021;5(5):e28668

URL: <https://formative.jmir.org/2021/5/e28668>

doi: [10.2196/28668](https://doi.org/10.2196/28668)

PMID: [33970120](https://pubmed.ncbi.nlm.nih.gov/33970120/)

©Serena Petrocchi, Chiara Filippini, Giacomo Montagna, Marta Bonollo, Olivia Pagani, Francesco Meani. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 10.05.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Video Consultation as an Adequate Alternative to Face-to-Face Consultation in Continuous Positive Airway Pressure Use for Newly Diagnosed Patients With Obstructive Sleep Apnea: Randomized Controlled Trial

Laura Kooij^{1,2}, MSc; Petra JE Vos³, MD, PhD; Antoon Dijkstra³, MD, PhD; Elisabeth A Roovers¹, PhD; Wim H van Harten^{1,2,4}, MD, PhD

¹Rijnstate, Arnhem, Netherlands

²Department of Health Technology and Services Research, University of Twente, Enschede, Netherlands

³Pulmonary Department, Rijnstate, Arnhem, Netherlands

⁴Division of Psychosocial Research and Epidemiology, Netherlands Cancer Institute, Amsterdam, Netherlands

Corresponding Author:

Wim H van Harten, MD, PhD

Rijnstate

Wagnerlaan 55

Arnhem

Netherlands

Phone: 31 088 0058888

Email: WvanHarten@Rijnstate.nl

Abstract

Background: The effectiveness of continuous positive airway pressure (CPAP) is dependent on the degree of use, so adherence is essential. Cognitive components (eg, self-efficacy) and support during treatment have been found to be important in CPAP use. Video consultation may be useful to support patients during treatment. So far, video consultation has rarely been evaluated in thorough controlled research, with only a limited number of outcomes assessed.

Objective: The aim of the study was to evaluate the superiority of video consultation over face-to-face consultation for patients with obstructive sleep apnea (OSA) on CPAP use (minutes per night), adherence, self-efficacy, risk outcomes, outcome expectancies, expectations and experiences with video consultation, and satisfaction of patients and nurses.

Methods: A randomized controlled trial was conducted with an intervention (video consultation) and a usual care group (face-to-face consultation). Patients with confirmed OSA (apnea-hypopnea index >15), requiring CPAP treatment, no history of CPAP treatment, having access to a tablet or smartphone, and proficient in the Dutch language were recruited from a large teaching hospital. CPAP use was monitored remotely, with short-term (weeks 1 to 4) and long-term (week 4, week 12, and week 24) assessments. Questionnaires were completed at baseline and after 4 weeks on self-efficacy, risk perception, outcome expectancies (Self-Efficacy Measure for Sleep Apnea), expectations and experiences with video consultation (covering constructs of the unified theory of acceptance and use of technology), and satisfaction. Nurse satisfaction was evaluated using questionnaires.

Results: A total of 140 patients were randomized (1:1 allocation). The use of video consultation for OSA patients does not lead to superior results on CPAP use and adherence compared with face-to-face consultation. A significant difference in change over time was found between groups for short-term (P -interaction=.008) but not long-term (P -interaction=.68) CPAP use. CPAP use decreased in the long term (P =.008), but no significant difference was found between groups (P =.09). Change over time for adherence was not significantly different in the short term (P -interaction=.17) or long term (P -interaction=.51). A relation was found between CPAP use and self-efficacy (P =.001), regardless of the intervention arm (P =.25). No significant difference between groups was found for outcome expectancies (P =.64), self-efficacy (P =.41), and risk perception (P =.30). The experiences were positive, and 95% (60/63) intended to keep using video consultation. Patients in both groups rated the consultations on average with an 8.4. Overall, nurses (n =3) were satisfied with the video consultation system.

Conclusions: Support of OSA patients with video consultation does not lead to superior results on CPAP use and adherence compared with face-to-face consultation. The findings of this research suggest that self-efficacy is an important factor in improving

CPAP use and that video consultation may be a feasible way to support patients starting CPAP. Future research should focus on blended care approaches in which self-efficacy receives greater emphasis.

Trial Registration: Clinicaltrials.gov NCT04563169; <https://clinicaltrials.gov/show/NCT04563169>

(*JMIR Form Res* 2021;5(5):e20779) doi:[10.2196/20779](https://doi.org/10.2196/20779)

KEYWORDS

video consultation; eHealth; obstructive sleep apnea; continuous positive airway pressure; randomized controlled trial

Introduction

Telemedicine is increasingly used to support self-management in chronic diseases and is defined as the use of information and communication technology to deliver health care at a distance [1], but so far we see little evidence in this field. Nevertheless, telemedicine solutions are used for patients with obstructive sleep apnea (OSA) for example, for monitoring, education, and consultation [2]. OSA is considered a chronic disease [1,3]; it is a sleep disorder that affects at least 2% to 4% of the adult population [4] and is characterized by repeated episodes of full or partial occlusion of the upper airway during sleep [4,5]. This condition can have multiple effects on patients' health such as cognitive dysfunction [4], decrease in health-related quality of life [4,6], increase in cardiovascular disease risk, and sleepiness during the daytime [6]. The severity is often determined with the apnea-hypopnea index (AHI) [4], which represents the number of apneas and hypopneas per hour [4] and is classified as mild (5 to 15 per hour), moderate (15 to 30 per hour) or severe (>30 per hour) [7]. Continuous positive airway pressure (CPAP) is the preferred treatment [6], especially for moderate to severe OSA [5]. CPAP prevents the airway from narrowing or collapsing by applying a positive pressure via a nasal mask during sleep [8] and is tailored to each patient [9]. As the effectiveness of CPAP is dependent on use [5,10], treatment adherence is essential. Cognitive components, mainly based on the social cognitive theory [11], are becoming increasingly important in predicting CPAP use [12-14]. Support during treatment [15], tailored interventions [16], and closer follow-up [17] can also positively affect adherence.

Video consultation may be a useful way to support patients [1,17,18] during treatment and is defined as a "technology used to realize a real-time visual and audio patient assessment at a distance" [19]. Video consultation has been beneficial in chronic conditions (eg, diabetes [20,21] and cancer [19,22]) and in care for OSA patients [17,18]. The use for OSA patients may be promising, especially since physical examination is not always needed [1], and CPAP use can already be monitored remotely [23]. However, the evidence on the effectiveness for OSA patients is still limited [24]. Previous studies were narrowly focused, with mainly adherence [18,25] and satisfaction [17,18,26] being assessed. Although cognitive components, (eg, self-efficacy and outcome expectancies) are found to be important elements for CPAP use [13,14,27], there is a lack of evidence about these effects on video consultation for OSA patients. Previous research on OSA patients also mainly evaluated the use of video consultation for initial contact with health care professionals focused on diagnosis, treatment plans [18,26], or for training purposes [17]. The use of video

consultation may be particularly relevant during follow-up (after an initial face-to-face contact) for newly diagnosed patients, since support during treatment is important [15] and successful CPAP use is often determined at an early stage of treatment [28].

Only a limited number of randomized controlled trials (RCTs) were conducted [17,25,26,29], with only one fully powered trial [29]. In a study by Smith et al [25], video consultation was used by nurses for patients who were nonadherent during the first 3 months of treatment. One group of patients received specific information (n=10) about CPAP and one group (n=9) generic information. Both adherence and satisfaction were higher in the intervention group ($P=.003$). Isetta et al [29] conducted a multicenter RCT with patients receiving access to either a telemedicine program (n=69) with video consultations or usual care (hospital visits, n=70). Although the telemedicine approach was assumed to be more cost-effective, CPAP adherence was equivalent after 6 months [29]. Video consultation was also used for initial contact before starting treatment, with mixed results. The use of video consultation for training purposes did not lead to a difference in knowledge [17]. Also, no significant differences in satisfaction and CPAP adherence were found after 14 days for new OSA patients starting CPAP treatment [18]. Adherence rates were found to be higher after 6 months for patients who received their initial consultation face-to-face than via video consultation. However, statistically significant difference was not reported [26].

Video consultation is often found to be as effective as face-to-face consultation in terms of CPAP use [18,29]. Previous studies often focused on newly diagnosed patients before the start of treatment [17,18,26], with generally small sample sizes [17,25,26]. Patients are satisfied with video consultation [17,18,25], and it may be a promising way to deliver more convenient care with indirect benefits for patients (eg, less travel time) [24]. Additionally, remote monitoring [30] and patient support treatment [31] can positively affect CPAP use [30,31]. Therefore, it may be expected that video consultation in combination with remotely monitoring CPAP use, consultation with nurses, and the indirect benefits of video consultation (eg, less travel time) [24] may improve CPAP use. Cognitive components (eg, self-efficacy) are also found to be important elements for CPAP use [13,14,27], but evaluation in combination with video consultation is lacking [24]. More evidence about the technology being used and health care professionals' perceptions is also needed to ensure successful implementations [17]. Such knowledge is essential because the use of video consultation is increasing, but evidence is still lacking and powered studies are needed [24].

Therefore, the objective of this paper is to evaluate the superiority of video consultation versus face-to-face consultation for patients with OSA on CPAP use (minutes per night), CPAP adherence, self-efficacy, risk perception, outcome expectancy, video consultation expectations and experiences with technology, and the satisfaction of patients and nurses.

Methods

Study Design

We conducted a nonblinded RCT with an intervention group (video consultation) and a usual care group (face-to-face consultation), with 1:1 allocation.

Recruitment and Participants

Patients were recruited from a large teaching hospital (Rijnstate, Arnhem). To be eligible to participate, patients had to be older than 18 years, be diagnosed with moderate or severe OSA (AHI >15), require CPAP treatment, have no history of CPAP treatment, have access to a tablet or smartphone, and be proficient in the Dutch language. Exclusion criteria were having a psychiatric or cognitive disorder.

Study Process

Prior to the study, a letter was sent to patients to confirm their appointments (eg, sleep study and consultation with the pulmonologist) including information about the study. During the first face-to-face consultation with the pulmonologist, patients received their treatment plan and information about the study (including information letter and informed consent form). This was followed by instruction about their CPAP treatment. After this consultation, the researcher provided patients with additional information about the study, and they were asked to sign the informed consent form. For reasons of clinical necessity, patients started treatment the same day.

Randomization

After patients signed informed consent and completed the baseline questionnaire, they were randomized by the researcher to the intervention or usual care group using the software program Research Manager (Cloud9 Software) with block size of 10. The researcher informed the patients about their allocation, and the intervention group received additional information about the video consultation app (Facetalk, Qconferencing) [32]. All participants received a copy of the informed consent form, and a follow-up appointment was planned directly.

Intervention

The video consultation app Facetalk [32] could be downloaded (for free) from Google Play [33] or the App Store [34]. The first video consultation with a nurse was planned for 1 week after the start of CPAP. Patients received an email with the date, time, and a link to start the video consultation in the app. Three focus points were discussed during the consultations: (1) adherence (>6 hours per night), (2) rest AHI <5 (or <10 if age over 70 years), and (3) (improvements in) symptoms. If these objectives were achieved after 1 week, a new consultation was planned for 3 weeks later (4 weeks after the start). If these

objectives were not achieved, video consultations were planned for weekly (until 4 weeks after starting CPAP treatment). After 4 weeks, patients received a questionnaire. See [Multimedia Appendix 1](#) for the study process.

Usual Care

The usual care group followed the same care process but with face-to-face consultation instead of video consultation. Patients received a confirmation letter with the day and time of their next consultation.

Outcome Measures

Primary Outcome

The primary outcome was CPAP use (minutes per night), monitored remotely with Encore Anywhere (Philips). Conforming to the initial protocol, CPAP use was assessed during the first 4 weeks (short-term). Additionally, we assessed CPAP use after week 4, week 12, and week 24 (long-term).

Secondary Outcomes

CPAP Adherence

CPAP adherence was defined as CPAP use for at least 5 nights per week for at least 4 hours per night [15,35] and was assessed during the first 4 weeks (short-term) and week 4, week 12, and week 24 (long-term).

Treatment Self-Efficacy, Risk Perception, and Outcome Expectancies

The Self-Efficacy Measure for Sleep Apnea (SEMSA) [13] was used to measure cognitive components: self-efficacy, risk perception, and outcome expectancies. The SEMSA is a 26-item scale [13] with subscales: self-efficacy and outcome expectancies each have 9 questions rated on a 4-point scale from not at all true to very true and risk perception has 8 questions rated on a 4-point scale from very low to very high. The mean of the nonmissing item responses was calculated for risk perception, outcome expectancies, and self-efficacy. For the purpose of this study, the SEMSA was translated back (from English into Dutch) and forth (from Dutch into English) by Taalcentrum-VU [36]. In this study, the statements from the published paper were used [13].

Relation Between Self-Efficacy, Risk Perception, Outcome Expectancies, and CPAP Use

The relations between CPAP use and self-efficacy, risk perception, and outcome expectancies were assessed. Also, the differences between the intervention and usual care group were analyzed.

Expectations and Experiences With Video Consultation

Questions covering constructs of the unified theory of acceptance and use of technology (UTAUT) model [37] were used to measure expectations and experiences with the use of the video consultation system. The UTAUT consists of 4 constructs that influence behavioral intention and behavior—performance expectancy, effort expectancy, social influence, and facilitating conditions [37]. A total of 9 questions were rated on a 7-point scale (1=totally disagree to 7=totally agree).

Satisfaction

Patient satisfaction was evaluated with questions about the consultations and information received. Additionally, the intervention group answered questions about the video consultation system. All questions were rated on a 5-point scale (from 1=totally disagree to 5=totally agree). Nurses' experiences were evaluated using a questionnaire with questions about the video consultation system, satisfaction, and organizational benefits (eg, time and efficiency).

Other Parameters

Patient age, marital status, education, experience with internet and internet use, tablet or smartphone skills, and support (with tablet or smartphone use) were assessed via a questionnaire at baseline. Data about comorbidities, AHI, number of consultations, symptoms, and results of the Epworth Sleepiness Scale [38] were obtained from the electronic medical record. This scale is a self-administered questionnaire to examine the perception of daytime sleepiness that has 8 questions about how likely it is to doze off in different situations ranging from 0 to 3. A total score for this scale is calculated by taking the sum of the 8 items. A total of 11 to 12 is considered mild, 13 to 15 moderate, and 16 to 24 severe excessive daytime sleepiness [39]. In this study, a total score of >10 is considered excessive daytime sleepiness.

Sample Size Calculation

Since there is no determined clinically relevant difference for CPAP use [40], we assumed that a difference of 1 (SD 2.0) hour per day of average CPAP use (primary outcome) is clinically significant [13,29]. Using a *t* test, alpha of .05, and 80% power, 63 subjects per group (a total of 126) were needed. Correcting for 10% dropout, 70 patients were recruited for each group.

Statistical Analysis

Data analysis was performed using SPSS (version 22.0, IBM Corp). Descriptive statistics were used to report the baseline characteristics, experiences, expectations, and satisfaction. Linear mixed models were used to analyze differences in CPAP use over time for the intervention and usual care group (interaction term: time \times group). All available CPAP use data were used in the analysis, according to the intention-to-treat principle. Differences in adherence over time between groups was analyzed using generalized estimating equations. The

relation between CPAP use and risk perception, outcome expectancies, and self-efficacy was analyzed with a linear regression. Normally distributed variables were reported as mean and standard deviation, and statistical differences were tested using an independent samples *t* test. Nonnormally distributed data were reported with medians and interquartile range (25th to 75th percentiles), and differences between groups were analyzed with Mann-Whitney *U* tests.

Approval and Ethical Considerations

All participants signed a written informed consent form prior to inclusion in the study. The study was approved by the regional medical research ethics committee Commissie Mensgebonden Onderzoek Arnhem–Nijmegen and registered at Clinicaltrials.gov [NCT04563169].

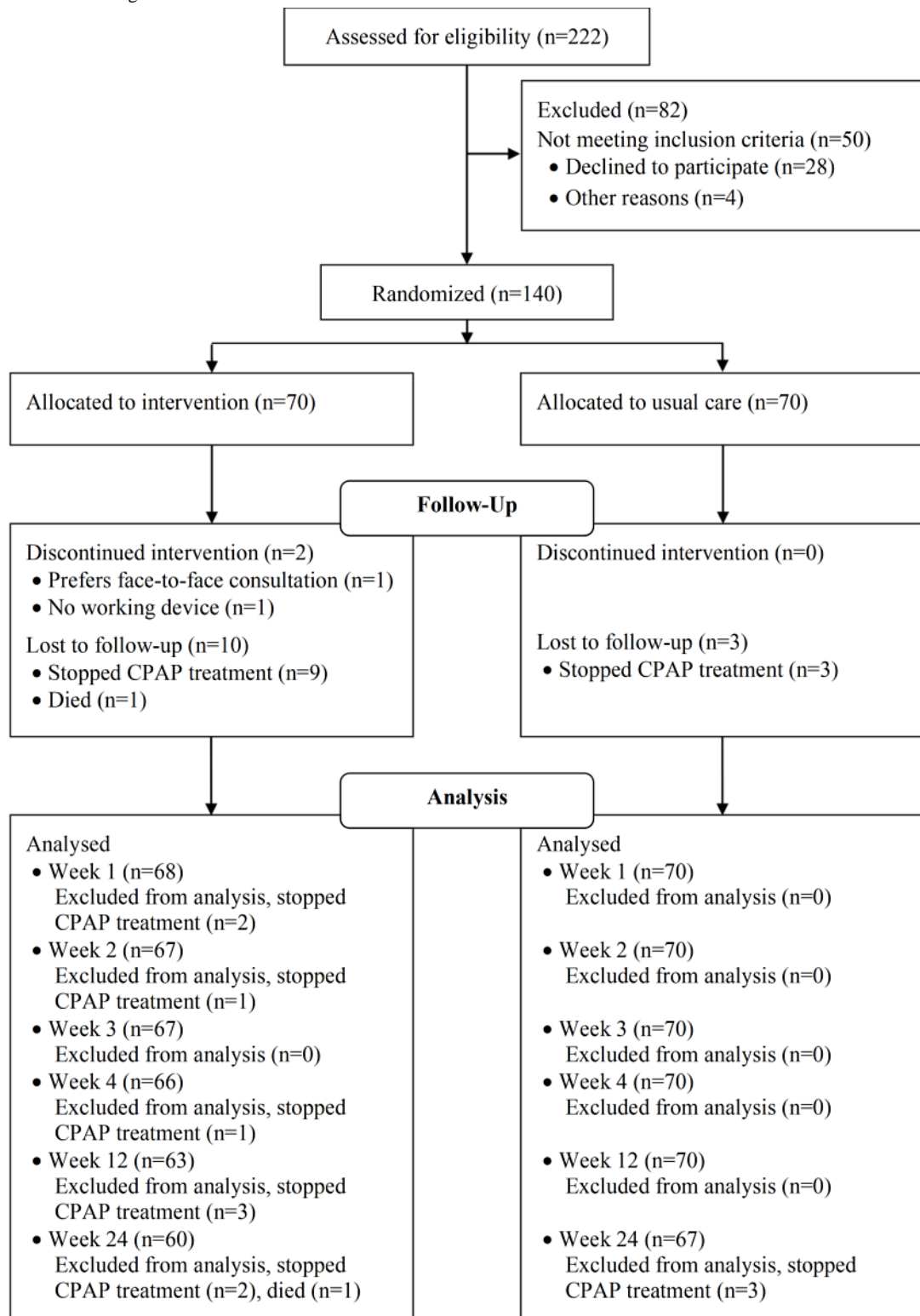
Results

Recruitment and Participants

Patients were included from January 2, 2019, until June 26, 2019. In total, 222 patients were screened for eligibility, and 50 patients did not meet the inclusion criteria: no tablet or smartphone ($n=17$), no proficiency in the Dutch language ($n=10$), AHI <15 ($n=10$), history of CPAP treatment ($n=5$), no OSA ($n=4$), psychiatric or cognitive disorder ($n=3$), and age <18 years ($n=1$). In total, 28 patients declined to participate, and 4 patients were not informed about the study for other reasons: 2 patients were not referred to the researcher due to logistical errors, 1 patient followed a different care process (there was no consultation with the pulmonologist that same day), and 1 patient had had CPAP for try out for a short period.

In total, 140 patients were randomized, and 70 patients were allocated to the intervention group and 70 patients to the usual care group. During the intervention period, 2 patients discontinued the intervention: 1 preferred face-to-face consultation, and 1 had no working device. Four patients stopped CPAP treatment during the intervention period (first 4 weeks). In total, 10 patients were lost to follow-up in the intervention group ($n=9$ stopped CPAP treatment and $n=1$ died) and 3 in the usual care group ($n=3$ stopped CPAP treatment). See [Figure 1](#) for the CONSORT (Consolidated Standards of Reporting Trials) flow diagram.

Figure 1. CONSORT flow diagram.



Baseline Characteristics

Both groups had similar baseline characteristics (Table 1), only outcome expectancies ($P=.048$) and risk perception ($P=.02$) appeared to be significantly different between groups.

Table 1. Baseline characteristics (n=140).

Characteristics	All patients (n=140)	Intervention (n=70)	Usual care (n=70)	P value
Gender, women, n (%)	29 (21)	12 (17)	17 (24)	.30
Age (years), mean (SD)	53.3 (12.1)	52.3 (12.4)	54.3 (11.9)	.40
AHI ^a , median (IQR)	31.0 (21.5-45.0)	31.0 (22.0-46.0)	30.5 (20.0-42.0)	.96
Living with a partner, n (%)	110 (79)	59 (84)	51 (73)	.10
Education, n (%)	— ^b	—	—	.22
Low	8 (6)	3 (4)	5 (7)	—
Middle	89 (64)	41 (59)	48 (69)	—
High	43 (31)	26 (37)	17 (24)	—
Internet use: duration, n (%)	—	—	—	>.99
<6 months	3 (2)	1 (1)	2 (3)	—
1-2 years	1 (1)	1 (1)	0 (0)	—
>2 years	1 (1)	1 (1)	0 (0)	—
>3 years	135 (96)	67 (96)	68 (97)	—
Internet use: frequency, n (%)	—	—	—	.31
(almost) every day	128 (91)	66 (94)	62 (89)	—
Multiple days per week	9 (6)	4 (6)	5 (7)	—
≤1 day per week	3 (2)	0 (0)	3 (4)	—
Tablet or smartphone skills, n (%)	—	—	—	.91
Quite bad or bad	5 (4)	2 (3)	3 (4)	—
Not good or not bad	23 (16)	11 (16)	12 (17)	—
Quite good	27 (19)	14 (20)	13 (19)	—
Good	55 (39)	26 (37)	29 (41)	—
Very good	30 (21)	17 (24)	13 (19)	—
Expects to need help with tablet or smartphone use, n (%)	26 (19)	11 (16)	15 (22)	.41
Comorbidities, n (%)	—	—	—	—
Obesity (BMI >30)	97 (69)	51 (73)	46 (66)	.36
Hypertension	48 (34)	24 (34)	24 (34)	>.99
Hypercholesterolemia	21 (15)	8 (11)	13 (19)	.24
Heart disease	20 (14)	11 (16)	9 (13)	.63
Diabetes	14 (10)	7 (10)	7 (10)	>.99
ESS^c score, n (%)	—	—	—	.19
Total score ≤10	105 (79)	56 (84)	49 (74)	—
Total score >10	28 (21)	11 (16)	17 (26)	—
SEMSA^d constructs	—	—	—	—
Outcome expectancies, mean (SD)	2.78 (0.62)	2.88 (0.57)	2.67 (0.65)	.048
Self-efficacy, median (IQR)	3.00 (2.56-3.56)	3.00 (2.56-3.33)	3.00 (2.56-3.67)	.40
Risk perception, median (IQR)	2.00 (1.54-2.50)	2.31 (1.63-2.63)	1.88 (1.50-2.31)	.02

^aAHI: apnea-hypopnea index.^bNot applicable.^cESS: Epworth Sleepiness Scale.^dSEMSA: Self-Efficacy Measure for Sleep Apnea.

CPAP Use

The use of video consultation does not lead to superior results on CPAP use compared with face-to-face consultation. A significant difference in change over time was found between groups for short-term (weeks 1 through 4) CPAP use (P -interaction=.008). However, the specific time points (week 1: P =.62; week 2: P =.15; week 3: P =.33, and week 4: P =.20) were not significantly different. See [Multimedia Appendix 2](#)

and [Multimedia Appendix 3](#) for more detailed information on short-term CPAP use.

No significant difference in change over time for long-term CPAP use (week 4, week 12, and week 24) was found between groups (P -interaction=.68). CPAP use decreased for both groups in the long term (P =.008), but no significant difference was found between the intervention and usual care group (P =.09). See [Table 2](#) and [Figure 2](#) for change in CPAP use over time (week 4, week 12, and week 24).

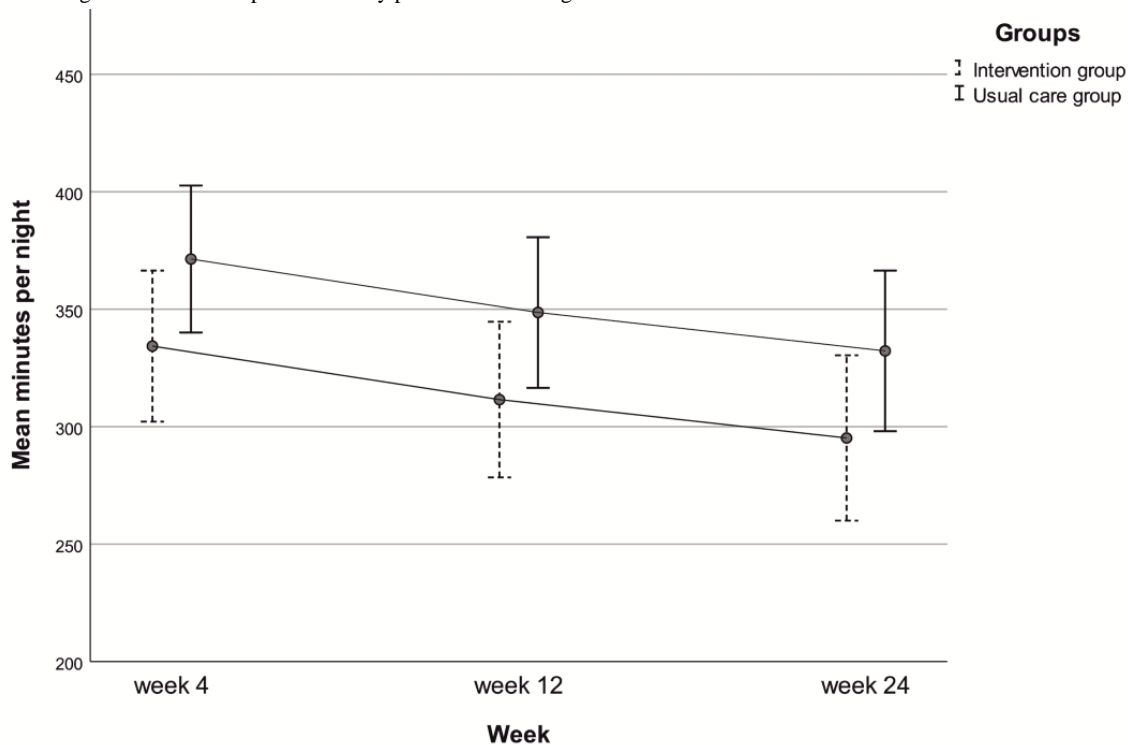
Table 2. Long-term continuous positive airway pressure use (minutes per night).

Week ^a	Intervention		Usual care	
	EMM ^b (SE)	95% CI	EMM(SE)	95% CI
Week 4	334.3 (16.3)	302.1-366.5	371.4 (15.8)	340.1-402.7
Week 12	311.5 (16.8)	278.4-344.6	348.6 (16.2)	316.5-380.7
Week 24	295.2 (17.8)	260.0-330.4	332.7 (17.3)	298.1-366.5

^aLinear mixed model.

^bEMM: estimated marginal mean.

Figure 2. Long-term continuous positive airway pressure use: change over time.



CPAP Adherence

The use of video consultation does not lead to superior results on CPAP adherence compared with face-to-face consultation. No significant difference was found between both groups for short-term (P =.95) and long-term (P =.12) CPAP adherence. Also, no significant difference in change over time between the intervention and usual care group was found for short-term (P -interaction=.17) and long-term (P -interaction=.51) CPAP adherence. See [Multimedia Appendix 4](#) and [Multimedia Appendix 5](#) for the short-term and long-term adherence rates per week.

Self-Efficacy, Risk Outcomes, and Outcome Expectancies

No significant difference between groups was found for the SEMSA constructs: outcome expectancies (P =.64), self-efficacy (P =.41), and risk perception (P =.30). See [Multimedia Appendix 6](#).

Relation Between Self-Efficacy, Risk Perception, Outcome Expectancies, and CPAP Use

After 4 weeks, a relation was found between CPAP use and self-efficacy (P =.001), meaning that patients with higher levels

of self-efficacy showed higher CPAP use. There was no relation between CPAP use and risk perception ($P=.34$) or outcome expectancies ($P=.76$). Also, the difference between the intervention and usual care group was not significant ($P=.25$).

Expectations and Experiences With Video Consultation

Patients expressed positive expectations for the use of video consultation. After 4 weeks, 76% (48/63) indicated that video consultation had a positive effect on control over their treatment, and 75% (47/63) indicated that it positively affected the treatment itself. The majority (58/63, 92%) implied it did not cost them effort, 95% (60/63) reported that they had enough skills to use a tablet or smartphone and that they received enough support (53/63, 84%). Although, 64% (44/69) expected to be stimulated by people in their direct environment to use video consultation, only 25% (16/63) were actually stimulated. Almost all patients (60/63, 95%) intended to keep using video consultation. See [Multimedia Appendix 7](#).

Satisfaction With Consultation

Patients in both groups were satisfied with the consultations. On average, the intervention group rated the consultations with an 8.5 and the usual care group with an 8.3 on a scale of 1 to 10 (1=not at all satisfied to 10=very satisfied). Patients indicated (intervention group versus usual care group) that health care professionals understood their problems (59/63, 94%, vs 58/68, 85%) and listened to them (60/63, 95%, vs 61/68, 90%). Almost all patients understood the content of the consultation (61/63, 97%, vs 62/68, 91%), could easily express their feelings (59/63, 94%, vs 62/68, 91%), and were satisfied with the information they received (58/63, 92%, vs 60/68, 88%). However, more patients with video consultation reported that they did not miss important information (56/63, 89%, vs 43/68, 63%). See [Multimedia Appendix 8](#).

Satisfaction With Video Consultation

The majority (56/63, 89%) of the patients were very satisfied with video consultation, the quality of the video (50/63, 79%), and sound of the system (45/63, 71%). It also saved them time (61/63, 97%) and provided better access to health care professionals (43/63, 68%). Almost all patients felt safe about their privacy and confidentiality (61/63, 97%) and preferred a video consultation over a face-to-face consultation (51/63, 81%). According to almost half (28/63, 44%) the patients, face-to-face consultation can be replaced by video consultation. See [Multimedia Appendix 9](#).

Nurse Satisfaction

Nurses ($n=3$) rated the use of video consultation on average with a 7.3 (SD .57) on a scale of 1 to 10 (1=not at all satisfied to 10=very satisfied). They were all satisfied with privacy and confidentiality and quality of the sound and video and would recommend its use to colleagues and patients. Two nurses agreed that its use fits in their work process. However, only one nurse was completely satisfied with the information she could provide. They did not think that the use of video consultation helped them save time or work more efficiently.

The nurses reported that use of video consultation is not suitable for new patients, and they prefer to use it during follow-up:

It is not suitable for a first consultation after starting CPAP because you cannot provide enough information.

Not for new patients because providing information and checking the device and sleep mask is difficult using video consultation.

The nurses also experienced some technical problems:

Sometimes there were log-in problems and I had to call the patient first by phone.

Sometimes it took long before there was a connection. This costs more time.

They also provided suggestions for improvement and described advantages of video consultations:

Plan the video consultations one after the other and not alternating with face-to-face consultations.

It is a good alternative for follow-up consultations. It is more patient friendly than a face-to-face consultation.

Saves time for patients.

Discussion

Principal Findings

In this RCT, we evaluated the superiority of video consultation over face-to-face consultation for newly diagnosed OSA patients. For CPAP use, we found a significant difference in change over time between groups in the short term (P -interaction=.008). However, the specific time points (week 1: $P=.62$; week 2: $P=.15$; week 3: $P=.33$, and week 4: $P=.20$) were not significantly different. No significant difference in change over time was found for long-term CPAP use (P -interaction=.68). No significant difference in change over time between groups was found for short-term (P -interaction=.17) or long-term (P -interaction=.51) CPAP adherence. Self-efficacy appeared to have a statistically significant effect on CPAP use in both groups ($P=.001$) regardless of the intervention arm ($P=.25$). No significant difference between groups was found for outcome expectancies ($P=.64$), self-efficacy ($P=.41$), or risk perception ($P=.30$). The experiences with video consultation were very positive. Almost all patients (60/63, 95%) intended to keep using video consultation. Patients in both groups rated the consultations on average with an 8.4. All nurses ($n=3$) were satisfied with privacy and confidentiality aspects and quality of the sound and video. However, they expressed some recommendations for improvement (eg, to use video consultation only in follow-up).

Comparison With Prior Work

Unfortunately, change over time was not evaluated in previous controlled studies [18,26,29], but this evaluation is as such a likely pattern. In our study, a significant difference in CPAP use between video consultation and face-to-face consultation was not found. Parikh et al [18] reported statistically equivalent CPAP use for new OSA patients (mean average use minutes per day 305.31 vs 340.55, $P=.15$). In a multicenter RCT, no statistically significant difference was found for CPAP use after 6 months (telemedicine mean use 4.4 [SD 2.0] hours per day

vs face-to-face 4.2 [SD 2.0] hours per day, $P=.83$) and adherence (telemedicine 65% vs usual care 57% compliance, $P=.33$) [29]. Based on these findings, it appears that CPAP use is equivalent to using video consultation.

Where previous studies mainly focused on CPAP use, adherence, and satisfaction with video consultation [17,18,25,26,29], we additionally evaluated the combination of cognitive components (self-efficacy, outcome expectancies, and risk perception), experience with the technology (using the UTAUT model), and satisfaction of patients and nurses. This combination of outcomes has received little attention until now. Cognitive components are found to be increasingly important in predicting CPAP use [13,14,27]. Our results show that use of CPAP is higher in patients with high levels of self-efficacy ($P=.001$) regardless of the intervention arm ($P=.25$). In order to improve self-efficacy, it is necessary to positively influence patient perceptions. Patients may benefit from a self-management approach [27,41,42] with tailored education to change their perceptions about CPAP use and subsequently improve self-efficacy [43]. Lai et al [44] provided patients with additional education to enhance, for example, self-efficacy. This increased CPAP use compared with patients receiving usual care ($P<.001$). Stepnowsky et al [41] showed that a self-management program with information about OSA- and CPAP-related issues led to high self-efficacy scores (4.5 [SD 0.6]; scale 0 to 5) and CPAP adherence (5.5 [SD 2.3] mean hours per night). Because self-efficacy scores can be affected by the time that patients are treated, scores should be assessed regularly in order to be useful in clinical practice [14].

However, limited evidence was available about the effect of video consultation for newly diagnosed patients starting CPAP. Most previous RCTs were small, with sample sizes varying from 19 to 40 patients [17,25]. Only Isetta et al [29] evaluated CPAP compliance with a fully powered sample size. Although almost half of the patients (40%) in this study had insufficient digital skills, technology aspects were not evaluated [29]. In our study, 9% (20/222) were unable to participate because of lack of access to a mobile device or due to psychiatric or cognitive disorder. During the intervention, 2 patients (2/70, 3%) discontinued the video consultation intervention because of preference for face-to-face consultation or problems with their mobile device. The use of video consultation is evolving rapidly in clinical practice, but digital services are not applicable to all patients and digital health literacy remains a challenge [45]. This is especially due to lack of awareness or knowledge or unwillingness to change [46] and emphasizes the importance of personalized interventions rather than a one-size-fits-all approach.

The assessment of UTAUT components and self-efficacy can also be used to indicate technology use [47]. To our knowledge, no previous studies have identified technology acceptance for OSA patients using video consultation. Patients in our study had positive experiences with the use of video consultation and were satisfied with the video consultation system and consultations in general. Previous studies also reported high satisfaction scores [17,18,25,26], mostly regarding communication with a health care professional [18] and privacy and security factors [17]. Although most patients would

recommend the use of video consultations to others, not all patients in our study are convinced that all visits can be replaced by video consultations. This is in line with findings from previous research [17].

The involvement of health care professionals is essential to achieve successful implementation of technology [48], but this is often not evaluated [17]. We found that nurses ($n=3$) preferred to start with a face-to-face consultation because education about the sleep mask and adjustments are often required during the first follow-up appointment with the nurse. The applicability of technology use may be dependent on the population [49], and for OSA patients, the use of video consultation in a blended care setting might therefore be beneficial. We found that the nurses were satisfied with video consultation and especially with the quality of the system, privacy and confidentiality. They would recommend it to colleagues and patients. Nurses also reported technical problems (eg, problems with Wi-Fi connections). Technological issues are often seen as a barrier [50], and it is important to take technical elements into account [48,51,52] during implementation. Another point for improvement is integration in existing health care processes (eg, planning). To achieve successful implementation, it can be beneficial to involve professionals during the implementation process itself [50].

Video consultation can be seen as a promising app to support OSA patients during treatment. Still, evidence was lacking and previous research was not strong enough in design or focused on a limited number of outcomes. With the evaluation of a broad range of outcomes affecting CPAP use and implementation of video consultation in clinical practice, this RCT adds value to current knowledge.

However, proper evaluation in this field is challenging because research often lags behind the rapid development of technology [53]. The use of pragmatic trials may be promising [54] to evaluate different elements of eHealth solutions in a hospital setting and can, for example, be used to get (more) rapid insights in relevant implementation outcomes such as feasibility, impact on an organization, and acceptance and adoption by health care professionals and patients. Future research should focus on blended care approaches in which self-efficacy especially receives greater emphasis. For organizations to be able to implement video consultation on a larger scale, integration in existing health care processes and technology acceptance by patients and professionals is necessary.

Limitations

Several limitations should be considered. Risk perception and outcome expectancies were significantly different at baseline, despite randomization. For a limited number of patients (7/66, 11%, in the intervention group and 6/70, 9%, in the control group), video consultations or face-to-face consultations were replaced with a telephonic consultation due to technical problems in the intervention group and because patients in the control group could not come to the hospital. The protocol process were not strictly followed because patients failed to attend their scheduled appointment (no show, sick, on holiday) or there were organizational inaccuracies such as wrongly scheduled appointments. The percentage of patients that

followed the process exactly as described (Multimedia Appendix 1) was higher in the intervention group (approximately half) than in the usual care group (approximately one-third). However, all patients received the intervention (type of consultation) they were allocated to except for the 2 patients who discontinued the intervention (Figure 1). Another limitation is that only 3 nurses were involved in the evaluation. Therefore, a firm conclusion on professional aspects cannot be drawn.

Conclusion

Support of OSA patients with video consultation does not lead to superior results on CPAP use and adherence compared with face-to-face consultation. The findings of this research show that a significant difference in change over time was found between groups for short-term CPAP use (but not on specific time points), but not for long-term CPAP use. Levels of

self-efficacy were positively related to CPAP use in both groups. Patients were very satisfied with video consultation and reported positive experiences.

Therefore, the findings of this research suggest that self-efficacy is an important factor in improving CPAP use and that video consultation may be a feasible way to support patients starting CPAP. The integration in health care processes and tailoring video consultation use to patient and professional needs is essential to ensure successful use. A blended care setting, in which an initial video consultation is combined with face-to-face consults, may be beneficial. To our knowledge, this is the first RCT that examined the effects of video consultation on CPAP use over time for newly diagnosed OSA patients in combination with cognitive components and experience with technology use. Future research should focus on blended care approaches in which self-efficacy receives greater emphasis.

Acknowledgments

The authors thank Els Fikkers (nurse practitioner, pulmonology) for her assistance with the study.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Study process.

[PNG File, 61 KB - [formative_v5i5e20779_app1.png](#)]

Multimedia Appendix 2

Short-term CPAP use.

[DOC File, 30 KB - [formative_v5i5e20779_app2.doc](#)]

Multimedia Appendix 3

Short-term CPAP use: change over time.

[PNG File, 70 KB - [formative_v5i5e20779_app3.png](#)]

Multimedia Appendix 4

Short-term CPAP adherence.

[DOC File, 29 KB - [formative_v5i5e20779_app4.doc](#)]

Multimedia Appendix 5

Long-term CPAP adherence.

[DOC File, 29 KB - [formative_v5i5e20779_app5.doc](#)]

Multimedia Appendix 6

Self-Efficacy Measure for Sleep Apnea constructs: self-efficacy, risk perception, and outcome expectancies.

[DOC File, 29 KB - [formative_v5i5e20779_app6.doc](#)]

Multimedia Appendix 7

Expectations and experiences with video consultation.

[DOC File, 38 KB - [formative_v5i5e20779_app7.doc](#)]

Multimedia Appendix 8

Patient satisfaction with consultation.

[DOC File, 36 KB - [formative_v5i5e20779_app8.doc](#)]

Multimedia Appendix 9

Patient satisfaction with video consultation.

[\[DOC File, 36 KB - formative_v5i5e20779_app9.doc\]](#)

Multimedia Appendix 10

CONSORT-eHEALTH checklist (V 1.6.1).

[\[PDF File \(Adobe PDF File\), 412 KB - formative_v5i5e20779_app10.pdf\]](#)

References

1. Verbraecken J. Telemedicine applications in sleep disordered breathing: thinking out of the box. *Sleep Med Clin* 2016 Dec;11(4):445-459 [FREE Full text] [doi: [10.1016/j.jsmc.2016.08.007](https://doi.org/10.1016/j.jsmc.2016.08.007)] [Medline: [28118869](https://pubmed.ncbi.nlm.nih.gov/28118869/)]
2. Bruyneel M. Telemedicine in the diagnosis and treatment of sleep apnoea. *Eur Respir Rev* 2019 Mar 31;28(151):180093 [FREE Full text] [doi: [10.1183/16000617.0093-2018](https://doi.org/10.1183/16000617.0093-2018)] [Medline: [30872397](https://pubmed.ncbi.nlm.nih.gov/30872397/)]
3. Epstein LJ, Kristo D, Strollo PJ, Friedman N, Malhotra A, Patil SP, Adult Obstructive Sleep Apnea Task Force of the American Academy of Sleep Medicine. Clinical guideline for the evaluation, management and long-term care of obstructive sleep apnea in adults. *J Clin Sleep Med* 2009 Jun 15;5(3):263-276 [FREE Full text] [Medline: [19960649](https://pubmed.ncbi.nlm.nih.gov/19960649/)]
4. Punjabi NM. The epidemiology of adult obstructive sleep apnea. *Proc Am Thorac Soc* 2008 Feb 15;5(2):136-143 [FREE Full text] [doi: [10.1513/pats.200709-155MG](https://doi.org/10.1513/pats.200709-155MG)] [Medline: [18250205](https://pubmed.ncbi.nlm.nih.gov/18250205/)]
5. Kushida CA, Littner MR, Hirshkowitz M, Morgenthaler TI, Alessi CA, Bailey D, American Academy of Sleep Medicine. Practice parameters for the use of continuous and bilevel positive airway pressure devices to treat adult patients with sleep-related breathing disorders. *Sleep* 2006 Mar;29(3):375-380. [doi: [10.1093/sleep/29.3.375](https://doi.org/10.1093/sleep/29.3.375)] [Medline: [16553024](https://pubmed.ncbi.nlm.nih.gov/16553024/)]
6. Gay P, Weaver T, Loube D, Iber C, Positive Airway Pressure Task Force, Standards of Practice Committee, American Academy of Sleep Medicine. Evaluation of positive airway pressure treatment for sleep related breathing disorders in adults. *Sleep* 2006 Mar;29(3):381-401. [doi: [10.1093/sleep/29.3.381](https://doi.org/10.1093/sleep/29.3.381)] [Medline: [16553025](https://pubmed.ncbi.nlm.nih.gov/16553025/)]
7. American Academy of Sleep Medicine Task Force. Sleep-related breathing disorders in adults: recommendations for syndrome definition and measurement techniques in clinical research. *Sleep* 1999 Aug 01;22(5):667-689. [Medline: [10450601](https://pubmed.ncbi.nlm.nih.gov/10450601/)]
8. Gibson GJ. Obstructive sleep apnoea syndrome: underestimated and undertreated. *Br Med Bull* 2004;72:49-65. [doi: [10.1093/bmb/ldh044](https://doi.org/10.1093/bmb/ldh044)] [Medline: [15798032](https://pubmed.ncbi.nlm.nih.gov/15798032/)]
9. Sullivan CE, Issa FG, Berthon-Jones M, Eves L. Reversal of obstructive sleep apnoea by continuous positive airway pressure applied through the nares. *Lancet* 1981 Apr 18;1(8225):862-865. [doi: [10.1016/s0140-6736\(81\)92140-1](https://doi.org/10.1016/s0140-6736(81)92140-1)] [Medline: [6112294](https://pubmed.ncbi.nlm.nih.gov/6112294/)]
10. Stuck BA, Leitzbach S, Maurer JT. Effects of continuous positive airway pressure on apnea-hypopnea index in obstructive sleep apnea based on long-term compliance. *Sleep Breath* 2012 Jun;16(2):467-471. [doi: [10.1007/s11325-011-0527-8](https://doi.org/10.1007/s11325-011-0527-8)] [Medline: [21590521](https://pubmed.ncbi.nlm.nih.gov/21590521/)]
11. Bandura A. Self-efficacy mechanism in human agency. *Am Psychol* 1982;37(2):122-147. [doi: [10.1037/0003-066X.37.2.122](https://doi.org/10.1037/0003-066X.37.2.122)]
12. Stepnowsky CJ, Marler MR, Ancoli-Israel S. Determinants of nasal CPAP compliance. *Sleep Med* 2002 May;3(3):239-247. [doi: [10.1016/s1389-9457\(01\)00162-9](https://doi.org/10.1016/s1389-9457(01)00162-9)] [Medline: [14592213](https://pubmed.ncbi.nlm.nih.gov/14592213/)]
13. Weaver TE, Maislin G, Dinges DF, Younger J, Cantor C, McCloskey S, et al. Self-efficacy in sleep apnea: instrument development and patient perceptions of obstructive sleep apnea risk, treatment benefit, and volition to use continuous positive airway pressure. *Sleep* 2003 Sep;26(6):727-732. [doi: [10.1093/sleep/26.6.727](https://doi.org/10.1093/sleep/26.6.727)] [Medline: [14572127](https://pubmed.ncbi.nlm.nih.gov/14572127/)]
14. Aloia MS, Arnedt JT, Stepnowsky C, Hecht J, Borrelli B. Predicting treatment adherence in obstructive sleep apnea using principles of behavior change. *J Clin Sleep Med* 2005 Oct 15;1(4):346-353. [Medline: [17564399](https://pubmed.ncbi.nlm.nih.gov/17564399/)]
15. Sawyer AM, Gooneratne NS, Marcus CL, Ofer D, Richards KC, Weaver TE. A systematic review of CPAP adherence across age groups: clinical and empiric insights for developing CPAP adherence interventions. *Sleep Med Rev* 2011 Dec;15(6):343-356 [FREE Full text] [doi: [10.1016/j.smrv.2011.01.003](https://doi.org/10.1016/j.smrv.2011.01.003)] [Medline: [21652236](https://pubmed.ncbi.nlm.nih.gov/21652236/)]
16. Mehrtash M, Bakker JP, Ayas N. Predictors of continuous positive airway pressure adherence in patients with obstructive sleep apnea. *Lung* 2019 Apr;197(2):115-121. [doi: [10.1007/s00408-018-00193-1](https://doi.org/10.1007/s00408-018-00193-1)] [Medline: [30617618](https://pubmed.ncbi.nlm.nih.gov/30617618/)]
17. Isetta V, León C, Torres M, Embid C, Roca J, Navajas D, et al. Telemedicine-based approach for obstructive sleep apnea management: building evidence. *Interact J Med Res* 2014 Feb 19;3(1):e6 [FREE Full text] [doi: [10.2196/ijmr.3060](https://doi.org/10.2196/ijmr.3060)] [Medline: [24554392](https://pubmed.ncbi.nlm.nih.gov/24554392/)]
18. Parikh R, Touville MN, Wang H, Zallek SN. Sleep telemedicine: patient satisfaction and treatment adherence. *Telemed J E Health* 2011 Oct;17(8):609-614. [doi: [10.1089/tmj.2011.0025](https://doi.org/10.1089/tmj.2011.0025)] [Medline: [21859348](https://pubmed.ncbi.nlm.nih.gov/21859348/)]
19. Kitamura C, Zurawel-Balaura L, Wong RKS. How effective is video consultation in clinical oncology? A systematic review. *Curr Oncol* 2010 Jun;17(3):17-27 [FREE Full text] [doi: [10.3747/co.v17i3.513](https://doi.org/10.3747/co.v17i3.513)] [Medline: [20567623](https://pubmed.ncbi.nlm.nih.gov/20567623/)]
20. Fatehi F, Martin-Khan M, Smith AC, Russell AW, Gray LC. Patient satisfaction with video teleconsultation in a virtual diabetes outreach clinic. *Diabetes Technol Ther* 2015 Jan;17(1):43-48. [doi: [10.1089/dia.2014.0159](https://doi.org/10.1089/dia.2014.0159)] [Medline: [25296189](https://pubmed.ncbi.nlm.nih.gov/25296189/)]

21. Rasmussen OW, Lauszus FF, Loekke M. Telemedicine compared with standard care in type 2 diabetes mellitus: a randomized trial in an outpatient clinic. *J Telemed Telecare* 2016 Sep;22(6):363-368. [doi: [10.1177/1357633X15608984](https://doi.org/10.1177/1357633X15608984)] [Medline: [26468213](https://pubmed.ncbi.nlm.nih.gov/26468213/)]
22. Sabesan S, Simcox K, Marr I. Medical oncology clinics through videoconferencing: an acceptable telehealth model for rural patients and health workers. *Intern Med J* 2012 Jul;42(7):780-785. [doi: [10.1111/j.1445-5994.2011.02537.x](https://doi.org/10.1111/j.1445-5994.2011.02537.x)] [Medline: [21627743](https://pubmed.ncbi.nlm.nih.gov/21627743/)]
23. Slaapmasker. Rijnstate. URL: <https://www.rijnstate.nl/aandoening-en-behandeling/slaapapneu/behandeling/slaapmasker/> [accessed 2020-10-24]
24. Murphie P, Little S, McKinsty B, Pinnock H. Remote consulting with telemonitoring of continuous positive airway pressure usage data for the routine review of people with obstructive sleep apnoea hypopnoea syndrome: A systematic review. *J Telemed Telecare* 2019 Jan;25(1):17-25. [doi: [10.1177/1357633X17735618](https://doi.org/10.1177/1357633X17735618)] [Medline: [28990455](https://pubmed.ncbi.nlm.nih.gov/28990455/)]
25. Smith CE, Daut ER, Clements F, Puno FN, Cook D, Doolittle G, et al. Telehealth services to improve nonadherence: a placebo-controlled study. *Telemed J E Health* 2006 Jun;12(3):289-296. [doi: [10.1089/tmj.2006.12.289](https://doi.org/10.1089/tmj.2006.12.289)] [Medline: [16796496](https://pubmed.ncbi.nlm.nih.gov/16796496/)]
26. Coma-Del-Corral MJ, Alonso-Álvarez ML, Allende M, Cordero J, Ordax E, Masa F, et al. Reliability of telemedicine in the diagnosis and treatment of sleep apnea syndrome. *Telemed J E Health* 2013 Jan;19(1):7-12 [FREE Full text] [doi: [10.1089/tmj.2012.0007](https://doi.org/10.1089/tmj.2012.0007)] [Medline: [23186084](https://pubmed.ncbi.nlm.nih.gov/23186084/)]
27. Stepnowsky CJ, Marler MR, Ancoli-Israel S. Determinants of nasal CPAP compliance. *Sleep Med* 2002 May;3(3):239-247. [doi: [10.1016/s1389-9457\(01\)00162-9](https://doi.org/10.1016/s1389-9457(01)00162-9)] [Medline: [14592213](https://pubmed.ncbi.nlm.nih.gov/14592213/)]
28. Weaver TE, Kribbs NB, Pack AI, Kline LR, Chugh DK, Maislin G, et al. Night-to-night variability in CPAP use over the first three months of treatment. *Sleep* 1997 Apr;20(4):278-283. [doi: [10.1093/sleep/20.4.278](https://doi.org/10.1093/sleep/20.4.278)] [Medline: [9231953](https://pubmed.ncbi.nlm.nih.gov/9231953/)]
29. Isetta V, Negrín MA, Monasterio C, Masa JF, Feu N, Álvarez A, SPANISH SLEEP NETWORK. A Bayesian cost-effectiveness analysis of a telemedicine-based strategy for the management of sleep apnoea: a multicentre randomised controlled trial. *Thorax* 2015 Nov;70(11):1054-1061. [doi: [10.1136/thoraxjnl-2015-207032](https://doi.org/10.1136/thoraxjnl-2015-207032)] [Medline: [26310452](https://pubmed.ncbi.nlm.nih.gov/26310452/)]
30. Chen C, Wang J, Pang L, Wang Y, Ma G, Liao W. Telemonitor care helps CPAP compliance in patients with obstructive sleep apnea: a systemic review and meta-analysis of randomized controlled trials. *Ther Adv Chronic Dis* 2020;11:2040622320901625 [FREE Full text] [doi: [10.1177/2040622320901625](https://doi.org/10.1177/2040622320901625)] [Medline: [32215196](https://pubmed.ncbi.nlm.nih.gov/32215196/)]
31. Bouloukaki I, Giannadaki K, Mermigkis C, Tzanakis N, Mauroudi E, Moniaki V, et al. Intensive versus standard follow-up to improve continuous positive airway pressure compliance. *Eur Respir J* 2014 Nov;44(5):1262-1274 [FREE Full text] [doi: [10.1183/09031936.00021314](https://doi.org/10.1183/09031936.00021314)] [Medline: [24993911](https://pubmed.ncbi.nlm.nih.gov/24993911/)]
32. Facetalk. URL: <https://facetalk.nl/> [accessed 2021-04-24]
33. Facetalk. Google Play. URL: <https://play.google.com/store/apps/details?id=nl.facetalk.mobile&hl=nl&gl=US> [accessed 2020-10-24]
34. Facetalk. App Store. URL: <https://apps.apple.com/nl/app/facetalk/id1178990857> [accessed 2020-10-24]
35. Weaver TE, Sawyer AM. Adherence to continuous positive airway pressure treatment for obstructive sleep apnoea: implications for future interventions. *Indian J Med Res* 2010 Feb;131:245-258 [FREE Full text] [Medline: [20308750](https://pubmed.ncbi.nlm.nih.gov/20308750/)]
36. Taalcentrum-VU. Vertaalbureau Amsterdam. URL: <https://www.taalcentrum-vu.nl/vertaalbureau/> [accessed 2020-10-24]
37. Venkatesh, Morris, Davis, Davis. User acceptance of information technology: toward a unified view. *MIS Quarterly* 2003;27(3):425. [doi: [10.2307/30036540](https://doi.org/10.2307/30036540)]
38. Johns MW. A new method for measuring daytime sleepiness: the Epworth sleepiness scale. *Sleep* 1991 Dec;14(6):540-545. [Medline: [1798888](https://pubmed.ncbi.nlm.nih.gov/1798888/)]
39. Johns M. The Epworth Sleepiness Scale. URL: <https://epworthsleepinessscale.com/about-the-ess/> [accessed 2019-12-03]
40. Aardoom JJ, Loheide-Niesmann L, Ossebaard HC, Riper H. Effectiveness of eHealth interventions in improving treatment adherence for adults with obstructive sleep apnea: meta-analytic review. *J Med Internet Res* 2020 Feb 18;22(2):e16972 [FREE Full text] [doi: [10.2196/16972](https://doi.org/10.2196/16972)] [Medline: [32130137](https://pubmed.ncbi.nlm.nih.gov/32130137/)]
41. Stepnowsky CJ, Palau JJ, Gifford AL, Ancoli-Israel S. A self-management approach to improving continuous positive airway pressure adherence and outcomes. *Behav Sleep Med* 2007;5(2):131-146. [doi: [10.1080/15402000701190622](https://doi.org/10.1080/15402000701190622)] [Medline: [17441783](https://pubmed.ncbi.nlm.nih.gov/17441783/)]
42. Barlow J, Wright C, Sheasby J, Turner A, Hainsworth J. Self-management approaches for people with chronic conditions: a review. *Patient Educ Couns* 2002;48(2):177-187. [Medline: [12401421](https://pubmed.ncbi.nlm.nih.gov/12401421/)]
43. Sawyer AM, Canamucio A, Moriarty H, Weaver TE, Richards KC, Kuna ST. Do cognitive perceptions influence CPAP use? *Patient Educ Couns* 2011 Oct;85(1):85-91 [FREE Full text] [doi: [10.1016/j.pec.2010.10.014](https://doi.org/10.1016/j.pec.2010.10.014)] [Medline: [21071166](https://pubmed.ncbi.nlm.nih.gov/21071166/)]
44. Lai AYK, Fong DYT, Lam JCM, Weaver TE, Ip MSM. The efficacy of a brief motivational enhancement education program on CPAP adherence in OSA: a randomized controlled trial. *Chest* 2014 Sep;146(3):600-610. [doi: [10.1378/chest.13-2228](https://doi.org/10.1378/chest.13-2228)] [Medline: [24810282](https://pubmed.ncbi.nlm.nih.gov/24810282/)]
45. Smith B, Magnani JW. New technologies, new disparities: the intersection of electronic health and digital health literacy. *Int J Cardiol* 2019 Oct 01;292:280-282. [doi: [10.1016/j.ijcard.2019.05.066](https://doi.org/10.1016/j.ijcard.2019.05.066)] [Medline: [31171391](https://pubmed.ncbi.nlm.nih.gov/31171391/)]
46. Almathami HKY, Win KT, Vlahu-Gjorgievska E. Barriers and facilitators that influence telemedicine-based, real-time, online consultation at patients' homes: systematic literature review. *J Med Internet Res* 2020 Feb 20;22(2):e16407 [FREE Full text] [doi: [10.2196/16407](https://doi.org/10.2196/16407)] [Medline: [32130131](https://pubmed.ncbi.nlm.nih.gov/32130131/)]

47. Kohnke A, Cole ML, Bush R. Incorporating UTAUT predictors for understanding home care patients' and clinician's acceptance of healthcare telemedicine equipment. *J Technol Manag Innov* 2014 Jul;9(2):29-41. [doi: [10.4067/S0718-27242014000200003](https://doi.org/10.4067/S0718-27242014000200003)]
48. Hennemann S, Beutel ME, Zwerenz R. Ready for eHealth? Health professionals' acceptance and adoption of ehealth interventions in inpatient routine care. *J Health Commun* 2017 Mar;22(3):274-284. [doi: [10.1080/10810730.2017.1284286](https://doi.org/10.1080/10810730.2017.1284286)] [Medline: [28248626](https://pubmed.ncbi.nlm.nih.gov/28248626/)]
49. Johansson AM, Lindberg I, Söderberg S. The views of health-care personnel about video consultation prior to implementation in primary health care in rural areas. *Prim Health Care Res Dev* 2014 Apr;15(2):170-179 [FREE Full text] [doi: [10.1017/S1463423613000030](https://doi.org/10.1017/S1463423613000030)] [Medline: [23402617](https://pubmed.ncbi.nlm.nih.gov/23402617/)]
50. de Veer AJE, Fleuren MAH, Bekkema N, Francke AL. Successful implementation of new technologies in nursing care: a questionnaire survey of nurse-users. *BMC Med Inform Decis Mak* 2011;11:67 [FREE Full text] [doi: [10.1186/1472-6947-11-67](https://doi.org/10.1186/1472-6947-11-67)] [Medline: [22032728](https://pubmed.ncbi.nlm.nih.gov/22032728/)]
51. Scott KC, Karem P, Shifflett K, Vegi L, Ravi K, Brooks M. Evaluating barriers to adopting telemedicine worldwide: a systematic review. *J Telemed Telecare* 2018 Jan;24(1):4-12 [FREE Full text] [doi: [10.1177/1357633X16674087](https://doi.org/10.1177/1357633X16674087)] [Medline: [29320966](https://pubmed.ncbi.nlm.nih.gov/29320966/)]
52. Gagnon MP, Orruño E, Asua J, Abdeljelil AB, Emparanza J. Using a modified technology acceptance model to evaluate healthcare professionals' adoption of a new telemonitoring system. *Telemed J E Health* 2012;18(1):54-59 [FREE Full text] [doi: [10.1089/tmj.2011.0066](https://doi.org/10.1089/tmj.2011.0066)] [Medline: [22082108](https://pubmed.ncbi.nlm.nih.gov/22082108/)]
53. Baker TB, Gustafson DH, Shah D. How can research keep up with eHealth? Ten strategies for increasing the timeliness and usefulness of eHealth research. *J Med Internet Res* 2014;16(2):e36 [FREE Full text] [doi: [10.2196/jmir.2925](https://doi.org/10.2196/jmir.2925)] [Medline: [24554442](https://pubmed.ncbi.nlm.nih.gov/24554442/)]
54. Peterson ED, Harrington RA. Evaluating health technology through pragmatic trials: novel approaches to generate high-quality evidence. *JAMA* 2018 Jul 10;320(2):137-138. [doi: [10.1001/jama.2018.8512](https://doi.org/10.1001/jama.2018.8512)] [Medline: [29998322](https://pubmed.ncbi.nlm.nih.gov/29998322/)]

Abbreviations

- AHI:** apnea-hypopnea index
CONSORT: Consolidated Standards of Reporting Trials
CPAP: continuous positive airway pressure
OSA: obstructive sleep apnea
RCT: randomized controlled trial
SEMSA: Self-Efficacy Measure for Sleep Apnea
UTAUT: unified theory of acceptance and use of technology

Edited by G Eysenbach; submitted 28.05.20; peer-reviewed by M Stommel; comments to author 04.07.20; revised version received 25.10.20; accepted 13.04.21; published 11.05.21.

Please cite as:

Kooij L, Vos PJE, Dijkstra A, Roovers EA, van Harten WH

Video Consultation as an Adequate Alternative to Face-to-Face Consultation in Continuous Positive Airway Pressure Use for Newly Diagnosed Patients With Obstructive Sleep Apnea: Randomized Controlled Trial

JMIR Form Res 2021;5(5):e20779

URL: <https://formative.jmir.org/2021/5/e20779>

doi: [10.2196/20779](https://doi.org/10.2196/20779)

PMID: [33973866](https://pubmed.ncbi.nlm.nih.gov/33973866/)

©Laura Kooij, Petra JE Vos, Antoon Dijkstra, Elisabeth A Roovers, Wim H van Harten. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 11.05.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Using Multimodal Assessments to Capture Personalized Contexts of College Student Well-being in 2020: Case Study

Jocelyn Lai¹, MA; Amir Rahmani^{2,3,4}, MBA, PhD; Asal Yunusova¹, BA; Alexander P Rivera¹, BA; Sina Labbaf², BSc; Sirui Hu^{5,6}, BA, BSc; Nikil Dutt^{2,4,7,8}, PhD; Ramesh Jain^{2,4}, PhD; Jessica L Borelli^{1,4}, PhD

¹UCI THRIVE Lab, Department of Psychological Science, University of California, Irvine, Irvine, CA, United States

²Department of Computer Science, University of California, Irvine, Irvine, CA, United States

³School of Nursing, University of California, Irvine, Irvine, CA, United States

⁴Institute for Future Health, University of California, Irvine, Irvine, CA, United States

⁵Department of Statistics, University of California, Irvine, Irvine, CA, United States

⁶Department of Economics, University of California, Irvine, Irvine, CA, United States

⁷Department of Electrical Engineering and Computer Science, University of California, Irvine, Irvine, CA, United States

⁸Department of Cognitive Science, Irvine, CA, United States

Corresponding Author:

Jocelyn Lai, MA

UCI THRIVE Lab

Department of Psychological Science

University of California, Irvine

4201 Social and Behavioral Sciences Gateway

Irvine, CA, 92697

United States

Phone: 1 4086565508

Email: jocelyn.lai@uci.edu

Abstract

Background: The year 2020 has been challenging for many, particularly for young adults who have been adversely affected by the COVID-19 pandemic. Emerging adulthood is a developmental phase with significant changes in the patterns of daily living; it is a risky phase for the onset of major mental illness. College students during the pandemic face significant risk, potentially losing several protective factors (eg, housing, routine, social support, job, and financial security) that are stabilizing for mental health and physical well-being. Individualized multiple assessments of mental health, referred to as *multimodal personal chronicles*, present an opportunity to examine indicators of health in an ongoing and personalized way using mobile sensing devices and wearable internet of things.

Objective: To assess the feasibility and provide an in-depth examination of the impact of the COVID-19 pandemic on college students through multimodal personal chronicles, we present a case study of an individual monitored using a longitudinal subjective and objective assessment approach over a 9-month period throughout 2020, spanning the prepandemic period of January through September.

Methods: The individual, referred to as Lee, completed psychological assessments measuring depression, anxiety, and loneliness across 4 time points in January, April, June, and September. We used the data emerging from the multimodal personal chronicles (ie, heart rate, sleep, physical activity, affect, behaviors) in relation to psychological assessments to understand patterns that help to explicate changes in the individual's psychological well-being across the pandemic.

Results: Over the course of the pandemic, Lee's depression severity was highest in April, shortly after shelter-in-place orders were mandated. His depression severity remained mildly severe throughout the rest of the months. Associations in positive and negative affect, physiology, sleep, and physical activity patterns varied across time periods. Lee's positive affect and negative affect were positively correlated in April ($r=0.53$, $P=.04$) whereas they were negatively correlated in September ($r=-0.57$, $P=.03$). Only in the month of January was sleep negatively associated with negative affect ($r=-0.58$, $P=.03$) and diurnal beats per minute ($r=-0.54$, $P=.04$), and then positively associated with heart rate variability (resting root mean square of successive differences between normal heartbeats) ($r=0.54$, $P=.04$). When looking at his available contextual data, Lee noted certain situations as supportive coping factors and other situations as potential stressors.

Conclusions: We observed more pandemic concerns in April and noticed other contextual events relating to this individual's well-being, reflecting how college students continue to experience life events during the pandemic. The rich monitoring data alongside contextual data may be beneficial for clinicians to understand client experiences and offer personalized treatment plans. We discuss benefits as well as future directions of this system, and the conclusions we can draw regarding the links between the COVID-19 pandemic and college student mental health.

(*JMIR Form Res* 2021;5(5):e26186) doi:[10.2196/26186](https://doi.org/10.2196/26186)

KEYWORDS

COVID-19; emerging adulthood; multimodal personal chronicles; case study; wearable internet of things; individualized mHealth; college students; mental health

Introduction

Emerging Adulthood and College

Young adults between the ages of 18-25 years report greater feelings of loneliness than other age groups [1]. Emerging adulthood is a developmental period marked by tremendous growth and opportunity but also by significant risk [2,3]. It is during this phase that the majority of major disease processes have their onset [4-7]; this is also a time when emerging adults develop habits for adaptive ways of managing stress and help-seeking behaviors [8]. The health habits formed during emerging adulthood have the potential to influence well-being throughout the lifespan [9], leading to the argument that this is a phase characterized by both risk and opportunity. The high level of plasticity during this phase can be attributed to various factors, including the lack of structure and increased self-directed activity [10,11], shifting roles and responsibilities [12], and relatively less institutional structure to support development [13]. A large proportion of these individuals pursue higher education. The college experience is a unique time in which this age range typically faces a myriad new experiences and opportunities for growth, development, and risk. As such, college students are an important subgroup of emerging adults that are worthy of investigation.

College students are a diverse cross-section of the population [14], coexisting in environments with less structure imposed by parents [15]. They are likely to be exposed to high doses of peer influence [16-18], less regulation in sleep and wake cycles [19,20], and less adult oversight of risk-taking behaviors [21]. Alongside developmental shifts, the college experience in combination with academic, career, and socioeconomic pressures may create a perfect storm for the onset of mental health problems. However, stability in academic, career, and financial livelihood during this time characterized by the exploration of one's self-identity—in the form of active college student life, social support, and financial aid—are considered protective factors that sustain well-being [22]. Studies suggest that college students with higher levels of stress often have poor mental health outcomes [23]. Furthermore, they are more likely to engage in maladaptive behaviors such as substance use [24] and less likely to engage in health-promoting and adaptive help-seeking behaviors like accessing mental health services [25], asking professors for help [26], exercising [27], healthy eating [28], and practicing sleep hygiene [29]. One way this has been addressed is through technology. Emerging adults are also the largest subgroup of the population to rely on technology

[30,31]. Over 30% of teens and young adults in the United States report using digital platforms as a method for health tracking and monitoring [32]. Together, these factors provide support for the real-world case and importance of stress management and mental health promotion among young adults.

The Year 2020 and the COVID-19 Pandemic

The year 2020 has largely been defined by the COVID-19 pandemic, which has disrupted the daily lives of people across the world. All of the aforementioned health-related concerns for college student mental health have been exacerbated by the COVID-19 pandemic [33,34]; universities across the states have converted into remote learning centers, sending college students home, disrupting their previous routines and daily rhythms, upending job prospects and the economy, and increasing uncertainty and anxiety across the United States [35-39]. Altogether, a tremendous shift has occurred from the excitement and opportunities of collegiate life to turbulence, trauma, and stress, relating to increased rates of mental health problems in this population [40-43]. The uncertainty with the ongoing pandemic renders many of the stabilizing factors in the college experience even harder to attain; therefore, there may be added stressors in maintaining social support networks and uncertainties with financial support for current college students, as well as worries about the quality of one's education and career prospects for students close to finishing their degree. In addition to the pandemic, sociopolitical unrest and a series of other potential stressors have arisen throughout 2020 (eg, the murder of George Floyd, the Black Lives Matter movement) that may have been of concern to students [44,45]. Because the state of the environment is intricately associated with daily life and well-being [46,47], it is important to contextualize the nuanced experiences of mental health with relevant immediate and larger, externally occurring events.

Multimodal Assessments and Personalized Approaches of Well-being

Examining student wellness in the context of their experiences may help to inform programming and intervention-based efforts to maintain student mental health, a key priority of campus wellness. From a biopsychosocial framework, wellness not only embodies internal functioning but also the social and external environment [48,49]. For each individual, the interaction of their own genetic and biological systems combines with their psychological and social systems [48,49]; the resultant psychopathology may differ. Thus, assessing numerous systems

and factors are necessary to fully capture an individual's mental health trajectory.

Precision medicine and idiographic approaches to health have gained traction to address the heterogeneity of symptomatology [50,51]. With advancements in technology, researchers have utilized and are continuing to refine mobile, passive sensing to best capture lived experiences and improve ecological validity at the individual level [52]. Furthermore, these technological and methodological advancements have permitted researchers to pursue individualized approaches in examining mental health and informing personalized treatment [51,53-55]. These methods allow for continuous measurement of health markers such as physiology, sleep, and physical activity, contributing to rich and large amounts of data at the individual level. Use of wearable and mobile devices to capture data in daily life, known as the internet of things (IoT) and wearable IoT (WIoT) can help process and synthesize multiple forms of assessments (ie, geographical, physiological, behavioral, subjective, and contextual data) that may aid a health professional in offering guidance to an individual seeking mental health treatment [56,57]. Therefore, understanding well-being and the contexts relevant to the trajectory of an individual's mental health may benefit from a whole-systems approach using WIoT and ecological momentary assessments. In combination with active sensing, multimodal personal chronicles, as we define it, incorporate these multiple forms of assessment (eg, physiology, subjective experiences, contexts) as a way to capture unique, personal contexts of an individual's daily lived experiences. Young adults are avid users of technology, being more dependent on their smartphones and using social media compared to adults ≥ 30 years of age [30,31]. As such, we note that young adults, and college students in particular, may be an ideal population to examine the feasibility of the multimodal personal chronicles system as a method for examining individualized approaches.

This Study

Given the novelty of the COVID-19 pandemic, we reasoned that the use of a case study was an important step in the scientific investigation of this phenomenon, one that would highlight the potential of the multimodal personal chronicles. The purpose of this case study is to provide a nuanced view of one individual's subjective experience and objective bodily responses during a historic and unprecedented event, while underscoring the potential of a research methodology for assessing the heterogeneity of well-being among college students. Case studies are effective methods for exploring contextual phenomenon within a single or small unit (eg, $N=1$ or small sample) in a given spatial and temporal context; they are often used when examining a newly emerging or poorly understood phenomenon [58,59]. These approaches may be advantageous in enhancing our understanding of the pandemic experience and its relevance to college student well-being beyond that of just self-report. Furthermore, case studies allow researchers to assess the feasibility of novel approaches on a small scale [59]. Such approaches are particularly useful at the outset of a historical phenomenon or the beginning of a scientific investigation, when the accumulation of knowledge is in its infancy and a more fine-grained analysis may provide important

information that could be lost if compressed into a study involving many individuals.

Our recent work [55] has demonstrated the feasibility of estimating stress levels by monitoring physiological signs of sympathetic nervous system activity (eg, heart rate, heart rate variability, respiration rate, galvanic skin response, etc). However, this method of objective monitoring only captures changes in internal states and not contextual factors—such as mental activity and social interactions—that are critical for the diagnosis, treatment, and prevention of mental health problems that may follow from the long-term effects of stress. There is a fundamental need to conduct studies that not only assess individual subjective and objective physiological reporting of health but also capture higher-level life events and contextual information to enable root cause analysis for treatment and prevention. Our purpose is to thus incorporate these multimodal assessments together in studies examining contexts. As prior studies have documented evidence of the negative associations of the pandemic among college students [33,34,40-43], capturing daily life during the pandemic, shelter-in-place orders, social distancing, and remote learning among college students may help to better understand the individualized increased risks for student mental health. In turn, this may ultimately further expand our understanding of the pandemic experience and its association with college student well-being. We present a case study of one participant from our larger pilot study to illustrate both (1) the potential and feasibility of a multimodal personal chronicles approach toward well-being and (2) the fine-grained experiences of one individual college student during the pandemic. Thus, our research questions are as follows:

1. How does a multimodal (passive and active) sensing approach via multimodal personal chronicles capture individualized and contextual experiences of well-being?
2. What is the experience of daily living for a college student in the context of the COVID-19 pandemic?

Methods

Procedure

We discuss the procedures for the larger study in which our case study is derived from. In January 2020, we began an Institutional Review Board (IRB)-approved (#2019-5153) investigation to pilot and examine the utility of this multimodal personal chronicles system in understanding changes in mental health symptoms (eg, depression, anxiety) and general psychological distress (eg, loneliness, negative emotion) over time among emerging adults in college. Participants completed a comprehensive psychological battery at study intake (eg, depression, anxiety, loneliness) and used devices within the multimodal system (ie, Oura ring, Samsung Gear Sport, ecological momentary assessments, and Personicle [60-62]) that assessed physiology, sleep, physical activity, step count, and affect. They then received another comprehensive battery 3 months later in April. Originally, we intended to end the pilot study after 3 months; however, concurrent to our original cohort's 3-month participation period, the COVID-19 pandemic worsened in the United States and elsewhere. This pandemic had a profound impact on the livelihood of individuals across

the world [63]; this prompted our investigative team to feel that it was important to capture the lived experiences of college students as it was occurring and to examine the association between the pandemic and student well-being from an intensive, longitudinal approach. At the onset of COVID-19 and the shutdown associated with this global pandemic, we had enrolled 13 college students into our immersive pilot; however, one of those participants was discontinued and two voluntarily withdrew early, resulting in a total sample size of 10. As a result, we extended our study, continuing participants' data collection and incorporating in-depth psychological assessments once at

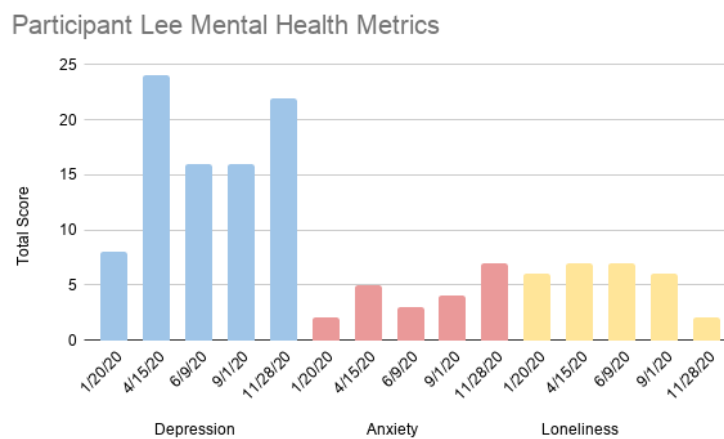
2 months (June) following the original study duration and then at 3 months (September) thereafter. Participants enrolled had the option to continue or end the study as planned.

Measures

Psychological Battery

The battery of psychological and well-being-related questionnaires were administered at baseline in person and during subsequent follow-up assessments (see Figure 1 for a timeline) as an online survey link.

Figure 1. Lee's depression, anxiety, and loneliness scores at 4 different time points in the study. Note: the sum of anxiety scores is reported rather than the means to show score distribution over time.



Depression

Participants completed the Beck Depression Index II (BDI-II) [64] to assess depressive symptoms. This is a 21-item, well-validated measure rated on a 4-point scale (0-3; $\alpha=.90-.91$) [65,66] that allows individuals to indicate to what degree they are experiencing a feeling, for example, "I do not feel sad" (0) to "I am so sad or unhappy that I can't stand it" (3) with the exception of 2 items using a 6-point scale to reflect different degrees of behavioral changes. Scores are calculated as a total, which indicates the range in depression from mild (14-19) to severe (29-63) [65]. If participants scored above 24 and/or endorsed suicidality at intake or any follow-up assessments, the principal investigator (author JLB, a licensed psychologist) reached out to these individuals to determine symptom severity and if the participant should continue with or be withdrawn from the study. These individuals were then provided additional mental health resources.

Anxiety

The Brief Symptom Inventory (BSI) [67], a 6-item anxiety subscale, was used to measure anxiety. The BSI is a commonly used measure that assesses a wide array of psychological symptom dimensions, including depression, somatization, and anxiety on a 5-point Likert-type scale (0=not at all, 4=extremely; $\alpha=.81-.86$) [68,69]. Anxiety scores are typically calculated as a mean of all items (eg, "During the past 7 days how often were you distressed by nervousness or shakiness inside?"); for our purposes, we used the sum of all items to observe the distribution of participants' anxiety score over a period of time

alongside other psychological constructs of interest. A higher score indicates greater anxiety. Normative ratings among nonpsychiatric patients average around 0.35 (SD 0.45) [69].

Loneliness

Participants were administered the UCLA (University of California, Los Angeles) 3-item Loneliness Scale, a frequently used measure of perceived loneliness (eg, "How often do you feel you are left out?"; $\alpha=.72$) [70]. The UCLA Loneliness Scale has been used widely in the literature to examine how loneliness relates to psychological and physiological outcomes (eg, mood, depression) [70]. Items are totaled and rated on a 3-point Likert-type scale (1=hardly ever, 3=often), where higher scores reflect greater perceived loneliness.

Multimodal Personal Chronicles

As part of the intensive active and passive data sensing, over the course of the study, we tracked participants' emotional states, physiological patterns, and behavioral habits through WIoT devices and ecological momentary assessments.

Physiology, Sleep, and Behavioral Patterns

Participants were instructed to wear 2 devices (Oura ring and Samsung Gear Sport smartwatch) as well as download the corresponding Oura and Samsung Android mobile apps to measure and store their physiology, sleep, and behavioral patterns. The Oura ring [71] measures a wide array of personal variables across numerous categories, including heart rate, heart rate variability, respiration rate, skin temperature, sleep, and activity level (majority of these are monitored only overnight). For our investigation, we assessed sleep quality using the Oura

ring's personalized sleep score that uses a combination of individual body metrics (ie, height, weight, age, and gender) with a validated system that detects time spent in different stages of sleep [60,61]. The Samsung smartwatch tracks several user activities and physiological variables 24/7 (for more details on how these devices were used for our study, see Yunusova et al [72]). We were able to use the watch to assess participants' heart rate variability (root mean square of successive differences between normal heartbeats [RMSSD]) and beats per minute (BPM) during diurnal periods (RMSSD and day BPM), and during the entire 24-hour period (RMSSD and resting BPM).

Lastly, the Personicle app [62] was downloaded onto each participant's phone to observe behavioral trends. Personicle tracks users' changes in location through the Google GPS application programming interface. For the purpose of this study, the participants' step count was assessed through this app.

Ecological Momentary Assessments of Affect and Context

Participants completed daily assessments of positive and negative affect using the Positive and Negative Affect Schedule (PANAS), a widely used scale that has high internal consistency, validity, and reliability (for positive affect, $\alpha=.85$; for negative affect, $\alpha=.91$) [73]. This scale is highly correlated with measures of psychopathology and distress such as depression and anxiety [74,75]. In clinical samples, the PANAS scale can differentiate between anxiety and depression among individuals, with positive affect being highly correlated with depression and negative affect being highly correlated with both depression and anxiety [76,77]. Example items of positive affect items include "excited" and "inspired" while negative affect include items such as "afraid" and "nervous," all rated on a sliding scale from "very slightly" (0) to "extremely" (100), where higher scores reflect higher positive or negative affect. This was sent as a survey prompt to participants' mobile phone using a separately downloaded app that was designed specifically for the survey to be taken on, and was completed once a day between the hours of 8 pm and 2 am. Additionally, participants completed a weekly "feel-in" every Sunday afternoon between the hours of 8 am and 3 pm, where they were instructed to answer an open-ended question about the quality of their week (ie, "Please write about your high points and low points this week. Please try to be as detailed as possible").

Data Analytic Plan

We followed participants from January through September, in which each individual completed a total of 4 battery assessments of psychological well-being (across roughly 3-month time frames—January, April, June, and September) in addition to intensive longitudinal assessments of heart rate, heart rate variability measures, respiration rate, skin temperature, sleep quality, step count, and affect throughout the study period. We focused on heart rate (BPM), heart rate variability (RMSSD), sleep quality, step count, and affect for this case study.

For our data analytic plan, we will use a descriptive approach to illustrate the data over the course of 9 months, examining any changes in participants' depression, anxiety, and loneliness ratings at each of the 4 assessments as well as their BPM, RMSSD, sleep quality, step count, and positive and negative affect aggregated across the 2-week period after completion of the psychological assessment. Then, we will run correlations to examine associations between the aggregated physiological, sleep, behavioral, and affect measures at each time point. Doing so offers a broader perspective of how these multimodal assessments relate to one another and how they may differ over the course of the pandemic.

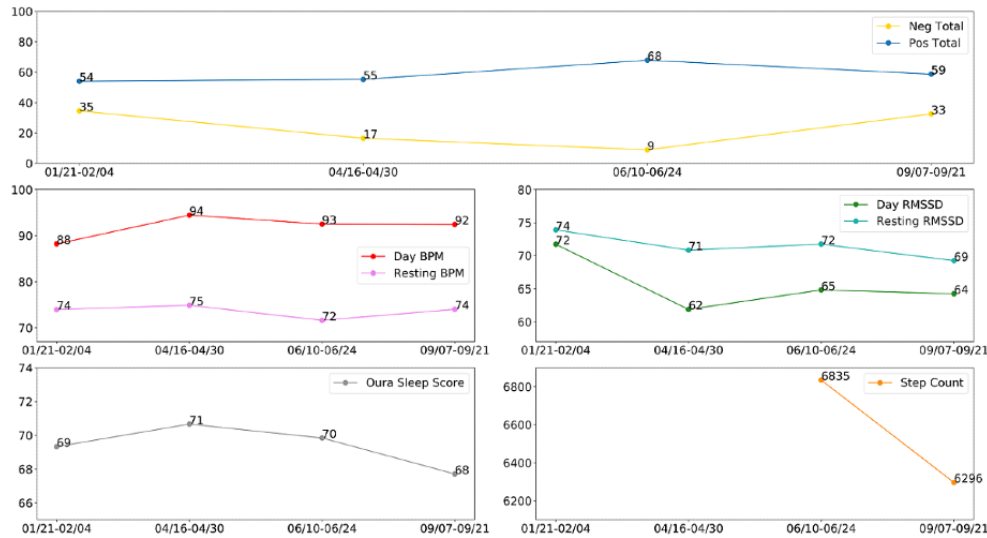
We will then narrow down to explore 2-week periods of data to better understand contextual factors potentially relevant to the fluctuations in their emotional and physiological patterns over the 2-week period (ie, data points for each day for 14 days). Focusing on the occurrences at 4 different, 2-week time points may further help us observe relationships across multimodal measures at the daily level rather than aggregates across months. We note that the adjustments made to the study in response to the pandemic were approved by the IRB in April, thus limiting the available self-reported contextual data we can examine. Therefore, while we present the 2-week data for each of the 4 time points, we only describe in detail the contextual data for the periods of June and September. Finally, we will offer clinical recommendations based on the data derived from the multimodal personal chronicles as potential suggestions for the single subject this case study focuses on.

Results

For the purposes of this case study, we focused on the results of one individual. Like many other participants in our study, this individual exhibited a decline in mental health as the pandemic-enforced period of social isolation extended. In order to preserve the confidentiality of this participant, we have modified details regarding this participant's case.

"Lee," as we shall refer to this individual, is an Asian-American male student in the middle of his sophomore year of college. We report Lee's pandemic narrative in chronological order since the start of his participation in the study in January through to September. Lee began the study with mental health symptoms that were low and not within the clinically significant range. For instance, his depression score at intake was 8, below the cut-off for mild depression (see Figure 1 for levels of depression, anxiety, and loneliness across each assessment time point). Figure 2 visually displays his 2-week aggregate score for all multimodal measures at each time point. Over the course of the pandemic, we observed that Lee reported relatively low levels of negative affect compared to his levels of positive affect, despite his higher depressive scores in April through September.

Figure 2. Lee's multimodal assessments at the 4 different assessment time points in the study. Note: we did not collect data on step count through the mobile phone app until May, thus limiting our interpretation and examination of step count during the January and April periods. Neg: negative, pos: positive, BPM: beats per minute, RMSSD: root mean square of successive differences between normal heartbeats.



By the second time point in April, compared to his baseline assessment in January, his depression score had increased to 24—well within the moderate depression range. His anxiety and perceived feelings of loneliness also slightly increased (Figure 1). Because Lee had shown such an increase in his depression, roughly 3 times his initial levels in January, author JLB conducted a phone screening to assess his mental health. When JLB spoke to Lee, she learned that since the onset of COVID-19 toward the end of March, Lee had experienced more stress—a family member of his had experienced a significant stressor due to COVID-19, which in turn also impacted him. Lee's mother had been living abroad in [country] as a visiting professor and was unable to return home before travel restrictions were imposed, resulting in her being unable to leave [country] for a period of months. This created a great deal of anxiety for everyone in Lee's family—having Lee's mother be separated from the family, apart from everyone during a global crisis, with no certainty regarding when she would be able to return, was highly distressing. During this time, Lee moved back home to live with his family, disrupting the social support and routine he had established in his college environment. Lee reported that his mood was low but that he was not feeling suicidal. JLB provided him with referrals and informed him that he could reach out to her at any time should he need additional support. Consistent with the study protocol, the research team and JLB continued to track Lee's mood, sleep, physiology, and physical activity throughout the study.

Since his assessment in April, Lee's depression remained in the mildly severe range in June and September. His reported anxiety had lowered since, but we noted that his loneliness was consistent between the periods of April and September. From his wearable sensor and ecological momentary assessment data, we noted that there were fairly minimal fluctuations in his physiology and sleep. Although his BPM was highest in April,

it was not largely different from his BPM at other time points. His sleep was lower in the January and September periods, which reflect the time of the start of the winter and fall quarter, possibly relating to these differences than the April (spring quarter) and June (close to summer) periods. We report correlations of the measures at each time point in Multimedia Appendix 1 (Tables S1-S4). Based on the correlational data, associations varied across time periods. For example, Lee's positive and negative affect were positively correlated in April (eg, high positive and high negative affect; $r=0.53$, $P=.04$) whereas they were negatively correlated in September (eg, high positive and low negative affect; $r=-0.57$, $P=.03$). BPM and RMSSD were relatively correlated throughout each time point (Multimedia Appendix 1, Tables S1-S4). Only in the month of January was sleep negatively associated with negative affect ($r=-0.58$, $P=.03$) and day BPM ($r=-0.54$, $P=.04$), and then positively associated with resting RMSSD ($r=0.54$, $P=.04$).

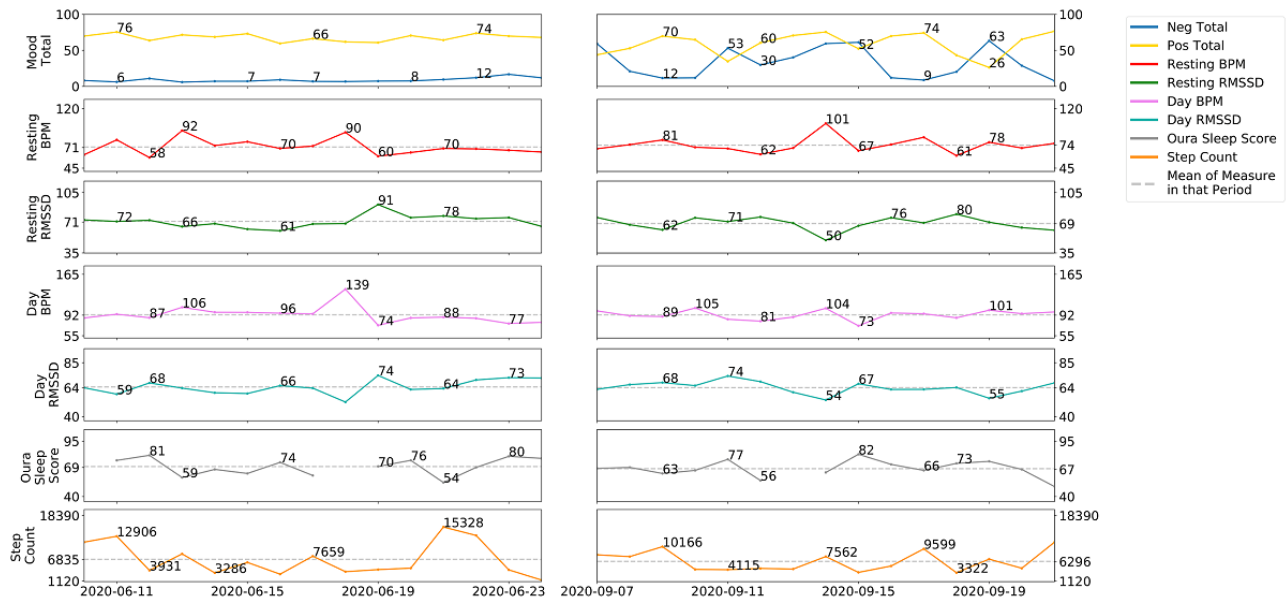
To further illustrate the potential of a personalized approach in helping to inform our understanding of college student mental health, we focused on the June and September 2-week time periods. Table 1 displays contextual data from Lee's weekly responses, and Figure 3 highlights the trends in the multimodal assessments. More specifically, the figure displays the daily monitoring of Lee's positive and negative affect, physiology, sleep, and physical activity. Based on Figure 3, in June, Lee experienced greater distinctions in his positive and negative affect, where he reported consistently higher levels of positive affect and consistently lower levels of negative affect. When observing Lee's data during this time, his subjective weekly reporting of his high and low points (Table 1) suggested that he was socially engaged with peers (ie, hiking with friends, participating in the Black Lives Matter social movement with peers) and had a sense of purpose (eg, making social change) as well as accomplishment (eg, finishing a year of school).

Table 1. Lee's weekly "feel-in" responses to weekly highs and lows across the June and September time frames.

Time frame and time stamp	Response	LIWC ^a analysis
Time 1 (June 10-24)		Negative emotion word frequency=2.63; positive emotion word frequency=5.26
Sun, Jun 7, 2020; 09:56:00	"Some high points this week are leading a talk on the Black Lives Matter movement with my [group of people] and [group of people] which was really successful. Another high point is completing three of my courses. A low point is having to listen to racist comments and people that don't support the Black Lives Matter movement. Also, I still have to study for two more finals and a term paper"	
Sun, Jun 14, 2020; 11:53:57	"High points include being done with [a year of school]. I went running with friends and had so much fun. I got to see the night sky and great views in [location redacted]. Low points include listening to my friends' problems and trying to help them both through it. Another low point is hanging out with a friend and another person joined us who I don't like, so I didn't really enjoy myself anymore."	
Sun, Jun 21, 2020; 12:45:24	"High points in my week are hiking and playing [game] with friends. I did not have any low points. This has been my best week in months"	
Sun, Jun 28, 2020; 14:55:40	"Some high points include starting a [program] that will help me work on a project for the future as well as going to the beach with friends. Some low points include having to help friends through a difficult time, the large amount of work I have to take care of, and beginning to delay responsibilities and need to fix my schedule."	
Time 2 (September 7-21)		Negative emotion word frequency=2.84; positive emotion word frequency=5.88
Sun, Aug 30, 2020; 12:44:13	"The biggest high point of my week was being able to spend some more time with the [person] I like. We were able to eat dinner, talk, and walk on the boardwalk. It felt like nothing else really mattered and I let go of my responsibilities and things that I had to do. I also got to lead a group meant to hear [opinions] about career resources for the project that I am involved in. It felt good to hear from [group of people] and I want to be able to help them with their concerns. Low points this week was having to revise my [project] again because it didn't meet certain expectations. It's irritating having to change it multiple times even though I already address the revisions that my [supervisor] requires."	
Sun, Sep 6, 2020; 14:16:27	"The high points in my week are my hikes with people. It feels good to hike and being active again since I cannot play games. I also was able to complete a project for a [group of people] that I hope gets approved by the [group of people]. These past few days have really sucked because I wasn't able to work on my [project] and being stuck. It's been frustrating and I have a difficult time when I'm not able to do work and I've been having [sic] feelings of inadequacy. It's been hard but one of my best friends has been really helpful because they always reassure me and give me words of affirmation."	
Sun, Sep 13, 2020; 11:55:19	"This past week has had such ups and downs with so many emotions. I was sad to discover that my cousin, who was assaulted last year at [name of a different college campus], has been going through a difficult time and [they've] been experiencing triggers. I got angry with myself for not being able to physically support [them] last year when she needed it most. I was able to talk on the phone and text them now which was helpful but I wanted to be there in person to show that I'm there for [them]. Another low point is studying for an upcoming [a test]. The grade isn't that important to me but I want to do good. A couple high points is having a virtual get together. It was great to hear their voices and see their smiles. I also got to spend a day with my friend and had a lot of fun. I also finished another draft of my [project] and that was giving me headaches all week."	
Sun, Sep 20, 2020; 16:10:00	"This week had it's [sic] highs and lows. A high is finishing the [test] which felt great. Another high is being accepted into [a program] but the low part of that is finalizing edits which has been really hard to get a start on. Another low is finding out my friends [relative] passed away in [country] and helping [them] through it. It's been hard for [them], especially with [their] other [relative] passing away recently."	

^aLIWC: Linguistic Inquiry Word Count System.

Figure 3. Raw numbers of all measures during the 2-week period on June 10-24 and September 7-21. Note: negative (neg) total, positive (pos) total, and Oura sleep score range between 0 and 100. Mood total: blue line signifies negative emotion and yellow line signifies positive emotion; resting beats per minute (BPM): sampled 15 minutes every 2 hours, including sleep; resting RMSSD (root mean square of successive differences between normal heartbeats): index of heart rate variability, sampled 15 minutes every 2 hours, including sleep; day (resting) BPM: sampled 15 minutes every 2 hours, not including sleep; day RMSSD: sampled 15 minutes every 2 hours, not including sleep; Oura sleep score: summary score of overall sleep quality, a weighted sum of sleep contributors, which combines sleep latency, onset, restfulness, rapid eye movement, deep sleep, timing, and efficiency; step counts: generated from the watch, total number of steps taken per day.



In contrast, during September, Lee exhibited more erratic mood states, which we interpret to potentially be an indication of low well-being. Lee's negative and positive affect show greater variations from day to day. Compared to when Lee's positive affect was more distinct from his negative affect in June, his negative affect was nearly as high as his positive affect in September. Further, in September, he experienced days in which his negative affect was higher than his positive affect. Between the two time periods, there were relatively similar fluctuations in his physical activity, heart rate levels, heart rate variability (as indexed by the RMSSD score), sleep, and step count (Figure 3). He also had fewer social contact, as reflected in his weekly self-reported high and low points (Table 1). As an additional exploratory analysis, we analyzed the narrative responses of Lee during his weekly "feel-ins" using the Linguistic Inquiry Word Count System [78]; our findings revealed that Lee used more negative ($t_1=26.05$, $P=.02$) than positive words ($t_1=17.97$, $P=.04$) in September than in June, further reflecting greater variation in his affect.

Discussion

Principal Results

The use of a case study enabled us to closely examine the experience and daily life of Lee across distant time points within the past year. Lee's depressive symptoms increased sharply from January to April, a time coinciding with the onset of the COVID-19 pandemic and supporting prior evidence for increased depressive symptoms and distress among young adults during this period [38,39,47,48]. When examining his data more closely at a 2-week period with his contextual responses, Lee reported greater negative affect and use of negative emotion words in the month of September than in June. As a rising

college senior, Lee may have felt the winding down of the previous academic year in the month of June while feeling the build-up of an eventful start to the academic year in the month of September. Courses were transitioned remotely at the end of March due to COVID-19; while this may have resulted in increased stress and uncertainty in April, students may have felt some relief at the anticipation of summer, which heralds a reduction in academic stressors. With regard to September, students may have felt a rise in stress and uncertainty with what the upcoming school year will look like. To inquire about student concerns, one department at the university had conducted a series of undergraduate focus groups in the month of September. Students overall had expressed unease regarding the prolonged period of remote learning and wondered whether instructors would expect students to have adjusted and thus no longer offer flexibility in their courses in the upcoming quarter. Much like the students in the focus group, Lee may have felt increased stress surrounding the uncertainty of the upcoming school year, which could be associated with increases in his experiences of negative affect.

We are able to understand Lee's unique pandemic experiences from this longitudinal assessment. Although his depression lowered from severe to mildly severe in the following months from April, Lee's loneliness score remained consistent between the months of April through September, possibly relating to the continued social distancing and having to rely on alternative methods of keeping in contact with friends (eg, online chatting and hangouts). We also observed relatively higher levels of positive affect and lower levels of negative affect throughout Lee's participation in the study. His negative affect was lowest in June, which is reflected in the 2-week period we examined more closely. While he may have experienced external events that affected his depressive symptoms (eg, his mother's

situation), there were potentially supportive factors (eg, social networks, meaningful work) that sustained his positive affect while maintaining lower negative affect during the pandemic. Thus, the meaning and significance of the multimodal data becomes most clear when aligned with the contextual data.

With regard to the pandemic, the approach allowed us to understand fluctuations in health and well-being that are linked to the unique experiences an individual may encounter during the pandemic. For example, the situation with Lee's mother was a significant stressor for him, and this was only possible to document with the multiple assessments of depression and having this data accessible to a clinical psychologist to follow up with him regarding his well-being. Alternatively, we also recognize that there were some fluctuations that were relatively low (eg, physiology across the 4 time periods as well as in the 2-week period of June and September). This may also point to important coping and resiliency factors that allow individuals like Lee to navigate this stressful pandemic. With any kind of stressful event, Lee may be using his current resources and available coping strategies to self-regulate during the pandemic.

The case study highlights preliminary data from our larger ongoing study, which in turn enables researchers and clinicians to better understand the context and unique experiences of individuals and how these factors relate to well-being. The utility of multimodal personal chronicles for mental health may be clearest when it is used in conjunction with the collaborative care of a clinician. From a health navigation framework [79], health professionals can access this data, synthesize it with their current knowledge of the individual, and offer informed guidance and treatment recommendations [80]. Researchers may also consider using advanced modeling or network analyses [81,82] to examine this intensive, longitudinal approach as a way to understand the interconnection of biological, psychological, and social factors relevant to well-being.

Recommendations

Access to such rich amounts of personalized data can help clinicians or therapists offer evidence-based treatments or recommendations. If the provider in the loop were to offer recommendations to Lee based on his pattern of behaviors, it would be to increase and maintain a moderate to ideal level of physical and interpersonal activity with his support networks. As noted, based on his weekly open-ended responses, many of his high points included spending time with friends or enjoying time outdoors doing either physical activity or enjoying nature with others. Interpersonal support and interactions were particularly relevant, and thus seeking strategies to ensure consistent social connectedness with others may be helpful for Lee in sustaining well-being. Given that public health recommendations of social distancing preclude in-person social interactions, Lee may wish to savor memories of times when he has enjoyed spending time with other people (eg, relational savoring [83]), which has been shown to increase positive emotion and reduce depression. Similarly, attending to and prolonging the positive emotions associated with one's romantic partner has shown to have a positive impact on those in long-distance relationships [84], suggesting that this technique may also assist people during periods of social distancing.

Furthermore, Lee noted being involved in purposeful work facilitating discussions relating to sociopolitical movements and supporting wellness efforts on campus. Creating meaning and finding a sense of purpose are linked to subjective well-being [85,86], and so continuing to find meaning in his work or savoring these feelings that coincide with the work he finds meaningful may also be particularly helpful in sustaining mental health. Additional recommendations could potentially include engaging in positive self-reflections, such as savoring interpersonal or achievement moments throughout the week. These reflections may also be helpful in the form of gratitude journaling or writing down 3 things they were grateful for that day. Lee's low points were often linked to academic and interpersonal stressors; focusing on his small achievements can help Lee focus on his competency and ability to be resilient despite in-the-moment stressors. While each week may include life challenges, Lee can recall his own strengths and resiliency to stay motivated and overcome various stressors.

Limitations

Limitations to our study include the fact that, as with many case studies, it is primarily exploratory in our understanding of a phenomenon in a specific space and time. As we are considering the feasibility of multimodal personal chronicles, we restrain from making causal claims regarding the experiences of Lee. Additionally, self-reported emotion data were only captured once a day toward the evening, which allows for recall bias. Recent studies are now suggesting that the PANAS scale is in need of improvement [87] while also pointing out that positive and negative affect are not completely independent dimensions [88]. Furthermore, due to a data collection error, we did not have weekly self-reported "feel-in" data available for Lee until approximately 1 month after the initial shelter-in-place order. Thus, we were only able to examine the contextual data alongside the active and passive sensing data for June onwards. Because we selected a 2-week period for the time points, there were limited data to examine when conducting statistical analyses. Thus, the case study may benefit from including more observations and a larger study including more participants. Once the full study is complete, we intend to use analytical techniques (eg, structural equation modeling), which will permit us to use a larger number of observations in our data to understand our participants' mental health trajectories.

Comparison With Prior Work and Future Directions

Emerging adults navigate a number of stressors during this developmental period; experiencing all of this while facing a global pandemic may exacerbate psychopathology or detriments to psychological well-being among college students. For example, students may have to juggle negative impacts of the pandemic on their family while also adjusting to online and remote learning. Previous studies examining student well-being during the pandemic have largely conducted single time-point assessments [38,45,46,48] or are limited to self-report assessments [39,47]; our study utilizes an intensive longitudinal case study to showcase how the year throughout the pandemic has not only affected student mental health but also potentially through indirect means like student livelihood and daily experiences. The study contributes to our understanding of how

college students may cope with a global stressor, in that some may navigate stress similarly to how they cope with other life stressors. Indeed, much of the current literature on the pandemic has focused on adverse health outcomes, and so future investigations may consider coping and resiliency among individuals through this period.

The feasibility of multimodal personal chronicles to assess mental health may help to inform what types of interventions may be best for individuals. Future studies may consider assessing the feasibility of the multimodal personal chronicles in other populations (eg, clinical samples, across cultures) and incorporating interventions or guidance in the process of data collection as a way to examine how self-monitoring of health alongside guidance may then have trickle-down effects on improving health. In addition, this approach offers the potential for a system in which individuals can tweak their own health behaviors and see its associations with changes in their health trajectory [89,90]. Health professionals may also consider how

different types of strategies may be beneficial at different times, depending on the context and experiences of the individual [84]. Indeed, a personalized approach grants flexibility in treatment and interventions in ways that address the dynamic nature of health and well-being.

Conclusion

From our ongoing study, we present results from a case study examining 1 participant and how the trajectory of his well-being may be assessed through a holistic multimodal personal chronicles system that captures personalized experiences. Case studies have the benefit of presenting in-depth information regarding a single individual, enabling a fine-grained analysis of the circumstances, characteristics, and experiences for that person at a specific point in time. Future studies may consider the multimodal personal chronicles approach and how it may inform treatment and intervention to mitigate psychopathology and aid in maintaining well-being.

Acknowledgments

Preparation for this case study occurred in collaboration with the School of Nursing, the Department of Computer Sciences, and the Department of Psychological Science at the University of California, Irvine.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Correlation tables of the 2-week assessments for January, April, June, and September.

[[DOCX File, 21 KB - formative_v5i5e26186_app1.docx](#)]

References

- Barreto M, Victor C, Hammond C, Eccles A, Richins MT, Qualter P. Loneliness around the world: Age, gender, and cultural differences in loneliness. *Pers Individ Dif* 2021 Feb 01;169:110066 [FREE Full text] [doi: [10.1016/j.paid.2020.110066](https://doi.org/10.1016/j.paid.2020.110066)] [Medline: [33536694](https://pubmed.ncbi.nlm.nih.gov/33536694/)]
- Hoyt LT, Chase-Lansdale PL, McDade TW, Adam EK. Positive youth, healthy adults: does positive well-being in adolescence predict better perceived health and fewer risky health behaviors in young adulthood? *J Adolesc Health* 2012 Jan;50(1):66-73 [FREE Full text] [doi: [10.1016/j.jadohealth.2011.05.002](https://doi.org/10.1016/j.jadohealth.2011.05.002)] [Medline: [22188836](https://pubmed.ncbi.nlm.nih.gov/22188836/)]
- Steger MF, Oishi S, Kashdan TB. Meaning in life across the life span: Levels and correlates of meaning in life from emerging adulthood to older adulthood. *The Journal of Positive Psychology* 2009 Jan;4(1):43-52. [doi: [10.1080/17439760802303127](https://doi.org/10.1080/17439760802303127)]
- Auerbach RP, Mortier P, Bruffaerts R, Alonso J, Benjet C, Cuijpers P, WHO WMH-ICS Collaborators. WHO World Mental Health Surveys International College Student Project: Prevalence and distribution of mental disorders. *J Abnorm Psychol* 2018 Oct;127(7):623-638 [FREE Full text] [doi: [10.1037/abn0000362](https://doi.org/10.1037/abn0000362)] [Medline: [30211576](https://pubmed.ncbi.nlm.nih.gov/30211576/)]
- de Girolamo G, Dagani J, Purcell R, Cocchi A, McGorry PD. Age of onset of mental disorders and use of mental health services: needs, opportunities and obstacles. *Epidemiol Psychiatr Sci* 2012 Mar 13;21(1):47-57. [doi: [10.1017/s2045796011000746](https://doi.org/10.1017/s2045796011000746)] [Medline: [22670412](https://pubmed.ncbi.nlm.nih.gov/22670412/)]
- Kalaydjian A, Swendsen J, Chiu W, Dierker L, Degenhardt L, Glantz M, et al. Sociodemographic predictors of transitions across stages of alcohol use, disorders, and remission in the National Comorbidity Survey Replication. *Compr Psychiatry* 2009;50(4):299-306 [FREE Full text] [doi: [10.1016/j.comppsy.2008.09.012](https://doi.org/10.1016/j.comppsy.2008.09.012)] [Medline: [19486727](https://pubmed.ncbi.nlm.nih.gov/19486727/)]
- Twenge JM, Cooper AB, Joiner TE, Duffy ME, Binau SG. Age, period, and cohort trends in mood disorder indicators and suicide-related outcomes in a nationally representative dataset, 2005-2017. *J Abnorm Psychol* 2019 Apr;128(3):185-199. [doi: [10.1037/abn0000410](https://doi.org/10.1037/abn0000410)] [Medline: [30869927](https://pubmed.ncbi.nlm.nih.gov/30869927/)]
- Lawrence EM, Mollborn S, Hummer RA. Health lifestyles across the transition to adulthood: Implications for health. *Soc Sci Med* 2017 Nov;193:23-32 [FREE Full text] [doi: [10.1016/j.socscimed.2017.09.041](https://doi.org/10.1016/j.socscimed.2017.09.041)] [Medline: [28992538](https://pubmed.ncbi.nlm.nih.gov/28992538/)]
- Howard AL, Galambos NL, Krahn HJ. Paths to success in young adulthood from mental health and life transitions in emerging adulthood. *International Journal of Behavioral Development* 2010 Jul 05;34(6):538-546. [doi: [10.1177/0165025410365803](https://doi.org/10.1177/0165025410365803)]

10. Arnett JJ. Emerging adults in America: Coming of age in the 21st century. In: Arnett JJ, Tanner JL, editors. *Emerging Adulthood: Understanding the New Way of Coming of Age*. Washington, DC: American Psychological Association; 2006:3-19.
11. Schwartz SJ, Côté JE, Arnett JJ. Identity and Agency in Emerging Adulthood. *Youth & Society* 2016 Aug 19;37(2):201-229. [doi: [10.1177/0044118x05275965](https://doi.org/10.1177/0044118x05275965)]
12. Guan SA, Fuligni AJ. Changes in Parent, Sibling, and Peer Support During the Transition to Young Adulthood. *J Res Adolesc* 2015 Jan 08;26(2):286-299. [doi: [10.1111/jora.12191](https://doi.org/10.1111/jora.12191)]
13. Arnett JJ. Socialization in emerging adulthood. In: Grusec JE, Hastings PD, editors. *Handbook of Socialization: Theory and Research*. New York, NY: The Guildford Press; 2007:208-239.
14. Mayhew MJ, Rockenbach AN, Bowman NA, Seifert TA, Wolniak GC. *How College Affects Students: 21st Century Evidence that Higher Education Works*. San Francisco, CA: John Wiley & Sons; 2016.
15. Mulder CH, Clark WAV. Leaving Home for College and Gaining Independence. *Environ Plan A* 2016 Dec 01;34(6):981-999. [doi: [10.1068/a34149](https://doi.org/10.1068/a34149)]
16. Mason MJ, Zaharakis N, Benotsch EG. Social networks, substance use, and mental health in college students. *J Am Coll Health* 2014 Sep 16;62(7):470-477. [doi: [10.1080/07448481.2014.923428](https://doi.org/10.1080/07448481.2014.923428)] [Medline: [24848433](https://pubmed.ncbi.nlm.nih.gov/24848433/)]
17. Patterson MS, Goodson P. Social network analysis for assessing college-aged adults' health: A systematic review. *J Am Coll Health* 2019 Jan 31;67(1):59-67. [doi: [10.1080/07448481.2018.1462820](https://doi.org/10.1080/07448481.2018.1462820)] [Medline: [29652600](https://pubmed.ncbi.nlm.nih.gov/29652600/)]
18. Rinker DV, Krieger H, Neighbors C. Social Network Factors and Addictive Behaviors among College Students. *Curr Addict Rep* 2016 Dec 14;3(4):356-367 [FREE Full text] [doi: [10.1007/s40429-016-0126-7](https://doi.org/10.1007/s40429-016-0126-7)] [Medline: [28580226](https://pubmed.ncbi.nlm.nih.gov/28580226/)]
19. Becker SP, Jarrett MA, Luebke AM, Garner AA, Burns GL, Kofler MJ. Sleep in a large, multi-university sample of college students: sleep problem prevalence, sex differences, and mental health correlates. *Sleep Health* 2018 Apr;4(2):174-181 [FREE Full text] [doi: [10.1016/j.sleh.2018.01.001](https://doi.org/10.1016/j.sleh.2018.01.001)] [Medline: [29555131](https://pubmed.ncbi.nlm.nih.gov/29555131/)]
20. Hershner S, Chervin R. Causes and consequences of sleepiness among college students. *NSS* 2014 Jun;73. [doi: [10.2147/nss.s62907](https://doi.org/10.2147/nss.s62907)]
21. Stone AL, Becker LG, Huber AM, Catalano RF. Review of risk and protective factors of substance use and problem use in emerging adulthood. *Addict Behav* 2012 Jul;37(7):747-775. [doi: [10.1016/j.addbeh.2012.02.014](https://doi.org/10.1016/j.addbeh.2012.02.014)] [Medline: [22445418](https://pubmed.ncbi.nlm.nih.gov/22445418/)]
22. Grant-Vallone E, Reid K, Umali C, Pohlert E. An Analysis of the Effects of Self-Esteem, Social Support, and Participation in Student Support Services on Students' Adjustment and Commitment to College. *Journal of College Student Retention: Research, Theory & Practice* 2016 Jul 20;5(3):255-274. [doi: [10.2190/c0t7-yx50-f71v-00cw](https://doi.org/10.2190/c0t7-yx50-f71v-00cw)]
23. Beiter R, Nash R, McCrady M, Rhoades D, Linscomb M, Clarahan M, et al. The prevalence and correlates of depression, anxiety, and stress in a sample of college students. *J Affect Disord* 2015 Mar 01;173:90-96. [doi: [10.1016/j.jad.2014.10.054](https://doi.org/10.1016/j.jad.2014.10.054)] [Medline: [25462401](https://pubmed.ncbi.nlm.nih.gov/25462401/)]
24. Wei M, Russell DW, Zakalik RA. Adult Attachment, Social Self-Efficacy, Self-Disclosure, Loneliness, and Subsequent Depression for Freshman College Students: A Longitudinal Study. *Journal of Counseling Psychology* 2005 Oct;52(4):602-614. [doi: [10.1037/0022-0167.52.4.602](https://doi.org/10.1037/0022-0167.52.4.602)]
25. Eisenberg D, Golberstein E, Gollust SE. Help-seeking and access to mental health care in a university student population. *Med Care* 2007 Jul;45(7):594-601. [doi: [10.1097/MLR.0b013e31803bb4c1](https://doi.org/10.1097/MLR.0b013e31803bb4c1)] [Medline: [17571007](https://pubmed.ncbi.nlm.nih.gov/17571007/)]
26. Eisenberg D, Downs MF, Golberstein E, Zivin K. Stigma and help seeking for mental health among college students. *Med Care Res Rev* 2009 Oct 19;66(5):522-541. [doi: [10.1177/1077558709335173](https://doi.org/10.1177/1077558709335173)] [Medline: [19454625](https://pubmed.ncbi.nlm.nih.gov/19454625/)]
27. Keating XD, Guan J, Piñero JC, Bridges DM. A meta-analysis of college students' physical activity behaviors. *J Am Coll Health* 2005 Sep;54(2):116-125. [doi: [10.3200/JACH.54.2.116-126](https://doi.org/10.3200/JACH.54.2.116-126)] [Medline: [16255324](https://pubmed.ncbi.nlm.nih.gov/16255324/)]
28. Munt AE, Partridge SR, Allman-Farinelli M. The barriers and enablers of healthy eating among young adults: a missing piece of the obesity puzzle: A scoping review. *Obes Rev* 2017 Jan 20;18(1):1-17. [doi: [10.1111/obr.12472](https://doi.org/10.1111/obr.12472)] [Medline: [27764897](https://pubmed.ncbi.nlm.nih.gov/27764897/)]
29. Brown FC, Buboltz WC, Soper B. Relationship of sleep hygiene awareness, sleep hygiene practices, and sleep quality in university students. *Behav Med* 2002;28(1):33-38. [doi: [10.1080/08964280209596396](https://doi.org/10.1080/08964280209596396)] [Medline: [12244643](https://pubmed.ncbi.nlm.nih.gov/12244643/)]
30. Demographics of mobile device ownership and adoption in the United States. Pew Research Center. 2021 Apr 7. URL: <https://www.pewresearch.org/internet/fact-sheet/mobile/> [accessed 2021-04-07]
31. Social media fact sheet. Pew Research Center. 2021 Apr 7. URL: <https://www.pewresearch.org/internet/fact-sheet/social-media/> [accessed 2021-04-07]
32. Rideout V, Fox S, Well Being Trust. Digital health practices, social media use, and mental well-being among teens and young adults in the U.S. Articles, Abstracts, and Reports. 2018. URL: <https://digitalcommons.psjhealth.org/publications/1093> [accessed 2021-04-26]
33. Holman EA, Thompson RR, Garfin DR, Silver RC. The unfolding COVID-19 pandemic: A probability-based, nationally representative study of mental health in the United States. *Sci Adv* 2020 Oct 18;6(42):eabd5390 [FREE Full text] [doi: [10.1126/sciadv.abd5390](https://doi.org/10.1126/sciadv.abd5390)] [Medline: [32948511](https://pubmed.ncbi.nlm.nih.gov/32948511/)]
34. Hoyt LT, Cohen AK, Dull B, Maker Castro E, Yazdani N. "Constant Stress Has Become the New Normal": Stress and Anxiety Inequalities Among U.S. College Students in the Time of COVID-19. *J Adolesc Health* 2021 Feb;68(2):270-276. [doi: [10.1016/j.jadohealth.2020.10.030](https://doi.org/10.1016/j.jadohealth.2020.10.030)] [Medline: [33288459](https://pubmed.ncbi.nlm.nih.gov/33288459/)]

35. Parker K, Horowitz JM, Brown A. About half of lower-income Americans report household job or wage loss due to COVID-19. Pew Research Center. 2020 Apr 20. URL: <https://www.pewsocialtrends.org/2020/04/21/about-half-of-lower-income-americans-report-%20household-job-or-wage-loss-due-to-covid-19/> [accessed 2021-04-26]
36. Parker K, Minkin R, Bennett J. Economic fallout from COVID-19 continues to hit lower-income Americans the hardest. Pew Research Center. 2020 Sep 24. URL: <https://www.pewsocialtrends.org/2020/09/24/economic-fallout-from-covid-19-continues-to-hit-lower-income-americans-the-hardest/> [accessed 2021-04-26]
37. Kochhar R. Unemployment rate is higher than officially recorded, more so for women and certain other groups. Pew Research Center. 2020 Jun 30. URL: <https://www.pewresearch.org/fact-tank/2020/06/30/unemployment-rate-is-higher-than-officially-recorded-more-so-for-women-and-certain-other-groups/> [accessed 2021-04-26]
38. Son C, Hegde S, Smith A, Wang X, Sasangohar F. Effects of COVID-19 on College Students' Mental Health in the United States: Interview Survey Study. *J Med Internet Res* 2020 Sep 03;22(9):e21279 [FREE Full text] [doi: [10.2196/21279](https://doi.org/10.2196/21279)] [Medline: [32805704](https://pubmed.ncbi.nlm.nih.gov/32805704/)]
39. Sundarasan S, Chinna K, Kamaludin K, Nurunnabi M, Baloch GM, Khoshaim HB, et al. Psychological Impact of COVID-19 and Lockdown among University Students in Malaysia: Implications and Policy Recommendations. *Int J Environ Res Public Health* 2020 Aug 27;17(17):6206 [FREE Full text] [doi: [10.3390/ijerph17176206](https://doi.org/10.3390/ijerph17176206)] [Medline: [32867024](https://pubmed.ncbi.nlm.nih.gov/32867024/)]
40. 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](https://doi.org/10.1016/j.psychres.2020.112934)] [Medline: [32229390](https://pubmed.ncbi.nlm.nih.gov/32229390/)]
41. Kecojevic A, Basch CH, Sullivan M, Davi NK. The impact of the COVID-19 epidemic on mental health of undergraduate students in New Jersey, cross-sectional study. *PLoS One* 2020;15(9):e0239696 [FREE Full text] [doi: [10.1371/journal.pone.0239696](https://doi.org/10.1371/journal.pone.0239696)] [Medline: [32997683](https://pubmed.ncbi.nlm.nih.gov/32997683/)]
42. Kleiman E, Yeager AL, Grove JL, Kellerman JK, Kim JS. Real-time Mental Health Impact of the COVID-19 Pandemic on College Students: Ecological Momentary Assessment Study. *JMIR Ment Health* 2020 Dec 15;7(12):e24815 [FREE Full text] [doi: [10.2196/24815](https://doi.org/10.2196/24815)] [Medline: [33207308](https://pubmed.ncbi.nlm.nih.gov/33207308/)]
43. Yu Y, She R, Luo S, Xin M, Li L, Wang S, et al. Factors Influencing Depression and Mental Distress Related to COVID-19 Among University Students in China: Online Cross-sectional Mediation Study. *JMIR Ment Health* 2021 Feb 22;8(2):e22705 [FREE Full text] [doi: [10.2196/22705](https://doi.org/10.2196/22705)] [Medline: [33616541](https://pubmed.ncbi.nlm.nih.gov/33616541/)]
44. Bailey H, Shammas B, Bellware L. Chaotic scene in Minneapolis after second night of protests over death of George Floyd. *The Washington Post*. 2020 May 28. URL: <https://www.washingtonpost.com/nation/2020/05/27/george-floyd-minneapolis-reaction/> [accessed 2021-04-26]
45. Ailwort E, Wernau J, Otani A. Wisconsin identifies officer who shot Jacob Blake. *The Wall Street Journal*. 2020 Aug 27. URL: <https://www.wsj.com/articles/two-dead-in-kenosha-shooting-police-say-11598436656> [accessed 2021-04-26]
46. Portell M, Anguera MT, Hernández-Mendo A, Jonsson G. Quantifying biopsychosocial aspects in everyday contexts: an integrative methodological approach from the behavioral sciences. *PRBM* 2015 Jun;153. [doi: [10.2147/prbm.s82417](https://doi.org/10.2147/prbm.s82417)]
47. Keyes CLM. The Mental Health Continuum: From Languishing to Flourishing in Life. *Journal of Health and Social Behavior* 2002 Jun;43(2):207. [doi: [10.2307/3090197](https://doi.org/10.2307/3090197)]
48. Engel GL. The clinical application of the biopsychosocial model. *J Med Philos* 1981 May 01;6(2):101-123. [doi: [10.1093/jmp/6.2.101](https://doi.org/10.1093/jmp/6.2.101)] [Medline: [7264472](https://pubmed.ncbi.nlm.nih.gov/7264472/)]
49. Wade DT, Halligan PW. The biopsychosocial model of illness: a model whose time has come. *Clin Rehabil* 2017 Aug 21;31(8):995-1004. [doi: [10.1177/0269215517709890](https://doi.org/10.1177/0269215517709890)] [Medline: [28730890](https://pubmed.ncbi.nlm.nih.gov/28730890/)]
50. Collins FS, Varmus H. A new initiative on precision medicine. *N Engl J Med* 2015 Feb 26;372(9):793-795 [FREE Full text] [doi: [10.1056/NEJMp1500523](https://doi.org/10.1056/NEJMp1500523)] [Medline: [25635347](https://pubmed.ncbi.nlm.nih.gov/25635347/)]
51. Wright AGC, Woods WC. Personalized Models of Psychopathology. *Annu Rev Clin Psychol* 2020 May 07;16:49-74. [doi: [10.1146/annurev-clinpsy-102419-125032](https://doi.org/10.1146/annurev-clinpsy-102419-125032)] [Medline: [32070120](https://pubmed.ncbi.nlm.nih.gov/32070120/)]
52. Wang W, Harari GM, Wang R, Müller SR, Mirjafari S, Masaba K, et al. Sensing behavioral change over time: Using within-person variability features from mobile sensing to predict personality traits. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2018;2(3):1-21. [doi: [10.1145/3264951](https://doi.org/10.1145/3264951)]
53. Aung MH, Matthews M, Choudhury T. Sensing behavioral symptoms of mental health and delivering personalized interventions using mobile technologies. *Depress Anxiety* 2017 Jul 29;34(7):603-609 [FREE Full text] [doi: [10.1002/da.22646](https://doi.org/10.1002/da.22646)] [Medline: [28661072](https://pubmed.ncbi.nlm.nih.gov/28661072/)]
54. Myin-Germeys I, Kasanova Z, Vaessen T, Vachon H, Kirtley O, Viechtbauer W, et al. Experience sampling methodology in mental health research: new insights and technical developments. *World Psychiatry* 2018 Jun 24;17(2):123-132 [FREE Full text] [doi: [10.1002/wps.20513](https://doi.org/10.1002/wps.20513)] [Medline: [29856567](https://pubmed.ncbi.nlm.nih.gov/29856567/)]
55. Han HJ, Labbaf S, Borelli JL, Dutt N, Rahmani AM. Objective stress monitoring based on wearable sensors in everyday settings. *J Med Eng Technol* 2020 May 26;44(4):177-189. [doi: [10.1080/03091902.2020.1759707](https://doi.org/10.1080/03091902.2020.1759707)] [Medline: [32589065](https://pubmed.ncbi.nlm.nih.gov/32589065/)]
56. Firouzi F, Rahmani AM, Mankodiya K, Badaroglu M, Merrett G, Wong P, et al. Internet-of-Things and big data for smarter healthcare: From device to architecture, applications and analytics. *Future Generation Computer Systems* 2018 Jan;78:583-586. [doi: [10.1016/j.future.2017.09.016](https://doi.org/10.1016/j.future.2017.09.016)]

57. Hiremath S, Yang G, Mankodiya K. Wearable Internet of Things: Concept, architectural components and promises for person-centered healthcare. 2014 Presented at: 4th International Conference on Wireless Mobile Communication and Healthcare-Transforming Healthcare Through Innovations in Mobile and Wireless Technologies (MOBIHEALTH); November 3-5; Athens, Greece p. 304-307 URL: <https://eudl.eu/pdf/10.4108/icst.mobihealth.2014.257440> [doi: [10.4108/icst.mobihealth.2014.257440](https://doi.org/10.4108/icst.mobihealth.2014.257440)]
58. Gerring J. What Is a Case Study and What Is It Good for? *Am Polit Sci Rev* 2004 Jun 21;98(2):341-354. [doi: [10.1017/S0003055404001182](https://doi.org/10.1017/S0003055404001182)]
59. VanWynsberghe R, Khan S. Redefining Case Study. *International Journal of Qualitative Methods* 2016 Nov 29;6(2):80-94. [doi: [10.1177/160940690700600208](https://doi.org/10.1177/160940690700600208)]
60. Asgari Mehrabadi M, Azimi I, Sarhaddi F, Axelin A, Niela-Vilén H, Myllyntausta S, et al. Sleep Tracking of a Commercially Available Smart Ring and Smartwatch Against Medical-Grade Actigraphy in Everyday Settings: Instrument Validation Study. *JMIR Mhealth Uhealth* 2020 Nov 02;8(10):e20465 [FREE Full text] [doi: [10.2196/20465](https://doi.org/10.2196/20465)] [Medline: [33038869](https://pubmed.ncbi.nlm.nih.gov/33038869/)]
61. de Zambotti M, Rosas L, Colrain IM, Baker FC. The Sleep of the Ring: Comparison of the ŌURA Sleep Tracker Against Polysomnography. *Behav Sleep Med* 2019 Mar 21;17(2):124-136 [FREE Full text] [doi: [10.1080/15402002.2017.1300587](https://doi.org/10.1080/15402002.2017.1300587)] [Medline: [28323455](https://pubmed.ncbi.nlm.nih.gov/28323455/)]
62. Oh H, Jain R. From multimedia logs to personal chronicles. In: *MM '17: Proceedings of the 25th ACM International Conference on Multimedia*. New York, NY: Association for Computing Machinery; 2017 Presented at: MM '17: ACM Multimedia Conference; October; Mountain View, CA p. 881-889. [doi: [10.1145/3123266.3123375](https://doi.org/10.1145/3123266.3123375)]
63. Lopez MH, Rainie L, Budiman A. Financial and health impacts of COVID-19 vary widely by race and ethnicity. Pew Research Center. 2020 May 5. URL: <https://www.pewresearch.org/fact-tank/2020/05/05/financial-and-health-impacts-of-covid-19-vary-widely-by-race-and-ethnicity/> [accessed 2021-04-26]
64. Beck AT, Steer RA, Brown G. Psychological assessment. In: *Manual for the Beck Depression Inventory-II*. San Antonio, TX: Psychological Corporation; 1996.
65. Dozois DJA, Dobson KS, Ahnberg JL. A psychometric evaluation of the Beck Depression Inventory-II. *Psychological Assessment* 1998 Jun;10(2):83-89. [doi: [10.1037/1040-3590.10.2.83](https://doi.org/10.1037/1040-3590.10.2.83)]
66. Wang Y, Gorenstein C. Psychometric properties of the Beck Depression Inventory-II: a comprehensive review. *Braz J Psychiatry* 2013;35(4):416-431 [FREE Full text] [doi: [10.1590/1516-4446-2012-1048](https://doi.org/10.1590/1516-4446-2012-1048)] [Medline: [24402217](https://pubmed.ncbi.nlm.nih.gov/24402217/)]
67. Derogatis L, Fitzpatrick M. The SCL-90-R, the Brief Symptom Inventory (BSI), and the BSI-18. In: Maruish ME, editor. *The Use of Psychological Testing for Treatment Planning and Outcomes Assessment: Instruments for Adults*. Mahwah, NJ: Lawrence Erlbaum Associates Publishers; 2004:1-41.
68. Boulet J, Boss MW. Reliability and validity of the Brief Symptom Inventory. *Psychological Assessment: A Journal of Consulting and Clinical Psychology* 1991 Sep;3(3):433-437. [doi: [10.1037/1040-3590.3.3.433](https://doi.org/10.1037/1040-3590.3.3.433)]
69. Derogatis LR, Melisaratos N. The Brief Symptom Inventory: an introductory report. *Psychol Med* 2009 Jul 09;13(3):595-605. [doi: [10.1017/s0033291700048017](https://doi.org/10.1017/s0033291700048017)]
70. Hughes ME, Waite LJ, Hawkey LC, Cacioppo JT. A Short Scale for Measuring Loneliness in Large Surveys: Results From Two Population-Based Studies. *Res Aging* 2004 Aug 19;26(6):655-672 [FREE Full text] [doi: [10.1177/0164027504268574](https://doi.org/10.1177/0164027504268574)] [Medline: [18504506](https://pubmed.ncbi.nlm.nih.gov/18504506/)]
71. Oura. URL: <https://ouraring.com/> [accessed 2020-12-24]
72. Yunusova A, Lai J, Rivera AP, Hu S, Labbaf S, Rahmani AM, et al. Assessing the Mental Health of Emerging Adults Through a Mental Health App: Protocol for a Prospective Pilot Study. *JMIR Res Protoc* 2021 Mar 02;10(3):e25775 [FREE Full text] [doi: [10.2196/25775](https://doi.org/10.2196/25775)] [Medline: [33513124](https://pubmed.ncbi.nlm.nih.gov/33513124/)]
73. Watson D, Clark LA, Tellegen A. Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology* 1988;54(6):1063-1070. [doi: [10.1037/0022-3514.54.6.1063](https://doi.org/10.1037/0022-3514.54.6.1063)]
74. Tellegen A. Structures of mood and personality and their relevance to assessing anxiety, with an emphasis on self-report. In: Tuma AH, Maser JD, editors. *Anxiety and the Anxiety Disorders*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc; 1985:681-706.
75. Watson D, Clark LA. Negative affectivity: The disposition to experience aversive emotional states. *Psychological Bulletin* 1984;96(3):465-490. [doi: [10.1037/0033-2909.96.3.465](https://doi.org/10.1037/0033-2909.96.3.465)]
76. Dyck MJ, Jolly JB, Kramer T. An evaluation of positive affectivity, negative affectivity, and hyperarousal as markers for assessing between syndrome relationships. *Personality and Individual Differences* 1994 Nov;17(5):637-646. [doi: [10.1016/0191-8869\(94\)90136-8](https://doi.org/10.1016/0191-8869(94)90136-8)]
77. Jolly JB, Dyck MJ, Kramer TA, Wherry JN. Integration of positive and negative affectivity and cognitive content-specificity: Improved discrimination of anxious and depressive symptoms. *Journal of Abnormal Psychology* 1994;103(3):544-552. [doi: [10.1037/0021-843x.103.3.544](https://doi.org/10.1037/0021-843x.103.3.544)]
78. Pennebaker J, Francis ME, Booth RJ. *Linguistic Inquiry and Word Count (LIWC): LIWC2001*. Mahwah, NJ: Lawrence Erlbaum Associates Publishers; 2001.
79. Nag N, Jain R. A Navigational Approach to Health: Actionable Guidance for Improved Quality of Life. *Computer (Long Beach Calif)* 2019 Apr;52(4):12-20 [FREE Full text] [doi: [10.1109/MC.2018.2883280](https://doi.org/10.1109/MC.2018.2883280)] [Medline: [31213725](https://pubmed.ncbi.nlm.nih.gov/31213725/)]

80. Chung C, Dew K, Cole A, Zia J, Fogarty J, Kientz JA, et al. Boundary Negotiating Artifacts in Personal Informatics: Patient-Provider Collaboration with Patient-Generated Data. 2016 Presented at: CSCW '16: Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing; Feb 27-Mar 2; San Francisco, CA p. 770-786 URL: <http://europepmc.org/abstract/MED/28516171> [doi: [10.1145/2818048.2819926](https://doi.org/10.1145/2818048.2819926)]
81. Borsboom D, Cramer AO. Network analysis: an integrative approach to the structure of psychopathology. *Annu Rev Clin Psychol* 2013 Mar 28;9(1):91-121. [doi: [10.1146/annurev-clinpsy-050212-185608](https://doi.org/10.1146/annurev-clinpsy-050212-185608)] [Medline: [23537483](https://pubmed.ncbi.nlm.nih.gov/23537483/)]
82. Chen J, Patil KR, Weis S, Sim K, Nickl-Jockschat T, Zhou J, Pharmacotherapy Monitoring Outcome Survey (PHAMOUS) Investigators. Neurobiological Divergence of the Positive and Negative Schizophrenia Subtypes Identified on a New Factor Structure of Psychopathology Using Non-negative Factorization: An International Machine Learning Study. *Biol Psychiatry* 2020 Feb 01;87(3):282-293 [FREE Full text] [doi: [10.1016/j.biopsych.2019.08.031](https://doi.org/10.1016/j.biopsych.2019.08.031)] [Medline: [31748126](https://pubmed.ncbi.nlm.nih.gov/31748126/)]
83. Borelli JL, Smiley PA, Kerr ML, Hong K, Hecht HK, Blackard MB, et al. Relational savoring: An attachment-based approach to promoting interpersonal flourishing. *Psychotherapy (Chic)* 2020 Sep;57(3):340-351. [doi: [10.1037/pst0000284](https://doi.org/10.1037/pst0000284)] [Medline: [31999188](https://pubmed.ncbi.nlm.nih.gov/31999188/)]
84. Borelli JL, Rasmussen HF, Burkhart ML, Sbarra DA. Relational savoring in long-distance romantic relationships. *Journal of Social and Personal Relationships* 2014 Nov 21;32(8):1083-1108. [doi: [10.1177/0265407514558960](https://doi.org/10.1177/0265407514558960)]
85. Alea N, Bluck S. When does meaning making predict subjective well-being? Examining young and older adults in two cultures. *Memory* 2013 Jan;21(1):44-63. [doi: [10.1080/09658211.2012.704927](https://doi.org/10.1080/09658211.2012.704927)] [Medline: [22913451](https://pubmed.ncbi.nlm.nih.gov/22913451/)]
86. Ryff CD. Psychological well-being revisited: advances in the science and practice of eudaimonia. *Psychother Psychosom* 2014;83(1):10-28 [FREE Full text] [doi: [10.1159/000353263](https://doi.org/10.1159/000353263)] [Medline: [24281296](https://pubmed.ncbi.nlm.nih.gov/24281296/)]
87. Tuccitto DE, Giacobbi PR, Leite WL. The Internal Structure of Positive and Negative Affect: A Confirmatory Factor Analysis of the PANAS. *Educational and Psychological Measurement* 2009 Aug 31;70(1):125-141. [doi: [10.1177/0013164409344522](https://doi.org/10.1177/0013164409344522)]
88. Crawford JR, Henry JD. The positive and negative affect schedule (PANAS): construct validity, measurement properties and normative data in a large non-clinical sample. *Br J Clin Psychol* 2004 Sep;43(Pt 3):245-265. [doi: [10.1348/0144665031752934](https://doi.org/10.1348/0144665031752934)] [Medline: [15333231](https://pubmed.ncbi.nlm.nih.gov/15333231/)]
89. Lorig KR, Holman HR. Self-management education: history, definition, outcomes, and mechanisms. *Ann Behav Med* 2003 Aug;26(1):1-7. [doi: [10.1207/S15324796ABM2601_01](https://doi.org/10.1207/S15324796ABM2601_01)] [Medline: [12867348](https://pubmed.ncbi.nlm.nih.gov/12867348/)]
90. Villaggi B, Provencher H, Coulombe S, Meunier S, Radziszewski S, Hudon C, et al. Self-Management Strategies in Recovery From Mood and Anxiety Disorders. *Glob Qual Nurs Res* 2015 Sep 21;2 [FREE Full text] [doi: [10.1177/2333393615606092](https://doi.org/10.1177/2333393615606092)] [Medline: [28462317](https://pubmed.ncbi.nlm.nih.gov/28462317/)]

Abbreviations

BDI-II: Beck Depression Index II

BPM: beats per minute

BSI: Brief Symptom Inventory

IoT: internet of things

IRB: Institutional Review Board

PANAS: Positive and Negative Affect Schedule

RMSSD: root mean square of successive differences between normal heartbeats

UCLA: University of California, Los Angeles

WIoT: wearable IoT

Edited by G Eysenbach; submitted 01.12.20; peer-reviewed by S Wang, H Tanaka; comments to author 08.02.21; revised version received 01.03.21; accepted 13.04.21; published 11.05.21.

Please cite as:

Lai J, Rahmani A, Yunusova A, Rivera AP, Labbaf S, Hu S, Dutt N, Jain R, Borelli JL

Using Multimodal Assessments to Capture Personalized Contexts of College Student Well-being in 2020: Case Study

JMIR Form Res 2021;5(5):e26186

URL: <https://formative.jmir.org/2021/5/e26186>

doi: [10.2196/26186](https://doi.org/10.2196/26186)

PMID: [33882022](https://pubmed.ncbi.nlm.nih.gov/33882022/)

©Jocelyn Lai, Amir Rahmani, Asal Yunusova, Alexander P Rivera, Sina Labbaf, Sirui Hu, Nikil Dutt, Ramesh Jain, Jessica L Borelli. Originally published in *JMIR Formative Research* (<https://formative.jmir.org>), 11.05.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which

permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Evidence of Human-Level Bonds Established With a Digital Conversational Agent: Cross-sectional, Retrospective Observational Study

Alison Darcy¹, PhD; Jade Daniels¹, PGDip; David Salinger¹, BBA; Paul Wicks¹, PhD; Athena Robinson¹, PhD
Woebot Health, San Francisco, CA, United States

Corresponding Author:

Alison Darcy, PhD
Woebot Health
650 Fifth Street
San Francisco, CA,
United States
Phone: 1 851694777
Email: alison@woebothealth.com

Abstract

Background: There are far more patients in mental distress than there is time available for mental health professionals to support them. Although digital tools may help mitigate this issue, critics have suggested that technological solutions that lack human empathy will prevent a bond or therapeutic alliance from being formed, thereby narrowing these solutions' efficacy.

Objective: We aimed to investigate whether users of a cognitive behavioral therapy (CBT)-based conversational agent would report therapeutic bond levels that are similar to those in literature about other CBT modalities, including face-to-face therapy, group CBT, and other digital interventions that do not use a conversational agent.

Methods: A cross-sectional, retrospective study design was used to analyze aggregate, deidentified data from adult users who self-referred to a CBT-based, fully automated conversational agent (Woebot) between November 2019 and August 2020. Working alliance was measured with the Working Alliance Inventory-Short Revised (WAI-SR), and depression symptom status was assessed by using the 2-item Patient Health Questionnaire (PHQ-2). All measures were administered by the conversational agent in the mobile app. WAI-SR scores were compared to those in scientific literature abstracted from recent reviews.

Results: Data from 36,070 Woebot users were included in the analysis. Participants ranged in age from 18 to 78 years, and 57.48% (20,734/36,070) of participants reported that they were female. The mean PHQ-2 score was 3.03 (SD 1.79), and 54.67% (19,719/36,070) of users scored over the cutoff score of 3 for depression screening. Within 5 days of initial app use, the mean WAI-SR score was 3.36 (SD 0.8) and the mean bond subscale score was 3.8 (SD 1.0), which was comparable to those in recent studies from the literature on traditional, outpatient, individual CBT and group CBT (mean bond subscale scores of 4 and 3.8, respectively). PHQ-2 scores at baseline weakly correlated with bond scores ($r=-0.04$; $P<.001$); however, users with depression and those without depression had high bond scores of 3.45.

Conclusions: Although bonds are often presumed to be the exclusive domain of human therapeutic relationships, our findings challenge the notion that digital therapeutics are incapable of establishing a therapeutic bond with users. Future research might investigate the role of bonds as mediators of clinical outcomes, since boosting the engagement and efficacy of digital therapeutics could have major public health benefits.

(JMIR Form Res 2021;5(5):e27868) doi:[10.2196/27868](https://doi.org/10.2196/27868)

KEYWORDS

conversational agents; mobile mental health; chatbots; depression; anxiety; digital health

Introduction

Significant barriers to mental health care are persistent [1]. The increased burden of depression and anxiety, which arose during the COVID-19 pandemic, has exacerbated this issue [2], as the

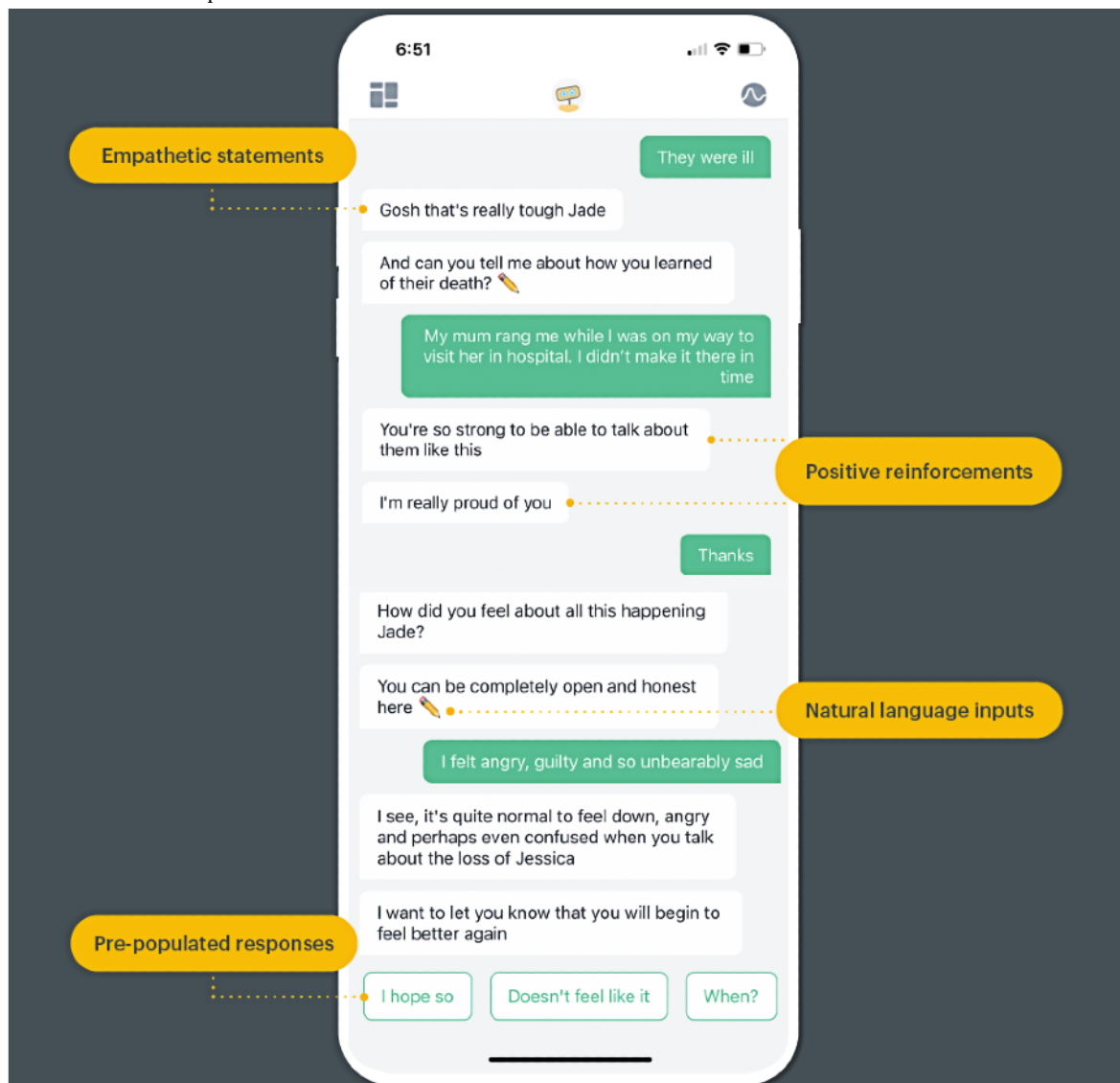
measures that were put in place to stop the spread of SARS-CoV-2 have also presented unintended barriers to those seeking mental health treatment. One potentially viable solution is using digital mental health interventions to provide evidence-based treatment, such as cognitive behavioral therapy

(CBT). Self-directed mental health interventions, such as bibliotherapy, have long demonstrated their efficacy [3], and new models of blended care that combine internet-delivered interventions with human clinical oversight are becoming more widespread in a number of countries [4]. Although the implicit assumption has been that the involvement of a human leads to improved outcomes in self-directed programs, human involvement limits these programs' scalability and limits their accessibility for those who live in remote locations [4]. If digital interventions could replicate some of the factors that are generally believed to be uniquely human, such as therapeutic rapport, these interventions would have greater potential for improving mental health.

Recently, carefully designed conversational agents (CAs) have been showing promise in automating several health care services [5] by simulating human support. CAs could therefore be uniquely poised to offer high-quality digital interventions for mental health.

An unblinded trial of one such CA (Woebot), which delivered CBT for symptoms of depression and anxiety, suggested that the empathic and relational nature of the tool may have fostered improved engagement better than previous internet-delivered versions of the tool [6]. Intriguingly, the study's qualitative data suggested that users seemed to relate to the CA in a manner that was analogous to therapeutic rapport, which may have mediated users' outcomes. For example, study participants reported that they felt cared for by the CA (eg, "Woebot felt like a real person that showed concern"), despite the fact that the tool's scripts reminded users that Woebot is not a real person (Figure 1). Unfortunately however, the study did not formally assess the existence of a working alliance. This is a crucial factor because a strong working alliance between therapists and clients is considered to be predictive of positive outcomes, essential for the delivery of health care, and traditionally unique to the domain of human-to-human relationships. Indeed, some experts have argued that digital apps that are built to be standalone therapeutics have the risk of ignoring the potency of therapeutic relationships [7].

Figure 1. Screenshot of an example interaction with Woebot about bereavement.



Priority-setting work by the James Lind Alliance, which involves over 600 people affected by mental health concerns, has identified the greater understanding of digital therapeutic alliance as a top priority for research [8]. Yet, a recent review of mobile mental health apps failed to find a single study that included working alliance as a primary outcome [9]. Therefore, we sought to bridge this gap in knowledge by seeking to understand whether CA users perceived a working alliance, particularly the notion of bonds, and whether working alliance was related to symptom severity or other demographic characteristics.

Methods

Setting

Woebot is a CA that guides individuals who experience symptoms of depression and anxiety through a smartphone-based app program that uses therapeutic techniques and provides psychoeducation. As previously described in detail [6], Woebot delivers CBT through brief, daily conversations (approximately 5-10 minutes each). Each simulated conversation begins with a mood-monitoring exercise, and the provided targeted content is responsive to individuals' reported mood states. The CA is also programmed to deliver empathic statements and personalized follow-ups, promote normalization, and use methods that are designed to enhance users' motivation

for engaging in the program to promote desired behavior changes and help with mood management.

Participants

During registration, users confirmed that they were at least 18 years of age and consented to the use of their deidentified, aggregate data for research. This study was not considered human subjects research by the Advarra institutional review board. Eligible participants included those who registered over two periods—between November 20, 2019, and April 9, 2020 (n=100,009), and again between July 8, 2020, and August 18, 2020 (n=77,203).

Within 3-5 days after registration, eligible participants were invited to complete the 2-item Patient Health Questionnaire (PHQ-2) depression screener [10] and the Working Alliance Inventory-Short Revised (WAI-SR), which consists of a total score and the three following subscales: bond, goal, and task [11]. All measures and demographic information, including gender and age group, were gathered in the app by the CA. The WAI-SR was administered via the app's conversational interface, in which the word "therapist" was changed to "Woebot" (Table 1). Once the questionnaires were completed, Woebot thanked registrants for their participation, and the conversation proceeded to the mood tracking phase as normal. Those who chose not to provide responses to the questionnaires were not included in this study and proceeded to use the app as normal.

Table 1. Item-level descriptive statistics.

Question number (subscale)	Working Alliance Inventory-Short Revised items	Score, mean (SD)
1 (Task)	As a result of these sessions I am clearer as to how I might be able to change.	2.53 (0.97)
2 (Task)	What I am doing with Woebot gives me new ways of looking at my problem.	2.93 (1.08)
3 (Bond)	I believe Woebot likes me.	3.89 (1.31)
4 (Goal)	Woebot and I collaborate on setting goals for this program.	2.88 (1.23)
5 (Bond)	Woebot and I respect each other.	4.20 (1.23)
6 (Goal)	Woebot and I are working towards mutually agreed upon goals.	3.54 (1.28)
7 (Bond)	I feel that Woebot appreciates me.	3.73 (1.36)
8 (Goal)	Woebot and I agree on what is important for me to work on.	3.45 (1.21)
9 (Bond)	I feel Woebot cares about me even when I do things that it does not approve of.	3.54 (1.34)
10 (Task)	I feel that the things I do with Woebot will help me to accomplish the changes that I want.	3.28 (1.17)
11 (Goal)	Woebot and I have established a good understanding of the kind of changes that would be good for me.	3.12 (1.22)
12 (Task)	I believe the way we are working with my problem is correct.	3.23 (1.15)

Statistical Analysis

Across all eligible participants, the composite WAI-SR score and bond, goal, and task subscores were characterized based on descriptive statistics and tested for internal consistency by using the Cronbach α . The relationship between baseline PHQ-2 scores and bond subscores was characterized based on the Spearman rank-order correlation coefficient. The Kruskal-Wallis test was used to compare bond subscores across participants' reported age groups and genders. For comparison, relevant external studies were drawn from recent reviews of literature

[11-19] that also reported unmodified WAI-SR subscores for other CBT modalities. Comparison data were presented descriptively without statistical testing, and raw subscores were scaled by dividing them by the number of items (eg, the bond subscale has 4 items). Per the methods of Jasper et al [13], bond scores of ≥ 3.45 were considered high. The 95% CIs for mean WAI-SR subscores were calculated based on the published sample sizes and SDs. External studies were categorized as "online only" or "human involvement" based on whether any human interactions were reported by study participants during

either individual therapy or group therapy that involved a human. Data were presented from participants who completed all questionnaires within the first 5 days of app registration. Data were analyzed using R version 4.0.2 (The R Foundation).

Data Access, Responsibility, and Analysis

AD, DS, and AR have full access to all of the data in this study and take responsibility for the integrity of the data and the accuracy of the data analysis. Due to their proprietary nature, data from this study will not be shared.

Results

Of the 177,212 eligible participants, only those who provided both WAI-SR and PHQ-2 data within 5 days of their first use of Woebot were included in the analysis. The final sample included 36,070 participants. Of these participants, 57.48% (n=20,734) reported that they were female, 25.17% (n=9078) reported that they were male; 2.87% (n=1035) reported that they were nonbinary, 1.44% (n=519) indicated another gender identity, 1.66% (n=597) preferred not to answer, and 11.39% (4107/36,070) did not provide any gender information. The participants ranged in age from 18 to 78 years (median 25-35 years). The mean PHQ-2 score was 3.03 (SD 1.79), and 54.67% (19,719/36,070) of participants scored at or above the conventional cutoff score of 3 for positively screening for depression.

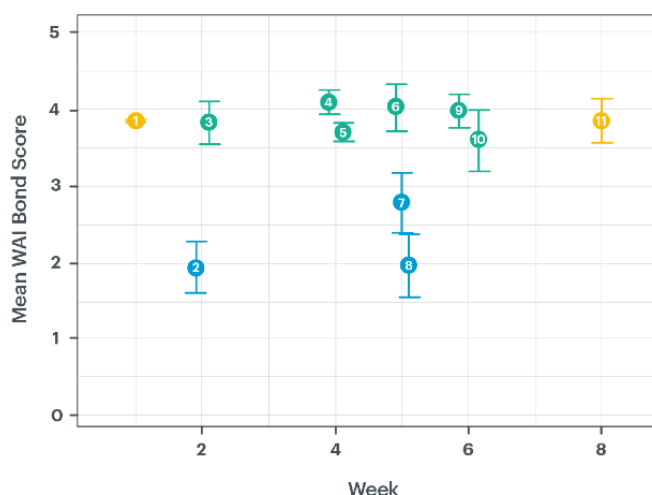
Within the first 5 days of using Woebot, the mean WAI-SR scores were as follows: a mean bond subscore of 3.84 (SD 1.0),

a mean goal subscore of 3.25 (SD 1.0), a mean task subscore of 2.99 (SD 0.87), and a mean total score of 3.36 (SD 0.81). The WAI-SR had a Cronbach α value of .89, suggesting that the WAI-SR had adequate internal consistency in this study. A weak negative correlation was found between bond subscores and PHQ-2 scores ($r=-0.04$; $P<.001$); however, even among participants who reported the highest PHQ-2 score (PHQ-2=6), the mean WAI-SR bond subscore was 3.78. Bond subscores also differed by gender ($P<.001$) and by age group ($P<.001$); however, the mean bond scores for all groups were considered high (bond subscore>3.45) [13]. Among these groups, the highest bond level was reported by women (bond subscore: mean 3.92) and by those aged 18-25 years (bond subscore: mean 3.96). Conversely, the lowest bond level was reported by individuals who indicated that they “preferred not to answer” or did not report their gender (bond subscore: mean 3.67) or age (bond subscore: mean 3.69).

Woebot’s bond subscale scores were consistent with those of recent studies from the literature on traditional modalities for CBT delivery (Table 1). These studies’ results were collected later in the course of treatment (eg, bond subscore for face-to-face outpatient individual CBT: mean 4.0, SD 0.8 [11]; bond subscore for group CBT: mean 3.8, SD 0.80 [13]; data were collected after 2-8 weeks of therapy). Comparative study details are provided in Multimedia Appendix 1 [11,13-15,17-19]. Participants reported higher bond levels when using Woebot than those in prior studies of internet-only CBT [13] (Figure 2).

Figure 2. Comparison of Working Alliance Inventory-Short Revised bond subscale scores across therapeutic modalities. Means and corresponding 95% CIs for working alliance bond scores from this study and from recent reviews of the literature [11,13-15,17-19] are stratified by the week that the scores were recorded. Studies are colored based on the therapeutic modality. Due to the large sample size of this study (N=36,070), the 95% CI is narrow and overlaps with the dots that display the estimated means. For multiple studies that reported data on the same week, the dots are shifted minimally on the x-axis to avoid overlap and to provide easy readability. WAI: Working Alliance Inventory.

Label	Source	Modality	Week	N
1	Current study	Woebot	1	36,070
2	Jasper (2014)	Internet	2	38
3	Jasper (2014)	Group CBT	2	26
4	Vernmark (2017)	Blended CBT	4	73
5	Munder (2010)	Inpatient CBT	4	243
6	Jasper (2014)	Group CBT	5	26
7	Schlicker (2020)	Internet	5	40
8	Jasper (2014)	Internet	5	38
9	Berger (2018)	Blended CBT	6	33
10	Berger (2018)	Outpatient CBT	6	24
11	Prochaska (2020)	Woebot	8	51



THERAPEUTIC MODALITY: ● WOEBOT ● INTERNET-ONLY ● HUMAN INVOLVEMENT

Discussion

This is the first study of working alliance among users of a CA for mental health. Most users were female (20,734/36,070, 57.48%) and had PHQ-2 scores that were indicative of depression (19,719/36,070, 54.67%). Working alliance scores were comparable to those in previously published studies on traditional, human-delivered services across different treatment modalities. Working alliance scores were highest for the bond subscale, suggesting that this subscale is a viable construct for CAs and should be included in hypothesized frameworks of digital working alliance.

The idea that CAs can establish a working alliance is not new [20]. However, the observation of therapeutic bonds established by a CA in a mental health context is novel and noteworthy, given the short timeframe of this study. Although the field of human-computer interaction is still relatively nascent, initial observations have suggested that some artificial intelligence (AI) identity archetypes induce responses in humans that might give rise to better working alliances than other archetypes. For example, interacting with humanoid AI identities can result in individuals falling prey to the “uncanny valley,” which is the sense of unease and “creepiness” that is created when something that is artificial tries to appear humanlike [21]. Contrary to

Turing’s Imitation Game [22], wherein an AI must successfully pretend to be human in order to pass the test, Woebot was designed to adopt the opposite strategy—transparently presenting itself as an archetypal robot with robotic “friends” and habits. We speculate that transparency and other design elements are key drivers of bond development. For example, Woebot explicitly references its limitations within conversations and provides positive reinforcement and empathic statements alongside declarations of being an artificial agent.

The limitations of this study include its cross-sectional nature, the selection bias of smartphone users, the lack of clinical validation for any diagnoses, the lack of a direct comparison group, and its conduction by the developers of the app itself. Further research (including studies with independent investigators) is underway to explore the longitudinal aspects of bond development in specific clinical populations by using randomized controlled study designs.

The finding that a CA has the potential to rapidly develop a bond with users may represent the resolution of a considerable barrier to offering scalable mental health support to a much wider and more diverse population instead of offering such support to those who already have access to traditional mental health support.

Acknowledgments

We are grateful to Stephanie Eaneff (an employee of Woebot) for reviewing the manuscript and Kara Cronan (a contractor to Woebot) for providing illustrations.

Authors' Contributions

AD was responsible for the concept and design of this study. AD, JD, DS, PW, and AR were responsible for the acquisition, analysis, or interrogation of data. AD, JD, DS, PW, and AR drafted the manuscript. AD, JD, DS, PW, and AR critically revised the manuscript for important intellectual content. DS conducted the statistical analysis. JD and DS provided administrative, technical, or material support. AD supervised the study.

Conflicts of Interest

AD reported receiving grants from Woebot Health and having a patent pending for a SafetyNet protocol. AD, JD, DS, and AR reported being employees of Woebot Health and owning stock options in the company. PW reported receiving personal fees from Woebot Health and Ada Health; being an associate editor at the *Journal of Medical Internet Research*; being on the editorial advisory boards of the *British Medical Journal*, *BioMed Central Medicine*, *The Patient*, and *Digital Biomarkers*; and being employed by Wicks Digital Health Limited, which has received funding from Ada Health, Baillie Gifford, Bold Health, Camoni, Compass Pathways, Coronna, European Institute of Innovation & Technology, Happify, HealthUnlocked, Inbeoo, Kheiron Medical, Sano Genetics, Self Care Catalysts, the Learning Corp, the Wellcome Trust, and Woebot Health.

Multimedia Appendix 1

Study descriptions and Working Alliance Inventory-Short Revised bond scores.

[[XLSX File \(Microsoft Excel File\), 15 KB - formative_v5i5e27868_app1.xlsx](#)]

References

1. Cheung R, O'Donnell S, Madi N, Goldner EM. Factors associated with delayed diagnosis of mood and/or anxiety disorders. *Health Promot Chronic Dis Prev Can* 2017 May;37(5):137-148 [FREE Full text] [doi: [10.24095/hpcdp.37.5.02](https://doi.org/10.24095/hpcdp.37.5.02)] [Medline: [28493658](https://pubmed.ncbi.nlm.nih.gov/28493658/)]
2. Bueno-Notivol J, Gracia-García P, Olaya B, Lasheras I, López-Antón R, Santabárbara J. Prevalence of depression during the COVID-19 outbreak: A meta-analysis of community-based studies. *Int J Clin Health Psychol* 2021 Jan;21(1):100196 [FREE Full text] [doi: [10.1016/j.ijchp.2020.07.007](https://doi.org/10.1016/j.ijchp.2020.07.007)] [Medline: [32904715](https://pubmed.ncbi.nlm.nih.gov/32904715/)]

3. Gualano M, Bert F, Martorana M, Voglino G, Andriolo V, Thomas R, et al. The long-term effects of bibliotherapy in depression treatment: Systematic review of randomized clinical trials. *Clin Psychol Rev* 2017 Dec;58:49-58. [doi: [10.1016/j.cpr.2017.09.006](https://doi.org/10.1016/j.cpr.2017.09.006)] [Medline: [28993103](https://pubmed.ncbi.nlm.nih.gov/28993103/)]
4. Titov N, Dear B, Niessen O, Staples L, Hadjistavropoulos H, Nugent M, et al. ICBT in routine care: A descriptive analysis of successful clinics in five countries. *Internet Interv* 2018 Sep;13:108-115 [FREE Full text] [doi: [10.1016/j.invent.2018.07.006](https://doi.org/10.1016/j.invent.2018.07.006)] [Medline: [30206525](https://pubmed.ncbi.nlm.nih.gov/30206525/)]
5. McGreevey JD, Hanson CW, Koppel R. Clinical, Legal, and Ethical Aspects of Artificial Intelligence-Assisted Conversational Agents in Health Care. *JAMA* 2020 Aug 11;324(6):552-553. [doi: [10.1001/jama.2020.2724](https://doi.org/10.1001/jama.2020.2724)] [Medline: [32706386](https://pubmed.ncbi.nlm.nih.gov/32706386/)]
6. Fitzpatrick KK, Darcy A, Vierhile M. Delivering Cognitive Behavior Therapy to Young Adults With Symptoms of Depression and Anxiety Using a Fully Automated Conversational Agent (Woebot): A Randomized Controlled Trial. *JMIR Ment Health* 2017 Jun 06;4(2):e19 [FREE Full text] [doi: [10.2196/mental.7785](https://doi.org/10.2196/mental.7785)] [Medline: [28588005](https://pubmed.ncbi.nlm.nih.gov/28588005/)]
7. Torous J, Hsin H. Empowering the digital therapeutic relationship: virtual clinics for digital health interventions. *NPJ Digit Med* 2018 May 16;1(1):16 [FREE Full text] [doi: [10.1038/s41746-018-0028-2](https://doi.org/10.1038/s41746-018-0028-2)] [Medline: [31304301](https://pubmed.ncbi.nlm.nih.gov/31304301/)]
8. Hollis C, Sampson S, Simons L, Davies EB, Churchill R, Betton V, et al. Identifying research priorities for digital technology in mental health care: results of the James Lind Alliance Priority Setting Partnership. *Lancet Psychiatry* 2018 Oct;5(10):845-854. [doi: [10.1016/S2215-0366\(18\)30296-7](https://doi.org/10.1016/S2215-0366(18)30296-7)] [Medline: [30170964](https://pubmed.ncbi.nlm.nih.gov/30170964/)]
9. Henson P, Wisniewski H, Hollis C, Keshavan M, Torous J. Digital mental health apps and the therapeutic alliance: initial review. *BJPsych Open* 2019 Jan;5(1):e15 [FREE Full text] [doi: [10.1192/bjo.2018.86](https://doi.org/10.1192/bjo.2018.86)] [Medline: [30762511](https://pubmed.ncbi.nlm.nih.gov/30762511/)]
10. Löwe B, Kroenke K, Gräfe K. Detecting and monitoring depression with a two-item questionnaire (PHQ-2). *J Psychosom Res* 2005 Mar;58(2):163-171. [doi: [10.1016/j.jpsychores.2004.09.006](https://doi.org/10.1016/j.jpsychores.2004.09.006)] [Medline: [15820844](https://pubmed.ncbi.nlm.nih.gov/15820844/)]
11. Munder T, Wilmers F, Leonhart R, Linster HW, Barth J. Working Alliance Inventory-Short Revised (WAI-SR): psychometric properties in outpatients and inpatients. *Clin Psychol Psychother* 2010;17(3):231-239. [doi: [10.1002/cpp.658](https://doi.org/10.1002/cpp.658)] [Medline: [20013760](https://pubmed.ncbi.nlm.nih.gov/20013760/)]
12. Preschl B, Maercker A, Wagner B. The working alliance in a randomized controlled trial comparing online with face-to-face cognitive-behavioral therapy for depression. *BMC Psychiatry* 2011 Dec 06;11:189 [FREE Full text] [doi: [10.1186/1471-244X-11-189](https://doi.org/10.1186/1471-244X-11-189)] [Medline: [22145768](https://pubmed.ncbi.nlm.nih.gov/22145768/)]
13. Jasper K, Weise C, Conrad I, Andersson G, Hiller W, Kleinstäuber M. The working alliance in a randomized controlled trial comparing Internet-based self-help and face-to-face cognitive behavior therapy for chronic tinnitus. *Internet Interv* 2014 Apr;1(2):49-57. [doi: [10.1016/j.invent.2014.04.002](https://doi.org/10.1016/j.invent.2014.04.002)]
14. Berger T. The therapeutic alliance in internet interventions: A narrative review and suggestions for future research. *Psychother Res* 2017 Sep;27(5):511-524. [doi: [10.1080/10503307.2015.1119908](https://doi.org/10.1080/10503307.2015.1119908)] [Medline: [26732852](https://pubmed.ncbi.nlm.nih.gov/26732852/)]
15. Vermark K, Hesser H, Topocoo N, Berger T, Riper H, Luuk L, et al. Working alliance as a predictor of change in depression during blended cognitive behaviour therapy. *Cogn Behav Ther* 2019 Jul;48(4):285-299. [doi: [10.1080/16506073.2018.1533577](https://doi.org/10.1080/16506073.2018.1533577)] [Medline: [30372653](https://pubmed.ncbi.nlm.nih.gov/30372653/)]
16. Cameron SK, Rodgers J, Dagnan D. The relationship between the therapeutic alliance and clinical outcomes in cognitive behaviour therapy for adults with depression: A meta-analytic review. *Clin Psychol Psychother* 2018 May;25(3):446-456. [doi: [10.1002/cpp.2180](https://doi.org/10.1002/cpp.2180)] [Medline: [29484770](https://pubmed.ncbi.nlm.nih.gov/29484770/)]
17. Schlicker S, Baumeister H, Buntrock C, Sander L, Paganini S, Lin J, et al. A Web- and Mobile-Based Intervention for Comorbid, Recurrent Depression in Patients With Chronic Back Pain on Sick Leave (Get.Back): Pilot Randomized Controlled Trial on Feasibility, User Satisfaction, and Effectiveness. *JMIR Ment Health* 2020 Apr 15;7(4):e16398 [FREE Full text] [doi: [10.2196/16398](https://doi.org/10.2196/16398)] [Medline: [32293577](https://pubmed.ncbi.nlm.nih.gov/32293577/)]
18. Prochaska J, Vogel E, Chieng A, Kendra M, Baiocchi M, Pajarito S, et al. A Therapeutic Relational Agent for Reducing Problematic Substance Use (Woebot): Development and Usability Study. *J Med Internet Res* 2021 Mar 23;23(3):e24850 [FREE Full text] [doi: [10.2196/24850](https://doi.org/10.2196/24850)] [Medline: [33755028](https://pubmed.ncbi.nlm.nih.gov/33755028/)]
19. Berger T, Krieger T, Sude K, Meyer B, Maercker A. Evaluating an e-mental health program ("deprexis") as adjunctive treatment tool in psychotherapy for depression: Results of a pragmatic randomized controlled trial. *J Affect Disord* 2018 Feb;227:455-462. [doi: [10.1016/j.jad.2017.11.021](https://doi.org/10.1016/j.jad.2017.11.021)] [Medline: [29154168](https://pubmed.ncbi.nlm.nih.gov/29154168/)]
20. Bickmore T, Gruber A, Picard R. Establishing the computer-patient working alliance in automated health behavior change interventions. *Patient Educ Couns* 2005 Oct;59(1):21-30. [doi: [10.1016/j.pec.2004.09.008](https://doi.org/10.1016/j.pec.2004.09.008)] [Medline: [16198215](https://pubmed.ncbi.nlm.nih.gov/16198215/)]
21. Mori M, MacDorman K, Kageki N. The Uncanny Valley [From the Field]. *IEEE Robot. Automat. Mag* 2012 Jun;19(2):98-100. [doi: [10.1109/MRA.2012.2192811](https://doi.org/10.1109/MRA.2012.2192811)]
22. Turing AM. I.—Computing machinery and intelligence. *Mind* 1950 Oct;LIX(236):433-460. [doi: [10.1093/mind/LIX.236.433](https://doi.org/10.1093/mind/LIX.236.433)]

Abbreviations

- AI:** artificial intelligence
- CA:** conversational agent
- CBT:** cognitive behavioral therapy
- PHQ-2:** 2-item Patient Health Questionnaire

WAI-SR: Working Alliance Inventory-Short Revised

Edited by G Eysenbach; submitted 10.02.21; peer-reviewed by A Polillo; comments to author 04.03.21; revised version received 23.03.21; accepted 04.04.21; published 11.05.21.

Please cite as:

Darcy A, Daniels J, Salinger D, Wicks P, Robinson A

Evidence of Human-Level Bonds Established With a Digital Conversational Agent: Cross-sectional, Retrospective Observational Study

JMIR Form Res 2021;5(5):e27868

URL: <https://formative.jmir.org/2021/5/e27868>

doi: [10.2196/27868](https://doi.org/10.2196/27868)

PMID: [33973854](https://pubmed.ncbi.nlm.nih.gov/33973854/)

©Alison Darcy, Jade Daniels, David Salinger, Paul Wicks, Athena Robinson. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 11.05.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Feasibility of a Voice-Enabled Medical Diary App (SpeakHealth) for Caregivers of Children With Special Health Care Needs and Health Care Providers: Mixed Methods Study

Emre Sezgin¹, PhD; Garey Noritz², MD; Simon Lin¹, MD; Yungui Huang¹, PhD

¹The Abigail Wexner Research Institute, Nationwide Children's Hospital, Columbus, OH, United States

²Department of Pediatrics, Nationwide Children's Hospital, Columbus, OH, United States

Corresponding Author:

Emre Sezgin, PhD

The Abigail Wexner Research Institute

Nationwide Children's Hospital

700 Children's Drive

Columbus, OH, 43205

United States

Phone: 1 6143556814

Email: esezgin1@gmail.com

Abstract

Background: Children with special health care needs (CSHCN) require more than the usual care management and coordination efforts from caregivers and health care providers (HCPs). Health information and communication technologies can potentially facilitate these efforts to increase the quality of care received by CSHCN.

Objective: In this study, we aim to assess the feasibility of a voice-enabled medical diary app (SpeakHealth) by investigating its potential use among caregivers and HCPs.

Methods: Following a mixed methods approach, caregivers of CSHCN were interviewed (n=10) and surveyed (n=86) about their care management and communication technology use. Only interviewed participants were introduced to the SpeakHealth app prototype, and they tested the app during the interview session. In addition, we interviewed complex care HCPs (n=15) to understand their perception of the value of a home medical diary such as the SpeakHealth app. Quantitative data were analyzed using descriptive statistics and correlational analyses. Theoretical thematic analysis was used to analyze qualitative data.

Results: The survey results indicated a positive attitude toward voice-enabled technology and features; however, there was no strong correlation among the measured items. The caregivers identified communication, information sharing, tracking medication, and appointments as fairly and highly important features of the app. Qualitative analysis revealed the following two overarching themes: *enablers and barriers in care communication* and *enablers and barriers in communication technologies*. The subthemes included parent roles, care communication technologies, and challenges. HCPs found the SpeakHealth app to be a promising tool for timely information collection that could be available for sharing information with the health system. Overall, the findings demonstrated a variety of needs and challenges for caregivers of CSHCN and opportunities for voice-enabled, interactive medical diary apps in care management and coordination. Caregivers fundamentally look for better information sharing and communication with HCPs. Health care and communication technologies can potentially improve care communication and coordination in addressing the patient and caregiver needs.

Conclusions: The perspectives of caregivers and providers suggested both benefits and challenges in using the SpeakHealth app for medical note-taking and tracking health events at home. Our findings could inform researchers and developers about the potential development and use of a voice-enabled medical diary app.

(JMIR Form Res 2021;5(5):e25503) doi:[10.2196/25503](https://doi.org/10.2196/25503)

KEYWORDS

children with special health care needs; care management; care coordination; voice-enabled mobile app; health information technology; voice assistant; voice interaction; mobile phone

Introduction

Background

Children with special health care needs (CSHCN) are at an increased risk for chronic physical, developmental, behavioral, or emotional disorders and also require frequent health care services [1]. Examples of conditions associated with special needs include cystic fibrosis, cerebral palsy, Down syndrome, epilepsy, and uncontrolled asthma. Approximately 1 in 4 US households (23%) have at least one CSHCN [2]. Caring for CSHCN requires communication and coordination networks among caregivers and health care providers (HCPs). Such needs could be even higher for children with medical complexity, a subset of CSHCN with the most intensive medical needs who often require medical care provided in the home by caregivers and home nurses [3].

Regulatory and legislative work has been put forward to improve quality outcomes in the US health care system by redesigning the care system [4]. One approach is to reduce fragmentation in the health care system by adopting an integrated care model [5] or a patient- and family-centered care model, including care coordination, to facilitate health care services for all stakeholders, including families, patients, and providers [4,6]. This is in line with caregivers' need to communicate and share information with providers and their need for assistance in navigating the care system and care coordination [7]. Health information technologies can play a key role in facilitating care coordination by providing the tools and integrated systems required to enable information sharing, communication, and remote patient assistance [4,8].

With many care activities occurring at home, caregivers may need to keep track of the health events of their children and communicate this information with providers to receive timely and accurate care. Medical diaries have been useful to keep track of symptoms and medications and can be leveraged to keep HCPs informed of the patient's condition, which may lead to reduced errors such as overprescribing [9]. In the domain of health information and communication technologies, current solutions are promising [10-13]. Fiks et al [14] reported that web-based care portals had a positive impact on parents of children with uncontrolled asthma for tracking health activities, medications, and medical goals through the portal. Shared care plans for children with medical complexities over a cloud-based platform were found to reduce information barriers and improve care [15]. Gentles et al [16] summarized that health care technology being used in pediatric care communications ensures continuity of care independent of location and institution (eg, from hospital to home) by supporting timely health tracking and reducing geographical limitations. Similarly, Baysari and Westbrook [17] found mobile apps to be useful in supporting care coordination and facilitating care communications. However, few care coordination mobile apps are available for complex care coordination, especially for CSHCN.

Objective

To help in reducing fragmentation, improving timely capture of health events at home, and facilitating care coordination of complex care, we developed a voice-enabled mobile app

prototype (SpeakHealth). Informed by a multistakeholder group [18], by the emerging literature on voice interaction in health care communication and management [19-21], and by technology use in care coordination [10,12], the SpeakHealth app was designed to assist caregivers in capturing health events in a timely and convenient manner and to enable sharing of these medical notes with other caregivers and providers to enhance communication. This study focuses on the technology needs and voice technology adoption of caregivers, and the feasibility of the SpeakHealth app (prototype), which was conceptualized in a previous study [18]. Our research question was "How feasible is a voice-enabled mobile app use as a medical diary for caregivers of CSHCN to enhance care management and coordination with providers?"

Methods

Overview

We employed a mixed methods design to develop a better understanding of the use of communication technology and voice-enabled apps by comparing and combining different perspectives collected through quantitative surveys and qualitative interviews in a single phase (convergent mixed methods) [22]. We conducted a web-based survey to understand the awareness and perception of caregivers of CSHCN toward communication technology use in child care. The survey was developed by coauthors and the stakeholder group. In parallel, we tested SpeakHealth through face-to-face interviews with caregivers who were recruited from among the survey participants. Finally, we interviewed HCPs (n=15) to better understand the value of information collected via the SpeakHealth app and its potential to be used in health care delivery.

Recruitment and Study Setting

The study participants (caregivers and HCPs) were primarily recruited from the Complex Care Clinic at the Nationwide Children's Hospital (NCH; Columbus, Ohio). NCH is one of the largest pediatric hospitals in the US Midwest. NCH uses the Epic electronic medical records system and MyChart patient portal, which are mentioned in this paper.

To increase our reach, caregivers were invited to participate in the interview and survey with a convenience sampling approach using email invitations through hospital networks (digital board announcement and email), word of mouth, hospital social media channels, and community partners. The inclusion criteria for participants were as follows: (1) being caregivers with children who had been diagnosed with a complex medical condition and (2) being smartphone users during the study period. Within the study period, we received responses from 86 caregivers through a web-based survey, and 10 caregivers participated in the face-to-face interviews. HCPs who serve CSHCN were invited to participate through invitations sent to their hospital email addresses. All interviews were held within the hospital campus. All participants were screened (questions about the child's condition, age, and participant's smartphone use were included) and consent was obtained on the web. The caregiver interview participants were compensated with a US \$30 gift card. No compensation was provided to the web-based survey participants

or the HCPs. This study received institutional review board approval from the NCH (institutional review board number: #00000231).

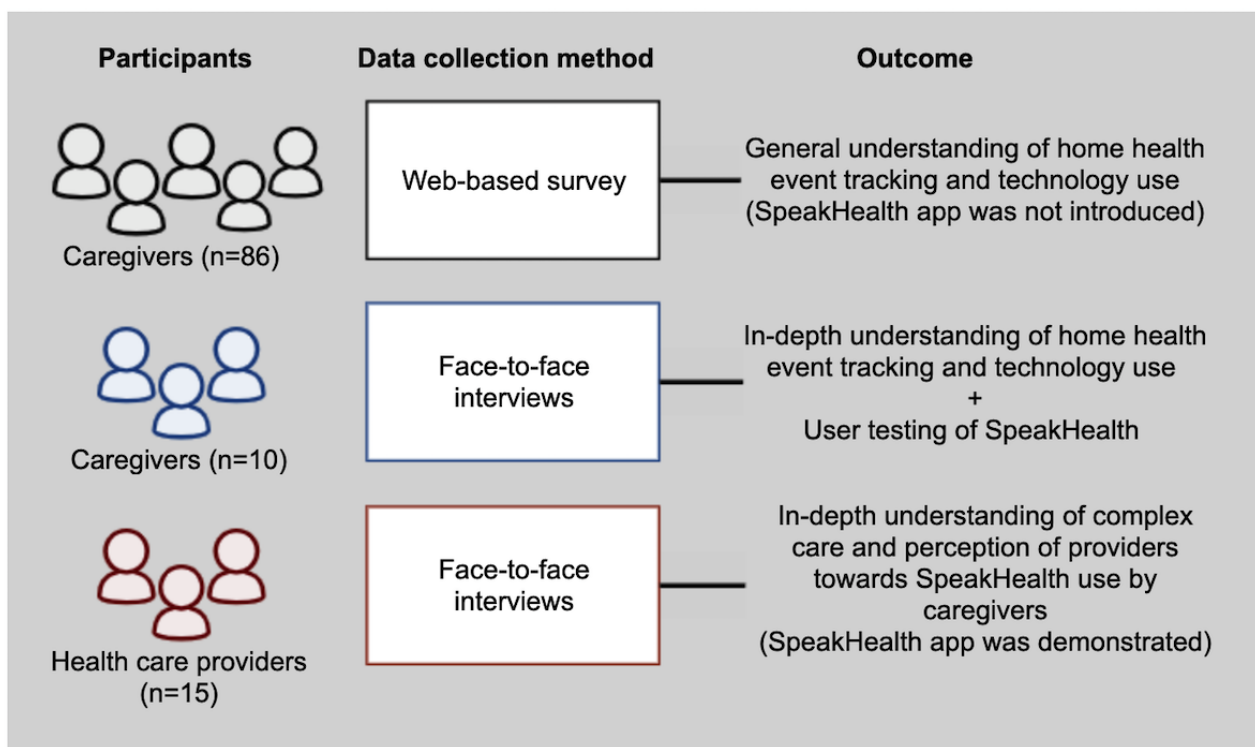
Data Collection

After caregivers were screened and they provided consent, they completed a web-based survey about demographics and technology awareness (June-August 2019). Subsequently, they were invited to a face-to-face interview. A semistructured interview framework was followed (a sample set of guiding questions of the interviews is available in [Multimedia Appendix 1](#)). The interview protocol focused on daily life and communication technology use, app testing scenarios, and user perception assessment questions. The user perception assessment questions were informed by the technology acceptance model [23]. Overall, 10 caregivers participated in interviews that took approximately 1 hour. Furthermore, 2 researchers participated in the interviews: one moderated the interviews and the other operated as a note taker. All interviews with caregivers were audio recorded. In addition, observational notes were taken during the interviews by the moderating researchers.

Along with the interviews, we assessed the awareness of caregivers regarding communication technology through a broader perspective by conducting a web-based survey. Overall, 86 participants completed the survey, which included the face-to-face interview participants.

HCPs were invited to a face-to-face interview where they were introduced to the project and the SpeakHealth app. The interviews were guided by two questions: (1) “What are your current experiences in complex care clinic?” (which focused on routine tasks, information needed, collected and missed information, and challenges) and (2) “How do you think the SpeakHealth app could help care management?” (which was used to understand if the information collected by the app fits the current care flow). Interviews with HCPs took approximately 40 minutes to complete. In total, 15 HCPs participated in the face-to-face interviews, which included physicians, nurses, care coordinators, and dietitians who have experience in the pediatric complex care clinic at the hospital where children or adolescents with chronic medical conditions (including neurodevelopmental disabilities and diseases affecting multiple organ systems) are being treated [24]. [Figure 1](#) summarizes the study design and data collection process.

Figure 1. Study design and data collection.

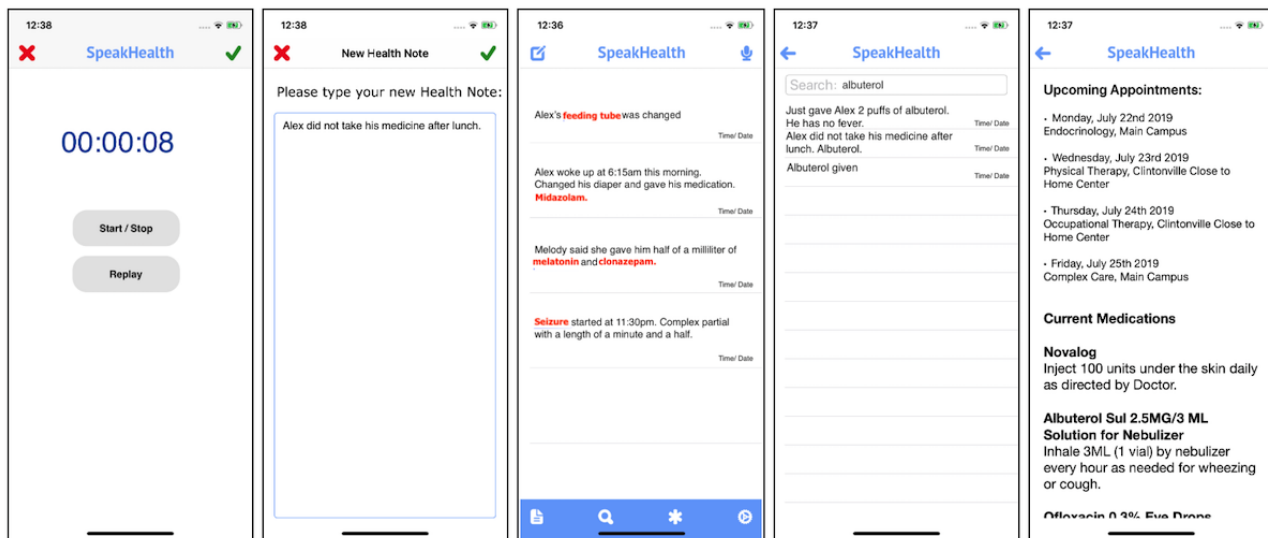


App Prototype Development and Testing

The app prototype was designed based on the feedback of a multistakeholder group (including physicians, nurses, caregivers, care coordinators, digital health scientists and clinical informaticians, developers, and designers) at our hospital [18]. The prototype ([Figure 2](#)) was developed using Swift as an Apple (iOS) mobile app [25]. We used Amazon Web Services (AWS)

application programmable interfaces (APIs) for the backend infrastructure and capabilities such as voice recognition and transcription (AWS Transcribe) [26]. The app leveraged existing voice technology components through iOS and AWS, and none of the voice interactive components or conversational agents were developed by the study team. The AWS infrastructure flowchart is presented in [Multimedia Appendix 2](#).

Figure 2. Screenshots of the SpeakHealth app. From left to right: voice interaction for note entry, note entry by text, notes overview and highlighted words, searching, and integration of medical information pulled from electronic health records.



Users were able to start the app through Siri (even when the phones were locked) by saying: “Hey Siri. Start SpeakHealth.” Then, SpeakHealth immediately started listening. After a speech segment was completed and collected, it was sent to AWS, transcribed, and sent back to the app and shared with the user. Users were also able to type notes or edit the transcription text.

As a part of the mock-up for planned features (not implemented with the prototype), users were able to see electronic health records (EHRs; eg, appointments and medications) and were able to access some of the prescribed medical care plans (eg, diabetes care guidelines) on the web. In addition, as a mock-up, important words and segments were highlighted in the notes (eg, medications and dosage). As they are not functional, we

explained these features verbally to obtain participants’ feedback. However, we plan to implement EHR integration through patient portals and available APIs through EHR system providers to pull medications, appointments, and prescribed care plans. Note that highlighting will be implemented through natural language processing methods leveraging medical vocabulary and ontologies (eg, SNOMED [Systemized Nomenclature of Medicine] [27]).

Textbox 1 outlines the functional features (participants were able to test through the app) and introduced features (participants were informed about). Participants followed predefined scenarios to test the app (**Textbox 2**). We plan to include all the introduced features in the final production of the app.

Textbox 1. Tested and introduced features of SpeakHealth app.

Functional features

- Voice-to-text note entry
- Text entry and editing
- Sham electronic health record (EHR) showing upcoming appointments and prescribed medications
- Preloaded information and guidelines
- Highlights of predefined notes

Introduced (mock-up) features

- Real-time note highlights
- Creating reports and sharing notes with health care providers
- Integration with EHRs
- Prescribed medical information and care plans

Textbox 2. App testing scenarios.**Scenario 0**

- This scenario is for the participants who did not use voice assistant before.
- Please say “Hey, Siri, how is the weather today?”, “How is traffic in downtown Columbus?”

Scenario 1

- Please say “Hey Siri, start SpeakHealth.” And, when you see the app running, say “just gave Alex 2 puffs of albuterol. He has no fever.”

Scenario 2

- Select one of the notes and edit the text. Then, please enter a new note with using “enter note” button, and type “Alex did not take his medicine after lunch.”

Scenario 3

- Check the last entered note. This is transcribed from your voice input. Please see the highlighted words. Those are identified as important to be shared with providers.

Scenario 4

- Use the search button to search for *albuterol* and check the results. You can see all your records where you administered albuterol.

Scenario 5

- See the user information content at the bottom left menu, find information about diabetes, and take a look at the content

Scenario 6

- See the medical record content at the bottom right menu and take a look at the content. Please find and check the next medical appointment date.

Data Analysis**Qualitative Analysis**

Audio recordings were transcribed verbatim and analyzed by the researcher, who moderated the interviews using Microsoft Office products. Observational notes taken during the interviews were used to supplement the transcripts, and these notes were incorporated into the transcriptions. Theoretical thematic analysis was used to systematically analyze the qualitative data [28]. Braun and Clarke [28] defined theoretical thematic analysis as a deductive method that is driven by the researcher’s theoretical or analytic interest in the research domain, and coding may focus on research questions.

In the initial step, the researchers read and reread the transcriptions to familiarize themselves with the data. During the reading iterations, instances were identified and labeled, and eventually coded. Codes were reviewed and gathered under potential themes, and an initial thematic map was created. Themes and codes were reviewed and the data were reviewed. The code groups and themes were revised, and the thematic map was updated. The codes and themes were reviewed by at least one researcher, and disagreements were resolved by a consensus among the researchers. Finally, the findings were reported and interpreted based on the literature findings. Coding and analysis of the qualitative data were completed using the Microsoft Office suite.

Quantitative Analysis

Survey responses were collected through the REDCap (Research Electronic Data Capture) web-based platform using Health Insurance Portability and Accountability Act–compliant secure services [29]. We completed a descriptive analysis of demographics and survey responses and conducted a correlational analysis. Data cleaning and verification were completed using Microsoft Excel, and the data were then analyzed using IBM SPSS software, version 26 (IBM Corporation).

Results**Survey Results**

A total of 86 parents responded to the web-based survey. The responses are presented in Table 1. The survey participants were parents of at least one CSHCN with an average age of 11 (SD 7.4) years. Most children were diagnosed with developmental delay, speech and vision problems, epilepsy, intellectual disability, and learning disability. Most of them used daily prescribed medications (72/86, 86%); attended occupational, behavioral, or speech therapy (62/86, 74%); and used medical devices such as a feeding tube (45/86, 53%). Parents preferred to use a notebook (30/86, 35%) or paper to take notes when tracking symptoms, health events, and activities (29/86, 34%). The use of generic note-taking mobile apps was relatively low (13/86, 15%). Parents mainly tracked appointments (77/86, 90%) and symptoms (49/86, 57%) and set up reminders (54/86, 63%). Most parents reported that they always (25/86, 30%) or often (21/86, 38%) tracked all symptoms and health events.

Table 1. Survey responses.^a

Variables	Values
Child's age (years; n=84)	
Range	0-31
Mean (SD)	11.1 (7.4)
Child's diagnosed conditions (n=86), n (%)	
Developmental delay	56 (65)
Speech problems	47 (55)
Vision problems	44 (51)
Seizures	38 (44)
Intellectual disability	37 (43)
Learning disability	36 (42)
Cerebral palsy	33 (38)
Joint or muscle problems	31 (36)
Anxiety problems	23 (27)
Hearing problems	23 (27)
Asthma	22 (26)
Behavioral problems	19 (22)
Brain injury	17 (20)
ADHD ^b	16 (19)
Asperger, autism spectrum	11 (13)
Depression	9 (10)
Diabetes	3 (3)
Other ^c	52 (60)
What treatments or devices does your child use? (n=84), n (%)	
Daily prescribed medications	72 (86)
Occupational, behavioral, or speech therapy	62 (74)
Feeding tube	45 (53)
Tracheostomy	18 (21)
Other ^d	49 (58)
How do you track symptoms, health events, and care activities at home currently? (n=85), n (%)	
Dedicated notebook	30 (35)
Notes on paper or cards	29 (34)
Health apps ^e	14 (16)
Generic note-taking app	13 (15)
I do not track	9 (10)
What kind of care activities do you do for tracking? (n=86), n (%)	
Following appointments	77 (90)
Setting up reminders	54 (63)
Tracking symptoms	49 (57)
Other ^f	17 (20)
How often do you track symptoms and events at home in a day? (n=82), n (%)	
I always track and record all symptoms and events	25 (30)

Variables	Values
I often track and record symptoms and events	31 (38)
I rarely track and record symptoms and events	16 (20)
I do not track and record symptoms and events	10 (12)
What do you think is the ideal tool or technology to use for tracking the symptoms and events at home? (n=85), n (%)	
Mobile phone and apps	65 (76)
Pen and paper or notebook	38 (45)
Tablet PC or iPad	31 (36)
Laptop or PC	19 (22)
Voice assistant (Amazon Alexa or Google Home)	13 (15)

^aResponses to "other" are grouped by the most encounters that are not listed on the existing responses.

^bADHD: attention-deficit/hyperactivity disorder.

^cChronic lung disease, rare genetic diseases (DDX3X and complex vertebral malformation), neurodevelopmental diseases (hydrocephalus and bilateral schizencephaly), metabolic disorders, and heart disease.

^dG tube, hearing aids, ventilation, ventriculoperitoneal shunt, wheelchair, nebulizer, suction machine, adaptive communication glucose and ketone monitor, and rescue inhaler.

^eCalendar, Dexcom, seizure app, MyChart, Apple Health, and reminder app.

^fTube feed, medications, urine and bowel movements, blood glucose levels, seizures, medical changes, new words, and vitals.

Parents were also asked to identify the ideal communication technology or tool that they would prefer to use for tracking symptoms and events at home (Table 2). In contrast to the common method that they used (dedicated notebook and notes on a paper), most of them preferred to use mobile phones and apps for tracking (65/86, 76%) in addition to pen and paper (38/86, 45%). These were followed by tablet PCs (31/86, 36%) and laptops or PCs (19/86, 22%). The number of parents preferring voice technology as an ideal tool for tracking symptoms was relatively low (13/86, 15%). Finally, parents

were asked to identify important features that a health app should have. Most of them found communication, sharing information, and tracking medication and appointments to be fairly important to very important (range 4.43-4.74). Health education and coaching for home care and voice interaction were found to be moderately important (range 3.28-3.72). Overall, parents found all the features to be important to very important and stated that such features would be a good fit to their current care routines and lifestyle (mean 4.6, SD 0.626).

Table 2. Ideal health app features.

How important is it for a health app to have the following features?	Participant, n (%)	Importance ^a , mean (SD)
Tracking medications	80 (93)	4.49 (0.994)
Tracking appointments	81 (94)	4.74 (0.787)
Hands-free and voice interactive engagement with the app	83 (97)	3.28 (1.355)
Keeping medical diary	83 (97)	4.25 (1.124)
Recording of audio or video to be shared with provider	83 (97)	3.96 (1.273)
Relevant health education or coaching contents for at-home care	83 (97)	3.72 (1.262)
Guidelines for coordinating care activities with other caregivers	82 (95)	4 (1.227)
Visualizing and summarizing the history of symptoms with statistics	81 (94)	4.22 (0.949)
Summary of home events between clinical visits to share with the provider	83 (97)	4.47 (0.738)
Interacting with personal health records at your provider (for instance, using the MyChart app)	82 (95)	4.57 (0.754)
Providing a communication channel to connect to the providers	82 (95)	4.43 (0.903)
Regarding your current lifestyle, habits and routines, what would be the impact of such a tool or technology if it had the features you ranked as <i>important</i> or above	82 (95)	4.6 (0.626)

^a1: not at all important; 2: slightly important; 3: important; 4: fairly important; and 5: very important.

Correlational Analysis

We used the Spearman nonparametric correlation test to explore if health condition, treatment, symptom tracking activity,

tracking frequency, and expected app features were correlated. Our hypothesis was that having more health conditions and more treatments may positively correlate with tracking more health activities and an increase in the frequency of tracking.

Similarly, they all may positively correlate with parents' perceptions of more significant features in an ideal app. The variables tested for correlation were the child's age, health condition (aggregated score), treatment (aggregated score), tracked activities (aggregated score), tracking frequency, and expected features (aggregated score). The missing data were excluded from pairwise deletion. However, we were unable to identify any strong correlations. There was a weak and positive correlation between the number of diagnosed conditions and total treatment ($r=0.263$; $P=.02$), tracking activities and total treatment ($r=0.226$; $P=.04$), and the child's age and number of diagnosed conditions ($r=0.247$; $P=.02$).

Interview Results

The interview participants were predominantly young (25-34 years: 4/10, 40%; 35-44 years: 4/10, 40%; and 45 or older: 2/10, 20%) and were predominantly part of a two-parent household (7/10, 70%) and White (10/10, 100%). Most households had 1 to 2 children (7/10, 70%), at least one of which had special needs (mean age 10, SD 8 years). The children had conditions with special health care needs for more than a year. Most children had multiple conditions. These conditions included the following: learning disability (4/10, 40%), speech problems (7/10, 70%), ADHD (1/10, 10%), depression (1/10, 10%), anxiety (3/10, 30%), diabetes (1/10, 10%), autism spectrum disorder (2/10, 20%), seizures (4/10, 40%), behavioral problems (2/10, 20%), hearing problems (1/10, 10%), developmental delay (6/10, 60%), vision problems (4/10, 40%), intellectual disability (6/10, 60%), joint or muscle problems (4/10, 40%), cerebral palsy (4/10, 40%), brain injury (3/10, 30%), genetic syndrome (2/10, 20%), and nonalcoholic fatty liver disease (1/10, 10%).

The children received daily prescribed medications (9/10, 90%), behavioral or speech therapy (6/10, 60%), feeding tube (5/10, 50%), tracheostomy (1/10, 10%), physical and occupational therapy (2/10, 20%), ventriculoperitoneal shunt (1/10, 10%), vagal nerve stimulator (1/10, 10%), and baclofen pump (1/10, 10%). Half of the parents often reported tracking health events at home, including activities such as tracking symptoms (8/10, 80%), follow-up appointments (10/10, 100%), and setting up reminders (6/10, 60%).

Themes

Summary of Themes

We identified the following two overarching themes: *enablers and barriers in care communication* and *enablers and barriers in communication technologies*. These themes represent the perspectives of parent roles, needs, and communication dynamics with or without the technology component. The first theme, *enablers and barriers in care communication*, outlined the communication perspective in two subthemes. The first subtheme, *parent roles in care*, was an *enabler* theme, and it outlined the state of the parent-child relationship in terms of care delivery and management. This study aimed to understand the dynamics between parents and children. The second subtheme, *care communication challenges*, was a *barrier*, and it covered challenges encountered during child care. The second theme, *enablers and barriers in communication technologies*,

outlined the communication perspective with the presence of technology in two subthemes. The first subtheme, *communication technology use in care*, was an *enabler*, and it explained the use of reported communication technologies (including voice-enabled technology) by caregivers. The second subtheme, *communication technology limitations*, was a *barrier*, and it was about the limitations of reported communication technologies (including voice-enabled technology).

Theme 1: Enablers and Barriers in Care Communication

Subtheme 1.1: Parent Roles in Care (Enablers)

Parents play a major role in the hands-on care of their children. They reported having scheduled tasks throughout the day, including feeding; administering medications; assisting in training and education; and scheduling medical care services, appointments, and activities. Most of the parents tracked medication supply, food intake, therapies, and medical appointments and set up reminders for recurring activities such as medication administration:

So, I usually get up in the morning and make sure her pills are out for her to take as she wakes up with some sort of food, because if she'll take them on an empty stomach and then get sick.... [Parent 1]

...pretty much his biggest area is development, so just get on to therapies and things like that. Tracking just progress and you know, how much time in therapy we did each day is what we do. [Parent 2]

For some parents, the frequency of each care activity and time spent may differ depending on the child's condition. Some parents may need more guidance, assistance, and care coordination than others. For instance, some parents may require more time to prepare for feeding a child through a feeding tube. Other parents may need to help their children with after-school physical exercise. Similarly, a child's age can potentially affect the effort dedicated by the parents. Children might be more independent regarding self-care as they age. However, their medical conditions may still affect their independence:

...we're up at seven with my daughter. We hook her up to a feeding pump through a feeding tube and she gets several medicines in the morning...she gets four feeds throughout the day about every three hours. She also receives afternoon meds about three o'clock with her feed. At night, she gets several more medications with her evening feed. [Parent 3]

Subtheme 1.2: Care Communication Challenges (Barriers)

Depending on the child's condition and severity, families may encounter challenges that could lead to frustration and stress for the family. One of the overarching challenges is the *frequent need for medical communication*. Parents may need to communicate with multiple providers in a day and repeat the same information in each call:

He (child) probably has a dozen different care providers. So it's never ending talking to all of them, keeping them up to date. [Parent 6]

(Calling HCPs)...maybe a half dozen on average a week. However, there are some days, especially when

he's (child) sick, or we are transitioning or something. It (calling HCPs) could be six per day. I'd say six in a week to six in a day. Those busy days also involve like supplies and stuff too when we are calling to make sure food gets here and whatever. [Parent 6]

In addition, parents reported the *difficulty of tracking* and the constant need to remember everything related to the child's care. They may keep copies of health records and periodically keep notes on symptoms and medications. However, this may create some burden, and some information could still be missed. Parents come up with their own strategies to deal with this; however, the inefficiency of the tools used and methods for tracking and using information in communication were common challenges for parents:

...every time she eats she has to bolus for something or you know, every night she has to take a shot for a different type of shot to last throughout the day. So it's just being able just to record quickly, we can record the sugars, but we can't record how much we take. [Parent 5]

Similar concerns were expressed by separated parents in tracking and sharing a child's condition. Two of the parents, who were divorced, underlined the struggle of keeping all information at one place, as their child was not always staying at one home:

...me and her father are divorced. So if there was a way for us to both be able to log in and while she was gone, you know. If she complained of something, or if he had to take her to the ER, and I wasn't there right away just to have who her doctors are, what her meds are, like basic major things that we could at least view as well as my husband. So kind of just like having that like information across the parents. [Parent 7]

Parents understand the need and benefit of care coordination and communication; however, *gathering the information* to be communicated, such as the child's health care notes and activities, was found to be difficult to keep up with:

...just tracking illness and stuff is just very difficult between different caregivers, different areas that he's in. Care coordination is great in the fact that it gives appointments scheduled, but actually tracking symptoms of care is pretty difficult. [Parent 10]

Children who need care assistance are not always compliant or cooperative with their caregivers. They may have trouble *gaining self-management skills* or may prefer not to keep track of health events and act based on personal preferences instead of meeting essential care needs, especially during the adolescent years:

They [HCPs] want to know so much information...telling a 14-year-old to write stuff down on paper. It's not gonna happen. [Parent 5]

Theme 2: Enablers and Barriers in Communication Technologies

Subtheme 2.1: Communication Technology Use in Care (Enablers)

Almost all parents had tablet computers or laptops with internet access at home; however, smartphones were the most frequently used personal technology. A couple of children used medical-grade technologies, such as glucose monitoring (Dexcom) and insulin devices (Omnipod), that parents are able to track through apps. However, some parents may prefer to use a pen and paper to keep track of health activities, such as medication refills, therapy appointments, and symptoms. This method is also regarded as convenient, with the use of color codes to follow multiple tasks. One of the methods to share care notes with other family members and home nurses is to use a calendar on a fridge:

Usually taking notes on calendars like...the one in the fridge. Sometimes I've heard her she feels weird or something write it on the fridge calendar. [Parent 4]

Some parents mentioned the use of consumer-grade apps for tracking and note-taking, such as calendar apps and pharmacy apps. One parent kept a Microsoft Excel sheet to track and remember the data. As an mHealth app, some parents used MyChart (patient portal and mobile app of Epic EHR system) to check appointments, send messages, and check laboratory test results. Parents usually find apps through word of mouth, social media communities, and web-based searches. Similarly, information-seeking activities for child health were performed through web searches and community social media pages:

...have some of the Facebook groups that parent groups...what apps work for them? You know, and then that's pretty much how I go about figuring things out. Sometimes I'll go and look just like an app store search for some...[Parent 2]

there's a lot of that diabetic Facebook groups that I'm in...I'm a kind of like a reader...[Parent 9]

With older children, parents expected their children to gain self-management skills. Parents envision children's involvement in their medical care to be maintained with current communication technologies:

She (child) spends most of her time on her laptop, and her phone. I use my phone and laptop mostly. She also has an Apple Watch. And since she's older, I think she could somehow be able to just "oh, I feel a weird" thing and tap on something and she also could record information. [Parent 1]

Voice Interactions

Almost half of the parents reported that they were aware of and felt positive about using voice assistants via smartphones. Parents consider a mobile device to be an ideal tool for voice interaction and, when necessary, to track health events. Currently, some parents use voice interaction for basic tasks such as searching, sending reminders, using commands during driving, and using it for entertainment. Voice interaction was

perceived as easier to use, especially for searching, setting up timers, or playing music. The use of voice in health communication received mixed feedback. Voice, as an enabler, was found to be a promising medium to receive medical information together as a family, instead of one parent taking the lead to get the information and share the child's medical updates with the family. At home, medical communication through voice technology was not perceived as having a privacy risk:

...a lot of times what it is now is I'll go look up something in my chart, look out for what the result is. I'll go home tell whoever it is...tell her (child) about what they found or what they didn't find and then tell what they said what they found...So it's like cutting down like layers of happening to say the same information. Multiple times, if everyone's sitting there they can all hear it together. And then that way, we are all on the same page at the same time together instead of staggered. [Parent 9]

Subtheme 2.2: Communication Technology Limitations (Barriers)

Even though communication technologies have the potential to improve daily lives, finding the right one for an individual and their family, especially to meet the specific needs of health care could be burdensome. Parents mentioned problems in finding the right app, learning it, and using it regularly. Some parents stated that they did not have time to do so. In some cases, parents may have their own technological strategies:

I had to buy a little Bluetooth keyboard to go with the phone for when I do stuff on the phone, or the iPad. I was like, I gotta take this out. I can't use my thumbs for everything. [Parent 8]

Conversely, apps being used may not meet parents' needs. One parent mentioned communicating digital notes was difficult because sending notes to the doctors from an app was troublesome. MyChart meets the basic needs, but parents may not *rely on it*. Some of the reasons were the issues with reminders and laboratory results, messaging and access limitations, and issues with reaching the messaging history and searching for notes. However, the reason may also be that parents may not know the functionalities of the app:

...I noticed sometimes when they're (children) hospitalized, a lot of information doesn't make it into MyChart anyhow. I don't know why. [Parent 4]

I feel like it's like the one of the problems with MyChart is we have very little control. Yeah, I mean, would like to see there.... [Parent 1]

...here's his certain idiosyncrasies of this thing or this process. If care providers can receive those notes in a centralized way. Okay, maybe MyChart already does that? I don't know. [Parent 6]

Limitations of Voice Interaction

Even though voice interaction was found to be promising, the need for a network connection to use voice interaction may be problematic. One parent mentioned their concern with poor

internet connection in rural areas. Other concerns with voice interaction at home were proximity to the device (eg, being in the other room) and loud environments such as screaming children, playing instruments, or watching TV, which may disrupt voice interaction. In addition, individual technology interaction styles may differ and affect the extent of the use of voice, for example, one parent could prefer an auditory interaction, whereas another could prefer visual feedback.

Conversely, in terms of promoting a child's self-management, the voice assistant's capability in understanding commands was reported to be limited, especially in children with speech disorders:

...she has Down syndrome. Her speech is clear enough for you to understand but Siri or Alexa has a hard time understanding some things. [Parent 4]

App Feedback

Following the predefined scenarios (Textbox 2), participants tested the SpeakHealth app prototype and provided feedback. Parents mostly found voice-enabled interactive engagement to be useful for hands-free note-taking compared with typing. In particular, for parents who were not able to keep up with tracking health activities and taking notes, voice interaction was found to be promising to keep notes for use in care follow-ups:

...especially when sick, they want to know, okay, what is your sugar, when did she bolus and how much did she give? You know, what did she eat? And what time did they check it again? And how much [insulin] did you give again, and what's your correction factor and all this other stuff. So, I mean, for us it (voice interaction) would be a total game changer because then we could just easily keep track because we are poor paper. We are poor paper charters. We just can't do it. I don't know why it's just it's so taxing to sit down and try to write down just all throughout the day, every day three o'clock in the morning, you're not going to remember to write.... [Parent 5]

The main concerns about communicating personal health information through voice were as follows: (1) talking and listening out loud in public places and (2) voice recognition issues with understanding voice instructions and commands, especially with uncommon or medical words. An example of a misunderstanding of a command was experienced during the testing. While parents were narrating a medication name, *Albuterol* (an asthma medicine), it was understood as *computer-all* in two instances. However, interestingly, transcription errors were found to be acceptable as long as they were minimum:

...with using voice command, it's often a matter of like, it seemed like the benefit outweighed the risk of it not hearing your particular word...I will use voice over typing it (note) [Parent 6]

Parents also provided their feedback for the SpeakHealth app and voice interaction regarding how they would envision it to be used while caring for their children. Table 3 presents the breakdown of parents' feedback regarding app and voice interaction.

Table 3. Parents' feedback about the SpeakHealth app.

Feedback groups	App or voice interaction	Feedback summary
Sharing information	App	Parents like the idea of sharing notes with providers and significant others as well as easily notifying all health care providers in their circle in case of emergency. Shared accounts for everyone to access and edit notes received positive feedback.
Education	App	Educational links were found to be useful. Providing guidelines for first-time parents were suggested to be highly needed.
Integration	App	Electronic health record integration was well received. In addition, calendar integration was suggested. Parents would like to have more control over patient medical records in terms of access and transparency of all the notes.
Visualization	App	Parents would like to have a personalized dashboard, such as for epilepsy, seizure activities should be more visible. The highlighting ability was found to be useful.
Accessibility	Voice	Parents liked the idea of using voice while the phone was away and locked to enter notes. A parent suggested having the ability to navigate and search notes with voice.
Real-world use	Voice	Parents shared real-world scenarios to be used regarding their children's condition. These include seizure timer, reminder of medications and appointments, searching medical notes, calling providers, and feeding tracking.
Voice-over typing	Voice	Parents stated that voice interaction helps to enter notes easier and faster than typing. Misunderstandings are correctable through text editing. Spell check feature for transcription was suggested for auto-correction.
More with voice: navigation	Voice	A parent mentioned navigation of notes with voice, search with voice.

HCP Interviews

Most HCPs were pediatric physicians and nurses who were experienced complex care professionals. Table 4 summarizes HCP feedback.

Physicians in pediatric complex care frequently keep themselves up to date about their patient's condition and maintain ongoing communication with each of the patients and their families. This requires them to periodically review a patient's previous clinical notes, especially before visits. After visits, physicians may need 10-60 minutes to write patient visit notes. Physicians valued patient personal notes (eg, medical diaries and at-home care activities). They usually receive this information during visits or through phone triage notes. However, the integration of patient personal notes with the EHR requires a systematic effort to make information useful to physicians. In that regard, physicians would like to have this information *digestible* and *timely*, which means the presentation of the information should be succinct, organized, and accessible when it is needed. In addition, nurses emphasized communication efficiency. During periodic visits, emergency visits, or phone triage, the nurses need to receive information about the child's condition and medication. This requires repeated effort from both sides (nurses and caregivers). Therefore, the solution is expected to make information exchange easier and more accurate, including information such as home care details, medication intake, and medication names. Similar feedback was provided by care coordinators, underlining the need for integrated medical information collection and sharing.

In the interviews, one example brought up by HCPs was seizure events. For tracking seizures, the patient- or caregiver-reported data are crucial in clinical decision-making. The frequency, duration, date, and trend over time are important information that must be accurately shared with HCPs. Therefore, keeping the information timely, categorizing by date, visualization with occurrence, duration, and time are some of the expectations from the patient data that are suggested for the SpeakHealth app. HCPs would also need to send a patient seizure action plan back to the caregiver to guide them in case of emergency.

Another example is gastrointestinal disorders in patients diagnosed with cystic fibrosis. Regarding the severity of the condition, significant time can be spent on the phone for guidance; information sharing; and collection of patient condition; feeding tube practice; and getting updates on symptoms, food intake, and bowel movements. A medical team, including nurses, physicians, dietitians, and pharmacists, have roles in clinical decision-making. Health events may occur multiple times a day, and the timely capture of patient personal notes can facilitate the decision-making process in terms of providing more accurate and complete data for assessment. For example, infant data information is collected through parent reports. Eventually, this information is entered into the EHR through necessary forms and templates (eg, nutrition plan and intake form) during triage or visits.

The SpeakHealth app was found to be promising for timely information collection that could be available for sharing with HCPs. It is suggested to have the ability to send reminders to parents for taking personal notes and eventually mapping these notes to the EHR.

Table 4. Health care provider feedback.

Profession	Expertise or specialty	Key takeaway
Physicians (n=7)	Internal medicine, neurodevelopmental disabilities, emergency medicine in pediatrics, gastroenterology, hepatology, nutrition, and pulmonary medicine	<ul style="list-style-type: none"> Organized, curated, categorized Visualization of the data Using patient notes into clinical notes Accessible over electronic health records
Nurses (n=4), care coordinators (n=2), clinical dietitian (n=1), and patient care pharmacist (n=1)	Complex care, pulmonary, and pediatric care	<ul style="list-style-type: none"> Medication adherence tracking and reminders Efficient communication and data sharing Reducing missing information

Discussion

Principal Findings

Our study demonstrated that caregivers of CSHCN may engage in different care activities and have different care management strategies depending on their child's condition. Similar to the findings by Golden et al [7], caregivers desire better information sharing and communication with HCPs. Health care and communication technologies could potentially improve care communications and coordination [10]; however, a systematic approach is required to enable information to be conveniently captured at home and shared among caregivers and HCPs.

Our findings suggest that current communication technologies, especially mobile apps, could help to meet the fundamental care communication and coordination needs of at-home care, such as tracking medical symptoms, keeping up with medications and appointments, taking medical notes, accessing medical records, and communicating with HCPs. In line with this, voice interaction is perceived to be useful in taking medical notes. Even though voice technology adoption is still in the early stages of adoption by parents [30,31] and in health care [19], the use of voice in tandem with other methods (eg, audiovisual feedback through mobile phone enabling text entry) would increase voice technology adoption. As our findings suggest, taking medical notes is a significant part of care activity, and such efforts can facilitate care coordination activities by collecting necessary health information to be communicated with HCPs [7]. In that regard, having a medical diary conveniently at home is important as a part of the care strategy of parents. Considering the care activities that may keep a caregiver's hands full, voice interaction can enhance and facilitate convenient note-taking and health event tracking. Currently, voice assistants (eg, Google Assistant, Amazon Alexa, and Apple Siri) running on smartphones and smart speakers (eg, Google Home and Amazon Echo) may still need improvement in the comprehension of medical terminology [20,32], health information exchange [19,32,33], and privacy and security measures [34] to provide health communication services.

Convenient note-taking and health event tracking are important; thus, having a shared platform among caregivers and HCPs is a significant step toward enabling patient note use (by multiple caregivers and providers) in shared clinical decision-making and improving quality of care. HCPs emphasized the value of

timely health information collection, making it convenient for review through integration with medical records. Apps that are interoperable and designed to fit in the daily lifestyle of families and the clinical workflow of HCPs would help close the gap between communication and care coordination [13,17]. In the literature, integration of patient notes with the EHR has also been suggested to improve health outcomes [35]. Similarly, the OurNotes initiative (as a part of OpenNotes) values the ability of patients and caregivers to contribute to health records [36]. Therefore, the value and impact of medical notes kept at home would increase as they are shared with HCPs.

Limitations and Future Studies

Our study has several limitations. We were not able to identify condition-specific needs or patient-provider experiences, but we were able to capture a broader perspective of perceptions. In addition, the study did not focus on the experiences of participants from different socioeconomic statuses or demographics. The participants of interviews and surveys were limited in number, and they were residents of Ohio, which may not be representative of the population as a whole. We tested a prototype during the interview sessions; therefore, we were not able to capture real-world evidence of app use. The accuracy of speech to text, performance of voice assistant, and their impact on interactions were not assessed systematically but planned for future phases. Future studies are planned to address assessing the usability and feasibility in real-world settings with a larger user group and investigating condition-specific needs. In addition, investigating the health system and implementation pathways would help to understand the feasibility of integration in the current health system. Such efforts will fill the gap in the health system and in the literature on medical note-taking and remote health tracking [10]. In addition, it may improve care coordination in terms of reducing the time spent on documentation, meeting care coordination needs, long-term care planning, and improving shared decision-making [2,13].

Conclusions

In this paper, we present our findings on care management, coordination, and use of communication technology by caregivers at home, and we tested a voice-enabled medical diary app. Our study extends the literature in terms of understanding caregiver technology expectations and needs, as well as perceptions of HCPs. We believe the findings could further

inform researchers and developers and technology policy makers in shared decision-making and patient-reported outcomes.

Acknowledgments

The authors would like to acknowledge Lynne Fogel from OhioF2F for her support and help in caregiver recruitment; Tyler Lieser for his help in caregiver interviews; Jennifer Lee and Samuel Yang for their constructive feedback; Matt Bailey, Brannon Oiler, Brandon Abbott, and Dan Digby for their contribution to app development; and the advisory panel members of the SpeakHealth project for their support. This work was supported by the Health Resources and Services Administration Maternal and Child Health Bureau Grand Challenge for Care Coordination for CSHCN (grant 720467022000). The authors thank Carrie Robinson for proofreading the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Interview guidelines.

[PDF File (Adobe PDF File), 22 KB - [formative_v5i5e25503_app1.pdf](#)]

Multimedia Appendix 2

SpeakHealth Amazon Web Services structure.

[PDF File (Adobe PDF File), 267 KB - [formative_v5i5e25503_app2.pdf](#)]

References

1. McPherson M, Arango P, Fox H, Lauver C, McManus M, Newacheck PW, et al. A new definition of children with special health care needs. *Pediatrics* 1998 Jul;102(1 Pt 1):137-140. [doi: [10.1542/peds.102.1.137](#)] [Medline: [9714637](#)]
2. The national survey of children with special health care needs: chartbook 2009-2010. U.S. Department of Health and Human Services Health Resources and Services Administration. 2013. URL: <https://mchb.hrsa.gov/sites/default/files/mchb/Data/NSCH/nscshcn0910-chartbook-jun2013.pdf> [accessed 2021-01-01]
3. Cohen E, Kuo DZ, Agrawal R, Berry JG, Bhagat SK, Simon TD, et al. Children with medical complexity: an emerging population for clinical and research initiatives. *Pediatrics* 2011 Mar;127(3):529-538 [FREE Full text] [doi: [10.1542/peds.2010-0910](#)] [Medline: [21339266](#)]
4. Council on Children with Disabilities Medical Home Implementation Project Advisory Committee. Patient- and family-centered care coordination: a framework for integrating care for children and youth across multiple systems. *Pediatrics* 2014 May;133(5):1451-1460. [doi: [10.1542/peds.2014-0318](#)] [Medline: [24777209](#)]
5. Kodner DL, Spreeuwenberg C. Integrated care: meaning, logic, applications, and implications--a discussion paper. *Int J Integr Care* 2002;2:e12 [FREE Full text] [doi: [10.5334/ijic.67](#)] [Medline: [16896389](#)]
6. Antonelli R, McAllister J, Popp J. Making care coordination a critical component of the pediatric health system: a multidisciplinary framework. *The Commonwealth Fund*. 2009. URL: <https://www.commonwealthfund.org/publications/fund-reports/2009/may/making-care-coordination-critical-component-pediatric-health> [accessed 2021-01-01]
7. Golden SL, Nageswaran S. Caregiver voices: coordinating care for children with complex chronic conditions. *Clin Pediatr* 2012 Aug;51(8):723-729. [doi: [10.1177/0009922812445920](#)] [Medline: [22563062](#)]
8. McAllister JW, Presler E, Cooley WC. Practice-based care coordination: a medical home essential. *Pediatrics* 2007 Sep 31;120(3):723-733. [doi: [10.1542/peds.2006-1684](#)] [Medline: [17766512](#)]
9. Chew C, Hss AS, Chan H, Hassali MA. Medication Safety at Home: A Qualitative Study on Caregivers of Chronically Ill Children in Malaysia. *Hosp Pharm* 2020 Dec 06;55(6):405-411. [doi: [10.1177/0018578719851719](#)] [Medline: [33245714](#)]
10. Ranade-Kharkar P, Weir C, Norlin C, Collins SA, Scarton LA, Baker GB, et al. Information needs of physicians, care coordinators, and families to support care coordination of children and youth with special health care needs (CYSHCN). *J Am Med Inform Assoc* 2017 Sep 01;24(5):933-941. [doi: [10.1093/jamia/ocx023](#)] [Medline: [28371887](#)]
11. Sezgin E, Weiler M, Weiler A, Lin S. Proposing an ecosystem of digital health solutions for teens with chronic conditions transitioning to self-management and independence: exploratory qualitative study. *J Med Internet Res* 2018 Sep 06;20(9):e10285. [doi: [10.2196/10285](#)]
12. Looman WS, Antolick M, Cady R, Lunos S, Garwick A, Finkelstein S. Effects of a telehealth care coordination intervention on perceptions of health care by caregivers of children with medical complexity: a randomized controlled trial. *J Pediatr Health Care* 2015;29(4):352-363 [FREE Full text] [doi: [10.1016/j.pedhc.2015.01.007](#)] [Medline: [25747391](#)]
13. Desai AD, Wang G, Wignall J, Kinard D, Singh V, Adams S, et al. User-centered design of a longitudinal care plan for children with medical complexity. *J Am Med Inform Assoc* 2020 Dec 09;27(12):1860-1870. [doi: [10.1093/jamia/ocaa193](#)] [Medline: [33043368](#)]

14. Fiks AG, DuRivage N, Mayne SL, Finch S, Ross ME, Giacomini K, et al. Adoption of a portal for the primary care management of pediatric asthma: a mixed-methods implementation study. *J Med Internet Res* 2016 Jun 29;18(6):e172 [FREE Full text] [doi: [10.2196/jmir.5610](https://doi.org/10.2196/jmir.5610)] [Medline: [27357835](https://pubmed.ncbi.nlm.nih.gov/27357835/)]
15. Desai AD, Jacob-Files EA, Wignall J, Wang G, Pratt W, Mangione-Smith R, et al. Caregiver and health care provider perspectives on cloud-based shared care plans for children with medical complexity. *Hosp Pediatr* 2018 Jul 05;8(7):394-403 [FREE Full text] [doi: [10.1542/hpeds.2017-0242](https://doi.org/10.1542/hpeds.2017-0242)] [Medline: [29871887](https://pubmed.ncbi.nlm.nih.gov/29871887/)]
16. Gentles S, Lokker C, McKibbin K. Health information technology to facilitate communication involving health care providers, caregivers, and pediatric patients: a scoping review. *J Med Internet Res* 2010 Jun 18;12(2):e22 [FREE Full text] [doi: [10.2196/jmir.1390](https://doi.org/10.2196/jmir.1390)] [Medline: [20562092](https://pubmed.ncbi.nlm.nih.gov/20562092/)]
17. Baysari MT, Westbrook JI. Mobile applications for patient-centered care coordination: a review of human factors methods applied to their design, development, and evaluation. *Yearb Med Inform* 2018 Mar 10;24(01):47-54. [doi: [10.15265/iy-2015-011](https://doi.org/10.15265/iy-2015-011)]
18. Sezgin E, Noritz G, Elek A, Conkol K, Rust S, Bailey M, et al. Capturing at-home health and care information for children with medical complexity using voice interactive technologies: multi-stakeholder viewpoint. *J Med Internet Res* 2020 Feb 13;22(2):e14202 [FREE Full text] [doi: [10.2196/14202](https://doi.org/10.2196/14202)] [Medline: [32053114](https://pubmed.ncbi.nlm.nih.gov/32053114/)]
19. Sezgin E, Militello L, Huang Y, Lin S. A scoping review of patient-facing, behavioral health interventions with voice assistant technology targeting self-management and healthy lifestyle behaviors. *Transl Behav Med* 2020 Aug 07;10(3):606-628. [doi: [10.1093/tbm/ibz141](https://doi.org/10.1093/tbm/ibz141)] [Medline: [32766865](https://pubmed.ncbi.nlm.nih.gov/32766865/)]
20. Palanica A, Thommandram A, Lee A, Li M, Fossat Y. Do you understand the words that are comin outta my mouth? Voice assistant comprehension of medication names. *NPJ Digit Med* 2019 Jun 20;2(1):55 [FREE Full text] [doi: [10.1038/s41746-019-0133-x](https://doi.org/10.1038/s41746-019-0133-x)] [Medline: [31304401](https://pubmed.ncbi.nlm.nih.gov/31304401/)]
21. Cheng A, Raghavaraju V, Kanugo J, Handrianto Y, Shang Y. Development and evaluation of a healthy coping voice interface application using the Google home for elderly patients with type 2 diabetes. In: *Proceedings of the 15th IEEE Annual Consumer Communications & Networking Conference (CCNC)*. 2018 Presented at: 15th IEEE Annual Consumer Communications & Networking Conference (CCNC); January 12-15, 2018; Las Vegas, NV, USA. [doi: [10.1109/CCNC.2018.8319283](https://doi.org/10.1109/CCNC.2018.8319283)]
22. Creswell JW, Creswell JD. *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. Thousand Oaks, California, United States: SAGE Publications; 2018:1-304.
23. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Q* 1989 Sep;13(3):319-340. [doi: [10.2307/249008](https://doi.org/10.2307/249008)]
24. Complex health care program. Nationwide Children's Hospital. URL: <https://www.nationwidechildrens.org/specialties/complex-care> [accessed 2021-01-01]
25. Swift. Apple Developer. URL: <https://developer.apple.com/swift/> [accessed 2021-01-01]
26. Amazon Web Services (AWS). URL: https://aws.amazon.com/?nc2=h_lg [accessed 2021-01-01]
27. SNOMED CT. National Center for Biomedical Ontology (NCBO) BioPortal. 2021. URL: <https://bioportal.bioontology.org/ontologies/SNOMEDCT> [accessed 2021-01-01]
28. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol* 2006 Jan;3(2):77-101. [doi: [10.1191/1478088706qp0630a](https://doi.org/10.1191/1478088706qp0630a)]
29. Project REDCap. URL: <https://projectredcap.org/> [accessed 2021-01-01]
30. Kleinberg S. Why parents are voice-assistance power users. Think with Google. 2018. URL: <https://www.thinkwithgoogle.com/consumer-insights/voice-assistance-parent-users/> [accessed 2021-01-01]
31. Ilievski A, Dojchinovski D, Gusev M. Interactive voice assisted home healthcare systems. In: *Proceedings of the 9th Balkan Conference on Informatics*. 2019 Presented at: BCI'19: 9th Balkan Conference in Informatics; September, 2019; Sofia Bulgaria. [doi: [10.1145/3351556.3351572](https://doi.org/10.1145/3351556.3351572)]
32. Yang S, Lee J, Sezgin E, Bridge J, Lin S. Clinical advice by voice assistants on postpartum depression: cross-sectional investigation using Apple Siri, Amazon Alexa, Google Assistant, and Microsoft Cortana. *JMIR Mhealth Uhealth* 2021 Jan 11;9(1):e24045 [FREE Full text] [doi: [10.2196/24045](https://doi.org/10.2196/24045)] [Medline: [33427680](https://pubmed.ncbi.nlm.nih.gov/33427680/)]
33. Sezgin E, Huang Y, Ramtekkar U, Lin S. Readiness for voice assistants to support healthcare delivery during a health crisis and pandemic. *NPJ Digit Med* 2020 Sep 16;3(1):122 [FREE Full text] [doi: [10.1038/s41746-020-00332-0](https://doi.org/10.1038/s41746-020-00332-0)] [Medline: [33015374](https://pubmed.ncbi.nlm.nih.gov/33015374/)]
34. Hoy MB. Alexa, Siri, Cortana, and more: an introduction to voice assistants. *Med Ref Serv Q* 2018;37(1):81-88. [doi: [10.1080/02763869.2018.1404391](https://doi.org/10.1080/02763869.2018.1404391)] [Medline: [29327988](https://pubmed.ncbi.nlm.nih.gov/29327988/)]
35. Benhamou PY. Improving diabetes management with electronic health records and patients' health records. *Diabetes Metab* 2011 Dec;37 Suppl 4:53-56. [doi: [10.1016/S1262-3636\(11\)70966-1](https://doi.org/10.1016/S1262-3636(11)70966-1)] [Medline: [22208711](https://pubmed.ncbi.nlm.nih.gov/22208711/)]
36. Mafi JN, Gerard M, Chimowitz H, Anselmo M, Delbanco T, Walker J. Patients contributing to their doctors' notes: insights from expert interviews. *Ann Intern Med* 2017 Nov 14;168(4):302-305. [doi: [10.7326/m17-0583](https://doi.org/10.7326/m17-0583)]

Abbreviations

API: application programmable interface
AWS: Amazon Web Services
CSHCN: children with special health care needs
EHR: electronic health record
HCP: health care provider
NCH: Nationwide Children's Hospital
REDCap: Research Electronic Data Capture

Edited by G Eysenbach; submitted 04.11.20; peer-reviewed by J Piera-Jiménez, T Jadczyk; comments to author 23.11.20; revised version received 10.02.21; accepted 17.04.21; published 11.05.21.

Please cite as:

Sezgin E, Noritz G, Lin S, Huang Y

Feasibility of a Voice-Enabled Medical Diary App (SpeakHealth) for Caregivers of Children With Special Health Care Needs and Health Care Providers: Mixed Methods Study

JMIR Form Res 2021;5(5):e25503

URL: <https://formative.jmir.org/2021/5/e25503>

doi: [10.2196/25503](https://doi.org/10.2196/25503)

PMID: [33865233](https://pubmed.ncbi.nlm.nih.gov/33865233/)

©Emre Sezgin, Garey Noritz, Simon Lin, Yungui Huang. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 11.05.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Protocol

An App-Based Parenting Program to Promote Healthy Energy Balance–Related Parenting Practices to Prevent Childhood Obesity: Protocol Using the Intervention Mapping Framework

Levie T Karssen¹, MSc; Jacqueline M Vink¹, PhD; Carolina de Weerth², PhD; Roel C J Hermans^{3,4}, PhD; Carina P M de Kort⁵, MSc; Stef PJ Kremers³, PhD; Emilie L M Rutter⁶, MSc; Junilla K Larsen¹, PhD

¹Behavioural Science Institute, Radboud University, Nijmegen, Netherlands

²Donders Institute for Brain, Cognition and Behaviour, Radboud University Medical Center, Nijmegen, Netherlands

³NUTRIM School of Nutrition and Translational Research in Metabolism, Maastricht University Medical Center, Maastricht, Netherlands

⁴Netherlands Nutrition Centre, The Hague, Netherlands

⁵Institute of Health Studies, HAN University of Applied Sciences, Nijmegen, Netherlands

⁶Academic Collaborative Centre AMPHI, Primary and Community Care, Radboud University Medical Centre, Nijmegen, Netherlands

Corresponding Author:

Levie T Karssen, MSc
Behavioural Science Institute
Radboud University
Postbus 9104
Nijmegen, 6500 HE
Netherlands
Phone: 31 0640969305
Email: l.karssen@bsi.ru.nl

Abstract

Background: The family environment plays an important role in the development of children's energy balance–related behaviors. As a result, parents' energy balance–related parenting practices are important targets of preventive childhood obesity programs. Families with a lower socioeconomic position (SEP) may benefit from participating in such programs but are generally less well reached than families with a higher SEP.

Objective: This paper describes the application of the Intervention Mapping Protocol (IMP) for the development of an app-based preventive intervention program to promote healthy energy balance–related parenting practices among parents of children (aged 0–4 years) with a lower SEP.

Methods: The 6 steps of the IMP were used as a theory- and evidence-based framework to guide the development of an app-based preventive intervention program.

Results: In step 1, behavioral outcomes for the app-based program (ie, children have a healthy dietary intake, sufficient sleep, and restricted screen time and sufficient physical activity) and sociocognitive (ie, knowledge, attitudes, and self-efficacy) and automatic (ie, habitual behaviors) determinants of energy balance–related parenting were identified through a needs assessment. In step 2, the behavioral outcomes were translated into performance objectives. To influence these objectives, in step 3, theory-based intervention methods were selected for each of the determinants. In step 4, the knowledge derived from the previous steps allowed for the development of the app-based program *Samen Happie!* through a process of continuous cocreation with parents and health professionals. In step 5, community health services were identified as potential adopters for the app. Finally, in step 6, 2 randomized controlled trials were designed to evaluate the process and effects of the app among Dutch parents of infants (trial 1) and preschoolers (trial 2). These trials were completed in November 2019 (trial 1) and February 2020 (trial 2).

Conclusions: The IMP allowed for the effective development of the app-based parenting program *Samen Happie!* to promote healthy energy balance–related parenting practices among parents of infants and preschoolers. Through the integration of theory, empirical evidence, and data from the target population, as well as the process of continued cocreation, the program specifically addresses parents with a lower SEP. This increases the potential of the program to prevent the development of obesity in early childhood among families with a lower SEP.

Trial Registration: Netherlands Trial Register NL6727, <https://www.trialregister.nl/trial/6727>; Netherlands Trial Register NL7371, <https://www.trialregister.nl/trial/7371>.

(*JMIR Form Res* 2021;5(5):e24802) doi:[10.2196/24802](https://doi.org/10.2196/24802)

KEYWORDS

childhood obesity; preventive intervention; parenting practices; energy-balance related behavior; socio-economic position; mHealth; behavior change; mobile phone

Introduction

Background

Although childhood obesity rates have been reported to stabilize in many developed countries [1], the prevalence of overweight and obesity in young children is still high. In general, 8% of Dutch children around the age of 2 years were overweight or obese in 2018 [2]. However, these rates were considerably higher among children from families with a lower socioeconomic position (SEP) [2]. In addition, a growing body of research indicates that plateauing prevalence rates are predominantly evident among groups with a higher SEP [3]. In groups with a lower SEP, the rates are still rising [4], implying an increase in socioeconomic health disparities. To reduce SEP disparities in childhood obesity, preventive intervention programs should promote healthy energy balance-related behaviors (EBRBs; ie, dietary intake, sleep, physical activity, and screen time) before children have developed fixed patterns of EBRBs [5]. As patterns of EBRBs, such as dietary intake and screen time, develop in the first years of life [6], infancy and preschool (0-4) years present a critical window for childhood obesity prevention.

Parents are considered the main agents of change in effective preventive intervention programs for childhood obesity, especially in the first years of life [7]. Parents' behaviors, particularly their energy balance-related parenting practices (ie, energy balance-related, specific, discrete, and observable acts of parenting [8]), may significantly influence their children's EBRBs and body weight [9,10]. Recent reviews have shown that promoting responsive feeding guidance to teach maternal awareness and attention to children's hunger and satiety cues can support *normal* child body weight development [11]. Moreover, universal preventive intervention programs targeting early feeding and positive parenting skills, including programs that target sleep, show similar effects on the child's body weight [12]. In general, current evidence supports starting with promoting responsive feeding and parenting during infancy and incorporating the promotion of structure and rule setting in early childhood [13]. For physical activity parenting, support has been found for the potential importance of parental support and parents' own behaviors and role modeling [14,15]. However, it should be noted that only a few preventive intervention programs have targeted physical activity parenting [14]. More generally, preventive intervention programs for early childhood obesity that target parenting practices with respect to all relevant child EBRBs are scarce [12]. Moreover, parent-focused prevention programs have been mostly universal (ie, population based) in nature [12,16]. For instance, 3 of 4 parent-focused EPOCH (Early Prevention of Obesity in Children) trials that

commenced by early infancy were universal programs [17]. Our parent-focused program to prevent early childhood obesity adds to the literature by addressing parenting practices with respect to all the important child EBRBs [12,18] while simultaneously applying selective prevention to the subgroup of parents with a lower SEP, who generally display more problematic energy balance-related parenting compared with other SEP groups [19].

In addition, our program uses an innovative approach to address 2 frequently reported limitations of traditional (parent focused) obesity prevention programs: the costly, time-intensive face-to-face setting [16,20] and difficulties in reaching people with a lower SEP [21]. To successfully target and reach people with a lower SEP, intervention programs should address both practical (eg, time constraints, lack of transportation, and inflexible working hours [22-24]) and attitudinal (eg, irrelevant and not engaging program elements [25]) barriers for participation [26]. Practical barriers for people with a lower SEP could (at least partly) be overcome by delivering interventions through smartphones, including mobile apps [21,27]. Furthermore, attitudinal barriers can also be overcome because app-based interventions allow for options to increase program engagement through the presentation of bite-sized information in plain language that is accompanied by appealing visuals and through the possibility of personalizing the intervention and the ability for users to monitor their behavior [28,29]. For these reasons, app-based intervention programs have the potential to be cost-effective and reduce the gap in socioeconomic health disparities [29,30]. The apps used in these programs need to be high quality (eg, in terms of engagement, esthetics, and information quality [31]) and encourage behavior change (eg, providing knowledge and information as well as prompting goal setting and planning [32]). To achieve this, app-based interventions should be based on evidence, grounded in behavior change theory, and incorporate the needs and wishes of the target audience through formative research [33].

Objectives

Therefore, we developed the app-based prevention program *Samen Happie!* using the Intervention Mapping Protocol (IMP). The IMP is a widely used, standardized intervention planning format that helps intervention developers to incorporate empirical findings from the literature, effective behavioral change methods and their practical applications, and data collected in a representative population [34]. Although it is not specific to app-based program development, the IMP has been successfully applied to the development of digital (including mobile) interventions for youth health promotion [35-41]. The central goal of the app-based preventive parenting program *Samen Happie!* is to stimulate healthy energy balance-related

parenting practices to prevent early childhood obesity among children of families with a lower SEP. In this paper, we inform readers about the development of the *Samen Happie!* program based on the 6 steps of intervention mapping.

The Dutch title of the app-based parenting program (*Samen Happie!*) will be used throughout this paper. A possible translation is “Happy Together,” but this does not reflect the play on words the title indicates in Dutch.

Methods

Overview

The *Samen Happie!* program was developed through the 6 iterative and nonlinear steps of the IMP: (1) conducting a needs assessment; (2) preparing matrices of change objectives; (3) selecting theoretical methods and practical strategies; (4) developing the intervention program; (5) planning for adoption, implementation, and sustainability; and (6) planning the program evaluation [34]. We collaborated with a workgroup consisting of potential program implementers (ie, youth health care professionals of community health services) and users (ie, parents of young children with a lower SEP). Importantly, for the development and evaluation of the program, we asked only one parent-child dyad per family to participate, thereby focusing on the primary caregiver. In this section, the main tasks of each step of the IMP are explained and, when relevant, the role of the workgroup is exemplified. The outcomes of the IMP, including the choices and actions during each step, are described in the Results section.

Step 1: Conducting a Needs Assessment

The first step of the IMP was to conduct a needs assessment of our target group (ie, parents of children aged 0 to 4 years with a lower SEP) to build a logic model of the health problem [34]. Our needs assessment included a literature search, focus groups with parents with a lower SEP (N=16 mothers in total), and discussions with youth health care professionals (N=6 professionals in total). In the focus groups, the hindering and facilitating factors for healthy parenting and parenting practices in difficult parenting situations were discussed, with a focus on parenting practices regarding food and dietary intake. The focus groups, which were conducted until saturation was reached, were audio recorded, transcribed, and coded for themes and concepts using ATLAS.ti (ATLAS.ti Scientific Software Development GmbH). The discussions with youth health care professionals served to explore key parenting-related themes and issues that existed among the target group and as a sounding board for concrete questions during program development. For example, frequently reported parenting problems, (parental adherence to) national guidelines regarding child EBRBs, and effective, practical strategies to increase healthy child EBRBs through energy balance-related parenting practices were discussed. To build the logic model (based on the PRECEDE [Predisposing, Reinforcing, and Enabling Constructs in Educational Diagnosis and Evaluation] framework [42]), we identified the quality-of-life indicators associated with the health problem, behavioral and environmental risk behaviors for the health problem, and determinants related to these behaviors. On the basis of the knowledge derived from this assessment,

we selected behavioral outcomes for the program and formulated the program goal.

Step 2: Preparing Matrices of Change Objectives

In the second step of the IMP, the performance objectives were defined for the behavioral outcomes specified in step 1. These performance objectives constituted behaviors that are expected to contribute to achieving the program goal when performed by the target group [34]. By crossing the performance objectives with the determinants selected in step 1, the change objectives were specified. These change objectives indicate which actions are required to modify the determinants of the behavioral outcomes and reach the performance objectives [34].

Step 3: Selecting Theoretical Methods and Practical Strategies

The third step of the IMP evolved around the selection of theory- and evidence-based change methods to affect the determinants selected in step 1. We aimed to select a limited number of theoretical methods per determinant as interventions that use a small number of behavior change techniques are generally more effective for people with a lower SEP than interventions that use a larger number of techniques [43]. We then translated the selected methods into practical strategies through which they were delivered in the program [34].

Step 4: Developing the Intervention Program

The fourth step of the IMP involved building the program themes and components and drafting, pretesting, and producing the program materials based on the information gathered in the previous steps [34]. The development of the app-based parenting program consisted of 3 phases. In phase 1, qualitative user research was conducted to assess parents' (N=16 mothers) wishes regarding the content and functionality of the app. In phase 2, a prototype of the app structure, functionalities, content, and visual design was constructed in continuous cocreation with parents (N=4 mothers) and youth health care professionals (N=3) and pretested by parents (N=16 mothers). Finally, phase 3 involved building the final version of the app. For the development of the app, we collaborated with Dio Agency, an agency specialized in deploying software design to facilitate behavioral change.

Step 5: Planning for Adoption, Implementation, and Sustainability

The fifth step of the IMP involved the identification of potential program users (eg, implementers or adopters) and the design of a program implementation plan [34]. We planned to make the app-based intervention available free of charge after program evaluation; nevertheless, the excess supply of health-related apps minimizes the chances that parents will spontaneously find and download the *Samen Happie!* app. This is one of the reasons that digital health programs in particular need a delivery system (ie, a program adopter) to get the program to its intended participants [34].

Step 6: Planning the Program Evaluation

The sixth and last step of the IMP involved the development of a program evaluation plan [34]. We developed both process and effect evaluation plans to evaluate the quality of the

implementation and the effectiveness of the preventive intervention program.

Results

Step 1: Conducting a Needs Assessment

Our logic model of the health problem is presented in [Multimedia Appendix 1](#). The model displays the quality-of-life indicators (eg, cardiovascular diseases, depression, and risk of obesity in adulthood [44-46]) as correlates of health problems (ie, early childhood obesity). Moreover, it shows behavioral (eg, child dietary intake) and environmental (eg, unhealthy home environment) risk factors for health problems and determinants related to these factors (ie, parents' knowledge about healthy dietary intake). The parental determinants impact child risk behaviors through energy balance-related parenting practices. The child energy balance-related risk behaviors, role of parents, and determinants of healthy parenting are described hereafter.

Child Risk Behaviors and the Role of Parents

Ample research has established the intake of energy-rich foods and sugar-sweetened drinks [47-49] and longer screen time and shorter sleep duration [50-53] as key modifiable risk behaviors of childhood obesity. Notably, these unhealthy EBRBs are more common among children from lower SEPs than among those from families with a higher SEP [54-56]. This led us to select the following behavioral outcomes of our preventive intervention program: (1) children have a healthy dietary intake (ie, food and drinks), (2) children get sufficient sleep, and (3) children have a healthy balance between screen time and physical activity. Especially early in life, child EBRBs are largely shaped by parents and their parenting practices [57]. Hence, the overall program goal was to improve child EBRBs (ie, dietary intake, sleep, physical activity, and screen time) and subsequent body weight through the stimulation of healthy energy balance-related parenting practices.

Energy Balance-Related Parenting Practices

Parenting practices are broadly divided into 3 overarching dimensions of food parenting [10] that can also be observed in a wider range of energy balance-related parenting behaviors [58-62], namely, coercive control (ie, the use of pressure and dominance to control child behavior, such as restriction and threats [63]), structure (ie, the use of noncoercive forms of control, such as rules and routines [63]), and autonomy support (ie, the facilitation of children's independence, for instance through responsive feeding and praise [64]). Studies have shown that structured and autonomy-supportive parenting practices are mostly related to favorable child energy balance-related outcomes, whereas coercive controlling practices show unfavorable associations with children's EBRBs and body weight [61,65,66]. However, notably, as compared with parents with a higher SEP, parents with a lower SEP are more likely to use coercive control [56,67-72] and less likely to use structure-related [59,73-76] and autonomy-supportive parenting practices [70,77]. Hence, it is pivotal that preventive intervention programs for childhood obesity, particularly those targeting parents with a lower SEP, promote structure-related and

autonomy-supportive practices and discourage coercive controlling parenting practices regarding children's EBRBs.

General Parenting and Parental Well-being

Both general parenting and parental well-being may moderate the associations between parental energy balance-related parenting practices and child EBRBs. With respect to general parenting (ie, the broader emotional climate in which specific parenting practices are performed [78]), research has, for instance, demonstrated that the prospective associations between parental encouragement and covert control (eg, food availability) and dietary intake were stronger among children who were exposed to a positive parenting style [79]. Hence, desirable (ie, structured and autonomy-supportive) parenting practices performed in an authoritative parenting climate (ie, characterized by both demandingness and responsiveness [80]) might produce the largest intervention effects [9,59,81]. Moreover, research found reciprocal relationships between parental mental well-being and parenting, indicating that positive well-being among parents relates to higher parenting self-efficacy and more beneficial parenting practices [82]. Thus, improving parental well-being and stimulating authoritative parenting appear to be promising conditions for improving child EBRBs. Examining these moderators may not only provide more insight into potential differential intervention effects for subgroups of parents but it could also help identify eminent targets for future (personalized) obesity prevention programs.

Determinants of Healthy Parenting

For the selection of determinants, we were informed by the results of our focus groups and the empirical literature. Moreover, the I-Change model [83] was used as a basis to integrate influential theories on motivation and behavior change (ie, Theory of Planned Behavior [84], Social Cognitive Theory [85], the Transtheoretical Model [86], and the Health Belief Model [87]). The I-Change model explains how knowledge, attitudes, and self-efficacy play a role in a person's motivation and intention to perform health behaviors. This model also considers the gap between the intention to perform a behavior and actually performing the behavior (ie, the intention-behavior gap [88]). In the *Samen Happie!* program, with respect to performing healthy parenting practices related to children's EBRBs, the following determinants are targeted: parental knowledge, attitudes, self-efficacy, and habitual behavior (ie, habits).

Knowledge

Knowledge plays an important role in changing EBRBs and is a basic component of existing preventive intervention programs for childhood obesity [7]. It is particularly important to include remediation components in interventions for high-risk (eg, lower SEP) populations [89,90]. People with a lower SEP tend to have lower health literacy in general [91] and regarding healthy parenting in particular [92]. Illustratively, some parents in our focus groups held incorrect beliefs about the healthiness of drinks (eg, "Fruit juice is healthy because it contains vitamins"). Importantly, targeting knowledge may also indirectly change other sociocognitive determinants, including attitudes and self-efficacy [93].

Attitudes

Knowledge should be targeted by carefully considering the beliefs of the target group [94]. Some parents in our focus groups held negative attitudes toward specific energy balance–related parenting practices, such as providing water instead of sugar-sweetened drinks (eg, “Drinking water is for dogs”). These parental attitudes toward energy balance–related (parenting) behaviors can influence children’s behaviors, such as physical activity [95] and screen time [96,97]. When targeting attitudes, parents’ beliefs should be taken into consideration, as it has generally been acknowledged that it is difficult to change attitudes with a high affective component. Therefore, it is imperative to balance the stimulation of favorable attitudes about EBRBs with parents’ personal goals [98].

Self-efficacy

Self-efficacy refers to a parent’s beliefs in their capabilities to organize and execute a course of action (ie, performing energy balance–related parenting practices) in particular situations [99]. In general, children display more healthful behaviors if parents report higher self-efficacy [100], and improving parental self-efficacy also appears to be a promising approach to change young children’s EBRBs [101]. Enhancing self-efficacy might be especially important for parents with a lower SEP, as parents in our focus groups often felt insecure about their capabilities

to employ healthy energy balance–related parenting practices (eg, sticking to a maximum amount of screen time).

Habits

Parents often report a discrepancy between what they intend to do in terms of energy balance–related parenting practices and what they actually do [88]. For instance, parents in our focus groups found it difficult to form healthy habits and routines, especially when unhealthy habits were already established (eg, eating in front of the television). Habits influence health behaviors [102], and parental energy balance–related habits may impact energy balance–related parenting practices [88]. One previous intervention program that trained parents to perform healthy parenting habits proved to be promising [103]. Thus, targeting the development of healthy habits may assist parents in the long-term adherence to newly developed energy balance–related parenting practices and may bridge the gap between parenting intentions and behaviors [88].

Step 2: Preparing Matrices of Changes Objectives

Table 1 presents 3 examples of performance and change objectives for each behavioral outcome (ie, children have a healthy dietary intake, sufficient sleep, and restricted screen time and sufficient physical activity). Multimedia Appendix 2 [10] presents an overview of all performance objectives and change objectives specified for the app-based preventive intervention program.

Table 1. Change objectives for dietary intake, sleep, and restricted screen time and sufficient physical activity by crossing the determinants with the performance objectives.

Performance objectives	Determinants			
	Knowledge	Attitudes	Self-efficacy	Habits
Dietary intake				
Parents apply clear rules about the consumption of healthy and unhealthy food products and drinks.	Parents explain how they can apply clear rules about the consumption of healthy and unhealthy food products or drinks.	Parents express positive feelings toward having clear rules for the consumption of healthy food products or drinks.	Parents express confidence in applying clear rules about the consumption of healthy and unhealthy food products or drinks.	Parents consistently apply clear rules about the consumption of healthy and unhealthy food products or drinks.
Parents act as a role model by eating or drinking healthy food or drinks themselves.	Parents explain how they can act as positive role models by eating or drinking healthy food or drinks themselves.	Parents express positive feelings toward acting as a role model by eating or drinking healthy food or drinks themselves.	Parents express confidence in acting as a role model by eating or drinking healthy food or drinks themselves.	Parents consistently act as a role model by eating or drinking healthy food or drinks themselves.
Parents praise their child when he or she eats healthy food products or drinks water.	Parents explain how they can praise their child when he or she eats healthy food products or drinks water.	Parents express positive feelings toward praising their child when he or she eats healthy food products or drinks water.	— ^a	—
Sleep				
Parents apply clear rules about bed times.	Parents explain how they can apply clear rules about bed times.	Parents express positive feelings toward applying clear rules about bed times.	—	—
Parents make use of bedtime routines.	Parents explain how they can make use of bedtime routines.	Parents express positive feelings toward making use of bedtime routines.	Parents express confidence in making use of bedtime routines.	Parents consistently make use of bedtime routines.
Parents ensure a safe and quiet sleep environment for their child.	Parents explain how they can ensure a safe and quiet sleep environment for their child.	Parents express positive feelings about ensuring a safe and quiet sleep environment for their child.	—	—
Restricted screen time and sufficient physical activity				
Parents apply clear rules about screen time.	Parents explain how they can apply clear rules about screen time.	Parents express positive feelings toward applying clear rules about screen time.	Parents express confidence in applying clear rules about screen time.	Parents consistently apply clear rules about screen time.
Parents facilitate activities without the use of screens.	Parents explain how they can facilitate activities without the use of screens.	Parents express positive feelings toward facilitating activities without the use of screens.	Parents express confidence in facilitating activities without the use of screens.	Parents consistently facilitate activities without the use of screens.
Parents encourage their child to be physically active (eg, playing outside).	Parents explain how they can encourage their child to be physically active.	Parents express positive feelings toward encouraging their child to be physically active.	—	—

^aNot all performance objectives were translated into change objectives.

Step 3: Selecting Theoretical Methods and Practical Strategies

The theoretical methods we selected for each of the determinants (ie, knowledge, attitudes, self-efficacy, and habits) were derived from the taxonomies described by Kok et al [104] and Michie et al [105]. Table 2 presents an overview of the methods and examples of their practical application. This table shows that

we included both methods involving the provision of information (eg, consciousness raising and persuasive communication) and those that target more automatic processes (eg, implementation intentions and self-nudging). Automatic, *nonconscious* methods might be particularly effective for people with a lower SEP as these methods are less dependent on literacy capabilities [106].

Table 2. Theoretical methods that were selected to address the determinants and examples of how these methods were applied in the app-based program.

Determinant and method	Definition	Example of practical application in the app
Knowledge		
Consciousness raising	Giving information about the causes and consequences of a problem behavior and providing alternatives to substitute problem behaviors [104]	Parents are advised not to comfort or reward their children with food (ie, emotional and instrumental feeding). We explain that in doing so, children might learn to comfort themselves using food later in life or learn that they will be rewarded for demonstrating unwanted behavior. As alternatives for the problem behavior, we suggest to comfort or reward children with attention and affection (eg, giving a compliment, thumbs up, or hug)
Instruction on how to perform a behavior	Advise on how to perform the behavior [105]	We provide both simple and more elaborate advice on how to perform parenting behaviors. For instance, when parents want to encourage water consumption but their child is used to drinking fruit juice, we advise to gradually substitute parts of the fruit juice by water over the course of several weeks. A more elaborate advice is provided when parents want to decide which type of fruit juice is the healthier option. This advice involves 3 steps (ie, grab 2 drinks, turn them around and look at the food label, and pick the option with the least calories). For the second step (reading the food label), we advise parents to look at the amount of sugar and explain how they can calculate the amount of sugar per serving
Attitudes		
Persuasive communication	Guiding individuals toward the adoption of an idea, attitude, or action by using arguments or other means [104]	To encourage positive attitudes toward the consumption of water, we provide 3 benefits of drinking water (or tea): (1) it does not contain calories and contributes to a healthy body weight, (2) water and tea do not contain sugar, the teeth are not affected and cavities can be prevented, and (3) water supports the functioning of the body and can support learning and playing
Framing	Using gain-framed messages emphasizing the advantages of performing the healthy behavior [104]	We focused on providing gain-framed messages that emphasize the benefits for parents and/or children. Examples are “Did you know that you can save up to €150 euros per year if Maria drinks water instead of fruit juice?” and “Using a fixed bedtime routine can help Maria fall asleep faster”
Self-efficacy		
Verbal persuasion	Using messages that suggest that the participant possesses certain capabilities [104]	Before making an if-then plan (see <i>Implementation intentions</i>) to stimulate children’s water consumption, parents are asked whether they think drinking water is important and whether they feel confident in making their child drink (more) water. They answer on a scale from 0 (not important or confident at all) to 5 (very important or confident). On the basis of their answer, they receive an encouraging response (eg, “Many parents think they won’t be successful in getting their child to drink water. But you will be surprised to see that practice will eventually pay off. You can do it!”)
Action planning	Prompt detailed planning of performance of the behavior [105]	To regulate children’s screen time, parents are prompted to make a family screen time plan. For this plan, they first describe in which rooms screens can be used (eg, living room) and then at what times screens can be used (eg, before dinner). The plan should be focused on what <i>is</i> allowed, rather than what <i>is not</i> allowed (see also <i>Framing</i>). Parents are also prompted to write down this plan on paper and display it at a prominent place, and (if applicable) to discuss the plan with other caregivers
Habits		
Implementation intentions	Prompting making if-then plans that link situational cues with responses that are effective in attaining goals or desired outcomes [104]	Parents are prompted to make an if-then plan to stimulate their child’s physical activity. First, we explain what an if-then plan is and why it is important to make them. Next, 3 examples are presented (eg, “If we go outside together, then I will let Maria walk next to the stroller for a couple of minutes”). After that parents are asked to make their own plan by defining first <i>when</i> they can stimulate physical activity for their child (eg, if my child wakes up from her nap) and next <i>how</i> they will achieve this (eg, then I will play an active game with her)
Self-nudging	Making simple changes in the presentation of choice alternatives that make the desired choice the easy, automatic, or default choice [104]	We provide parents tips that can help make healthy eating at home the easy, automatic option: (1) buy mostly healthy food such as fruits and vegetables at the grocery store; (2) buy no or only a small amount of snacks, such as candy and chocolate; (3) display healthy foods in a way that they are easily noticed, eg, by presenting fruit in a bowl on the table; and (4) store unhealthy foods out of sight, for instance, keeping snacks at the back of the storage cabinet

In addition to potentially important theoretical methods, the form in which an intervention is delivered is also a key ingredient in behavior change interventions [107] and might be even more crucial for groups with a lower SEP in terms of

comprehensibility and engagement [108]. Considering the literacy capabilities of participants with a lower SEP, research has shown that people with literacy problems can remember written texts more easily when they are supported by (audio)visual aids [109,110]. Therefore, we included both written texts and supporting icons, images, videos, and voice-overs of important information in the app-based prevention program. Moreover, to ensure that the app was comprehensible for parents with varied literacy skills, all textual components were revised by a specialized organization to match language levels A2-B1 according to the Common European Framework of Reference [111]. In addition, the technological features of the app allowed us to tailor the program to individual participants, which might positively impact user engagement [112] and subsequent intervention retention. Illustratively, the app uses data about the name, sex, and birth date of the child (entered by the parent upon registration) to personalize texts with respect to names and pronouns (eg, “Set a good example for *Maria*. Try to eat healthy when *she* is around”). Finally, the data on children’s age were used to present parents with developmentally appropriate information via age-based modules (see *Step 4: Developing the Intervention Program*), which could enhance the perceived relevance of the information, thereby increasing engagement [113].

Step 4: Developing the Intervention Program

In phase 1 of the app development process, the results from our focus groups and discussions with youth health care professionals informed the development of the prototype of the app and the way in which we tailored the materials to our target group. For instance, based on parents’ negative affective attitudes toward water consumption, we asked them whether they would provide their children tea (without sugar) as an alternative, and they affirmed that they would do so. Therefore, we drafted the app content with “water or tea,” instead of focusing entirely on water consumption. Other examples of content that we based on the suggestions of parents from our focus groups included providing information about saving money on groceries and tips to stimulate vegetable consumption.

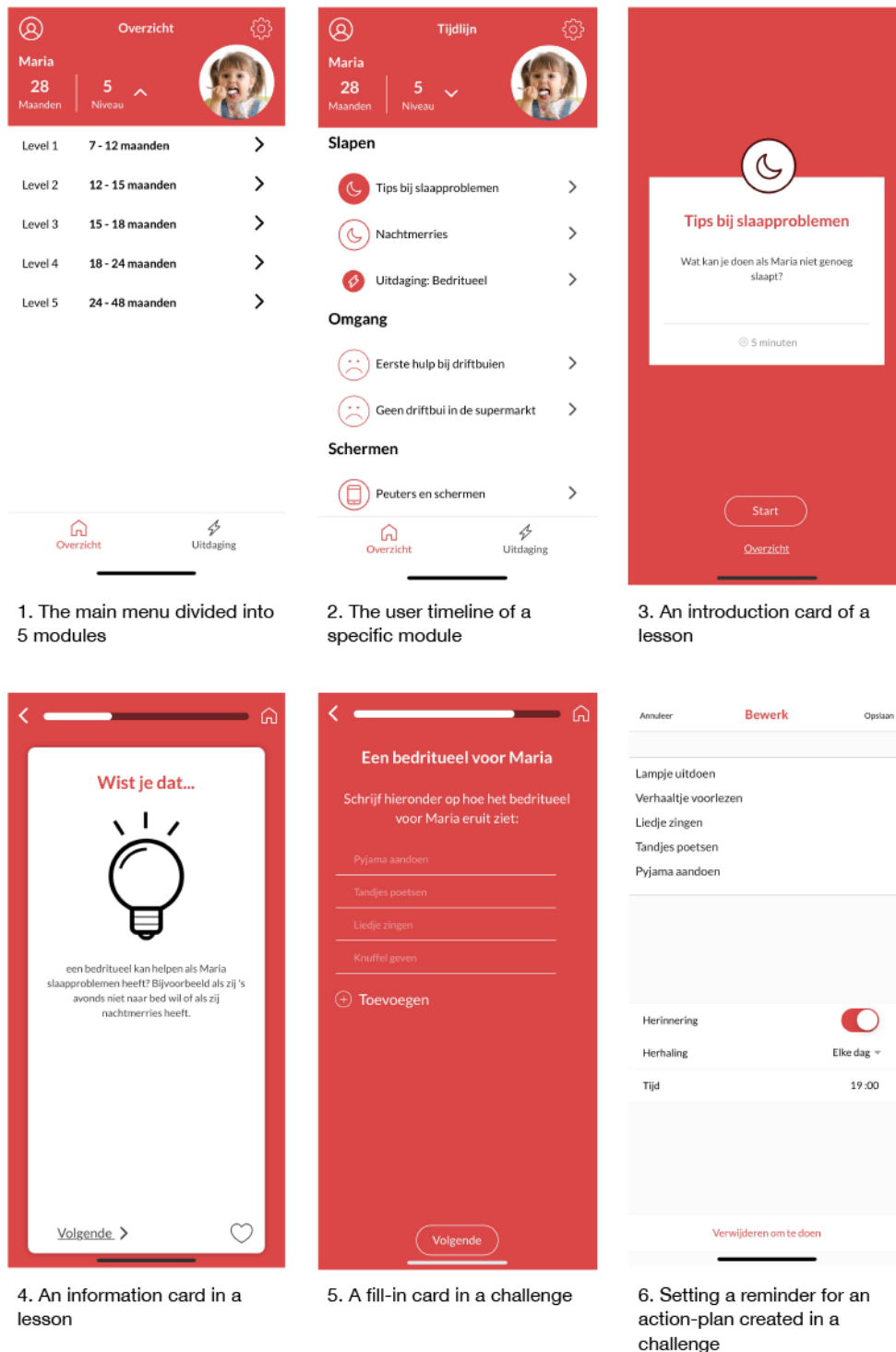
In phase 2, the content of the preventive intervention program was drafted based on national health care guidelines, relevant literature on parenting practices in relation to child EBRBs, and previous intervention projects [103,114]. For example, youth health care professionals suggested using videoclips that were specially designed by the Dutch Child and Family Center (Centrum voor Jeugd en Gezin) to stimulate health literacy among young (low literate) parents. The pretest of the prototype of the app yielded, among others, the following suggestions: the option to enlarge images, the possibility of replaying the instruction clip, a summary of the most important information per level, and the option to reread a lesson. These suggestions were incorporated into the final version of the app.

The final version of the app (phase 3) was launched in September 2018. Figure 1 presents 6 screenshots of the app (in Dutch). Screenshot 1 shows the main menu of the app. From here, parents could navigate to their profile, app settings, and

the actual app content. The content of the app was divided into 5 age-based modules ranging from 7 to 12 months (module 1) to 24 to 48 months (module 5). On the basis of their child’s birth date (entered upon registration), parents were granted access to the appropriate modules (eg, parents of children aged 14 months were granted access to the first 2 modules). When the child reached the minimum age of the subsequent module, the new module became unlocked. The content of each age-based module was divided into 6 themes reflecting all relevant child EBRBs (ie, eating, drinking, sleep, and physical activity and screen time) and themes on the well-being of the parent and temper tantrums (this last theme was only included in the modules for children aged 18 months and older).

Each of the 5 modules consisted of 2 types of activities: lessons and challenges. Screenshot 2 in Figure 1 shows the lessons (indicated by the larger icons) and challenges (indicated by the smaller icons with the thunderstroke) of the themes sleep and temper tantrums in module 5. Upon completion of a lesson or challenge, the icon of that lesson or challenge became filled (see screenshot 2 in Figure 1). The lessons and challenges were constructed using different types of cards. Every lesson or challenge started with an introduction card that showed the title and length of the lesson or challenge (in minutes; see screenshot 3 in Figure 1). The lessons consisted of multiple information cards that presented information on that specific theme in an engaging and easy-to-comprehend manner, for instance, through the use of facts (“Did you know...?”; see screenshot 4 in Figure 1), practical examples, tips, and quizzes and supported by icons or pictures. The challenges consisted of exercises that prompted parents to apply the information from the lessons. By employing techniques that tackle (unhealthy) automatic behaviors, parents were encouraged to implement (new) parenting skills as habits. The challenges were similar in their design. Each challenge started with providing information and examples of specific parenting target behaviors. Next, parents were asked whether they thought performing that behavior was important (ie, attitude) and whether they felt capable (ie, self-efficacy) of performing that behavior. On a slider card, they could indicate their responses on a scale from 0 (low importance or self-efficacy) to 5 (high importance or self-efficacy). Appropriate feedback on their responses was provided through a pop-up notification. Finally, parents created a personal goal or action plan in the context of the target behavior on a fill-in card (see screenshot 5 in Figure 1). For each goal or action plan, parents could set reminders to receive a notification at a date and time of their choice to help fulfill that goal (see screenshot 6 in Figure 1). An illustrated example of a lesson and the accompanying challenge within the theme Sleep can be found in Multimedia Appendix 3. The example includes screenshots of different types of cards within the lesson or challenge, a translation of the original Dutch text to English, and a reference to the theoretical method that formed the basis for that card. Moreover, an overview of the 5 modules and the corresponding lessons and challenges (specified per theme) is presented in Multimedia Appendix 4.

Figure 1. Screenshots of the *Samen Happie!* app showing (1) the main menu divided into 5 modules, (2) the user timeline of a specific module, (3) an introduction card of a lesson, (4) an information card in a lesson, (5) a fill-in card in a challenge, and (6) setting a reminder for an action plan created in a challenge.



Step 5: Planning for Adoption, Implementation, and Sustainability

We identified national community health services and child day care centers as potential adopters of the app-based program, as representatives of these organizations were involved in the development and execution of the program. The program

implementers will be youth health care professionals and pedagogical staff working at these organizations, and their tasks are to bring the app to the attention of parents (whose children they perceive as being at high risk for obesity) and motivate them to use the app. The implementation of the app-based program can support the daily practice of the program implementers in 2 ways. First, the app could function as an

addition to usual care, given that youth health care professionals indicated that standard consultations are generally too brief to give parents elaborate, well-rounded advice. In this sense, it is advantageous that the app is an easy-to-use product that does not require detailed instructions from a health care provider. Second, the app could function as an educational material that creates a legitimate opening to discuss topics such as food parenting and body weight [115], which health professionals often find to be difficult topics to address [67]. In addition, the active role of youth health care professionals in the referral of high-risk groups (in terms of obesity and other health problems) and a primary focus on prevention is in line with the national prevention agreement that the Dutch government issued in the first half of 2019 [116]. Moreover, the agreement calls for the inclusion of (potentially) effective (preventive) interventions for childhood obesity to be registered at the *Healthy Living Desk*, an intervention database of the Dutch ministry of Public Health, Welfare and Sports. To facilitate nationwide dissemination and aid in intervention sustainability, the *Samen Happie!* program was submitted to this database and accepted in September 2020.

Step 6: Planning the Program Evaluation

The process and effects of the app-based preventive intervention program were evaluated in 2 separate trials, the designs, eligibility criteria, procedures, and measures of which are explained hereafter.

Trial Designs

Both trials were randomized controlled trials (RCTs) with 2 parallel arms: an intervention condition in which parents received access to the *Samen Happie!* app and a waitlist control condition. In addition, trial 2 included a third condition in which parents received the app and 2 additional group sessions organized at locations where their child attended preschool. The third condition was a separate condition that did not interfere with the procedures of the RCT, as presented in this manuscript. Detailed information about the development and evaluation of these group sessions has been described elsewhere. Hence, this manuscript will concentrate on the intervention conditions in which parents received access to the app-only and control condition. The parents in both of these conditions completed a web-based baseline measurement (T0) and follow-up measurements at approximately 6 months (T1) and 12 months (T2). [Multimedia Appendix 5](#) provides a schematic overview of the trial flows and includes the exact timing of the measurements.

Sample Sizes

A power analysis using G*Power (version 3.1) indicated a minimum of 200 participants in trial 1. This calculation was based on child BMI as the outcome variable, which was assumed to have a mean of 16.67 (SD 1.70) [117]. We further assumed an effect size of 0.20, α of .05, and power of .80. We strived to recruit a minimum of 150 participants per condition (300 in total) to include a minimum of 50% (150/300) of participants with a lower SEP and to account for dropout over time. For trial 2, we recruited 70 participants per condition (140 in total for 2

conditions) based on previous research with a similar design including a face-to-face component [103].

In the end, the participants in trial 1 were 357 parents (346/357, 96.9% biological mothers, 7/357, 1.9% biological fathers, and 4/357, 1.1% nonbiological mothers or partners of the biological mother) of infants aged 5 to 15 months at baseline (ages corresponding to modules 1 and 2 of the app). Trial 2 was conducted among 153 parents (148/153, 96.7% biological mothers, 3/153, 1.9% biological fathers, and 2/153, 1.3% partners of the biological father or mother) with toddlers aged 18 to 55 months at baseline (ages correspond to modules 4 and 5).

Eligibility Criteria

To assess whether parents were eligible to participate in the trials, they completed a web-based screening that contained questions about their educational attainment and their child's age and health status. Parents were respectfully refused participation when their child was younger than 5 months or older than 15 months (trial 1), younger than 20 months or older than 55 months (trial 2), or when their child had a chronic disease or disability that affected normal development. Parents with multiple children could only participate with one child and only in 1 of the 2 trials. We strived to include at least 50% of parents with lower or medium-level SEP and used educational attainment as a proxy for SEP (ie, lower SEP was conceptualized as having completed no education, primary school education, or preparatory vocational education and medium-level SEP was conceptualized as having completed vocational education). Parents with higher educational attainment (ie, preuniversity or university degree) were not actively discouraged from participating in the trials.

Procedures

Parents were recruited offline (eg, through child day care centers and community health care centers for young children) and online (eg, through Facebook groups), for which we particularly considered locations that are often visited or used by parents with a lower SEP. Interested parents who fulfilled the eligibility criteria received an email in which they were asked to provide consent for their participation. After consenting, the parents were forwarded to the web-based baseline questionnaire. Randomization for each trial took place after the baseline measurement by means of a simple randomization procedure performed by an independent researcher using SPSS version 24. Among research with large sample sizes, this procedure can be trusted to produce equal samples in terms of numbers and covariates [118]. Participants were compensated for their time and effort with a €10 (US \$12) gift card (or a pack of diapers in trial 1) upon completion of the baseline questionnaire and the 2 follow-up measurements. Parents who were allocated to the intervention condition received a personal invitation code for the app to avoid contamination between the 2 conditions. Parents in the intervention condition received instructions on how to download and use the app. There were no instructions regarding the timing and frequency of the use of the app to stay as close as possible to app usage patterns in everyday life. After completing all 3 questionnaires, parents in the control condition were also granted access to the app. The procedures of the trials

were approved by the Ethics Committee of the Faculty of Social Sciences, Radboud University, the Netherlands (trial 1: ECSS-2017-013 and trial 2: ECSS-2018-084).

Measures

Process Evaluation

Two types of process evaluation data were collected: self-reported data and app user data.

Parent Self-Reports

We assessed parents' self-reported app use, their user experience of the app, and their suggestions for app improvement. In the 2 follow-up questionnaires (T1 and T2), we asked parents whether they downloaded the app (and why), whether they still had the app installed on their phone (and why), and how many times they used the app. Regarding user experience, we asked parents to rate several indicators of functionality (eg, ease of use), design, and content (eg, usefulness) on a scale from 1 (bad experience) to 7 (good experience). Parents also rated the app as a whole on a scale from 1 to 10, with higher scores indicating higher appreciation. Finally, we asked open-ended questions about the ways in which the app could be improved.

Preliminary Results of Parent Self-Reports

Preliminary analyses of parents' self-reported app evaluation data on the first follow-up measurement (T1) showed that most parents in the intervention condition of trial 1 (138/179, 77.1% parents) and almost half of the parents in the app-only intervention condition of trial 2 (33/76, 44% parents) reported that they downloaded the app. Most of these parents (127/179, 70.9% in trial 1 and 55/76, 72% in trial 2) indicated that they had used the app multiple times since installation but were not using it anymore at T1. Around a quarter of the parents (39/179, 21.7% in trial 1 and 21/76, 28% in trial 2) indicated that they still used the app multiple times per month. Parents in both trials generally appreciated the functionality, content, and design of the app. They graded the app with an average score of 6.7 (SD 1.45) in trial 1 and 7.2 (SD 1.05) in trial 2. The most important suggestions parents gave for improvement of the app included the incorporation of more detailed and elaborate parenting information, a clearer structure of the presented information (eg, based on weight-related themes instead of age), the option to look for specific information through a search function, and the integration of other parents' perspectives and experiences (eg, through personal accounts or online interactions).

User Data

In addition, to objectively assess parents' exposure to the preventive intervention program, their app usage was automatically monitored and collected in an online database. This database allowed us to examine the lessons and challenges that parents started and/or finished, the specific lesson cards they saved as *favorite*, and their answers to quiz questions.

Effect Evaluation

The primary outcome measures of the effect evaluation were EBRBs of the child (ie, dietary intake, sleep, and screen time), weight-for-height z scores (trial 1), and BMI z scores (trial 2) and parents' parenting practices related to their child's EBRBs.

The following secondary outcomes were assessed: parents' general parenting style, parental well-being (ie, depressive symptoms, life satisfaction, stress, and self-reported overall health), and parents' EBRBs (eg, snacking behavior and sugar-sweetened beverage consumption). An overview of the constructs, variables, and assessment points addressed in the evaluation of the program can be found in [Multimedia Appendix 6](#).

Child weight-for-height and BMI z scores were calculated using height and weight data reported by the parents in the questionnaires. We asked parents to draw this information from the measurement overview in the child's personal (digital) file, which is updated by the youth health care professional each time the parent and child visit the child health clinic. During the second follow-up questionnaire (T2), we additionally asked parents to send us a picture or screenshot of this measurement overview. This strategy not only allowed us to compare the information parents provided in the questionnaires with that in the child's file but also allowed us to collect more detailed anthropometric data as the overview contains height and weight measurements from the moment of birth to present day. Moreover, in the second follow-up questionnaire, we asked parents for their permission to be contacted again 12 months and 48 months after T2, so that they could send us a picture or screenshot of the updated measurement overview. This information allowed us to examine the effect of the preventive intervention program on the BMI of the child until approximately 2 years after the last follow-up measurement.

The trials were completed in November 2019 (trial 1) and February 2020 (trial 2). We are currently in the process of data cleaning. Effect analyses are thus underway, and the first results are expected to be submitted for publication in 2021.

Discussion

Principal Findings

The need for effective preventive intervention programs for childhood obesity is high, particularly among families with a lower SEP. The app-based parenting program *Samen Happie!* was developed primarily for these families and aims to stimulate healthy energy balance-related parenting practices from early childhood, before unhealthy energy balance-related habits have been established. More specifically, the program promotes both structured and autonomy-supportive practices and limits coercive controlling parenting practices with respect to all relevant energy balance-related determinants of childhood obesity (ie, dietary intake, sleep, physical activity, and screen time), with the ultimate goal of preventing children aged 0 to 4 years old from being overweight and obese. The successful development of the program was aided by the use of the IMP. The process of stepwise decision making made this large-scale and complex project manageable and contributed to thorough considerations. Regarding the selection of eminent EBRBs, for instance, we initially selected only parenting practices related to the children's dietary intake but we decided to also include practices relating to sleep, physical activity, and screen time after an extensive literature search and discussions with the target group. Moreover, insights from recent literature reviews

facilitated the selection of the most promising energy balance–related parenting practices (eg, parental support and modeling for physical activity). By facilitating a collaboration with experts and the target group, the IMP assured that the intended end users of the program were involved in multiple stages of program development. Overall, by integrating theory, empirical studies, professional knowledge, and the needs and preferences of the target group through continuous cocreation, we increased the chances of developing an effective preventive intervention program for childhood obesity [34].

Strengths, Limitations, and Directions for Future Research

Besides the use of the IMP, the app-based parenting program has several other notable strengths. First, previous digital preventive intervention programs for childhood obesity focused solely on the sociocognitive determinants of energy balance–related parenting [36]. A unique aspect of the *Samen Happie!* program is its focus on both sociocognitive (ie, knowledge, attitudes, and self-efficacy) and automatic (ie, habits) determinants, bridging the well-known gap between health parenting intentions and behaviors [88]. Furthermore, formulating long-term goals is potentially too proximal for parents with a lower SEP as their focus lies primarily on surviving in the here and now [119]. By using a self-regulatory planning approach with personally tailored prompts (eg, through implementation intentions), we facilitated the fulfillment of short-term goals. In addition, by assessing both general parenting style and parental mental well-being as potential moderators of the app-based prevention program, we might be able to identify groups of parents who might particularly benefit from the program, which could contribute to more personalized app usage. Finally, the app is an easy-to-use, stand-alone product that can potentially have a significant reach through its opportunities for widespread implementation.

One of the challenges of mobile health interventions is to keep users engaged for longer periods, which is particularly important for interventions targeting behavior change maintenance [120]. Although the preliminary results of the process evaluation indicated that parents appreciated the functionality, design, and content of the app to a reasonable degree, most parents who downloaded the app did not continue their app use over the course of several months. This might indicate that parents' information needs were fulfilled and new behavior patterns had been developed, explaining that further app use was no longer needed, but it could also suggest that the user engagement of the app should be increased. This is something that future research should consider. We will further develop the prevention program based on the results of the process evaluation and the input of potential program implementers (eg, youth health care professionals of community health services). On the basis of parents' suggestions for app improvement, the *Samen Happie!* website [121] has been developed (available in Dutch and English). This website contains more elaborate and structured parenting information and includes a search function, personal accounts of parents, and a forum on which parenting experiences

can be exchanged. Future research should examine whether parents' evaluations of the *Samen Happie!* program improve by offering them access to both the website and the app.

The app-based prevention program also has some limitations. With respect to the design of the trials, it was not possible to blind both participants and investigators to the allocation of conditions (ie, double blinding). As we used a waitlist control condition, the participating parents knew they would eventually receive an app about healthy parenting. However, neither the participants nor the investigators knew which trial condition the participants would be allocated to before randomization took place. Although double blinding in RCTs is generally recommended, methodological studies have shown that adequate allocation concealment is most important in minimizing bias [122,123]. Moreover, regarding our effect evaluations, we were unable to include in-depth measures of the 4 determinants that were targeted in the program (ie, knowledge, attitudes, self-efficacy, and habits), as it was imperative to keep our questionnaires short for our target group of parents with a lower SEP. Nevertheless, we included some questions that could serve as proxies for parental attitudes and self-efficacy. Even though this will give us some indication of the degree to which our preventive intervention program successfully targeted the selected determinants, future research should aim to include detailed measures of its program determinants to be able to examine the working mechanisms of the program.

In addition, future preventive intervention programs for childhood obesity should consider involving both caregivers. Although we intentionally targeted only primary caregivers (who turned out to be primarily mothers) for the recruitment, program materials, and questionnaires of our program, recent research has indicated that parents within a family differ in the energy balance–related parenting practices they perform [124]. This highlights the need for the inclusion of both parents in future preventive intervention programs targeting energy balance–related parenting. Finally, app-based parenting support might not be sufficient for the needs of parents of preschoolers entering their *terrible twos* and food neophobic phase [125,126], especially in the case of parents (with a lower SEP) who already experience parenting problems. To address the more extensive needs of these parents, future research should explore a combination of online tools with additional offline (group based [127]) counseling, which could provide a promising approach to change parenting attitudes and behaviors [128].

Conclusions

In conclusion, the IMP allowed for effective development of the app-based parenting program *Samen Happie!* to promote healthy energy balance–related parenting practices among parents of infants and preschoolers. By applying the IMP, including continued cocreation, the program specifically addressed the needs of parents with a lower SEP through a tailored program content and through theory-based behavior change techniques. This increases the potential of the program to prevent the development of obesity in early childhood among families with a lower SEP.

Acknowledgments

The authors want to thank Dio Agency for building the *Samen Happie!* app, the youth health care professionals and parents who were involved in the development of the program materials for their valuable input, the Netherlands Nutrition Centre for providing the information and visuals, Hannah Niermann for translating one of the questionnaires, and all students who were involved in the project for assisting with participant recruitment or otherwise. This project was funded by a grant from Fonds NutsOhra awarded to JKL (100.939). CW's work was supported by a Jacobs Foundation Advanced Research Fellowship and a grant from the Netherlands Organization for Scientific Research (016.Vici.185.038). The funding agencies were not involved in the design and execution of the trial and will not be involved in the analysis or interpretation of the data or in the decision to publish the results. More information about the *Samen Happie!* program can be found at the *Samen Happie!* website [121] (available in Dutch and English).

Authors' Contributions

JKL wrote the funding application. JKL, LTK, JMV, and CW designed the study. JKL, CPMK, RCJH, ELMR, LTK, JMV, and CW developed the preventive intervention program and questionnaires. LTK and JKL drafted the manuscript. JKL, JMV, CW, SPJK, and ELMR supervised the project. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Logic model of the health problem addressed in the Samen Happie! program.

[DOCX File, 2075 KB - [formative_v5i5e24802_app1.docx](#)]

Multimedia Appendix 2

Overview of the performance objectives (Tables a-c) and change objectives (Tables d-f) specified for the Samen Happie! program.

[DOCX File, 21 KB - [formative_v5i5e24802_app2.docx](#)]

Multimedia Appendix 3

Illustrated examples of a lesson and a challenge in the Samen Happie! program.

[PDF File (Adobe PDF File), 3543 KB - [formative_v5i5e24802_app3.pdf](#)]

Multimedia Appendix 4

Overview of the modules and corresponding lessons and challenges of the Samen Happie! app.

[DOCX File, 16 KB - [formative_v5i5e24802_app4.docx](#)]

Multimedia Appendix 5

Flowchart of the design and timelines for both trials of the Samen Happie! program.

[DOCX File, 40 KB - [formative_v5i5e24802_app5.docx](#)]

Multimedia Appendix 6

Overview of constructs, variables, and assessment points included in the evaluation of the Samen Happie! program.

[DOCX File, 15 KB - [formative_v5i5e24802_app6.docx](#)]

References

1. Wabitsch M, Moss A, Kromeyer-Hauschild K. Unexpected plateauing of childhood obesity rates in developed countries. *BMC Med* 2014 Jan 31;12:17 [FREE Full text] [doi: [10.1186/1741-7015-12-17](#)] [Medline: [24485015](#)]
2. RIVM monitor kansrijke start 2019. 2019. URL: <https://www.kansrijkestart.nl/actieprogramma-kansrijke-start/documenten/publicaties/2019/12/11/rivm-monitor-kansrijke-start-2019> [accessed 2021-04-13]
3. Chung A, Backholer K, Wong E, Palermo C, Keating C, Peeters A. Trends in child and adolescent obesity prevalence in economically advanced countries according to socioeconomic position: a systematic review. *Obes Rev* 2016;17(3):276-295. [doi: [10.1111/obr.12360](#)] [Medline: [26693831](#)]
4. Trend overgewicht naar opleiding. CBS. 2019. URL: <https://www.volksgezondheidenzorg.info/onderwerp/overgewicht/cijfers-context/trends#node-trend-overgewicht-kinderen-naar-leeftijd> [accessed 2020-08-03]
5. Taveras EM, Gillman MW, Kleinman KP, Rich-Edwards JW, Rifas-Shiman SL. Reducing racial/ethnic disparities in childhood obesity: the role of early life risk factors. *JAMA Pediatr* 2013 Aug 01;167(8):731-738 [FREE Full text] [doi: [10.1001/jamapediatrics.2013.85](#)] [Medline: [23733179](#)]

6. Lioret S, Campbell KJ, McNaughton SA, Cameron AJ, Salmon J, Abbott G, et al. Lifestyle patterns begin in early childhood, persist and are socioeconomically patterned, confirming the importance of early life interventions. *Nutrients* 2020 Mar 09;12(3):724 [FREE Full text] [doi: [10.3390/nu12030724](https://doi.org/10.3390/nu12030724)] [Medline: [32182889](https://pubmed.ncbi.nlm.nih.gov/32182889/)]
7. Skouteris H, McCabe M, Swinburn B, Newgreen V, Sacher P, Chadwick P. Parental influence and obesity prevention in pre-schoolers: a systematic review of interventions. *Obes Rev* 2011 May;12(5):315-328. [doi: [10.1111/j.1467-789X.2010.00751.x](https://doi.org/10.1111/j.1467-789X.2010.00751.x)] [Medline: [20492538](https://pubmed.ncbi.nlm.nih.gov/20492538/)]
8. Power TG, Sleddens EF, Berge J, Connell L, Govig B, Hennessy E, et al. Contemporary research on parenting: conceptual, methodological, and translational issues. *Child Obes* 2013 Aug;9 Suppl:87-94 [FREE Full text] [doi: [10.1089/chi.2013.0038](https://doi.org/10.1089/chi.2013.0038)] [Medline: [23944927](https://pubmed.ncbi.nlm.nih.gov/23944927/)]
9. Larsen JK, Hermans RC, Sleddens EF, Engels RC, Fisher JO, Kremers SP. How parental dietary behavior and food parenting practices affect children's dietary behavior. Interacting sources of influence? *Appetite* 2015 Jun;89:246-257. [doi: [10.1016/j.appet.2015.02.012](https://doi.org/10.1016/j.appet.2015.02.012)] [Medline: [25681294](https://pubmed.ncbi.nlm.nih.gov/25681294/)]
10. Vaughn AE, Ward DS, Fisher JO, Faith MS, Hughes SO, Kremers SP, et al. Fundamental constructs in food parenting practices: a content map to guide future research. *Nutr Rev* 2016 Feb;74(2):98-117 [FREE Full text] [doi: [10.1093/nutrit/nuv061](https://doi.org/10.1093/nutrit/nuv061)] [Medline: [26724487](https://pubmed.ncbi.nlm.nih.gov/26724487/)]
11. Spill MK, Callahan EH, Shapiro MJ, Spahn JM, Wong YP, Benjamin-Neelon SE, et al. Caregiver feeding practices and child weight outcomes: a systematic review. *Am J Clin Nutr* 2019 Mar 01;109(Suppl_7):990-1002. [doi: [10.1093/ajcn/nqy276](https://doi.org/10.1093/ajcn/nqy276)] [Medline: [30982865](https://pubmed.ncbi.nlm.nih.gov/30982865/)]
12. George SM, Agosto Y, Rojas LM, Soares M, Bahamon M, Prado G, et al. A developmental cascade perspective of paediatric obesity: a systematic review of preventive interventions from infancy through late adolescence. *Obes Rev* 2020 Feb;21(2):e12939 [FREE Full text] [doi: [10.1111/obr.12939](https://doi.org/10.1111/obr.12939)] [Medline: [31808277](https://pubmed.ncbi.nlm.nih.gov/31808277/)]
13. Balantekin KN, Anzman-Frasca S, Francis LA, Ventura AK, Fisher JO, Johnson SL. Positive parenting approaches and their association with child eating and weight: a narrative review from infancy to adolescence. *Pediatr Obes* 2020 Oct;15(10):e12722 [FREE Full text] [doi: [10.1111/ijpo.12722](https://doi.org/10.1111/ijpo.12722)] [Medline: [32881344](https://pubmed.ncbi.nlm.nih.gov/32881344/)]
14. Hutchens A, Lee RE. Parenting practices and children's physical activity: an integrative review. *J Sch Nurs* 2018 Feb;34(1):68-85. [doi: [10.1177/1059840517714852](https://doi.org/10.1177/1059840517714852)] [Medline: [28631518](https://pubmed.ncbi.nlm.nih.gov/28631518/)]
15. Petersen TL, Møller LB, Brønd JC, Jepsen R, Grøntved A. Association between parent and child physical activity: a systematic review. *Int J Behav Nutr Phys Act* 2020 May 18;17(1):67 [FREE Full text] [doi: [10.1186/s12966-020-00966-z](https://doi.org/10.1186/s12966-020-00966-z)] [Medline: [32423407](https://pubmed.ncbi.nlm.nih.gov/32423407/)]
16. Laws R, Campbell KJ, van der Pligt P, Russell G, Ball K, Lynch J, et al. The impact of interventions to prevent obesity or improve obesity related behaviours in children (0-5 years) from socioeconomically disadvantaged and/or indigenous families: a systematic review. *BMC Public Health* 2014 Aug 01;14:779 [FREE Full text] [doi: [10.1186/1471-2458-14-779](https://doi.org/10.1186/1471-2458-14-779)] [Medline: [25084804](https://pubmed.ncbi.nlm.nih.gov/25084804/)]
17. Seidler AL, Hunter KE, Johnson BJ, Ekambareshwar M, Taki S, Mauch CE, et al. Understanding, comparing and learning from the four EPOCH early childhood obesity prevention interventions: a multi-methods study. *Pediatr Obes* 2020 Nov;15(11):e12679. [doi: [10.1111/ijpo.12679](https://doi.org/10.1111/ijpo.12679)] [Medline: [32543054](https://pubmed.ncbi.nlm.nih.gov/32543054/)]
18. Ash T, Agaronov A, Young T, Aftosmes-Tobio A, Davison KK. Family-based childhood obesity prevention interventions: a systematic review and quantitative content analysis. *Int J Behav Nutr Phys Act* 2017 Aug 24;14(1):113 [FREE Full text] [doi: [10.1186/s12966-017-0571-2](https://doi.org/10.1186/s12966-017-0571-2)] [Medline: [28836983](https://pubmed.ncbi.nlm.nih.gov/28836983/)]
19. van Ansem WJ, van Lenthe FJ, Schrijvers CT, Rodenburg G, van de Mheen D. Socio-economic inequalities in children's snack consumption and sugar-sweetened beverage consumption: the contribution of home environmental factors. *Br J Nutr* 2014 Aug 14;112(3):467-476. [doi: [10.1017/S0007114514001007](https://doi.org/10.1017/S0007114514001007)] [Medline: [24833428](https://pubmed.ncbi.nlm.nih.gov/24833428/)]
20. Askie LM, Espinoza D, Martin A, Daniels LA, Mihrshahi S, Taylor R, et al. Interventions commenced by early infancy to prevent childhood obesity-The EPOCH Collaboration: an individual participant data prospective meta-analysis of four randomized controlled trials. *Pediatr Obes* 2020 Jun;15(6):e12618. [doi: [10.1111/ijpo.12618](https://doi.org/10.1111/ijpo.12618)] [Medline: [32026653](https://pubmed.ncbi.nlm.nih.gov/32026653/)]
21. Hayba N, Partridge SR, Nour MM, Grech A, Farinelli MA. Effectiveness of lifestyle interventions for preventing harmful weight gain among young adults from lower socioeconomic status and ethnically diverse backgrounds: a systematic review. *Obes Rev* 2018 Mar;19(3):333-346. [doi: [10.1111/obr.12641](https://doi.org/10.1111/obr.12641)] [Medline: [29178423](https://pubmed.ncbi.nlm.nih.gov/29178423/)]
22. Garcia-Dominic O, Wray LA, Treviño RP, Hernandez AE, Yin Z, Ulbrecht JS. Identifying barriers that hinder onsite parental involvement in a school-based health promotion program. *Health Promot Pract* 2010 Sep;11(5):703-713 [FREE Full text] [doi: [10.1177/1524839909331909](https://doi.org/10.1177/1524839909331909)] [Medline: [19339644](https://pubmed.ncbi.nlm.nih.gov/19339644/)]
23. Nagelhout GE, Abidi L, de Vries H. Reasons for (not) participating in a community-based health promotion program for low-income multi-problem households in the Netherlands: a qualitative study. *Health Soc Care Community* 2021 Jan;29(1):241-249 [FREE Full text] [doi: [10.1111/hsc.13087](https://doi.org/10.1111/hsc.13087)] [Medline: [32633021](https://pubmed.ncbi.nlm.nih.gov/32633021/)]
24. Winnail SD, Geiger BF, Macrina DM, Snyder S, Petri CJ, Nagy S. Barriers to parent involvement in middle school health education. *Am J Health Stud* 2000 May;16(4):193-198 [FREE Full text]
25. Smokowski P, Corona R, Bacallao M, Fortson BL, Marshall KJ, Yaros A. Addressing barriers to recruitment and retention in the implementation of parenting programs: lessons learned for effective program delivery in rural and urban areas. *J Child Fam Stud* 2018 Sep;27(9):2925-2942 [FREE Full text] [doi: [10.1007/s10826-018-1139-8](https://doi.org/10.1007/s10826-018-1139-8)] [Medline: [30100698](https://pubmed.ncbi.nlm.nih.gov/30100698/)]

26. Brannon EE, Kuhl ES, Boles RE, Aylward BS, Ratcliff MB, Valenzuela JM, et al. Strategies for recruitment and retention of families from low-income, ethnic minority backgrounds in a longitudinal study of caregiver feeding and child weight. *Child Health Care* 2013;42(3):198-213 [FREE Full text] [doi: [10.1080/02739615.2013.816590](https://doi.org/10.1080/02739615.2013.816590)] [Medline: [24078763](https://pubmed.ncbi.nlm.nih.gov/24078763/)]
27. Whittaker R, Merry S, Dorey E, Maddison R. A development and evaluation process for mHealth interventions: examples from New Zealand. *J Health Commun* 2012;17 Suppl 1:11-21. [doi: [10.1080/10810730.2011.649103](https://doi.org/10.1080/10810730.2011.649103)] [Medline: [22548594](https://pubmed.ncbi.nlm.nih.gov/22548594/)]
28. Tate EB, Spruijt-Metz D, O'Reilly G, Jordan-Marsh M, Gotsis M, Pentz MA, et al. mHealth approaches to child obesity prevention: successes, unique challenges, and next directions. *Transl Behav Med* 2013 Dec;3(4):406-415 [FREE Full text] [doi: [10.1007/s13142-013-0222-3](https://doi.org/10.1007/s13142-013-0222-3)] [Medline: [24294329](https://pubmed.ncbi.nlm.nih.gov/24294329/)]
29. Ondersteuning bij eHealth voor iedereen. Pharos. 2020. URL: <https://www.pharos.nl/wp-content/uploads/2020/01/Pharos-Zigzagfolder-Ondersteuning-bij-eHealth-voor-iedereen.pdf> [accessed 2021-04-13]
30. Fjeldsoe BS, Marshall AL, Miller YD. Behavior change interventions delivered by mobile telephone short-message service. *Am J Prev Med* 2009 Feb;36(2):165-173. [doi: [10.1016/j.amepre.2008.09.040](https://doi.org/10.1016/j.amepre.2008.09.040)] [Medline: [19135907](https://pubmed.ncbi.nlm.nih.gov/19135907/)]
31. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. *JMIR Mhealth Uhealth* 2015 Mar 11;3(1):e27 [FREE Full text] [doi: [10.2196/mhealth.3422](https://doi.org/10.2196/mhealth.3422)] [Medline: [25760773](https://pubmed.ncbi.nlm.nih.gov/25760773/)]
32. McKay FH, Slykerman S, Dunn M. The app behavior change scale: creation of a scale to assess the potential of apps to promote behavior change. *JMIR Mhealth Uhealth* 2019 Jan 25;7(1):e11130 [FREE Full text] [doi: [10.2196/11130](https://doi.org/10.2196/11130)] [Medline: [30681967](https://pubmed.ncbi.nlm.nih.gov/30681967/)]
33. Fjeldsoe BS, Miller YD, O'Brien JL, Marshall AL. Iterative development of MobileMums: a physical activity intervention for women with young children. *Int J Behav Nutr Phys Act* 2012 Dec 20;9:151 [FREE Full text] [doi: [10.1186/1479-5868-9-151](https://doi.org/10.1186/1479-5868-9-151)] [Medline: [23256730](https://pubmed.ncbi.nlm.nih.gov/23256730/)]
34. Eldredge LK, Markham CM, Rutter RA, Fernández ME, Kok G. Planning Health Promotion Programs: An Intervention Mapping Approach. San Francisco, CA: Jossey-Bass; 2016:1-704.
35. Chernick LS, Santelli J, Gonzalez AE, Mitchell JA, Ehrhardt AA, Bakken S, et al. The development of a theory-based, user-informed, digital intervention to promote pregnancy prevention among adolescent female emergency department patients. *J Adolesc Health* 2019 Feb 23;64(2):127-128 [FREE Full text] [doi: [10.1016/j.jadohealth.2018.10.267](https://doi.org/10.1016/j.jadohealth.2018.10.267)]
36. De Lepeleere S, Verloigne M, Brown HE, Cardon G, De Bourdeaudhuij I. Using the Intervention Mapping Protocol to develop an online video intervention for parents to prevent childhood obesity: movie models. *Glob Health Promot* 2018 Jun;25(2):56-66. [doi: [10.1177/1757975916658603](https://doi.org/10.1177/1757975916658603)] [Medline: [27503911](https://pubmed.ncbi.nlm.nih.gov/27503911/)]
37. DeSmet A, Van Cleemput K, Bastiaensens S, Poels K, Vandebosch H, Malliet S, et al. Bridging behavior science and gaming theory: using the Intervention Mapping Protocol to design a serious game against cyberbullying. *Comput Hum Behav* 2016 Mar;56:337-351. [doi: [10.1016/j.chb.2015.11.039](https://doi.org/10.1016/j.chb.2015.11.039)]
38. Jacobs NC, Völlink T, Dehue F, Lechner L. Online Pestkoppenstoppen: systematic and theory-based development of a web-based tailored intervention for adolescent cyberbully victims to combat and prevent cyberbullying. *BMC Public Health* 2014 Apr 24;14:396 [FREE Full text] [doi: [10.1186/1471-2458-14-396](https://doi.org/10.1186/1471-2458-14-396)] [Medline: [24758264](https://pubmed.ncbi.nlm.nih.gov/24758264/)]
39. Li DH, Moskowitz DA, Macapagal K, Saber R, Mustanski B. Using intervention mapping to developmentally adapt an online HIV risk reduction program for adolescent men who have sex with men. *Prev Sci* 2020 Oct;21(7):885-897. [doi: [10.1007/s11121-020-01148-w](https://doi.org/10.1007/s11121-020-01148-w)] [Medline: [32761287](https://pubmed.ncbi.nlm.nih.gov/32761287/)]
40. Scheerman JF, van Empelen P, van Loveren C, van Meijel B. A mobile app (WhiteTeeth) to promote good oral health behavior among Dutch adolescents with fixed orthodontic appliances: intervention mapping approach. *JMIR Mhealth Uhealth* 2018 Aug 17;6(8):e163 [FREE Full text] [doi: [10.2196/mhealth.9626](https://doi.org/10.2196/mhealth.9626)] [Medline: [30120085](https://pubmed.ncbi.nlm.nih.gov/30120085/)]
41. Voogt CV, Poelen EA, Kleinjan M, Lemmers LA, Engels RC. The development of a web-based brief alcohol intervention in reducing heavy drinking among college students: an Intervention Mapping approach. *Health Promot Int* 2014 Dec;29(4):669-679. [doi: [10.1093/heapro/dat016](https://doi.org/10.1093/heapro/dat016)] [Medline: [23525645](https://pubmed.ncbi.nlm.nih.gov/23525645/)]
42. Green LW, Kreuter M. Health Program Planning: An Educational and Ecological Approach. New York: McGraw-Hill Higher Education; 2005:1-600.
43. Michie S, Jochelson K, Markham WA, Bridle C. Low-income groups and behaviour change interventions: a review of intervention content, effectiveness and theoretical frameworks. *J Epidemiol Community Health* 2009 Aug;63(8):610-622. [doi: [10.1136/jech.2008.078725](https://doi.org/10.1136/jech.2008.078725)] [Medline: [19386612](https://pubmed.ncbi.nlm.nih.gov/19386612/)]
44. Monasta L, Batty GD, Cattaneo A, Lutje V, Ronfani L, Van Lenthe FJ, et al. Early-life determinants of overweight and obesity: a review of systematic reviews. *Obes Rev* 2010 Oct;11(10):695-708. [doi: [10.1111/j.1467-789X.2010.00735.x](https://doi.org/10.1111/j.1467-789X.2010.00735.x)] [Medline: [20331509](https://pubmed.ncbi.nlm.nih.gov/20331509/)]
45. Singh AS, Mulder C, Twisk JW, van Mechelen W, Chinapaw MJ. Tracking of childhood overweight into adulthood: a systematic review of the literature. *Obes Rev* 2008 Sep;9(5):474-488. [doi: [10.1111/j.1467-789X.2008.00475.x](https://doi.org/10.1111/j.1467-789X.2008.00475.x)] [Medline: [18331423](https://pubmed.ncbi.nlm.nih.gov/18331423/)]
46. Sahoo K, Sahoo B, Choudhury AK, Sofi NY, Kumar R, Bhadoria AS. Childhood obesity: causes and consequences. *J Family Med Prim Care* 2015;4(2):187-192 [FREE Full text] [doi: [10.4103/2249-4863.154628](https://doi.org/10.4103/2249-4863.154628)] [Medline: [25949965](https://pubmed.ncbi.nlm.nih.gov/25949965/)]

47. Malik VS, Pan A, Willett WC, Hu FB. Sugar-sweetened beverages and weight gain in children and adults: a systematic review and meta-analysis. *Am J Clin Nutr* 2013 Oct;98(4):1084-1102 [FREE Full text] [doi: [10.3945/ajcn.113.058362](https://doi.org/10.3945/ajcn.113.058362)] [Medline: [23966427](https://pubmed.ncbi.nlm.nih.gov/23966427/)]
48. Luger M, Lafontan M, Bes-Rastrollo M, Winzer E, Yumuk V, Farpour-Lambert N. Sugar-sweetened beverages and weight gain in children and adults: a systematic review from 2013 to 2015 and a comparison with previous studies. *Obes Facts* 2017;10(6):674-693 [FREE Full text] [doi: [10.1159/000484566](https://doi.org/10.1159/000484566)] [Medline: [29237159](https://pubmed.ncbi.nlm.nih.gov/29237159/)]
49. Pérez-Escamilla R, Obbagy JE, Altman JM, Essery EV, McGrane MM, Wong YP, et al. Dietary energy density and body weight in adults and children: a systematic review. *J Acad Nutr Diet* 2012 May;112(5):671-684. [doi: [10.1016/j.jand.2012.01.020](https://doi.org/10.1016/j.jand.2012.01.020)] [Medline: [22480489](https://pubmed.ncbi.nlm.nih.gov/22480489/)]
50. Chen X, Beydoun MA, Wang Y. Is sleep duration associated with childhood obesity? A systematic review and meta-analysis. *Obesity (Silver Spring)* 2008 Feb;16(2):265-274 [FREE Full text] [doi: [10.1038/oby.2007.63](https://doi.org/10.1038/oby.2007.63)] [Medline: [18239632](https://pubmed.ncbi.nlm.nih.gov/18239632/)]
51. Miller MA, Kruisbrink M, Wallace J, Ji C, Cappuccio FP. Sleep duration and incidence of obesity in infants, children, and adolescents: a systematic review and meta-analysis of prospective studies. *Sleep* 2018 Apr 01;41(4):- [doi: [10.1093/sleep/zsy018](https://doi.org/10.1093/sleep/zsy018)] [Medline: [29401314](https://pubmed.ncbi.nlm.nih.gov/29401314/)]
52. Fatima Y, Doi SA, Mamun AA. Longitudinal impact of sleep on overweight and obesity in children and adolescents: a systematic review and bias-adjusted meta-analysis. *Obes Rev* 2015 Feb;16(2):137-149. [doi: [10.1111/obr.12245](https://doi.org/10.1111/obr.12245)] [Medline: [25589359](https://pubmed.ncbi.nlm.nih.gov/25589359/)]
53. Fang K, Mu M, Liu K, He Y. Screen time and childhood overweight/obesity: a systematic review and meta-analysis. *Child Care Health Dev* 2019 Sep;45(5):744-753. [doi: [10.1111/cch.12701](https://doi.org/10.1111/cch.12701)] [Medline: [31270831](https://pubmed.ncbi.nlm.nih.gov/31270831/)]
54. Darmon N, Drewnowski A. Does social class predict diet quality? *Am J Clin Nutr* 2008 May;87(5):1107-1117. [doi: [10.1093/ajcn/87.5.1107](https://doi.org/10.1093/ajcn/87.5.1107)] [Medline: [18469226](https://pubmed.ncbi.nlm.nih.gov/18469226/)]
55. El-Sheikh M, Bagley EJ, Keiley M, Elmore-Staton L, Chen E, Buckhalt JA. Economic adversity and children's sleep problems: multiple indicators and moderation of effects. *Health Psychol* 2013 Aug;32(8):849-859 [FREE Full text] [doi: [10.1037/a0030413](https://doi.org/10.1037/a0030413)] [Medline: [23148451](https://pubmed.ncbi.nlm.nih.gov/23148451/)]
56. Tandon PS, Zhou C, Sallis JF, Cain KL, Frank LD, Saelens BE. Home environment relationships with children's physical activity, sedentary time, and screen time by socioeconomic status. *Int J Behav Nutr Phys Act* 2012 Jul 26;9:88 [FREE Full text] [doi: [10.1186/1479-5868-9-88](https://doi.org/10.1186/1479-5868-9-88)] [Medline: [22835155](https://pubmed.ncbi.nlm.nih.gov/22835155/)]
57. Lindsay AC, Sussner KM, Kim J, Gortmaker S. The role of parents in preventing childhood obesity. *Future Child* 2006;16(1):169-186. [doi: [10.1353/foc.2006.0006](https://doi.org/10.1353/foc.2006.0006)] [Medline: [16532663](https://pubmed.ncbi.nlm.nih.gov/16532663/)]
58. Gerards SM, Kremers SP. The role of food parenting skills and the home food environment in children's weight gain and obesity. *Curr Obes Rep* 2015 Mar;4(1):30-36 [FREE Full text] [doi: [10.1007/s13679-015-0139-x](https://doi.org/10.1007/s13679-015-0139-x)] [Medline: [25741454](https://pubmed.ncbi.nlm.nih.gov/25741454/)]
59. Larsen JK, Sleddens EF, Vink JM, Fisher JO, Kremers SP. General parenting styles and children's obesity risk: changing focus. *Front Psychol* 2018;9:2119 [FREE Full text] [doi: [10.3389/fpsyg.2018.02119](https://doi.org/10.3389/fpsyg.2018.02119)] [Medline: [30459686](https://pubmed.ncbi.nlm.nih.gov/30459686/)]
60. Patrick H, Hennessy E, McSpadden K, Oh A. Parenting styles and practices in children's obesogenic behaviors: scientific gaps and future research directions. *Child Obes* 2013 Aug;9 Suppl:73-86 [FREE Full text] [doi: [10.1089/chi.2013.0039](https://doi.org/10.1089/chi.2013.0039)] [Medline: [23944926](https://pubmed.ncbi.nlm.nih.gov/23944926/)]
61. Philips N, Sioen I, Michels N, Sleddens E, De Henauw S. The influence of parenting style on health related behavior of children: findings from the ChiBS study. *Int J Behav Nutr Phys Act* 2014 Jul 23;11:95 [FREE Full text] [doi: [10.1186/s12966-014-0095-y](https://doi.org/10.1186/s12966-014-0095-y)] [Medline: [25052905](https://pubmed.ncbi.nlm.nih.gov/25052905/)]
62. Rodenburg G, Oenema A, Kremers SP, van de Mheen D. Clustering of diet- and activity-related parenting practices: cross-sectional findings of the INPACT study. *Int J Behav Nutr Phys Act* 2013 Mar 25;10:36 [FREE Full text] [doi: [10.1186/1479-5868-10-36](https://doi.org/10.1186/1479-5868-10-36)] [Medline: [23531232](https://pubmed.ncbi.nlm.nih.gov/23531232/)]
63. Grolnick WS, Pomerantz EM. Issues and challenges in studying parental control: toward a new conceptualization. *Child Dev Perspect* 2009;3(3):165-170. [doi: [10.1111/j.1750-8606.2009.00099.x](https://doi.org/10.1111/j.1750-8606.2009.00099.x)]
64. Grolnick WS, Deci EL, Ryan RM. In: Grusec JE, Kuczynski L, editors. *Parenting and Children's Internalization of Values: A Handbook of Contemporary Theory*. Hoboken, New Jersey, United States: John Wiley & Sons Inc; 1997:1-439.
65. Paul IM, Bartok CJ, Downs DS, Stifter CA, Ventura AK, Birch LL. Opportunities for the primary prevention of obesity during infancy. *Adv Pediatr* 2009;56:107-133 [FREE Full text] [doi: [10.1016/j.yapd.2009.08.012](https://doi.org/10.1016/j.yapd.2009.08.012)] [Medline: [19968945](https://pubmed.ncbi.nlm.nih.gov/19968945/)]
66. Yee AZ, Lwin MO, Ho SS. The influence of parental practices on child promotive and preventive food consumption behaviors: a systematic review and meta-analysis. *Int J Behav Nutr Phys Act* 2017 Apr 11;14(1):47 [FREE Full text] [doi: [10.1186/s12966-017-0501-3](https://doi.org/10.1186/s12966-017-0501-3)] [Medline: [28399881](https://pubmed.ncbi.nlm.nih.gov/28399881/)]
67. Chamberlin LA, Sherman SN, Jain A, Powers SW, Whitaker RC. The challenge of preventing and treating obesity in low-income, preschool children: perceptions of WIC health care professionals. *Arch Pediatr Adolesc Med* 2002 Jul;156(7):662-668. [doi: [10.1001/archpedi.156.7.662](https://doi.org/10.1001/archpedi.156.7.662)] [Medline: [12090832](https://pubmed.ncbi.nlm.nih.gov/12090832/)]
68. Jain A, Sherman SN, Chamberlin LA, Carter Y, Powers SW, Whitaker RC. Why don't low-income mothers worry about their preschoolers being overweight? *Pediatrics* 2001 May;107(5):1138-1146. [doi: [10.1542/peds.107.5.1138](https://doi.org/10.1542/peds.107.5.1138)] [Medline: [11331699](https://pubmed.ncbi.nlm.nih.gov/11331699/)]

69. Clark HR, Goyder E, Bissell P, Blank L, Walters SJ, Peters J. A pilot survey of socio-economic differences in child-feeding behaviours among parents of primary-school children. *Public Health Nutr* 2008 Oct;11(10):1030-1036. [doi: [10.1017/S1368980007001401](https://doi.org/10.1017/S1368980007001401)] [Medline: [18093351](https://pubmed.ncbi.nlm.nih.gov/18093351/)]
70. Pinket A, De Craemer M, De Bourdeaudhuij I, Deforche B, Cardon G, Androustos O, et al. Can parenting practices explain the differences in beverage intake according to socio-economic status: the Toybox-study. *Nutrients* 2016 Sep 23;8(10):- [FREE Full text] [doi: [10.3390/nu8100591](https://doi.org/10.3390/nu8100591)] [Medline: [27669290](https://pubmed.ncbi.nlm.nih.gov/27669290/)]
71. Loth KA, MacLehose RF, Fulkerson JA, Crow S, Neumark-Sztainer D. Eat this, not that! Parental demographic correlates of food-related parenting practices. *Appetite* 2013 Jan;60(1):140-147 [FREE Full text] [doi: [10.1016/j.appet.2012.09.019](https://doi.org/10.1016/j.appet.2012.09.019)] [Medline: [23022556](https://pubmed.ncbi.nlm.nih.gov/23022556/)]
72. Cardel M, Willig AL, Dulin-Keita A, Casazza K, Beasley TM, Fernández JR. Parental feeding practices and socioeconomic status are associated with child adiposity in a multi-ethnic sample of children. *Appetite* 2012 Feb;58(1):347-353 [FREE Full text] [doi: [10.1016/j.appet.2011.11.005](https://doi.org/10.1016/j.appet.2011.11.005)] [Medline: [22100186](https://pubmed.ncbi.nlm.nih.gov/22100186/)]
73. Zarnowiecki DM, Dollman J, Parletta N. Associations between predictors of children's dietary intake and socioeconomic position: a systematic review of the literature. *Obes Rev* 2014 May;15(5):375-391. [doi: [10.1111/obr.12139](https://doi.org/10.1111/obr.12139)] [Medline: [24433310](https://pubmed.ncbi.nlm.nih.gov/24433310/)]
74. Anderson SE, Whitaker RC. Household routines and obesity in US preschool-aged children. *Pediatrics* 2010 Mar;125(3):420-428. [doi: [10.1542/peds.2009-0417](https://doi.org/10.1542/peds.2009-0417)] [Medline: [20142280](https://pubmed.ncbi.nlm.nih.gov/20142280/)]
75. Hale L, Berger LM, LeBourgeois MK, Brooks-Gunn J. Social and demographic predictors of preschoolers' bedtime routines. *J Dev Behav Pediatr* 2009 Oct;30(5):394-402 [FREE Full text] [doi: [10.1097/DBP.0b013e3181ba0e64](https://doi.org/10.1097/DBP.0b013e3181ba0e64)] [Medline: [19745760](https://pubmed.ncbi.nlm.nih.gov/19745760/)]
76. Jones CH, Ball H. Exploring socioeconomic differences in bedtime behaviours and sleep duration in English preschool children. *Infant Child Dev* 2014 Sep;23(5):518-531 [FREE Full text] [doi: [10.1002/icd.1848](https://doi.org/10.1002/icd.1848)] [Medline: [25598710](https://pubmed.ncbi.nlm.nih.gov/25598710/)]
77. Vereecken CA, Keukelier E, Maes L. Influence of mother's educational level on food parenting practices and food habits of young children. *Appetite* 2004 Aug;43(1):93-103. [doi: [10.1016/j.appet.2004.04.002](https://doi.org/10.1016/j.appet.2004.04.002)] [Medline: [15262022](https://pubmed.ncbi.nlm.nih.gov/15262022/)]
78. Darling N, Steinberg L. Parenting style as context: an integrative model. *Psychol Bull* 1993;113(3):487-496. [doi: [10.1037/0033-2909.113.3.487](https://doi.org/10.1037/0033-2909.113.3.487)]
79. Sleddens EF, Kremers SP, Stafleu A, Dagnelie PC, De Vries NK, Thijs C. Food parenting practices and child dietary behavior. Prospective relations and the moderating role of general parenting. *Appetite* 2014 Aug;79:42-50. [doi: [10.1016/j.appet.2014.04.004](https://doi.org/10.1016/j.appet.2014.04.004)] [Medline: [24727101](https://pubmed.ncbi.nlm.nih.gov/24727101/)]
80. Baumrind D. Current patterns of parental authority. *Dev Psychol* 1971 Jun;4(1):1-103. [doi: [10.1037/h0030372](https://doi.org/10.1037/h0030372)]
81. Gerards SM, Niermann C, Gevers DW, Eussen N, Kremers SP. Context matters! The relationship between mother-reported family nutrition climate, general parenting, food parenting practices and children's BMI. *BMC Public Health* 2016 Sep 27;16(1):1018 [FREE Full text] [doi: [10.1186/s12889-016-3683-8](https://doi.org/10.1186/s12889-016-3683-8)] [Medline: [27677380](https://pubmed.ncbi.nlm.nih.gov/27677380/)]
82. Kowalenko NM, Mares SP, Newman LK, Williams AE, Powrie RM, van Doesum KT. Family matters: infants, toddlers and preschoolers of parents affected by mental illness. *Med J Aust* 2013 Aug 05;199(3 Suppl):14-17. [doi: [10.5694/mja11.11285](https://doi.org/10.5694/mja11.11285)] [Medline: [25369842](https://pubmed.ncbi.nlm.nih.gov/25369842/)]
83. de Vries H. An integrated approach for understanding health behavior; the I-Change model as an example. *Psychol Behav Sci Int J* 2017 Mar 9;2(2):161-170 [FREE Full text] [doi: [10.19080/PBSIJ.2017.02.555585](https://doi.org/10.19080/PBSIJ.2017.02.555585)]
84. Ajzen I. The theory of planned behavior. *Organ Behav Hum Decis Process* 1991 Dec;50(2):179-211. [doi: [10.1016/0749-5978\(91\)90020-T](https://doi.org/10.1016/0749-5978(91)90020-T)]
85. Bandura A. Social cognitive theory: an agentic perspective. *Annu Rev Psychol* 2001;52:1-26. [doi: [10.1146/annurev.psych.52.1.1](https://doi.org/10.1146/annurev.psych.52.1.1)] [Medline: [11148297](https://pubmed.ncbi.nlm.nih.gov/11148297/)]
86. Prochaska J, Diclemente CC. Toward a comprehensive model of change. In: *Treating Addictive Behaviors*. Switzerland: Springer Nature; 1986:3-27.
87. Janz NK, Becker MH. The Health Belief Model: a decade later. *Health Educ Q* 1984 Jan 01;11(1):1-47. [doi: [10.1177/109019818401100101](https://doi.org/10.1177/109019818401100101)] [Medline: [6392204](https://pubmed.ncbi.nlm.nih.gov/6392204/)]
88. Larsen JK, Hermans RC, Sleddens EF, Vink JM, Kremers SP, Rutter EL, et al. How to bridge the intention-behavior gap in food parenting: automatic constructs and underlying techniques. *Appetite* 2018 Apr 01;123:191-200. [doi: [10.1016/j.appet.2017.12.016](https://doi.org/10.1016/j.appet.2017.12.016)] [Medline: [29277519](https://pubmed.ncbi.nlm.nih.gov/29277519/)]
89. Cluss PA, Ewing L, King WC, Reis EC, Dodd JL, Penner B. Nutrition knowledge of low-income parents of obese children. *Transl Behav Med* 2013 Jun;3(2):218-225 [FREE Full text] [doi: [10.1007/s13142-013-0203-6](https://doi.org/10.1007/s13142-013-0203-6)] [Medline: [24039639](https://pubmed.ncbi.nlm.nih.gov/24039639/)]
90. Vereecken C, Maes L. Young children's dietary habits and associations with the mothers' nutritional knowledge and attitudes. *Appetite* 2010 Feb;54(1):44-51. [doi: [10.1016/j.appet.2009.09.005](https://doi.org/10.1016/j.appet.2009.09.005)] [Medline: [19751782](https://pubmed.ncbi.nlm.nih.gov/19751782/)]
91. von Wagner C, Knight K, Steptoe A, Wardle J. Functional health literacy and health-promoting behaviour in a national sample of British adults. *J Epidemiol Community Health* 2007 Dec;61(12):1086-1090 [FREE Full text] [doi: [10.1136/jech.2006.053967](https://doi.org/10.1136/jech.2006.053967)] [Medline: [18000132](https://pubmed.ncbi.nlm.nih.gov/18000132/)]
92. McDowall PS, Galland BC, Campbell AJ, Elder DE. Parent knowledge of children's sleep: a systematic review. *Sleep Med Rev* 2017 Feb;31:39-47. [doi: [10.1016/j.smrv.2016.01.002](https://doi.org/10.1016/j.smrv.2016.01.002)] [Medline: [26899741](https://pubmed.ncbi.nlm.nih.gov/26899741/)]

93. Amin NAL, Tam WW, Shorey S. Enhancing first-time parents' self-efficacy: a systematic review and meta-analysis of universal parent education interventions' efficacy. *Int J Nurs Stud* 2018 Jun;82:149-162. [doi: [10.1016/j.ijnurstu.2018.03.021](https://doi.org/10.1016/j.ijnurstu.2018.03.021)] [Medline: [29656206](https://pubmed.ncbi.nlm.nih.gov/29656206/)]
94. Kreuter MW, McClure SM. The role of culture in health communication. *Annu Rev Public Health* 2004;25:439-455. [doi: [10.1146/annurev.publhealth.25.101802.123000](https://doi.org/10.1146/annurev.publhealth.25.101802.123000)] [Medline: [15015929](https://pubmed.ncbi.nlm.nih.gov/15015929/)]
95. Lindsay AC, Greaney ML, Wallington SF, Mesa T, Salas CF. A review of early influences on physical activity and sedentary behaviors of preschool-age children in high-income countries. *J Spec Pediatr Nurs* 2017 Jul;22(3):- . [doi: [10.1111/jspn.12182](https://doi.org/10.1111/jspn.12182)] [Medline: [28407367](https://pubmed.ncbi.nlm.nih.gov/28407367/)]
96. Asplund KM, Kair LR, Arain YH, Cervantes M, Oreskovic NM, Zuckerman KE. Early childhood screen time and parental attitudes toward child television viewing in a low-income latino population attending the special supplemental nutrition program for women, infants, and children. *Child Obes* 2015 Oct;11(5):590-599 [FREE Full text] [doi: [10.1089/chi.2015.0001](https://doi.org/10.1089/chi.2015.0001)] [Medline: [26390321](https://pubmed.ncbi.nlm.nih.gov/26390321/)]
97. Kostyrka-Allchorne K, Cooper NR, Simpson A. The relationship between television exposure and children's cognition and behaviour: a systematic review. *Develop Rev* 2017 Jun;44:19-58. [doi: [10.1016/j.dr.2016.12.002](https://doi.org/10.1016/j.dr.2016.12.002)]
98. Schuster RC, Szpak M, Klein E, Sklar K, Dickin KL. "I try, I do": child feeding practices of motivated, low-income parents reflect trade-offs between psychosocial- and nutrition-oriented goals. *Appetite* 2019 May 01;136:114-123. [doi: [10.1016/j.appet.2019.01.005](https://doi.org/10.1016/j.appet.2019.01.005)] [Medline: [30641158](https://pubmed.ncbi.nlm.nih.gov/30641158/)]
99. 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](https://doi.org/10.1037//0033-295x.84.2.191)] [Medline: [847061](https://pubmed.ncbi.nlm.nih.gov/847061/)]
100. Parekh N, Henriksson P, Delisle Nyström C, Silfvernagel K, Ruiz JR, Ortega FB, et al. Associations of parental self-efficacy with diet, physical activity, body composition, and cardiorespiratory fitness in Swedish preschoolers: results from the MINISTOP trial. *Health Educ Behav* 2018 Apr;45(2):238-246. [doi: [10.1177/1090198117714019](https://doi.org/10.1177/1090198117714019)] [Medline: [28629222](https://pubmed.ncbi.nlm.nih.gov/28629222/)]
101. Xu H, Wen LM, Rissel C. Associations of parental influences with physical activity and screen time among young children: a systematic review. *J Obes* 2015;2015:546925 [FREE Full text] [doi: [10.1155/2015/546925](https://doi.org/10.1155/2015/546925)] [Medline: [25874123](https://pubmed.ncbi.nlm.nih.gov/25874123/)]
102. Gardner B, de Bruijn G, Lally P. A systematic review and meta-analysis of applications of the Self-Report Habit Index to nutrition and physical activity behaviours. *Ann Behav Med* 2011 Oct;42(2):174-187. [doi: [10.1007/s12160-011-9282-0](https://doi.org/10.1007/s12160-011-9282-0)] [Medline: [21626256](https://pubmed.ncbi.nlm.nih.gov/21626256/)]
103. McGowan L, Cooke LJ, Gardner B, Beeken RJ, Croker H, Wardle J. Healthy feeding habits: efficacy results from a cluster-randomized, controlled exploratory trial of a novel, habit-based intervention with parents. *Am J Clin Nutr* 2013 Sep;98(3):769-777. [doi: [10.3945/ajcn.112.052159](https://doi.org/10.3945/ajcn.112.052159)] [Medline: [23864536](https://pubmed.ncbi.nlm.nih.gov/23864536/)]
104. Kok G, Gottlieb NH, Peters GY, Mullen PD, Parcel GS, Ruiter RA, et al. A taxonomy of behaviour change methods: an Intervention Mapping approach. *Health Psychol Rev* 2016 Sep;10(3):297-312 [FREE Full text] [doi: [10.1080/17437199.2015.1077155](https://doi.org/10.1080/17437199.2015.1077155)] [Medline: [26262912](https://pubmed.ncbi.nlm.nih.gov/26262912/)]
105. 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](https://doi.org/10.1007/s12160-013-9486-6)] [Medline: [23512568](https://pubmed.ncbi.nlm.nih.gov/23512568/)]
106. Hollands GJ, Marteau TM, Fletcher PC. Non-conscious processes in changing health-related behaviour: a conceptual analysis and framework. *Health Psychol Rev* 2016 Dec;10(4):381-394 [FREE Full text] [doi: [10.1080/17437199.2015.1138093](https://doi.org/10.1080/17437199.2015.1138093)] [Medline: [26745243](https://pubmed.ncbi.nlm.nih.gov/26745243/)]
107. Dombrowski SU, O'Carroll RE, Williams B. Form of delivery as a key 'active ingredient' in behaviour change interventions. *Br J Health Psychol* 2016 Nov;21(4):733-740. [doi: [10.1111/bjhp.12203](https://doi.org/10.1111/bjhp.12203)] [Medline: [27709824](https://pubmed.ncbi.nlm.nih.gov/27709824/)]
108. Bukman AJ, Teuscher D, Feskens EJ, van Baak MA, Meershoek A, Renes RJ. Perceptions on healthy eating, physical activity and lifestyle advice: opportunities for adapting lifestyle interventions to individuals with low socioeconomic status. *BMC Public Health* 2014 Oct 04;14:1036 [FREE Full text] [doi: [10.1186/1471-2458-14-1036](https://doi.org/10.1186/1471-2458-14-1036)] [Medline: [25280579](https://pubmed.ncbi.nlm.nih.gov/25280579/)]
109. Campbell FA, Goldman BD, Boccia ML, Skinner M. The effect of format modifications and reading comprehension on recall of informed consent information by low-income parents: a comparison of print, video, and computer-based presentations. *Patient Educ Couns* 2004 May;53(2):205-216. [doi: [10.1016/S0738-3991\(03\)00162-9](https://doi.org/10.1016/S0738-3991(03)00162-9)] [Medline: [15140461](https://pubmed.ncbi.nlm.nih.gov/15140461/)]
110. Choi J, Bakken S. Web-based education for low-literate parents in Neonatal Intensive Care Unit: development of a website and heuristic evaluation and usability testing. *Int J Med Inform* 2010 Aug;79(8):565-575 [FREE Full text] [doi: [10.1016/j.ijmedinf.2010.05.001](https://doi.org/10.1016/j.ijmedinf.2010.05.001)] [Medline: [20617546](https://pubmed.ncbi.nlm.nih.gov/20617546/)]
111. Council of Europe. Common European Framework of Reference for Languages: learning, teaching, assessment. Cambridge University Press. 2001. URL: <https://rm.coe.int/16802fc1bf> [accessed 2021-04-13]
112. Morrison LG, Yardley L, Powell J, Michie S. What design features are used in effective e-health interventions? A review using techniques from Critical Interpretive Synthesis. *Telemed J E Health* 2012 Mar;18(2):137-144. [doi: [10.1089/tmj.2011.0062](https://doi.org/10.1089/tmj.2011.0062)] [Medline: [22381060](https://pubmed.ncbi.nlm.nih.gov/22381060/)]
113. Short CE, Rebar AL, Plotnikoff RC, Vandelanotte C. Designing engaging online behaviour change interventions: a proposed model of user engagement. *Eur Health Psychol* 2015;17(1):32-38 [FREE Full text]

114. Denney-Wilson E, Laws R, Russell CG, Ong K, Taki S, Elliot R, et al. Preventing obesity in infants: the growing healthy feasibility trial protocol. *BMJ Open* 2015 Nov 30;5(11):e009258 [FREE Full text] [doi: [10.1136/bmjopen-2015-009258](https://doi.org/10.1136/bmjopen-2015-009258)] [Medline: [26621519](https://pubmed.ncbi.nlm.nih.gov/26621519/)]
115. Laws R, Campbell KJ, van der Pligt P, Ball K, Lynch J, Russell G, et al. Obesity prevention in early life: an opportunity to better support the role of Maternal and Child Health Nurses in Australia. *BMC Nurs* 2015;14:26 [FREE Full text] [doi: [10.1186/s12912-015-0077-7](https://doi.org/10.1186/s12912-015-0077-7)] [Medline: [25972765](https://pubmed.ncbi.nlm.nih.gov/25972765/)]
116. The National Prevention Agreement. Government of the Netherlands. 2019. URL: <https://www.government.nl/topics/overweight-and-obesity/documents/reports/2019/06/30/the-national-prevention-agreement> [accessed 2021-01-20]
117. Wen LM, Baur LA, Simpson JM, Rissel C, Wardle K, Flood VM. Effectiveness of home based early intervention on children's BMI at age 2: randomised controlled trial. *Br Med J* 2012 Jun 26;344:e3732 [FREE Full text] [doi: [10.1136/bmj.e3732](https://doi.org/10.1136/bmj.e3732)] [Medline: [22735103](https://pubmed.ncbi.nlm.nih.gov/22735103/)]
118. Suresh K. An overview of randomization techniques: an unbiased assessment of outcome in clinical research. *J Hum Reprod Sci* 2011 Jan;4(1):8-11 [FREE Full text] [doi: [10.4103/0974-1208.82352](https://doi.org/10.4103/0974-1208.82352)] [Medline: [21772732](https://pubmed.ncbi.nlm.nih.gov/21772732/)]
119. Warin M, Zivkovic T, Moore V, Ward PR, Jones M. Short horizons and obesity futures: disjunctures between public health interventions and everyday temporalities. *Soc Sci Med* 2015 Mar;128:309-315. [doi: [10.1016/j.socscimed.2015.01.026](https://doi.org/10.1016/j.socscimed.2015.01.026)] [Medline: [25645187](https://pubmed.ncbi.nlm.nih.gov/25645187/)]
120. Dennison L, Morrison L, Conway G, Yardley L. Opportunities and challenges for smartphone applications in supporting health behavior change: qualitative study. *J Med Internet Res* 2013 Apr 18;15(4):e86 [FREE Full text] [doi: [10.2196/jmir.2583](https://doi.org/10.2196/jmir.2583)] [Medline: [23598614](https://pubmed.ncbi.nlm.nih.gov/23598614/)]
121. *Samen Happie!* [Bites of Happiness]. URL: <https://www.samenhappie.nl/en/> [accessed 2021-04-27]
122. Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *J Am Med Assoc* 1995 Feb 01;273(5):408-412. [doi: [10.1001/jama.273.5.408](https://doi.org/10.1001/jama.273.5.408)] [Medline: [7823387](https://pubmed.ncbi.nlm.nih.gov/7823387/)]
123. Moher D, Pham B, Jones A, Cook DJ, Jadad AR, Moher M, et al. Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses? *Lancet* 1998 Aug 22;352(9128):609-613. [doi: [10.1016/S0140-6736\(98\)01085-X](https://doi.org/10.1016/S0140-6736(98)01085-X)] [Medline: [9746022](https://pubmed.ncbi.nlm.nih.gov/9746022/)]
124. Tugault-Lafleur C, González OD, O'Connor T, Masse L. Food parenting practices and children's eating behaviors: comparing mothers and fathers. *Curr Develop Nutri* 2020:1354 [FREE Full text] [doi: [10.1093/cdn/nzaa059_071](https://doi.org/10.1093/cdn/nzaa059_071)]
125. Dovey TM, Staples PA, Gibson EL, Halford JCG. Food neophobia and 'picky/fussy' eating in children: a review. *Appetite* 2008;50(2-3):181-193. [doi: [10.1016/j.appet.2007.09.009](https://doi.org/10.1016/j.appet.2007.09.009)] [Medline: [17997196](https://pubmed.ncbi.nlm.nih.gov/17997196/)]
126. Gallacher L. The terrible twos: gaining control in the nursery? *Children's Geographies* 2006 Aug 21;3(2):243-264. [doi: [10.1080/14733280500161677](https://doi.org/10.1080/14733280500161677)]
127. Coren E, Barlow J, Stewart-Brown S. The effectiveness of individual and group-based parenting programmes in improving outcomes for teenage mothers and their children: a systematic review. *J Adolesc* 2003 Feb;26(1):79-103. [doi: [10.1016/s0140-1971\(02\)00119-7](https://doi.org/10.1016/s0140-1971(02)00119-7)] [Medline: [12550823](https://pubmed.ncbi.nlm.nih.gov/12550823/)]
128. Nieuwboer CC, Fukkink RG, Hermanns JM. Online programs as tools to improve parenting: a meta-analytic review. *Child Youth Serv Rev* 2013 Nov;35(11):1823-1829. [doi: [10.1016/j.childyouth.2013.08.008](https://doi.org/10.1016/j.childyouth.2013.08.008)]

Abbreviations

EBRB: energy balance–related behavior

IMP: Intervention Mapping Protocol

RCT: randomized controlled trial

SEP: socioeconomic position

Edited by G Eysenbach; submitted 06.10.20; peer-reviewed by C Rissel, G ten Velde; comments to author 09.12.20; revised version received 02.02.21; accepted 15.03.21; published 14.05.21.

Please cite as:

Karszen LT, Vink JM, de Weerth C, Hermans RCJ, de Kort CPM, Kremers SPJ, Ruiters ELM, Larsen JK

An App-Based Parenting Program to Promote Healthy Energy Balance–Related Parenting Practices to Prevent Childhood Obesity: Protocol Using the Intervention Mapping Framework

JMIR Form Res 2021;5(5):e24802

URL: <https://formative.jmir.org/2021/5/e24802>

doi: [10.2196/24802](https://doi.org/10.2196/24802)

PMID: [33988510](https://pubmed.ncbi.nlm.nih.gov/33988510/)

©Levie T Karssen, Jacqueline M Vink, Carolina de Weerth, Roel C J Hermans, Carina P M de Kort, Stef PJ Kremers, Emilie L M Ruiters, Junilla K Larsen. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 14.05.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

The Adaptive GameSquad Xbox-Based Physical Activity and Health Coaching Intervention for Youth With Neurodevelopmental and Psychiatric Diagnoses: Pilot Feasibility Study

April B Bowling¹, MA, SCD; James Slavet², PhD; Chelsea Hendrick³, BA; Robbie Beyl³, PhD; Phillip Nauta³, MA; Marilyn Augustyn⁴, MD, PhD; Mediatrix Mbamalu⁴, MD; Carol Curtin^{5,6}, MSW, PhD; Linda Bandini^{5,6,7}, PhD; Aviva Must^{6,8}, MD, PhD; Amanda E Staiano³, PhD

¹Department of Public Health and Nutrition, School of Health Sciences, Merrimack College, North Andover, MA, United States

²Marblehead Public Schools, Marblehead, MA, United States

³Pennington Biomedical Research Center, Baton Rouge, LA, United States

⁴Developmental and Behavioral Pediatrics Clinic, Boston Medical Center, Boston, MA, United States

⁵Department of Family Medicine and Community Health, Eunice Kennedy Shriver Center, University of Massachusetts Medical School, Worcester, MA, United States

⁶Healthy Weight Research Network, Worcester, MA, United States

⁷Department of Health Sciences, Sargent College of Health and Rehabilitation Sciences, Boston University, Boston, MA, United States

⁸Department of Public Health and Community Medicine, Tufts University, Boston, MA, United States

Corresponding Author:

April B Bowling, MA, SCD

Department of Public Health and Nutrition

School of Health Sciences

Merrimack College

315 Turnpike Street

North Andover, MA, 01845

United States

Phone: 1 978 837 5187

Email: bowlinga@merrimack.edu

Abstract

Background: The prevalence of neurodevelopmental and psychiatric diagnoses (NPDs) in youth is increasing, and unhealthy physical activity (PA), diet, screen time, and sleep habits contribute to the chronic disease disparities and behavioral challenges this population experiences.

Objective: This pilot study aims to adapt a proven exergaming and telehealth PA coaching intervention for typically developing youth with overweight or obesity; expand it to address diet, screen, and sleep behaviors; and then test its feasibility and acceptability, including PA engagement, among youth with NPDs.

Methods: Participants (N=23; mean age 15.1 years, SD 1.5; 17 males, 9 people of color) recruited in person from clinic and special education settings were randomized to the Adaptive GameSquad (AGS) intervention or wait-list control. The 10-week adapted intervention included 3 exergaming sessions per week and 6 real-time telehealth coaching sessions. The primary outcomes included feasibility (adherence to planned sessions), engagement (uptake and acceptability as reported on process questionnaires), and PA level (combined light, moderate, and vigorous as measured by accelerometer). Descriptive statistics summarized feasibility and engagement data, whereas paired, two-tailed t tests assessed group differences in pre-post PA.

Results: Of the 6 coaching sessions, AGS participants (n=11; mean age 15.3 years, SD 1.2; 7 males, 4 people of color) completed an average of 5 (83%), averaging 81.2 minutes per week of exergaming. Of 9 participants who completed the exit questionnaire, 6 (67%) reported intention to continue, and 8 (89%) reported feeling that the coaching sessions were helpful. PA and sleep appeared to increase during the course of the intervention over baseline, video game use appeared to decrease, and pre-post intervention PA per day significantly decreased for the control (-58.8 min; $P=.04$) but not for the intervention group (-5.3 min; $P=.77$), despite potential seasonality effects. However, beta testers and some intervention participants indicated a need for reduced complexity of technology and more choice in exergames.

Conclusions: AGS shows promise in delivering a health behavior intervention remotely to youth with NPDs, but a full-scale efficacy trial with a larger sample size is needed to confirm this finding. On the basis of feedback from beta testers and intervention participants, the next steps should include reduced technology burden and increased exergame choice before efficacy testing.

Trial Registration: ClinicalTrials.gov NCT03665415; <https://clinicaltrials.gov/ct2/show/NCT03665415>.

(*JMIR Form Res* 2021;5(5):e24566) doi:[10.2196/24566](https://doi.org/10.2196/24566)

KEYWORDS

exercise; diet; sleep; mental health; children; adolescent; health promotion; telehealth; exergaming

Introduction

Background

Mental health and the resultant adverse chronic disease consequences among youth are growing concerns in the United States. Recent estimates suggest that pediatric psychiatric disorders occur in more than one-fourth of people aged <18 years [1], including diagnoses such as anxiety, mood, and psychotic disorders. Common comorbid diagnoses include neurodevelopmental disorders, such as autism spectrum disorder (ASD) and attention-deficit/hyperactivity disorder (ADHD), which are estimated to affect approximately 15% of US youth [2], leading to significant heterogeneity of symptom presentations.

The health disparities faced by children and youth with heterogeneous neurodevelopmental and psychiatric diagnoses (NPDs) are considerable and include a higher risk of obesity and type 2 diabetes [3], compared with typically developing youth. These disparities are at least partially attributable to unhealthy behavioral patterns established in childhood, patterns that often persist throughout life [4]. Several studies have found that children with ASD; ADHD; and bipolar, depressive, and anxiety disorders are at high risk of low physical activity (PA) levels [5,6], poor diet [7], disrupted sleep [8,9], and elevated screen time [10].

Unhealthy lifestyles among youth with NPDs are doubly unfortunate; they not only confer increased health risks across the life course but also exacerbate the challenges to cognitive and behavioral functioning experienced by this population. More than 25 published studies have documented associations between light-, moderate-, and vigorous-intensity PA and improvements to mood and executive functioning, such as the ability to focus and self-regulate, and meta-cognitive processes among children with NPDs [11]. For example, a recent study found that among children with moderate to severe NPDs who were exposed to cybercycling (ie, stationary bicycles that use immersive gaming features) during the school day, the odds of behavioral dysregulation declined by more than 70% relative to children who did not participate in exercise [12]. These results are consistent with many previous studies showing positive effects of exercise on mood and impulsivity; there is also evidence that these positive benefits may be most pronounced in children with NPDs [13]. Unfortunately, few interventions tested to date have shown effectiveness in increasing long-term engagement in PA among youth with NPDs, nor are many scalable, given their resource intensity.

Exergames have shown promise in promoting cost-effective engagement in light- to moderate-intensity PA in youth [14]. When integrated into theory-based interventions, they may also help improve the personal mediators of exercise engagement in children. For example, youth with ASD and major depressive disorder show high behavioral inhibition bias (BIS) [15], which indicates avoidance of novel or uncomfortable situations and which may act as a barrier to engagement in higher intensities of exercise or new exercise programming [16]. However, previous studies in typically developing youth have shown that acting on constructs of self-determination can help individuals with high BIS adopt long-term intrinsic motivation toward exercise [16]. Tailoring PA interventions to include lower-intensity exercise, agency in selection of exercise intensity, use of fun and noncompetitive exercise technology such as exergames, and monitoring of mood improvements after exercise may be particularly effective.

GameSquad is an intervention originally designed to improve PA among a socioeconomically and racially diverse population of typically developing children who meet the criteria for overweight or obesity [16]. GameSquad is an intervention delivered entirely remotely using exergaming and virtual health coaching (telehealth coaching delivered via video conference) as components within a behavior change intervention grounded in social cognitive theory, which frames behavioral change as the result of links among behaviors, environment, and psychosocial factors [16]. Exergames can be played with family members and friends, and social interaction during group-based exergame play has been identified as a key predictor of weight loss [16]. Exergames also use programmed features that can encourage exercise, such as motivational messaging during game play, to boost players' self-efficacy, which may translate to self-determination and predict exercise adherence and eventual engagement in nonscreen-based exercise modalities [17]. The GameSquad intervention directly emphasizes the element of social support by encouraging children to play with a family member or friend and by requiring children and parents to attend telehealth counseling sessions together from their home. These coaching sessions are delivered through the gaming console and are designed to promote self-efficacy and teach approaches and strategies to reduce perceived barriers to behavior change. A 6-month randomized controlled trial of the GameSquad intervention found a significant increase in moderate- to vigorous-intensity physical activity (MVPA) and a decrease in BMI, blood lipids, and blood pressure (BP) among participants [16].

Objective

Inclusion team science is an intervention development and testing framework that brings together disability researchers and intervention scientists to fast-track intervention adaptation for this underserved population [18]. Given GameSquad's initial efficacy in increasing MVPA in typically developing but difficult to engage youth and the potential of exergaming and virtual health coaching to engage youth with NPDs in positive health behaviors and as a tool to transition them to long-term intrinsic motivation to exercise, the aims of this study are as follows: (1) to adapt this intervention for use among youth with NPDs; (2) to expand the virtual coaching sessions to address dietary habits, screen time, and healthy sleep habits in addition to PA promotion; and (3) to pilot test the initial feasibility, engagement, and short-term changes in health behaviors. The primary outcomes examined during the pilot test included feasibility (adherence), engagement (uptake of PA during the intervention and acceptability), and change in objectively measured PA pre- and postintervention. The secondary outcomes included pre-post changes in self-reported exercise stage of change, sleep duration, hours per week of video game use, problematic mealtime behaviors, BMI, and BP.

Methods

Overview

All study procedures were approved by the Merrimack College, Pennington Biomedical Research Center, and Boston Medical Center Institutional Review Boards, and the study was registered as a clinical trial and is available on ClinicalTrials.gov (NCT03665415). As we adapted GameSquad to a population of youth with NPDs to create Adaptive GameSquad (AGS), we expanded the theoretical framework to include the Reserve Capacity Model [19] and the Family Ecological Model [20]. The purpose of this expansion was to ensure that intervention components addressed barriers to and used facilitators of health behaviors specific to the unique challenges faced by this population (eg, depleted caregiver reserve capacity) [21]. Given that remotely delivered home-based interventions are critical for upscaling but can also place a high burden on families, we used these models to act on specific constructs to improve participant and caregiver reserve capacity and downstream health behaviors and outcomes.

Adaptation was undertaken by an advisory team including a developmental psychologist, school psychologist, registered dietitian, clinical social worker, fitness coaches, and parents of

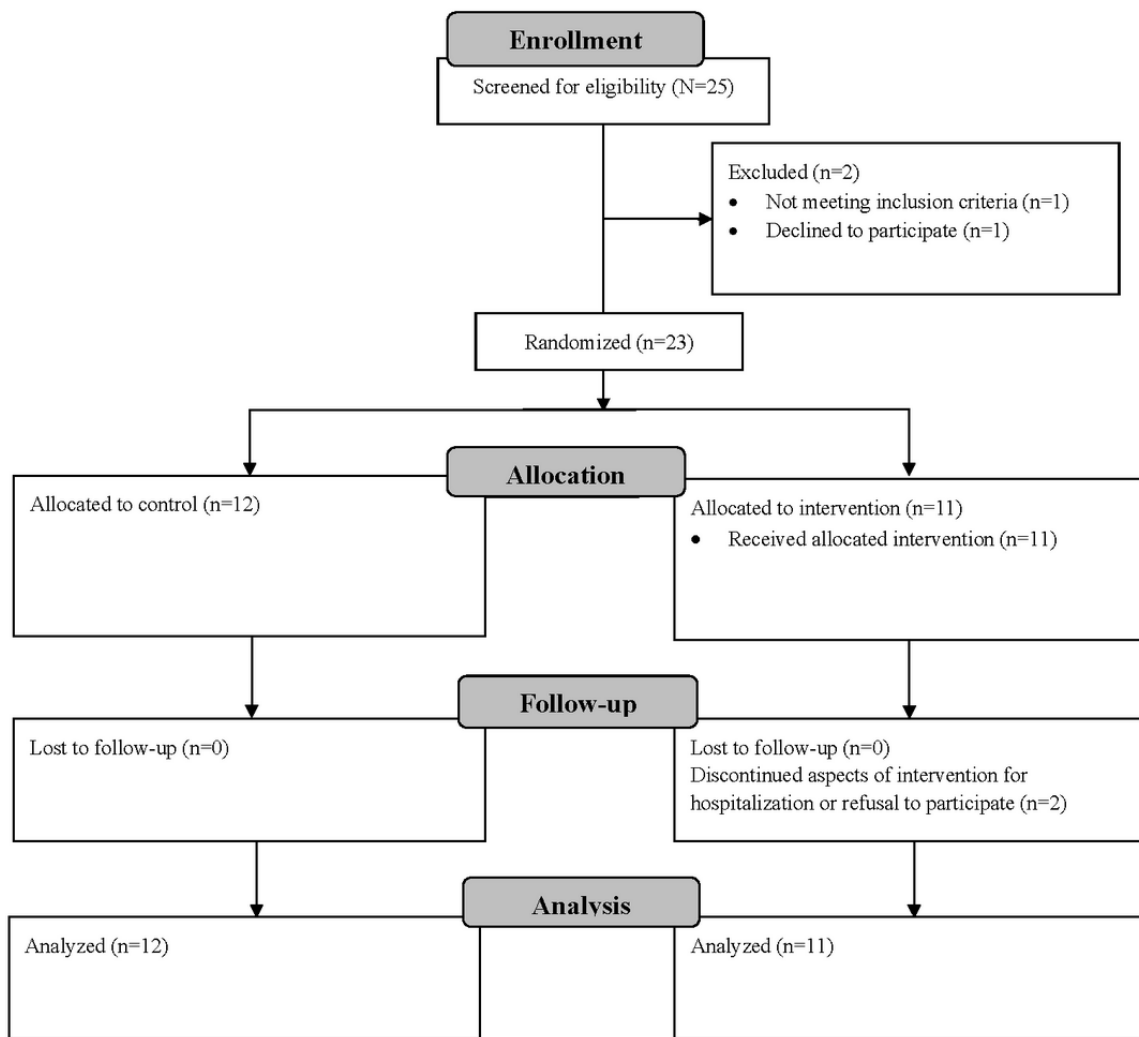
a child with NPD. As a result, we expanded caregiver support components, such as scheduling reminders; modified the original challenge booklet to increase MVPA goals more slowly; trained telehealth coaches in positive behavioral reinforcement techniques; and targeted exergames that were particularly engaging for the demographic profile of our target population. We also expanded coaching scripts to not only address PA but also to include health education and goal setting for dietary intake, sleep, and screen time habits. Coaches were also trained to work with youth with NPDs, particularly in positive reinforcement and de-escalation methods [22].

Finally, 3 beta testers and their caregivers were purposively recruited from the school study site for diversity of diagnosis, gender, and grade level (demographic information is not included here to protect them from inferred identification, given the small sample size). After written parental consent and child assent, the youth participated in a 4-week trial of the initial AGS intervention. We then used their feedback to make additional modifications to the telehealth coaching script and gaming menu, including increased agency in game selection and allowing for solitary game play. These modifications resulted in the final AGS intervention deployed for the pilot feasibility and engagement study.

Recruitment

Participants for the 10-week pilot study were recruited in person from October 15, 2018, to February 15, 2019, from either the Boston Medical Center Developmental or Behavioral Pediatrics Clinic from therapeutic programs at a large public middle and high school (Figure 1). This approach helped to ensure a diverse diagnostic sample.

The study was powered as a pilot feasibility and acceptability trial [23,24]. Of the 25 children whose parents completed a web-based screen survey within the recruiting time frame, 23 met the eligibility criteria and were randomly assigned in a 1-to-1 allocation ratio using the REDCap (Research Electronic Data Capture) software. The inclusion criteria included being in middle or high school, having a neurodevelopmental and/or psychiatric diagnosis, willingness to participate in exergaming 3 times weekly and meet with a telehealth coach every other week, and possessing the cognitive ability to understand gaming directions. Exclusion criteria included having impairments that prevented engagement in exergames, pregnancy, inability to speak or understand English, or having a caregiver unable or unwilling to attend the telehealth coaching sessions on a regular basis.

Figure 1. Study recruitment and enrollment flow diagram.

Procedures

Interested parents who met the criteria during the initial web-based screening then participated in an additional phone screening with research assistants and provided web-based written consent for their child to participate. Eligible children whose parents had consented were then assented in person during the initial data collection visit, during which height, weight, and BP, and psychosocial questionnaire data (see Measures section) were collected. Participants were given and instructed on the use of a hip-worn accelerometer (wGT3X-BT, ActiGraph). These accelerometers do not have digital readouts and thus minimize the effects of measurement on participants' normal PA patterns. Participants were provided with the accelerometer to wear for 7 days before randomization and again the week after the 10-week intervention, with a follow-up visit scheduled to collect the accelerometer. After the initial wear, the participants were randomized to the condition using the randomization algorithm in REDCap. Participants assigned to the wait-list control condition were asked to maintain their normal level of PA for 10 weeks but did not receive additional information or a PA tracker during the 10 weeks. They then received the same gaming equipment as the intervention

participants and 4-week of telehealth coaching after the final data collection. All data assessors and primary investigators were blinded to participant condition. Telehealth coaches and research assistants who installed equipment and training participants were not blinded.

All participants randomized to the intervention group received an Xbox One gaming console with a Kinect motion sensor (Microsoft), a 12-week Xbox Live subscription, and 3 exergames (Just Dance 3, Shape Up 3, and Kinect Sports Season 2). Two research assistants visited intervention participants in their homes to install the equipment and to train participants and caregivers on the use of the games, the Skype portal for coaching sessions, and the Fitbit for tracking of steps during the intervention. The AGS intervention asked participants to play exergames 3 days per week with a family member or friend, if possible (not required). On nonexergaming days, participants were asked to meet tailored and incrementally increasing minutes of PA (starting with 10 min in week 1 and increasing to as much as 40 min per day by week 4). The activities prescribed were laid out week by week in the intervention *challenge booklet* participants were provided. The targeted durations were below PA guidelines [25] but were selected to minimize frustration and improve self-efficacy; participants

were also encouraged and worked with coaches to brainstorm nonscreen-based physical activities that they could perform during each week. The challenge booklet provided an adaptable gameplay curriculum to play 3 challenges each week with increasing intensity, difficulty, and duration.

Participants and caregivers were asked to meet with a telehealth coach over a video chat using the Xbox console. Meetings were held in weeks 1, 2, 4, 6, 8, and 10 and were rescheduled as necessary. Participants were asked to wear a Fitbit Charge throughout the 10-week intervention period. They and their caregivers received charging and sync reminders 3 times per week. Steps per day could then be wirelessly and automatically uploaded and reviewed by the telehealth coach. Coaches followed a script for meetings that focused on reviewing the week's PA, praising progress, troubleshooting barriers, and then discussing a new healthy habit to try each week.

Measures

Feasibility and engagement measures were tracked by telehealth coaches in REDCap, including telehealth coaching session attendance (coach report), number of exergame sessions completed per week (participant log), minutes of exergaming per week (participant log), and steps per week (Fitbit Charge, Fitabase). We also administered exit surveys with participants and conducted semistructured interviews with beta testers and their caregivers to assess acceptability and elicit suggestions for improvement. Exit surveys were administered after all other follow-up data were collected. Two participants were unavailable for exit surveys; thus, there was a reduced sample size for acceptability measures only.

Participant weight was measured by trained research assistants in light clothing without shoes on a digital scale (Model 813, Seca), and height was measured using a clinical stadiometer (Model 217, Seca). Two measurements were taken for both weight and height, with the average of each used to calculate the BMI (kg/m²). After height and weight were recorded, appropriate cuff size was selected, and diastolic and systolic BP readings were taken, following the 2017 TRUE (International Consortium for Quality Research) Consortium guidelines [26]. To assess intrinsic changes to PA habits outside of active intervention support, time spent in MVPA before the intervention started and 1 week after it concluded was objectively measured with the ActiGraph wgt3x-bt accelerometer. Accelerometers were initialized to record data in 15-second epochs, and established pediatric cutoff points were used to estimate PA of light-, moderate-, and vigorous-intensity levels [27]. We used a 7-day weighted average of weekday and weekend activity counts to determine each participant's PA level, with a minimum requirement of 600 minutes of wear time that included at least one weekend day [28]. Participants were also asked to keep a log of the number and duration of exergaming sessions they completed per week as part of the challenge booklet they were provided.

Exercise stage of change (youth report) was measured using a pen and paper version of the Change of Stages of Exercise—University of Rhode Island Change Assessment, the third generation [29], a valid and reliable measure that captures both precontemplation nonbelief and belief conditions.

Participants also completed a video game use survey, which asked questions regarding their current video game use habits, including hours of nonexergame gaming per week. In addition to a demographic questionnaire, parents were asked to complete the Children's Sleep Habits Questionnaire [30], which includes duration of sleep, and the Meals In Our Household Questionnaire (MIOH) [31], which measures 6 domains, including the structure of family meals and problematic child mealtime behaviors. We extracted and used the Problem Mealtime Behaviors subscale of the MIOH to evaluate changes related to behaviors targeted by the intervention coaching sessions.

Statistical Analysis

The sample size was determined on the basis of the aims and design (feasibility and engagement pilot study), with a target of 50 participants but a minimum of 20, based on the current guidelines [23,24]. Descriptive statistics were used to summarize feasibility and engagement data. Two-sample, two-tailed *t* tests were used to assess potential baseline differences in average age, baseline minutes of sleep, minutes of PA and MVPA, hours of video game use, and problematic mealtime behavior scores between the intervention and wait-list control groups. Proportion tests were used to assess potential baseline differences by group in percent male, percent White, percent qualifying for free or reduced-price lunch, and percent taking psychiatric medication associated with weight gain. All analyses were performed in an intention-to-treat manner. Changes in the exercise stage of change by group were assessed using the nonparametric Wilcoxon signed-rank test. Paired *t* tests were used to assess significant differences in pre-post BMI, BP, total PA (light PA+MVPA), MVPA, sleep duration, problematic mealtime behavior score, and hours of nonexergaming video game use by group. STATA 13 was used for all analyses.

Results

Overview

Participant demographic information is presented in [Table 1](#). The average participant age was 15.1 years (range 12-17); 74% (17/23) of the participants were male, 40% (9/23) were people of color, and 35% (8/23) qualified for free or reduced-price lunch. ADHD was the most common diagnosis (13/23, 57%), followed by ASD (12/23, 52%), anxiety disorders (6/23, 26%), and depression (6/23, 26%). Nearly one-third (7/23, 30%) of the total sample reported taking psychiatric medication that was associated with weight gain. There were no significant differences in demographic characteristics between the control and intervention groups; however, given the small sample size, it is important to note that 5 out of 12 control participants were reported to be taking medication, whereas only approximately 2 out of 11 intervention participants were reported doing so. In addition, the type of medication is unknown, so the potential directionality of effects on PA, diet, and sleep is unclear ([Table 1](#)). However, the average daily MVPA, total PA, exercise stage of change, sleep duration, problematic mealtime behaviors score, and hours of video game use at baseline did not differ significantly between the control and intervention groups.

Table 1. Participant characteristics at baseline.

Characteristics	All participants (n=23)	Control group (n=12)	Intervention group (n=11)	Test statistic			P value
				t test (df) ^a	z test	Chi-square (df)	
Age (years), mean (SD)	15.1	14.9	15.3	-0.57 (20)	N/A ^b	N/A	.57
Sex (male), n (%)	17 (74)	10 (83)	7 (64)	N/A	-1.58	N/A	.12
Race or ethnicity, n (%)				N/A	N/A	1.2 (5)	.88
Asian	2 (9)	1 (8)	1 (9)				
Black	2 (9)	1 (8)	1 (9)				
Hispanic	3 (13)	2 (17)	1 (9)				
White	14 (60)	7 (58)	7 (64)				
Declined	2 (9)	1 (8)	1 (9)				
Qualifies for free or reduced-price lunch, n	8 (35)	3 (25)	5 (45)	N/A	N/A	1.5 (1)	.23
Diagnoses, n (%)^c				N/A	N/A	N/A	N/A
Autism spectrum disorder	12 (52)	7 (58)	5 (45)				
Attention-deficit/hyperactivity disorder	13 (56)	8 (67)	5 (45)				
Anxiety	6 (26)	5 (42)	1 (9)				
Depression	6 (26)	3 (25)	3 (27)				
Other	4 (17)	4 (33)	0 (0)				
Taking medication that causes weight gain, n (%)	7 (30)	5 (42)	2 (18)	N/A	N/A	5.5 (2)	.06

^aTwo-tailed *t* test (age), test of proportion (sex, race or ethnicity, free or reduced lunch, and medication use).

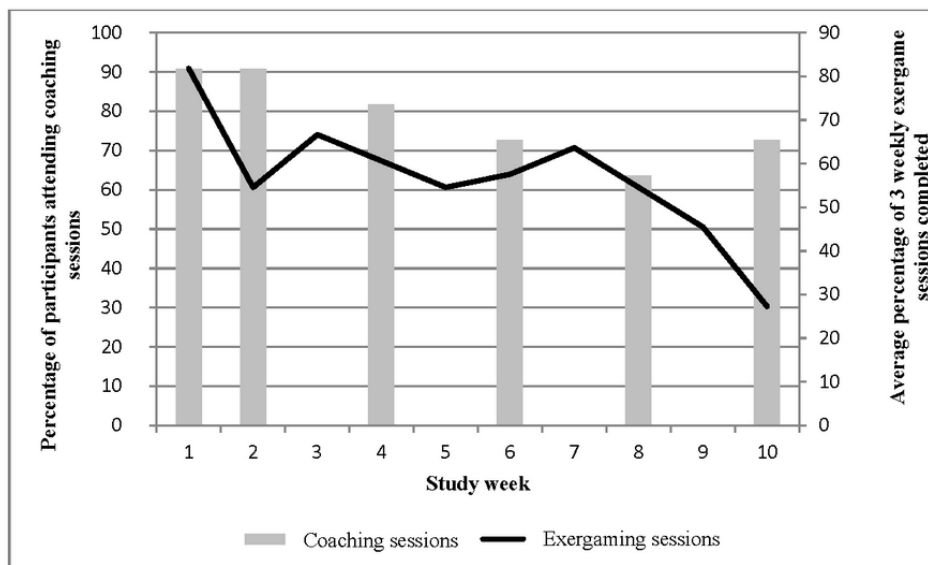
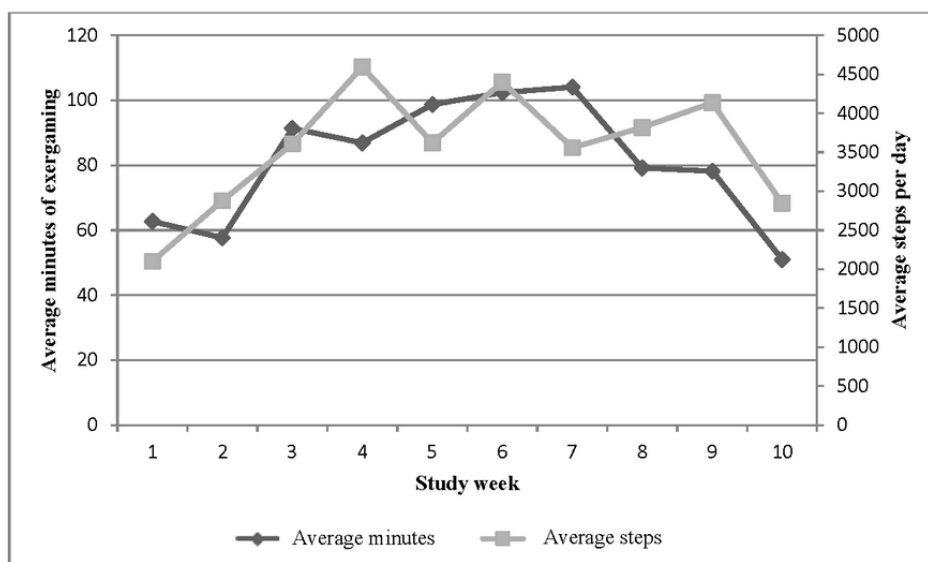
^bN/A: not applicable.

^cMore than 50% of participants had multiple diagnoses.

Primary Outcomes

Equipment was successfully installed and used by all intervention participants (n=11), who completed an average of 5 out of 6 possible coaching sessions (range 0-6). None of the participants were familiar with the 3 exergames used during the study before the start of the intervention. Including 1 participant who never exergamed or attended coaching sessions after the

initial home visit and training session, the intervention group executed an average of 1.7 out of 3 planned exergame sessions per week, averaging 81.2 minutes per week (SD 18.9) of self-reported exergaming and averaging 3559 steps per day over the course of the intervention. [Figure 2](#) shows trends in coaching session and exergame session adherence over time, whereas [Figure 3](#) shows the changes in exercise engagement over the course of the intervention.

Figure 2. Planned session adherence trends.**Figure 3.** Exercise engagement trends.

Considering the 11 participants, adherence to the coaching sessions exceeded 60% each week—from a high of 10 (91%) attending in weeks 1 and 2 to a low of 7 (64%) attending in week 8 (Figure 2). Exergaming adherence to the prescribed 3 times per week was highest in week 1, with 82% (27/33) of sessions completed; fluctuated between 55% (18/33) and 67% (22/33) in weeks 2 to 8; and then declined to only 27% (9/33) in week 10 (Figure 2). However, the total weekly duration of exergaming increased from about 60 minutes in week 1 to a peak at 104 minutes in week 7, before slowly declining to 51 minutes in week 10 (Figure 3). Finally, PA engagement measured by Fitbit increased from approximately 2100 steps per day in week 1 to 3500–4500 steps per day in weeks 3–9, before dropping to 2850 in week 10 (Figure 3). Of the 9 intervention participants who completed exit surveys, 6 reported intention to continue exergaming, 7 reported enjoying coaching sessions, and 8 reported feeling that the coaching sessions were

helpful and that the coach gave them tips they could use. Several caregivers noted during live telehealth sessions that the different technologies involved in the intervention (Wi-Fi, Fitbit, Xbox games, Kinect sensor, and Xbox Skype) presented a significant burden and expressed a desire for a more integrated technology delivery system such as a smartphone app.

The average daily MVPA measured by accelerometry declined between baseline and 1 week postintervention for both the control group (−9.3 min, SD 21.5; $P=.18$) and the intervention group (−1.4 min, SD 8.3; $P=.62$); neither change was statistically significant. However, although average daily total PA (MVPA+light PA) significantly decreased for the control group in postintervention period follow-up (−58.8 min, SD 85.7; $P=.04$), it did not significantly change for the intervention group (−5.3 min, SD 55.1; $P=.77$).

Secondary Outcomes

Secondary outcomes included pre-post changes in self-reported exercise stage of change, sleep duration, video game use, mealtime behaviors, BMI, and BP. The exercise stage of change improved significantly for the intervention group ($P=.02$) but not for the control group ($P=.99$). None of the treatment participants were in the preparation or action stages before intervention; postintervention, there were 1 participant and 5 participants in the preparation and action stages, respectively. In contrast, although no control group participants were in the preparation stage, 3 were in the action stage before the intervention; postintervention, there were still none in the preparation stage and only 2 in the action stage.

The average daily minutes of sleep decreased by 12 minutes for the control group ($P=.91$) but increased by 12.9 minutes for the intervention group ($P=.55$). The intervention group reported a 1.7-hour decrease in weekly nonexergaming video game use ($P=.28$); the control group reported no change. There were no significant changes in BMI, BP, or problematic meal behaviors in either group.

Discussion

Principal Findings

Youth with NPDs are at high risk of unhealthy lifestyle behaviors, including low PA levels [11], poor diet [32], high screen time [10], and poor sleep hygiene [9]. The consequences of these behaviors include elevated chronic disease risks, including obesity; cognitive impairment; and exacerbation of psychiatric symptoms. Barriers to engagement in lifestyle interventions are high among this population [21], and interventions demonstrating long-term engagement in improved health behaviors are scarce. Furthermore, as the COVID-19 pandemic has shown, a remotely delivered intervention using telehealth components can not only reduce barriers to initial participation but also allow flexibility for continued engagement during changing conditions [33], which are particularly important for this population.

The aim of this pilot study is to adapt and expand an existing, evidence-based exergaming and telehealth coaching intervention [16] to improve PA, diet, video game play time, and sleep habits among youth with a variety of mental health and neurodevelopmental disorders and to assess the program's initial feasibility and acceptability, including participants' engagement in PA. An expert working group made preliminary adaptations, which were then beta tested with the target population; feedback was used to finalize the intervention, and telehealth coaches were trained to use positive reinforcement and behavioral redirection techniques. Youth with relevant diagnoses ($n=23$) were recruited from both a therapeutic school and clinic setting and randomized to either the 10-week intervention or wait-list control.

Although not as high as the original GameSquad intervention that targeted a younger population (10-12 years vs 12-17 years), compliance with planned exergaming was good (participants completed an average of 57% of planned exergame sessions), and attendance at coaching sessions was excellent (participants

attended an average of 5 out of 6 coaching sessions), particularly given the unique barriers faced by participants with NPDs and the technological challenges noted by some caregivers.

Despite the short duration and small sample size that limited our power to detect intervention effects, the results indicated potential improvements to PA during the intervention over baseline and smaller declines in MVPA relative to control participants after the intervention had ended. Engagement in both exergame-based and nonscreen-based PA was good; however, it declined in weeks 9 and 10 of the pilot. When combined with a decline in MVPA in both the control and intervention groups after the intervention was completed, this may reflect a seasonality effect. Such an effect could have been because of the school year ending, discontinuation of physical education classes and school sports, or other external factors. However, the decline in duration of exergame sessions after week 7, combined with qualitative feedback from beta-testing participants, may suggest that the adolescent population in this study became bored with the limited menu of exergames available more quickly than the preadolescent population in the original GameSquad study. Alternatively, it may reflect frustration with or dislike of the increasing intensity demands of the exergaming menu as the intervention progressed. Further exploration of these factors is required to optimize long-term engagement.

Long-term implementation of AGS might reorient coaching to leverage the early weeks of exergame engagement into a greater emphasis on nonscreen-based PA modalities later in the intervention. Additional game choices and technologies should also be evaluated and included to maintain participants' interest and better meet the needs of this diverse population. The coaching sessions appeared to fill a psychosocial need for participants, independent of exergaming. Compliance with and acceptability of coaching sessions were higher than those of exergaming, with several participants repeatedly rescheduling sessions to enable them to meet with their coaches around significant clinical events such as inpatient hospital stays.

Secondary Outcome Results

Although the study was underpowered to evaluate secondary outcomes, participants reported increased sleep duration and decreased video game use time, despite the introduction of exergaming sessions. There were no improvements in problematic meal behaviors as a result of participation. Intervention participation also appeared to positively affect the exercise stage of change; although no treatment participants were in the active stage of exercise before the intervention, nearly 50% were in the active stage after the intervention conclusion. It is important to note, however, that although the exercise stage of change improved, the maintenance stage was not evaluated through long-term follow-up.

Limitations and Additional Research Needs

This pilot study has several limitations. The small sample size limits generalizability and decreased power to detect intervention effects; it also precludes stratified examination of differences in outcomes by subgroup. Although participants were randomized to condition, the sample size was small; thus, as

models were not adjusted for potential covariates and confounders, readers need to take caution in interpreting the results. In addition, although the strength of this study was that pre- and post-PA were objectively measured using accelerometers, several other measures were self-reported, including duration of exergaming sessions, sleep duration, and video game use. The 10-week intervention design and lack of long-term follow-up prevented the evaluation of sustained engagement and effects of the intervention; this must be the primary aim of any full-scale efficacy study.

This study has several notable strengths. We were able to include youth with a variety of mental health and neurodevelopmental disorders, recruiting participants from both clinical and therapeutic school settings. We believe this heterogeneity improves external generalizability and eventual translation to a variety of clinical and community venues. Next steps should include the development of a mobile health (mHealth) app to seamlessly deliver a wider variety of exergames, telehealth coaching sessions, and parental and participant reminders and integrate mood, health habit, and PA tracking, while eliminating the technical barriers associated with the Xbox-based approach. This will also help minimize disruptions to upscaling caused by changes in commercially available gaming technologies. Additional research should also be conducted with youth with NPDs to evaluate personal mediators of health behaviors, exercise preferences, and barriers to engagement that may inform mHealth app design and intervention optimization.

Conclusions

This pilot study is an innovative example of inclusion team science, a term coined by Rimmer and Vanderbom [18] in their 2016 call to action for health promotion research for children with disabilities. Their commentary urged greater collaboration between disability researchers and intervention scientists working with typically developing populations to more rapidly adapt existing interventions to meet the underserved needs of youth with disabilities [18]. In this study, we began with a proven intervention developed for typically developing youth with overweight and obesity and adapted it for youth with a variety of mental health and neurodevelopmental disorders using a process that included beta testing by the target population and adaptation of intervention components and methods based on their feedback. Although our sample size was too small to evaluate efficacy, the initial feasibility and acceptability results indicate that AGS may be a promising avenue for delivering health behavior interventions remotely to youth with NPDs, an increasingly critical need in light of the significant disruptions to in-person learning and clinical care caused by the COVID-19 pandemic. However, it is important to reduce the technological demands of the intervention on caregivers and increase the number and diversity of exergames used in the intervention to sustain engagement by participants. After additional intervention optimization, future efficacy testing must take place in a large, socioeconomically, racially or ethnically, and diagnostically diverse sample of youth.

Acknowledgments

Primary support for this study was provided by the Healthy Weight Research Network of the Maternal and Child Health Bureau, Health Resources and Services Administration (UA3MC25735). This study was partially supported by a Nutrition Obesity Research Center grant P30DK072476 entitled *Nutritional Programming: Environmental and Molecular Interactions* and grant 1U54GM104940 from the National Institute of General Medical Sciences of the National Institutes of Health, which funds the Louisiana Clinical and Translational Science Center. The authors would like to acknowledge the participants and their families, without whom this study and the knowledge it generated would not have been possible. The authors would also like to thank the research assistants, school staff, and clinicians who contributed to the execution of the study.

Conflicts of Interest

None declared.

References

1. Polanczyk GV, Salum GA, Sugaya LS, Caye A, Rohde LA. Annual research review: a meta-analysis of the worldwide prevalence of mental disorders in children and adolescents. *J Child Psychol Psychiatry* 2015 Mar 03;56(3):345-365. [doi: [10.1111/jcpp.12381](https://doi.org/10.1111/jcpp.12381)] [Medline: [25649325](https://pubmed.ncbi.nlm.nih.gov/25649325/)]
2. Boyle CA, Boulet S, Schieve L, Cohen RA, Blumberg SJ, Yeargin-Allsopp M, et al. Trends in the prevalence of developmental disabilities in US children, 1997-2008. *Pediatrics* 2011 Jun;127(6):1034-1042. [doi: [10.1542/peds.2010-2989](https://doi.org/10.1542/peds.2010-2989)] [Medline: [21606152](https://pubmed.ncbi.nlm.nih.gov/21606152/)]
3. Correll CU, Solmi M, Veronese N, Bortolato B, Rosson S, Santonastaso P, et al. Prevalence, incidence and mortality from cardiovascular disease in patients with pooled and specific severe mental illness: a large-scale meta-analysis of 3,211,768 patients and 113,383,368 controls. *World Psychiatry* 2017 Jun 12;16(2):163-180 [FREE Full text] [doi: [10.1002/wps.20420](https://doi.org/10.1002/wps.20420)] [Medline: [28498599](https://pubmed.ncbi.nlm.nih.gov/28498599/)]
4. Rimmer JH, Rowland JL, Yamaki K. Obesity and secondary conditions in adolescents with disabilities: addressing the needs of an underserved population. *J Adolesc Health* 2007 Sep;41(3):224-229. [doi: [10.1016/j.jadohealth.2007.05.005](https://doi.org/10.1016/j.jadohealth.2007.05.005)] [Medline: [17707291](https://pubmed.ncbi.nlm.nih.gov/17707291/)]

5. McCoy SM, Morgan K. Obesity, physical activity, and sedentary behaviors in adolescents with autism spectrum disorder compared with typically developing peers. *Autism* 2020 Feb 31;24(2):387-399. [doi: [10.1177/1362361319861579](https://doi.org/10.1177/1362361319861579)] [Medline: [31364386](https://pubmed.ncbi.nlm.nih.gov/31364386/)]
6. Cook BG, Li D, Heinrich KM. Obesity, physical activity, and sedentary behavior of youth with learning disabilities and ADHD. *J Learn Disabil* 2015 Jan 21;48(6):563-576. [doi: [10.1177/0022219413518582](https://doi.org/10.1177/0022219413518582)] [Medline: [24449262](https://pubmed.ncbi.nlm.nih.gov/24449262/)]
7. Howard AL, Robinson M, Smith GJ, Ambrosini GL, Piek JP, Oddy WH. ADHD is associated with a "Western" dietary pattern in adolescents. *J Atten Disord* 2011 Jul;15(5):403-411. [doi: [10.1177/1087054710365990](https://doi.org/10.1177/1087054710365990)] [Medline: [20631199](https://pubmed.ncbi.nlm.nih.gov/20631199/)]
8. Thomas S, Lycett K, Papadopoulos N, Sciberras E, Rinehart N. Exploring behavioral sleep problems in children with ADHD and comorbid autism spectrum disorder. *J Atten Disord* 2018 Aug 04;22(10):947-958. [doi: [10.1177/1087054715613439](https://doi.org/10.1177/1087054715613439)] [Medline: [26637841](https://pubmed.ncbi.nlm.nih.gov/26637841/)]
9. Gregory AM, Sadeh A. Annual research review: sleep problems in childhood psychiatric disorders—a review of the latest science. *J Child Psychol Psychiatry* 2016 Mar 28;57(3):296-317. [doi: [10.1111/jcpp.12469](https://doi.org/10.1111/jcpp.12469)] [Medline: [26412255](https://pubmed.ncbi.nlm.nih.gov/26412255/)]
10. Hoare E, Milton K, Foster C, Allender S. The associations between sedentary behaviour and mental health among adolescents: a systematic review. *Int J Behav Nutr Phys Act* 2016 Oct 08;13(1):108 [FREE Full text] [doi: [10.1186/s12966-016-0432-4](https://doi.org/10.1186/s12966-016-0432-4)] [Medline: [27717387](https://pubmed.ncbi.nlm.nih.gov/27717387/)]
11. Ash T, Bowling A, Davison K, Garcia J. Physical activity interventions for children with social, emotional, and behavioral disabilities—a systematic review. *J Dev Behav Pediatr* 2017;38(6):431-445. [doi: [10.1097/dbp.0000000000000452](https://doi.org/10.1097/dbp.0000000000000452)]
12. Bowling A, Slavet J, Miller DP, Haneuse S, Beardslee W, Davison K. Cybercycling effects on classroom behavior in children with behavioral health disorders: an RCT. *Pediatrics* 2017 Feb 09;139(2):e20161985 [FREE Full text] [doi: [10.1542/peds.2016-1985](https://doi.org/10.1542/peds.2016-1985)] [Medline: [28069663](https://pubmed.ncbi.nlm.nih.gov/28069663/)]
13. Pontifex MB, Saliba BJ, Raine LB, Picchietti DL, Hillman CH. Exercise improves behavioral, neurocognitive, and scholastic performance in children with attention-deficit/hyperactivity disorder. *J Pediatr* 2013 Mar;162(3):543-551 [FREE Full text] [doi: [10.1016/j.jpeds.2012.08.036](https://doi.org/10.1016/j.jpeds.2012.08.036)] [Medline: [23084704](https://pubmed.ncbi.nlm.nih.gov/23084704/)]
14. Quan M, Pope Z, Gao Z. Examining young children's physical activity and sedentary behaviors in an exergaming program using accelerometry. *J Clin Med* 2018 Sep 25;7(10):302 [FREE Full text] [doi: [10.3390/jcm7100302](https://doi.org/10.3390/jcm7100302)] [Medline: [30257418](https://pubmed.ncbi.nlm.nih.gov/30257418/)]
15. Malik AA, Williams CA, Weston KL, Barker AR. Influence of personality and self-efficacy on perceptual responses during high-intensity interval exercise in adolescents. *J Appl Sport Psychol* 2020 Feb 06:1-19. [doi: [10.1080/10413200.2020.1718798](https://doi.org/10.1080/10413200.2020.1718798)]
16. Staiano AE, Beyl RA, Guan W, Hendrick CA, Hsia DS, Newton RL. Home-based exergaming among children with overweight and obesity: a randomized clinical trial. *Pediatr Obes* 2018 Nov 20;13(11):724-733 [FREE Full text] [doi: [10.1111/ijpo.12438](https://doi.org/10.1111/ijpo.12438)] [Medline: [30027607](https://pubmed.ncbi.nlm.nih.gov/30027607/)]
17. Lubans DR, Foster C, Biddle SJ. A review of mediators of behavior in interventions to promote physical activity among children and adolescents. *Prev Med* 2008 Nov;47(5):463-470. [doi: [10.1016/j.ypmed.2008.07.011](https://doi.org/10.1016/j.ypmed.2008.07.011)] [Medline: [18708086](https://pubmed.ncbi.nlm.nih.gov/18708086/)]
18. Rimmer JH, Vanderbom KA. A call to action: building a translational inclusion team science in physical activity, nutrition, and obesity management for children with disabilities. *Front Public Health* 2016 Aug 10;4:164 [FREE Full text] [doi: [10.3389/fpubh.2016.00164](https://doi.org/10.3389/fpubh.2016.00164)] [Medline: [27559540](https://pubmed.ncbi.nlm.nih.gov/27559540/)]
19. Matthews KA, Gallo L, Taylor S. Are psychosocial factors mediators of socioeconomic status and health connections? A progress report and blueprint for the future. *Ann N Y Acad Sci* 2010 Feb;1186(1):146-173. [doi: [10.1111/j.1749-6632.2009.05332.x](https://doi.org/10.1111/j.1749-6632.2009.05332.x)] [Medline: [20201872](https://pubmed.ncbi.nlm.nih.gov/20201872/)]
20. Davison KK, Jurkowski JM, Lawson HA. Reframing family-centred obesity prevention using the Family Ecological Model. *Public Health Nutr* 2012 Oct 22;16(10):1861-1869. [doi: [10.1017/s1368980012004533](https://doi.org/10.1017/s1368980012004533)]
21. Bowling A, Blaine RE, Kaur R, Davison KK. Shaping healthy habits in children with neurodevelopmental and mental health disorders: parent perceptions of barriers, facilitators and promising strategies. *Int J Behav Nutr Phys Act* 2019 Jun 26;16(1):52 [FREE Full text] [doi: [10.1186/s12966-019-0813-6](https://doi.org/10.1186/s12966-019-0813-6)] [Medline: [31242904](https://pubmed.ncbi.nlm.nih.gov/31242904/)]
22. Nelson EL, Sharp TB. Telemental health competencies: training examples from a youth depression telemedicine clinic. In: *Technology Innovations for Behavioral Education*. New York: Springer; 2011:41-47.
23. Sim J, Lewis M. The size of a pilot study for a clinical trial should be calculated in relation to considerations of precision and efficiency. *J Clin Epidemiol* 2012 Mar;65(3):301-308. [doi: [10.1016/j.jclinepi.2011.07.011](https://doi.org/10.1016/j.jclinepi.2011.07.011)] [Medline: [22169081](https://pubmed.ncbi.nlm.nih.gov/22169081/)]
24. Julious SA. Sample size of 12 per group rule of thumb for a pilot study. *Pharmaceut Statist* 2005 Oct;4(4):287-291. [doi: [10.1002/pst.185](https://doi.org/10.1002/pst.185)]
25. Piercy KL, Troiano RP, Ballard RM, Carlson SA, Fulton JE, Galuska DA, et al. The physical activity guidelines for Americans. *J Am Med Assoc* 2018 Nov 20;320(19):2020-2028. [doi: [10.1001/jama.2018.14854](https://doi.org/10.1001/jama.2018.14854)] [Medline: [30418471](https://pubmed.ncbi.nlm.nih.gov/30418471/)]
26. TRUE Consortium. Recommended standards for assessing blood pressure in human research where blood pressure or hypertension is a major focus. *J Hum Hypertens* 2017 Aug 13;31(8):487-490. [doi: [10.1038/jhh.2017.10](https://doi.org/10.1038/jhh.2017.10)] [Medline: [28406234](https://pubmed.ncbi.nlm.nih.gov/28406234/)]
27. Puyau MR, Adolph AL, Vohra FA, Zakeri I, Butte NF. Prediction of activity energy expenditure using accelerometers in children. *Med Sci Sports Exerc* 2004 Sep;36(9):1625-1631. [Medline: [15354047](https://pubmed.ncbi.nlm.nih.gov/15354047/)]
28. Trost S, Loprinzi P, Moore R, Pfeiffer KA. Comparison of accelerometer cut points for predicting activity intensity in youth. *Med Sci Sports Exerc* 2011 Jul;43(7):1360-1368. [doi: [10.1249/MSS.0b013e318206476e](https://doi.org/10.1249/MSS.0b013e318206476e)] [Medline: [21131873](https://pubmed.ncbi.nlm.nih.gov/21131873/)]

29. Reed GR. Measuring stage of change for exercise. Open Access Dissertations, The University of Rhode Island 1995:1-150. [doi: [10.23860/diss-reed-gabrielle-1995](https://doi.org/10.23860/diss-reed-gabrielle-1995)]
30. Owens JA, Spirito A, McGuinn M. The Children's Sleep Habits Questionnaire (CSHQ): psychometric properties of a survey instrument for school-aged children. *Sleep Res Soc, Oxford Acad* 2000 Dec;23(8):1-9. [doi: [10.1093/sleep/23.8.1d](https://doi.org/10.1093/sleep/23.8.1d)] [Medline: [33886181](https://pubmed.ncbi.nlm.nih.gov/33886181/)]
31. Anderson SE, Must A, Curtin C, Bandini LG. Meals in our household: reliability and initial validation of a questionnaire to assess child mealtime behaviors and family mealtime environments. *J Acad Nutr Diet* 2012 Feb;112(2):276-284 [FREE Full text] [doi: [10.1016/j.jada.2011.08.035](https://doi.org/10.1016/j.jada.2011.08.035)] [Medline: [22741169](https://pubmed.ncbi.nlm.nih.gov/22741169/)]
32. O'Neil A, Quirk SE, Housden S, Brennan SL, Williams LJ, Pasco JA, et al. Relationship between diet and mental health in children and adolescents: a systematic review. *Am J Public Health* 2014 Oct;104(10):31-42. [doi: [10.2105/ajph.2014.302110](https://doi.org/10.2105/ajph.2014.302110)]
33. Madigan S, Racine N, Cooke JE, Korczak DJ. COVID-19 and telemental health: benefits, challenges, and future directions. *Can Psychol/Psychologie Canadienne* 2021 Feb;62(1):5-11. [doi: [10.1037/cap0000259](https://doi.org/10.1037/cap0000259)]

Abbreviations

ADHD: attention-deficit/hyperactivity disorder
AGS: Adaptive GameSquad
ASD: autism spectrum disorder
BIS: behavioral inhibition bias
BP: blood pressure
mHealth: mobile health
MIOH: Meals In Our Household Questionnaire
MVPA: moderate- to vigorous-intensity physical activity
NPD: neurodevelopmental and psychiatric diagnosis
PA: physical activity
REDCap: Research Electronic Data Capture

Edited by G Eysenbach; submitted 24.09.20; peer-reviewed by D Pfürringer, G Signorelli, X He, D Filos; comments to author 30.12.20; revised version received 18.01.21; accepted 13.04.21; published 14.05.21.

Please cite as:

Bowling AB, Slavet J, Hendrick C, Beyl R, Nauta P, Augustyn M, Mbamalu M, Curtin C, Bandini L, Must A, Staiano AE
The Adaptive GameSquad Xbox-Based Physical Activity and Health Coaching Intervention for Youth With Neurodevelopmental and Psychiatric Diagnoses: Pilot Feasibility Study
JMIR Form Res 2021;5(5):e24566
URL: <https://formative.jmir.org/2021/5/e24566>
doi: [10.2196/24566](https://doi.org/10.2196/24566)
PMID: [33988508](https://pubmed.ncbi.nlm.nih.gov/33988508/)

©April B Bowling, James Slavet, Chelsea Hendrick, Robbie Beyl, Phillip Nauta, Marilyn Augustyn, Mediatrix Mbamalu, Carol Curtin, Linda Bandini, Aviva Must, Amanda E Staiano. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 14.05.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

mHealth Interventions to Support Prescription Opioid Tapering in Patients With Chronic Pain: Qualitative Study of Patients' Perspectives

Michael Reece Magee¹, MPsych; Amy Gray McNeilage¹, BA, GradDipPsych; Nicholas Avery¹, MPE; Paul Glare¹, MD, MMed; Claire Elizabeth Ashton-James¹, PhD

Pain Management Research Institute, Faculty of Medicine and Health, The University of Sydney, St Leonards, Australia

Corresponding Author:

Claire Elizabeth Ashton-James, PhD
Pain Management Research Institute
Faculty of Medicine and Health
The University of Sydney
Lvl 2 Douglas Building
Royal North Shore Hospital
St Leonards, 2065
Australia
Phone: 61 294631528
Email: claire.ashton-james@sydney.edu.au

Abstract

Background: Patients with chronic pain who are tapering prescription opioids report a need for greater support for coping with symptoms of pain and withdrawal. Mobile health (mHealth) technologies (SMS text messaging– or app-based) have the potential to provide patients with educational, emotional, and motivational support for opioid tapering beyond what is offered by their health care provider. However, it is not known whether patients with chronic pain who are tapering opioids would be willing or able to engage with technology-based support.

Objective: This study aims to examine patients' use of mobile technologies in health care, interest in using mHealth support, preferences for the form and content of mHealth support, and potential barriers to and facilitators of engagement with mHealth support for opioid tapering.

Methods: A total of 21 patients (11 women and 10 men; age range 29–83 years) with chronic noncancer pain on long-term opioid therapy who had recently initiated a voluntary opioid taper were recruited from primary and tertiary care clinics in metropolitan and regional Australia for a larger study of patients' experiences of opioid tapering. Participants had been taking prescription opioids for a mean duration of 13 (SD 9.6; range 0.25–30) years at the time of the study. Survey items characterized participants' typical mobile phone use and level of interest in mobile technology–based support for opioid tapering. Semistructured interviews further explored patients' use of mobile technologies and their interest in, preferences for, and perspectives on potential barriers to and facilitators of engagement with mHealth support for opioid tapering. Two researchers collaborated to conduct a thematic analysis of the interview data.

Results: All participants reported owning and using a mobile phone, and most (17/21, 81%) participants reported using mobile apps. The majority of participants expressed interest in SMS text messaging–based (17/21, 81%) and app-based (15/21, 71%) support for opioid tapering. Participants expected that messages delivering both informational and socioemotional support would be helpful. Participants expected that access to technology, mobile reception, internet connectivity, vision impairment, and low self-efficacy for using apps may be barriers to user engagement. Patients expected that continuity of care from their health care provider, flexible message dosing, responsiveness, and familiarity with pain self-management strategies would increase user engagement.

Conclusions: The results of this study indicate that patients with chronic noncancer pain may be willing to engage with SMS text messaging–based and app-based mHealth interventions to support opioid tapering. However, the feasibility and acceptability of these interventions may depend on how patients' preferences for functionality, content, and design are addressed.

(JMIR Form Res 2021;5(5):e25969) doi:[10.2196/25969](https://doi.org/10.2196/25969)

KEYWORDS

prescription opioids; chronic noncancer pain; tapering; digital health; support; mobile health technology; SMS; mobile phone

Introduction

Background

Chronic pain is one of the leading causes of disability worldwide, affecting approximately 1 in 5 adults [1-4]. Historically, chronic pain management has relied heavily on prescription opioid analgesics. However, the use of opioids in the treatment of chronic noncancer pain (CNCN) remains controversial. Evidence suggests that long-term opioid therapy has limited effectiveness in the management of CNCN [5,6] and carries significant risks, including physical dependence, tolerance (and associated dose escalation), and hyperalgesia (increased pain sensitivity) [7,8].

Current guidelines recommend that patients on long-term opioid therapy for CNCN should be gradually tapered off these medications under the guidance of a health care provider [5,6]. However, many patients are reluctant to change the way they manage their pain, particularly after long-term opioid use. Prescription opioid tapering is notoriously difficult, but it may be particularly difficult for patients with CNCN [9-11]. After prolonged use, even gradual reductions in opioid dosage may lead to withdrawal symptoms, including disrupted appetite, sleep, and mood (eg, agitation, anxiety, and restlessness); abdominal dysfunction; nausea; and muscle aches [11-13]. In addition to these withdrawal symptoms, patients with CNCN may also experience a temporary *flare-up* of their pain condition [12]. Faced with the dual challenges of coping with pain and withdrawal symptoms, it is not surprising that many patients with CNCN find it difficult to maintain motivation to continue opioid tapering [7,9-11].

Emerging data suggest that when patients volunteer to taper, a patient-centered, multimodal approach involving pain education, nonpharmacological pain self-management strategies, routine follow-up with a trusted health care provider, and support for coping with pain and withdrawal symptoms may improve outcomes [9,14-18]. However, patients face a number of obstacles in accessing comprehensive opioid tapering support. In addition to cost issues, multidisciplinary pain services are limited, and primary care providers often lack confidence in delivering nonpharmacological pain management advice. Furthermore, patients may be reluctant to admit their need for support with opioid tapering for fear of being stigmatized [6,8,10-12,14,16,19].

Mobile health (mHealth) interventions may be an effective means of overcoming many of these barriers to opioid tapering in patients with CNCN [20]. mHealth interventions have been found to be effective in supporting people to manage a variety of complex chronic health conditions, including diabetes, cardiovascular disease, and obesity [21-25]. mHealth has also been used effectively to support the cessation of addictive substances [26-28], to improve people's ability to cope with chronic pain, and to enhance perceived quality of life [29-32]. One study found that a narrative (patient testimony) educational video about opioid tapering increased patients' tapering

self-efficacy to a greater degree than a pamphlet with the same information [33]. This study demonstrates how multimedia and digital platforms can play a role in increasing patients' confidence and motivation to taper opioid medications.

The potential benefit of SMS text messaging-based interventions to support opioid tapering in patients with CNCN is supported by studies investigating the use of telephone-delivered messages to support chronic pain management, on the one hand, and opioid tapering, on the other. First, a 2010 study investigated the effect of delivering automated telephone voice messages to patients to remind them to use pain self-management skills after they completed an 11-week group pain management program. The researchers found that those who received the messages used fewer opioid medications 8 months after the pain management program compared with participants who did not receive digital aftercare [34]. Second, a 2020 study investigated the impact of receiving SMS text messaging-delivered cognitive behavioral pain self-management strategies on pain intensity and opioid use after surgery. Participants who received the SMS text messaging-delivered support reported lower pain intensity 2 weeks after surgery compared with a control group who received treatment as usual [35].

Objective

Although these studies point to the potential of mHealth interventions to support prescription opioid tapering in individuals with chronic pain, there is currently no direct evidence that patients with chronic pain would be willing to engage with mHealth interventions as a means of support for opioid tapering and, if so, whether or not they would be able to engage with these technologies. This study aims to explore patients' access to and use of mobile technology, interest in mHealth support for opioid tapering, preferences for the form and content of mHealth support, and potential barriers to and facilitators of engagement with mHealth support for opioid tapering. Using a combination of survey and interview methods to achieve these objectives, we also aim to elicit consumer perspectives to inform the design, content, and functionality of an mHealth intervention.

Methods

Setting

Participants were recruited from a private primary care practice in regional New South Wales and a public tertiary pain clinic in a metropolitan area in Australia between September 2019 and March 2020. The Human Research Ethics Committee of the Northern Sydney Local Health District approved the study (study 2019/STE00599).

Inclusion and Exclusion Criteria

Eligible participants were those who were aged >18 years, were living with CNCN (pain persisting for ≥6 months), had been taking opioid medications for chronic pain for >3 months, and voluntarily commenced or planned to commence opioid tapering. Exclusion criteria included a history of opioid use

disorder, major psychiatric conditions, insufficient English proficiency, or inability to provide informed consent. Patients with comorbid opioid use disorder were excluded from the study because the challenges faced by these patients when tapering are potentially more complex, requiring greater medical management, and the needs of this group are not necessarily shared by most patients tapering opioids for chronic pain [36].

Design and Procedure

Data for this study were collected as part of a larger investigation into the experience of tapering in patients with CNCP. Patients who met the inclusion criteria were asked if they would be interested in hearing about the study. A researcher (AGM) contacted interested patients to obtain their informed consent to participate. Participants completed a very brief survey before or within the first few weeks of voluntarily tapering prescription opioids. Participants were later (during the first three months of their taper) interviewed (by AGM) and asked to elaborate upon their survey responses. A total of 2 independent researchers (MRM and CEAJ) analyzed the survey and interview data.

Survey

The survey was designed to gather descriptive data on the sample of participants' use of mobile phone technology (specifically SMS text messaging, apps, and health apps) and level of interest in engaging with mHealth support for opioid tapering (Multimedia Appendix 1). Specifically, participants were asked to respond to 2 key questions using a 5-point scale (1=no interest at all and 5=very interested): "How interested are you in receiving text messages with informative and supportive content while reducing your opioid medication?" and "How interested are you in using an app designed to support you in reducing opioid medication?" The survey also recorded demographic and clinical data, including age, gender, education, relationship status, duration of pain, and duration of opioid use.

Interview

Semistructured interview questions were designed to further explore patterns of mobile phone and app use among participants in this cohort (eg, "How often do you use your phone?" and "At what times of day do you typically use your phone?"), identify reasons why participants expressed an interest or lack of interest in mHealth support for opioid tapering (eg, "You've said you would prefer SMS-based support over app-based

support. Why is that?"), elicit participants' preferences for the form and content of an mHealth intervention ("Do you have any thoughts on what kind of information would be helpful?" and "What advice would be helpful for someone who was about to taper their opioid medication?"), and explore potential barriers to and facilitators of engagement with mHealth support for opioid tapering (eg, "Why do you feel it would not be helpful for you?" and "Is there a time of day you might be more or less likely to look at it?").

An interview guide was prepared as a prompt for interviewers to focus their discussion on issues pertinent to the aims of the study (Multimedia Appendix 1). However, interviewers were trained to use open-ended questions in response to participants' answers to the stem questions to encourage elaboration (eg, "Tell me more about why you think you would not be interested in this sort of app?"). Data saturation was reached when participants no longer had additional perspectives to share when prompted on the semistructured interview items or open-ended follow-up questions.

Participants

A total of 21 participants were recruited for this study. Participants were recruited until sample heterogeneity was achieved in terms of age, tapering experience, gender, education level, employment status, health care setting, pain duration, and opioid use duration. One participant withdrew from the study after completing the survey but before being interviewed for this study (discontinued tapering). Overall, 62% (13/21) of participants were recruited from a tertiary pain clinic, and 38% (8/21) of participants were recruited from primary care. Table 1 provides the demographic and clinical characteristics of the sample, including age, education, relationship status, employment status, pain duration, opioid medication duration, and oral morphine equivalent daily dose before tapering.

Most participants (17/21, 81%) reported prior experience with opioid tapering. Tapering experience was variable, including tapering from a very high dose to a moderate dose, with subsequent tapering to further reduce the dose. Others had successfully reduced or discontinued opioid medications in the past but had resumed opioid use after an accident or injury and were motivated to taper again. Participants were not recruited for this study based on their mHealth use or level of comfort with digital technology.

Table 1. Participant characteristics (N=21).

Characteristics	Values
Age (years), mean (SD; range)	55 (12.26; 29-83)
Gender, n (%)	
Female	11 (52)
Male	10 (48)
Recruitment setting, n (%)	
Tertiary pain clinic	13 (62)
Primary care	8 (38)
Education, n (%)	
High school graduate	3 (14)
Vocational training	6 (29)
Bachelor's degree	7 (33)
Postgraduate degree	2 (10)
Data missing or not reported	3 (14)
Relationship status, n (%)	
Single	2 (10)
Married	10 (48)
In a relationship	3 (14)
Widowed	2 (10)
Data missing or not reported	4 (19)
Employment status, n (%)	
Employed	4 (19)
Unemployed or not working	16 (78)
Data missing or not reported	1 (5)
Pain duration (years), mean (SD; range)	13 (9.6; 0.25-30)
Duration of opioid treatment (years), mean (SD; range)	9.3, (7.5; 0.25-30)
Oral morphine equivalent daily dose (mg), mean (SD; range)	150 (229.29; 20-1080)

Data Analytic Technique

Interview data were audio recorded and then transcribed verbatim into NVivo 12 (QSR International), a software platform for the qualitative coding and analysis of data [37]. A total of 2 researchers (MRM and CEAJ) conducted a mixed inductive-deductive thematic analysis. Thematic analysis was chosen as it provides a stepwise procedure for identifying the key features of a qualitative data set and is an appropriate methodology when there are clear research questions [37]. Furthermore, the data were considered to be generally of a manifest nature, in that the literal interpretations of the responses were deemed to accurately reflect the participants' perspectives. Interview excerpts were coded, and initial themes were identified. The researchers independently completed conceptual mapping and thematic grouping of the data and collaboratively identified emergent higher-order themes and subthemes. The findings were discussed with a wider research group, in

particular, as other members of the team had conducted the interviews and confirmed that accurate interpretation of results was imperative. Participants were not asked to review the results, as interview data were not of a complex phenomenological nature, and as such, it was considered that participant responses were reasonably clear and straightforward to interpret.

Results

Survey Results

All participants (21/21, 100%) reported using a mobile phone. Most participants (17/21, 81%) reported using mobile apps, including health and wellness apps (10/21, 47%). Overall, 81% (17/21) of participants reported being *interested* or *very interested* in SMS text messaging support for opioid tapering, and 71% (15/21) reported being *interested* or *very interested* in app support for opioid tapering (Table 2).

Table 2. Survey results.

Survey item and response option	Frequency, n (%)
Use a mobile phone	
Yes	21 (100)
No	0 (0)
Interest in SMS text messaging support for opioid tapering	
Not interested at all	0 (0)
Not interested	3 (14)
Neither interested nor disinterested	1 (5)
Interested	16 (76)
Very interested	1 (5)
Use of mobile apps	
Yes	17 (81)
No	4 (19)
Use of health and wellness apps	
Yes	10 (48)
No	9 (43)
Interest in app support for opioid tapering	
Not interested at all	2 (9)
Not interested	2 (9)
Neither interested nor disinterested	0 (0)
Interested	13 (62)
Very interested	2 (9)
Missing responses	2 (9)

Interview Results

Patterns of Mobile Phone Use

During interviews, participants revealed that they most commonly used their mobile phones to communicate with others and did so via phone calls, SMS text messaging, and apps (eg, WhatsApp and Facebook Messenger). Participants reported using apps, including internet browsers, calendars, web-based booking platforms, games, banking, news, and social media. One participant reported using apps to record and listen to sessions with a psychologist for pain management support.

The frequency of phone use varied. Participants indicated that they had access to their phones all the time (“They’re kind of attached to you,” P18, male, 41 years, pain clinic, regional area), even if they did not always use them (“It’s just for emergencies,” P13, male, 51 years, pain clinic, metropolitan area). Some participants reported using their phones mostly in the morning, whereas others reported more frequent phone and app use in the afternoons and evenings. One participant attributed reduced phone use in the morning to pain- and medication-related issues:

I struggle to get out of bed in the mornings because I don’t sleep well and that’s generally the point of time where the medication has worn off from the night

before. [P18, male, 41 years, pain clinic, regional area]

Interest in mHealth Support for Tapering

The proposed delivery of support via SMS text messaging or an app was perceived generally as positive:

I think just getting the support no matter what form it is in is helpful. [P08, female, 32 years, pain clinic, metropolitan area]

Some participants reported that electronic delivery of motivation and encouragement would be welcome:

I think any form of motivation is helpful. [P10, male, 64 years, pain clinic, metropolitan area]

A primary benefit of an mHealth intervention reported by participants was increased access to support for opioid tapering during periods when other services were not available (ie, out of office hours):

That’s where messages are going to work well...these [worries] can [come] at two in the morning and you need a response...It can be an automated thing. [P01, male, 60 years, pain clinic, regional area]

One participant reported that an mHealth intervention would be beneficial as she had “nobody checking in to see how I am” over the Christmas period:

I think the mobile system would be good for that because it feels like you have some way of reaching someone or some way of someone reaching you to check in. [P08, female, 32 years, pain clinic, metropolitan area]

Although participants were, in general, interested in using SMS text messaging–based or app-based support for opioid tapering, they also expressed a strong preference for clinician-delivered support either face-to-face or over the phone:

I like being able to I guess just talk...I tend to kind of ignore my texts sometimes if I'm busy. [P12, male, 46 years, pain clinic, metropolitan area]

No I think it's better to talk to somebody...No, I'd sooner talk over the phone. [P11, female, 83 years, primary care, regional area]

I think the phone call thing is better than any social media, better than anything you just type in or anything because it's human. [P04, male, 57 years, pain clinic, metropolitan area]

There were concerns about how this would be interpreted if it was clear that the support was computer delivered:

If it's just a text message, I'm not dumb, I know it's just a computer sending it. If it gets sent out the same time every day, it's a computer sending it. [P06, female, 59 years, pain clinic, metropolitan area]

Participants reported some skepticism that automated messages would feel genuine:

If I think I'm getting just some random "great you're doing well" and I know it's been automatically generated, I probably wouldn't be that tickled by it. [P02, female, 61 years, pain clinic, metropolitan area]

Participants who did not express an interest in an mHealth intervention to support opioid tapering reasoned either that they had a strong desire for face-to-face support ("I think personally you need to be human to human," P21, male, 53, primary care, regional area) or a preference for alternative sources of information and support:

I would probably prefer just reading information or just talking to someone before using an app. If there was a web page, I'd probably be likely to use the web page more. [P02, female, 61 years, pain clinic, regional area]

However, one participant reported a preference for SMS text messaging rather than a phone call or face-to-face check in:

The text messaging I think would be really good because the thing that probably scares me most is to speak to someone about it. [P08, female, 32 years, pain clinic, metropolitan area]

Preferences for mHealth Design and Functionality

In line with the preference for face-to-face and nonautomated services outlined earlier, several participants recommended that an option for personal follow-up be built into the service:

I think the mobile system would be good for that because it feels like you have some way of reaching someone or some way of someone reaching you to check in. [P08, female, 32 years, pain clinic, metropolitan area]

Most participants did not express a strong preference for either SMS text messaging–based or app-based support. Participants who expressed a preference for SMS text messaging–based support indicated a higher level of familiarity with text messaging:

I can use [SMS] easier...because I don't use my phone much for that sort of thing [apps]. [P09, male, 61 years, primary care, regional area]

Participants who expressed a preference for app-based support reasoned that apps may provide access to more content:

I'm guessing inside that app there would be probably more information than in the texts. [P18, male, 41 years, pain clinic, regional area]

One participant stated that, although they were willing to read texts, they did not like to respond to them:

I just hate texting. [P16, female, 46 years, pain clinic, metropolitan area]

Recommended Content of mHealth Support

Participants provided recommendations for content that they believed would be helpful in an mHealth intervention. Specifically, they expressed interest in receiving information about pain management during flare-ups and about withdrawal from opioid medications (informational support and reassurance) and messages of encouragement, motivation, and validation (socioemotional support and reassurance).

Several participants expressed a desire for information related to pain self-management strategies:

I think it would be really useful to have maybe different things on there like techniques to help. "If you've got really bad pain, why don't you try this?" [P08, female, 32 years, pain clinic, metropolitan area]

I think it would be nice to have like an app or texting or whatever that did embody like the mindfulness and the meditation and the feedback, so it's all in one place. [P20, female, 63 years, pain clinic, metropolitan area]

It's about thought and relaxation...thought management. [P03, female, 29 years, pain clinic, regional area]

One participant recommended tailoring the information to reflect the user's stage of tapering to reassure them that what they were experiencing was normal:

Maybe "You're on week 4 of your taper and you're on this dose. Be mindful that you may experience this, this, and this." So, give them a bit of warning. And "good ways to deal with those things are..." Then people know that they're normal. Once people go "oh okay this is normal," you take the fear out of it. [P08, female, 32 years, pain clinic, metropolitan area]

Another participant agreed that knowing what to expect, particularly with regard to withdrawal symptoms, would be empowering:

I didn't know what to expect. I thought it might just get worse and worse and worse for weeks. If I just had a time frame and it might be different...If I had been warned that you could get constipation and you'll need to take precautions. [P02, female, 61 years, pain clinic, regional area]

A number of participants commented more generally on the importance of social support when tapering, either from friends and family, a group within a treatment program, or clinicians. One participant recommended integrating electronic social support for patients who were tapering their opioid medication:

If you can push for a forum or something, I think that would be awesome. [P16, female, 46 years, pain clinic, metropolitan area]

Participants' views regarding the acceptability of mHealth interventions for opioid tapering were influenced by factors such as the frequency of the messages and pattern of delivery. One participant preferred receiving messages at night ("Probably early evening just as you're winding down, to help me to wind down properly," P04, male, 57 years, pain clinic, metropolitan area), whereas another was concerned about receiving too frequent or too repetitive messages. Their suggestion to counter this was for the digital service to be responsive to the needs of the user:

So maybe there is a text that's in response to a word we send out. Like we need some affirmation...It could be geared to that more. [P01, male, 60 years, pain clinic, regional area]

Potential Barriers to and Facilitators of mHealth Engagement

Several participants described perceived barriers toward phone-based digital interventions including limited phone reception or access to internet:

I don't have Wi-Fi. I can't afford it. [P06, female, 59 years, pain clinic, metropolitan area]

I don't think my old phone will support it. [P14, male, 56 years, primary care, regional area]

Some participants spoke about vision impairment:

I use glasses and my phone is a very cheap phone. If it's in the middle of the day I can't read them. [P14, male, 56 years, primary care, regional area]

For the older ones they can't see stuff on their phone properly. [P16, female, 46 years, pain clinic, metropolitan area]

Some participants reported that low confidence in using mobile devices and apps may be an obstacle to engagement with app-based interventions:

No, I wouldn't know how [to use apps]...I'm not tech savvy. [P09, Male, 61 years, primary care, regional area]

Relatedly, some participants suggested that ease of use and simplicity of content would facilitate engagement:

If it's not too time consuming and it's not complicated...your average person isn't [so tech savvy] so it can't be complicated. [P05, female, 58 years, pain clinic, metropolitan area]

At the same time, one participant suggested that their confidence in using mHealth interventions could be improved if provided with technical support (P20, female, 63 years, pain clinic, metropolitan area).

Participants in the study had varying levels of familiarity and experience with pain self-management strategies, and one participant suggested that it may be difficult to engage with digitally delivered pain self-management advice if one is not already experienced in using them:

You don't go driving on your own for the first time. You have to get the license. You need your lessons first...I would hate to have to do this [pain self-management] and not know how. [P01, male, 60 years, pain clinic, regional area]

Participants predicted that their level of interest and engagement with mHealth support may fluctuate over the course of their taper depending on their mood, pain, and state of mind:

Whether I looked at it or not would depend I guess on my mood—how I'm feeling, where I am. [P18, male, 41 years, pain clinic, regional area]

At the time you receive it you mightn't be that open to it but in a few hours' time when you're starting to struggle a bit or the next day it may be something you open up again and read through. [P10, male, 64 years, pain clinic, metropolitan area]

Whether the content was automated, personalized, and integrated with face-to-face treatment was identified as a potential barrier to engagement. Participants were concerned that receiving impersonal (automated and standardized) messages may not have the same impact as receiving individualized messages from a person in real time:

The text message, in reality that's still a computer sending it, there's not someone at the end of it that's sending it. I don't know whether it would just get irritating after a while because I know nothing is going to come of it, like there is no one at the other end. Or whether it would just jolt me out of feeling sorry for myself or something. I guess it depends how it's worded. [P06, female, 59 years, pain clinic, metropolitan area]

Discussion

Principal Findings

The results revealed that all participants owned and used mobile phones. Most participants reported using smartphone apps. The majority of participants expressed interest in SMS text messaging-based and app-based support for tapering prescription opioids. Mostly, participants held no strong preference for either SMS text messaging or app support.

Variables potentially influencing acceptability included message simplicity, messaging frequency, time of delivery, responsiveness to user needs, individual stage of tapering, and mental state. There was a strong preference for clinician-delivered, individualized, real-time support over automated and standardized mHealth support. At the same time, however, a perceived benefit of mHealth was increased accessibility of opioid tapering support. Content recommendations included pain self-management and opioid withdrawal information as well as encouragement, motivation, and validation. Perceived barriers included limited phone or internet reception and low confidence in using mobile devices and apps. These results suggest that SMS text messaging–delivered or app-delivered support may be acceptable, feasible, and even helpful to individuals living with CNCP who are tapering opioids. However, participants identified a number of barriers to engagement that may be critical to address in the development of such an mHealth intervention.

Implications for Design and Functionality

mHealth interventions are typically delivered via an app or SMS text messaging. Participants identified barriers such as access to technology, mobile reception, internet connectivity, vision impairment, and low confidence in using apps. Most of these obstacles can be overcome by using an SMS text messaging–based intervention rather than an app-based intervention. SMS text messaging–based interventions are very simple and easy to use and do not require an internet connection, and messages tend to be shorter and therefore less burdensome to read.

Participants suggested a number of factors that might influence engagement with mHealth interventions for prescription opioid tapering. In particular, participants expressed a desire to continue to receive care from their health care provider, pointing to the importance of ensuring that mHealth interventions are used as an adjunct to, rather than a replacement of, the patient-provider communication. Relatedly, some participants expressed concern that support received via text messages would feel impersonal or ingenuine. In response to this (common) concern, many mHealth interventions now personalize SMS text messaging content by, for example, addressing users by their preferred name or by tailoring messaging to the interests of the user (eg, “Hi Sam, doing things we enjoy helps us to feel good, so we feel less pain. Make plans to go for a hike, or perhaps meet a friend for coffee.”) These personalization strategies have been found to increase engagement with mHealth interventions [21,38].

Participants also indicated that their engagement with text messages may fluctuate and suggested that flexible message *dosing*—the ability to change the frequency of messages they receive depending on their mood, symptoms, coping, and need—may enhance the acceptability of an mHealth intervention for opioid tapering. It was also suggested that responsiveness—the ability to request support *on demand*—would enhance the acceptability of an mHealth intervention. These concerns are not uncommon, and research suggests that user engagement can be enhanced by allowing users a degree of control over the

frequency of messages they receive, either up front (initially) or after a standardized *loading* dose [21].

Another method of tailoring the frequency of messages to user demands is to allow users to text prompts to the program when they wish to receive a message. The content of *on-demand* messages can be determined by the user based on established keywords. For example, in response to texting *pain*, the user might receive a message offering pain coping strategies (eg, “breathe in slowly to a count of four, breathe out for six”) or informational reassurance (eg, “pain may flare up for many reasons—not necessarily due to tapering”). This functionality is common and feasible in mHealth technology [28,30].

Implications for mHealth Intervention Content

Participants made 2 key recommendations for the content of an mHealth intervention to support people similar to them who were tapering opioids that were prescribed for CNCP. First, they recommended that users would benefit from messages about chronic pain management, opioid tapering, and strategies for coping with pain and withdrawal symptoms. This perspective concurs with research demonstrating the positive association between giving patients informational or *cognitive* reassurance and pain management outcomes. The results of a systematic review found that patients who were provided with an explanation for their pain condition and reasons for their treatment advice experienced less distress and showed greater symptom improvement [39,40].

At the same time, however, participants also noted that it may be difficult to understand and engage with information about coping with pain and symptoms of withdrawal via text if one does not already have some familiarity with these concepts. This concept is in line with previous research that demonstrated reduced opioid medication use in a group of patients who received automated telephone voice messages following an 11-week group pain management program [34]. This highlights the value of mHealth reminders in using prelearned skills. This perspective also aligns with research demonstrating that people are more likely to remember messages that are familiar and messages that they agree with [41,42]. People tend to ignore information they do not understand or that conflicts with their existing beliefs and attitudes [43]. Hence, engagement with informational messages may require users to first have foundational pain and opioid tapering education covering concepts such as “what is pain?,” “how is pain managed?,” “reasons for tapering opioids,” and “what to expect.” Although there is some evidence that SMS text messaging–based mHealth support alone may reduce pain intensity and opioid use after surgery [35], it is likely that a pre-educational video would be of benefit [33].

Participants also suggested that messages offering social or emotional support may be beneficial and expressed a need for encouragement and validation. This is consistent with research demonstrating that validation—having one’s feelings and experience acknowledged without judgment—significantly reduces distress, elevates positive mood, and increases pain tolerance [44]. It is theorized that by reducing the levels of physiological arousal, validation sets the stage for behavior change [45]. Hence, the inclusion of messages of validation and

support for emotional distress in the content of an mHealth intervention may indirectly help to facilitate behavior change.

Limitations

A key limitation of this study is that participants reported their *expectations* of the feasibility and acceptability of mHealth interventions to support prescription opioid tapering rather than their actual *experiences* with an mHealth intervention. Likewise, the experience of tapering that an individual participant had, both their previous experiences of tapering and the stage of tapering when they were interviewed, may have influenced their responses regarding perceptions toward mHealth support. It is possible that participants' expectations are not congruent with their experience [46], and future research will be needed to evaluate the experience of using mHealth to support prescription opioid tapering in people with CNCP. However, as a first step, the results of this study suggest that in our sample of people living with chronic pain who have been on long-term opioid therapy, there is an interest in piloting this mHealth intervention to support them in their efforts to taper prescription opioids.

We acknowledge that the interview guide contained prompt questions that were closed ended (eg, "Would you be interested in receiving SMS messages with informative and supportive content when you are reducing your opioid medication?"). Although interviewers were trained to follow up any closed-ended questions with an open-ended question, it is possible that closed-ended or leading questions may have biased patients' responses.

This study was conducted within the context of a larger study, which determined the sample size of this study. Certainly, the results of our descriptive (quantitative) survey would be strengthened with further sampling. However, the goal of qualitative data collection was to canvas a variety of perspectives on mHealth support for opioid tapering rather than to evaluate the proportion of people living with chronic pain who share certain perspectives. For this purpose, our sample size was adequate.

This study was conducted before the COVID-19 pandemic in Australia. As a result of the pandemic and associated social distancing measures, Australians have become more familiar with digital health technologies (ie, telehealth). It is possible that in the interests of maintaining social distancing and leveraging some of the efficiencies of digitally delivered health care, Australians may be more willing to engage with digital health interventions now more than ever.

Future Directions

The results of this study will be used to inform the development of a user-centered mHealth intervention to support people who are tapering prescription opioids for CNCP. Consistent with the existing guidelines for the development of mHealth interventions, the next stage of our research will involve co-designing and testing of intervention content (messages) in partnership with clinical experts as well as end users [38]. [Table 3](#) summarizes the key findings regarding barriers to and facilitators of patients' engagement and suggests implications of the design, content, and functionality of an mHealth intervention to support prescription opioid tapering.

Table 3. Summary of perceived barriers to and facilitators of patients' engagement with implications for mobile health design, functionality, and content.

Theme	Implication	Recommendation
Barriers to engagement		
Internet access	Functionality	SMS text messaging–based intervention may be more feasible than app-based intervention for this population
Access to smartphone technology	Functionality	SMS text messaging–based intervention may be more feasible than app-based intervention for this population
Digital literacy or confidence in using SMS text messaging versus apps	Design and functionality	SMS text messaging–based intervention may be more feasible than app-based intervention for this population
Vision impairment	Design	SMS text messaging–based intervention may be more feasible than app-based intervention for this population
Reservation about app use	Design	SMS text messaging–based intervention may be more feasible than app-based intervention for this population
Facilitator of engagement		
Maintaining contact with health care provider	Functionality	Communicate to users that the purpose of the intervention is to provide support in addition to that delivered by the health care provider
Fluctuating dose (support) needs	Design	User controls frequency of messaging after an initial <i>standardized</i> dose
Fluctuating content needs	Functionality	User can request support with specific issues by texting keywords to the server (eg, “crave,” “anxious,” or “pain”)
Familiarity with pain management strategies	Design	Standardized pre-education may be needed to ensure that patients have a basic understanding of pain management and reasons for opioid tapering before they receive supportive messages
Desire for individualized care	Functionality and content	Personalization of SMS text messaging content. Use of name and message tailored to demographic details of user
Predictability	Functionality	Variability in the time of day the messages are sent can increase attention and engagement
Socioemotional reassurance	Content	Validating message content that aims to normalize concerns that participants may have in their tapering (eg, “It’s natural to worry about pain increasing”)
Informational reassurance	Content	Informational message content to educate on nonopioid pain management (eg, “opioids are proven to relieve acute pain but not chronic pain”)

Conclusions

The results of this study indicate that patients with CNCP expressed interest in engaging with an mHealth intervention to support prescription opioid tapering. Interviews with this diverse group of potential users revealed barriers to feasibility and

acceptability to be addressed and offered insights into factors that may increase engagement with this mHealth intervention. Future research will evaluate whether mHealth interventions improve patients' confidence in their ability to manage their pain and maintain their quality of life while tapering opioid medications.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Survey and semistructured interview guide.

[[DOCX File, 21 KB - formative_v5i5e25969_app1.docx](#)]

References

1. Blyth FM, March LM, Brnabic AJ, Jorm LR, Williamson M, Cousins MJ. Chronic pain in Australia: a prevalence study. *Pain* 2001 Jan;89(2-3):127-134. [doi: [10.1016/s0304-3959\(00\)00355-9](https://doi.org/10.1016/s0304-3959(00)00355-9)] [Medline: [11166468](https://pubmed.ncbi.nlm.nih.gov/11166468/)]
2. Dahlhamer J, Lucas J, Zelaya C, Nahin R, Mackey S, DeBar L, et al. Prevalence of chronic pain and high-impact chronic pain among adults - United States, 2016. *MMWR Morb Mortal Wkly Rep* 2018 Sep 14;67(36):1001-1006 [[FREE Full text](#)] [doi: [10.15585/mmwr.mm6736a2](https://doi.org/10.15585/mmwr.mm6736a2)] [Medline: [30212442](https://pubmed.ncbi.nlm.nih.gov/30212442/)]

3. Häuser W, Wolfe F, Henningsen P, Schmutzer G, Brähler E, Hinz A. Untying chronic pain: prevalence and societal burden of chronic pain stages in the general population - a cross-sectional survey. *BMC Public Health* 2014 May 13;14:352 [FREE Full text] [doi: [10.1186/1471-2458-14-352](https://doi.org/10.1186/1471-2458-14-352)] [Medline: [24725286](https://pubmed.ncbi.nlm.nih.gov/24725286/)]
4. Blyth FM, Schneider CH. Global burden of pain and global pain policy-creating a purposeful body of evidence. *Pain* 2018 Sep;159 Suppl 1:43-48. [doi: [10.1097/j.pain.0000000000001311](https://doi.org/10.1097/j.pain.0000000000001311)] [Medline: [30113946](https://pubmed.ncbi.nlm.nih.gov/30113946/)]
5. Dowell D, Haegerich TM, Chou R. CDC guideline for prescribing opioids for chronic pain--United States, 2016. *J Am Med Assoc* 2016 Apr 19;315(15):1624-1645. [doi: [10.1001/jama.2016.1464](https://doi.org/10.1001/jama.2016.1464)] [Medline: [26977696](https://pubmed.ncbi.nlm.nih.gov/26977696/)]
6. Preventing and managing problems with opioid prescribing for chronic non-cancer pain. NSW Therapeutic Advisory Group Inc. 2015. URL: <https://www.nswtag.org.au/wp-content/uploads/2017/07/pain-guidance-july-2015.pdf> [accessed 2021-05-08]
7. Sandhu H, Underwood M, Furlan AD, Noyes J, Eldabe S. What interventions are effective to taper opioids in patients with chronic pain? *Br Med J* 2018 Sep 27;362:k2990. [doi: [10.1136/bmj.k2990](https://doi.org/10.1136/bmj.k2990)] [Medline: [30262590](https://pubmed.ncbi.nlm.nih.gov/30262590/)]
8. Kaye AD, Jones MR, Kaye AM, Ripoll JG, Galan V, Beakley BD, et al. Prescription opioid abuse in chronic pain: an updated review of opioid abuse predictors and strategies to curb opioid abuse: Part 1. *Pain Physician* 2017 Feb;20(2S):93-109 [FREE Full text] [Medline: [28226333](https://pubmed.ncbi.nlm.nih.gov/28226333/)]
9. Frank JW, Lovejoy TI, Becker WC, Morasco BJ, Koenig CJ, Hoeffcker L, et al. Patient outcomes in dose reduction or discontinuation of long-term opioid therapy: a systematic review. *Ann Intern Med* 2017 Aug 01;167(3):181-191 [FREE Full text] [doi: [10.7326/M17-0598](https://doi.org/10.7326/M17-0598)] [Medline: [28715848](https://pubmed.ncbi.nlm.nih.gov/28715848/)]
10. Henry SG, Paterniti DA, Feng B, Iosif A, Kravitz RL, Weinberg G, et al. Patients' experience with opioid tapering: a conceptual model with recommendations for clinicians. *J Pain* 2019 Feb;20(2):181-191 [FREE Full text] [doi: [10.1016/j.jpain.2018.09.001](https://doi.org/10.1016/j.jpain.2018.09.001)] [Medline: [30243859](https://pubmed.ncbi.nlm.nih.gov/30243859/)]
11. Goesling J, DeJonckheere M, Pierce J, Williams DA, Brummett CM, Hassett AL, et al. Opioid cessation and chronic pain: perspectives of former opioid users. *Pain* 2019 May;160(5):1131-1145. [doi: [10.1097/j.pain.0000000000001493](https://doi.org/10.1097/j.pain.0000000000001493)] [Medline: [30889052](https://pubmed.ncbi.nlm.nih.gov/30889052/)]
12. Berna C, Kulich RJ, Rathmell JP. Tapering long-term opioid therapy in chronic noncancer pain: evidence and recommendations for everyday practice. *Mayo Clin Proc* 2015 Jun;90(6):828-842. [doi: [10.1016/j.mayocp.2015.04.003](https://doi.org/10.1016/j.mayocp.2015.04.003)] [Medline: [26046416](https://pubmed.ncbi.nlm.nih.gov/26046416/)]
13. Demidenko MI, Dobscha SK, Morasco BJ, Meath TH, Ilgen MA, Lovejoy TI. Suicidal ideation and suicidal self-directed violence following clinician-initiated prescription opioid discontinuation among long-term opioid users. *Gen Hosp Psychiatry* 2017 Jul;47:29-35. [doi: [10.1016/j.genhosppsych.2017.04.011](https://doi.org/10.1016/j.genhosppsych.2017.04.011)] [Medline: [28807135](https://pubmed.ncbi.nlm.nih.gov/28807135/)]
14. Recommendations regarding the use of opioid analgesics in patients with chronic non-cancer pain. Faculty of Pain Medicine of the Australian and New Zealand College of Anaesthetists. 2015. URL: <https://www.betterpainmanagement.com/catalog?pagename=Useful-Links> [accessed 2021-05-08]
15. Dowell D, Compton WM, Giroir BP. Patient-centered reduction or discontinuation of long-term opioid analgesics: the HHS guide for clinicians. *J Am Med Assoc* 2019 Oct 10:1-3. [doi: [10.1001/jama.2019.16409](https://doi.org/10.1001/jama.2019.16409)] [Medline: [31600366](https://pubmed.ncbi.nlm.nih.gov/31600366/)]
16. Darnall BD, Ziadni MS, Stieg RL, Mackey IG, Kao M, Flood P. Patient-centered prescription opioid tapering in community outpatients with chronic pain. *JAMA Intern Med* 2018 May 01;178(5):707-708 [FREE Full text] [doi: [10.1001/jamainternmed.2017.8709](https://doi.org/10.1001/jamainternmed.2017.8709)] [Medline: [29459978](https://pubmed.ncbi.nlm.nih.gov/29459978/)]
17. Darnall BD, Mackey SC, Lorig K, Kao M, Mardian A, Stieg R, et al. Comparative effectiveness of cognitive behavioral therapy for chronic pain and chronic pain self-management within the context of voluntary patient-centered prescription opioid tapering: the EMPOWER study protocol. *Pain Med* 2020 Aug 01;21(8):1523-1531 [FREE Full text] [doi: [10.1093/pm/pnz285](https://doi.org/10.1093/pm/pnz285)] [Medline: [31876947](https://pubmed.ncbi.nlm.nih.gov/31876947/)]
18. Nicholas MK, Asghari A, Sharpe L, Beeston L, Brooker C, Glare P, et al. Reducing the use of opioids by patients with chronic pain: an effectiveness study with long-term follow-up. *Pain* 2020 Mar;161(3):509-519. [doi: [10.1097/j.pain.0000000000001763](https://doi.org/10.1097/j.pain.0000000000001763)] [Medline: [31764391](https://pubmed.ncbi.nlm.nih.gov/31764391/)]
19. Nicola M, Correia H, Ditchburn G, Drummond P. Invalidation of chronic pain: a thematic analysis of pain narratives. *Disabil Rehabil* 2021 Mar;43(6):861-869. [doi: [10.1080/09638288.2019.1636888](https://doi.org/10.1080/09638288.2019.1636888)] [Medline: [31290347](https://pubmed.ncbi.nlm.nih.gov/31290347/)]
20. Ashton-James CE, Glare P, Darnall BD. Out of office hours: scalable, on-demand, digital support for patients tapering prescription opioids. *Pain* 2020 Oct;161(10):2252-2254. [doi: [10.1097/j.pain.0000000000001947](https://doi.org/10.1097/j.pain.0000000000001947)] [Medline: [32501860](https://pubmed.ncbi.nlm.nih.gov/32501860/)]
21. Head KJ, Noar SM, Iannarino NT, Grant HN. Efficacy of text messaging-based interventions for health promotion: a meta-analysis. *Soc Sci Med* 2013 Nov;97:41-48. [doi: [10.1016/j.socscimed.2013.08.003](https://doi.org/10.1016/j.socscimed.2013.08.003)] [Medline: [24161087](https://pubmed.ncbi.nlm.nih.gov/24161087/)]
22. Jones KR, Lekhak N, Kaewluang N. Using mobile phones and short message service to deliver self-management interventions for chronic conditions: a meta-review. *Worldviews Evid Based Nurs* 2014 Apr;11(2):81-88. [doi: [10.1111/wvn.12030](https://doi.org/10.1111/wvn.12030)] [Medline: [24597522](https://pubmed.ncbi.nlm.nih.gov/24597522/)]
23. Orr JA, King RJ. Mobile phone SMS messages can enhance healthy behaviour: a meta-analysis of randomised controlled trials. *Health Psychol Rev* 2015 Dec;9(4):397-416. [doi: [10.1080/17437199.2015.1022847](https://doi.org/10.1080/17437199.2015.1022847)] [Medline: [25739668](https://pubmed.ncbi.nlm.nih.gov/25739668/)]
24. Islam SM, Chow CK, Redfern J, Kok C, Rådholm K, Stepien S, et al. Effect of text messaging on depression in patients with coronary heart disease: a substudy analysis from the TEXT ME randomised controlled trial. *BMJ Open* 2019 Feb 20;9(2):e022637 [FREE Full text] [doi: [10.1136/bmjopen-2018-022637](https://doi.org/10.1136/bmjopen-2018-022637)] [Medline: [30787075](https://pubmed.ncbi.nlm.nih.gov/30787075/)]

25. Wang Y, Min J, Khuri J, Xue H, Xie B, Kaminsky LA, et al. Effectiveness of mobile health interventions on diabetes and obesity treatment and management: systematic review of systematic reviews. *JMIR Mhealth Uhealth* 2020 Apr 28;8(4):e15400 [FREE Full text] [doi: [10.2196/15400](https://doi.org/10.2196/15400)] [Medline: [32343253](https://pubmed.ncbi.nlm.nih.gov/32343253/)]
26. Spohr SA, Nandy R, Gandhiraj D, Vemulapalli A, Anne S, Walters ST. Efficacy of SMS text message interventions for smoking cessation: a meta-analysis. *J Subst Abuse Treat* 2015 Sep;56:1-10. [doi: [10.1016/j.jsat.2015.01.011](https://doi.org/10.1016/j.jsat.2015.01.011)] [Medline: [25720333](https://pubmed.ncbi.nlm.nih.gov/25720333/)]
27. Haskins BL, Lesperance D, Gibbons P, Boudreaux ED. A systematic review of smartphone applications for smoking cessation. *Transl Behav Med* 2017 Jun;7(2):292-299 [FREE Full text] [doi: [10.1007/s13142-017-0492-2](https://doi.org/10.1007/s13142-017-0492-2)] [Medline: [28527027](https://pubmed.ncbi.nlm.nih.gov/28527027/)]
28. Noonan D, Silva S, Fish LJ, Peter K, Conley C, Simmons LA, et al. Feasibility of a text-based reduction intervention in helping rural and underserved smokeless tobacco users quit. *Addict Behav* 2020 Sep;108:106434. [doi: [10.1016/j.addbeh.2020.106434](https://doi.org/10.1016/j.addbeh.2020.106434)] [Medline: [32361367](https://pubmed.ncbi.nlm.nih.gov/32361367/)]
29. Devan H, Farmery D, Peebles L, Grainger R. Evaluation of self-management support functions in apps for people with persistent pain: systematic review. *JMIR Mhealth Uhealth* 2019 Feb 12;7(2):e13080 [FREE Full text] [doi: [10.2196/13080](https://doi.org/10.2196/13080)] [Medline: [30747715](https://pubmed.ncbi.nlm.nih.gov/30747715/)]
30. Fritsch CG, Ferreira PH, Prior JL, McLachlan AJ, Ferreira ML. Effects of using text message interventions for the management of musculoskeletal pain: a systematic review. *Pain* 2020 Nov;161(11):2462-2475. [doi: [10.1097/j.pain.0000000000001958](https://doi.org/10.1097/j.pain.0000000000001958)] [Medline: [32569088](https://pubmed.ncbi.nlm.nih.gov/32569088/)]
31. Kristjánsdóttir OB, Fors EA, Eide E, Finset A, Stensrud TL, van Dulmen S, et al. A smartphone-based intervention with diaries and therapist-feedback to reduce catastrophizing and increase functioning in women with chronic widespread pain: randomized controlled trial. *J Med Internet Res* 2013;15(1):e5 [FREE Full text] [doi: [10.2196/jmir.2249](https://doi.org/10.2196/jmir.2249)] [Medline: [23291270](https://pubmed.ncbi.nlm.nih.gov/23291270/)]
32. Gamwell KL, Kollin SR, Gibler RC, Bedree H, Bieniak KH, Jagpal A, et al. Systematic evaluation of commercially available pain management apps examining behavior change techniques. *Pain* 2021 Mar 01;162(3):856-865. [doi: [10.1097/j.pain.0000000000002090](https://doi.org/10.1097/j.pain.0000000000002090)] [Medline: [33003110](https://pubmed.ncbi.nlm.nih.gov/33003110/)]
33. Feng B, Malloch YZ, Kravitz RL, Verba S, Iosif A, Slavik G, et al. Assessing the effectiveness of a narrative-based patient education video for promoting opioid tapering. *Patient Educ Couns* 2021 Mar;104(2):329-336. [doi: [10.1016/j.pec.2020.08.019](https://doi.org/10.1016/j.pec.2020.08.019)] [Medline: [32900605](https://pubmed.ncbi.nlm.nih.gov/32900605/)]
34. Naylor MR, Naud S, Keefe FJ, Helzer JE. Therapeutic Interactive Voice Response (TIVR) to reduce analgesic medication use for chronic pain management. *J Pain* 2010 Dec;11(12):1410-1419 [FREE Full text] [doi: [10.1016/j.jpain.2010.03.019](https://doi.org/10.1016/j.jpain.2010.03.019)] [Medline: [20620119](https://pubmed.ncbi.nlm.nih.gov/20620119/)]
35. Anthony CA, Rojas EO, Keffala V, Glass NA, Shah AS, Miller BJ, et al. Acceptance and commitment therapy delivered via a mobile phone messaging robot to decrease postoperative opioid use in patients with orthopedic trauma: randomized controlled trial. *J Med Internet Res* 2020 Jul 29;22(7):e17750 [FREE Full text] [doi: [10.2196/17750](https://doi.org/10.2196/17750)] [Medline: [32723723](https://pubmed.ncbi.nlm.nih.gov/32723723/)]
36. Bohnert AS, Ilgen MA. Understanding links among opioid use, overdose, and suicide. *N Engl J Med* 2019 Jan 03;380(1):71-79. [doi: [10.1056/NEJMr1802148](https://doi.org/10.1056/NEJMr1802148)] [Medline: [30601750](https://pubmed.ncbi.nlm.nih.gov/30601750/)]
37. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol* 2006 Jan;3(2):77-101. [doi: [10.1191/1478088706qp0630a](https://doi.org/10.1191/1478088706qp0630a)]
38. Abrams LC, Whittaker R, Free C, Mendel VA, Schindler-Ruwisch JM. Developing and pretesting a text messaging program for health behavior change: recommended steps. *JMIR Mhealth Uhealth* 2015 Dec 21;3(4):e107 [FREE Full text] [doi: [10.2196/mhealth.4917](https://doi.org/10.2196/mhealth.4917)] [Medline: [26690917](https://pubmed.ncbi.nlm.nih.gov/26690917/)]
39. Pincus T, Holt N, Vogel S, Underwood M, Savage R, Walsh DA, et al. Cognitive and affective reassurance and patient outcomes in primary care: a systematic review. *Pain* 2013 Nov;154(11):2407-2416. [doi: [10.1016/j.pain.2013.07.019](https://doi.org/10.1016/j.pain.2013.07.019)] [Medline: [23872104](https://pubmed.ncbi.nlm.nih.gov/23872104/)]
40. Holt N, Pincus T, Vogel S. Reassurance during low back pain consultations with GPs: a qualitative study. *Br J Gen Pract* 2015 Oct;65(639):692-701 [FREE Full text] [doi: [10.3399/bjgp15X686953](https://doi.org/10.3399/bjgp15X686953)] [Medline: [26412846](https://pubmed.ncbi.nlm.nih.gov/26412846/)]
41. Frost P, Casey B, Griffin K, Raymundo L, Farrell C, Carrigan R. The influence of confirmation bias on memory and source monitoring. *J Gen Psychol* 2015;142(4):238-252. [doi: [10.1080/00221309.2015.1084987](https://doi.org/10.1080/00221309.2015.1084987)] [Medline: [26649923](https://pubmed.ncbi.nlm.nih.gov/26649923/)]
42. Rajsic J, Wilson DE, Pratt J. Confirmation bias in visual search. *J Exp Psychol Hum Percept Perform* 2015 Oct;41(5):1353-1364. [doi: [10.1037/xhp0000090](https://doi.org/10.1037/xhp0000090)] [Medline: [26098120](https://pubmed.ncbi.nlm.nih.gov/26098120/)]
43. Nickerson RS. Confirmation bias: a ubiquitous phenomenon in many guises. *Rev Gen Psychol* 1998 Jun 01;2(2):175-220. [doi: [10.1037/1089-2680.2.2.175](https://doi.org/10.1037/1089-2680.2.2.175)]
44. Linton SJ, Boersma K, Vangronsveld K, Fruzzetti A. Painfully reassuring? The effects of validation on emotions and adherence in a pain test. *Eur J Pain* 2012 Apr;16(4):592-599. [doi: [10.1016/j.ejpain.2011.07.011](https://doi.org/10.1016/j.ejpain.2011.07.011)] [Medline: [22396087](https://pubmed.ncbi.nlm.nih.gov/22396087/)]
45. Linton SJ, McCracken LM, Vlaeyen JW. Reassurance: help or hinder in the treatment of pain. *Pain* 2008 Jan;134(1-2):5-8. [doi: [10.1016/j.pain.2007.10.002](https://doi.org/10.1016/j.pain.2007.10.002)] [Medline: [18035496](https://pubmed.ncbi.nlm.nih.gov/18035496/)]
46. Wilson TD. *Strangers to Ourselves – Discovering the Adaptive Unconscious*. Cambridge, United States: Harvard University Press; 2004:1-272.

Abbreviations**CNCP:** chronic noncancer pain**mHealth:** mobile health

Edited by G Eysenbach; submitted 23.11.20; peer-reviewed by H De Loof, P Cook, E Omaki; comments to author 05.02.21; revised version received 29.03.21; accepted 06.04.21; published 18.05.21.

*Please cite as:**Magee MR, McNeilage AG, Avery N, Glare P, Ashton-James CE**mHealth Interventions to Support Prescription Opioid Tapering in Patients With Chronic Pain: Qualitative Study of Patients' Perspectives**JMIR Form Res 2021;5(5):e25969*URL: <https://formative.jmir.org/2021/5/e25969>doi: [10.2196/25969](https://doi.org/10.2196/25969)PMID: [34003133](https://pubmed.ncbi.nlm.nih.gov/34003133/)

©Michael Reece Magee, Amy Gray McNeilage, Nicholas Avery, Paul Glare, Claire Elizabeth Ashton-James. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 18.05.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Optimization of Patient Flow in Urgent Care Centers Using a Digital Tool for Recording Patient Symptoms and History: Simulation Study

Maryam Montazeri¹, PhD; Jan Multmeier¹, PhD; Claire Novoro¹, MD, PhD; Shubhanan Upadhyay¹, MD; Paul Wicks¹, PhD; Stephen Gilbert¹, PhD

Ada Health GmbH, Berlin, Germany

Corresponding Author:

Stephen Gilbert, PhD

Ada Health GmbH

Karl-Liebknecht-Str 1

Berlin, 10178

Germany

Phone: 49 3040367390

Email: science@ada.com

Abstract

Background: Crowding can negatively affect patient and staff experience, and consequently the performance of health care facilities. Crowding can potentially be eased through streamlining and the reduction of duplication in patient history-taking through the use of a digital symptom-taking app.

Objective: We simulated the introduction of a digital symptom-taking app on patient flow. We hypothesized that waiting times and crowding in an urgent care center (UCC) could be reduced, and that this would be more efficient than simply adding more staff.

Methods: A discrete-event approach was used to simulate patient flow in a UCC during a 4-hour time frame. The baseline scenario was a small UCC with 2 triage nurses, 2 doctors, 1 treatment/examination nurse, and 1 discharge administrator in service. We simulated 33 scenarios with different staff numbers or different potential time savings through the app. We explored average queue length, waiting time, idle time, and staff utilization for each scenario.

Results: Discrete-event simulation showed that even a few minutes saved through patient app-based self-history recording during triage could result in significantly increased efficiency. A modest estimated time saving per patient of 2.5 minutes decreased the average patient wait time for triage by 26.17%, whereas a time saving of 5 minutes led to a 54.88% reduction in patient wait times. Alternatively, adding an additional triage nurse was less efficient, as the additional staff were only required at the busiest times.

Conclusions: Small time savings in the history-taking process have potential to result in substantial reductions in total patient waiting time for triage nurses, with likely effects of reduced patient anxiety, staff anxiety, and improved patient care. Patient self-history recording could be carried out at home or in the waiting room via a check-in kiosk or a portable tablet computer. This formative simulation study has potential to impact service provision and approaches to digitalization at scale.

(*JMIR Form Res* 2021;5(5):e26402) doi:[10.2196/26402](https://doi.org/10.2196/26402)

KEYWORDS

symptom assessment app; discrete event simulation; health care system; patient flow modeling; patient flow; simulation; urgent care; waiting times

Introduction

Background

Crowding in health care facilities occurs when the number of patients seeking care exceeds the care facility's capacity in a

given time period. Long queues of patients can lead to delayed care delivery, increased health risk for urgent cases, higher rates of hospital-borne infections, increased stress, and avoidable staff burden [1,2]. Crowding has also been associated with increased occurrence of preventable medical errors and with negative effects on clinical trial outcomes [3-5]. Many studies

have shown that crowding in emergency departments (EDs) lowers satisfaction of patients [6], increases stress on staff, leads to less adherence of staff to guidelines, leads to less rapport between patients and health care professionals, and ultimately to a less “soft” interaction between patients and health care professionals [7]. Health care system performance can be measured in terms of patients’ waiting time and quality of the service, among other variables such as cost [8]. One method that can help analyze the performance of the whole system is patient flow modeling, which can aid decision-making in planning capacity, resources, and appointment scheduling [9].

Methods to improve the flow of health care delivery include eliminating unnecessary and duplicate activities, performing activities in parallel, and identifying alternative process flows [9]. History-taking and recording of patients’ symptoms by skilled health care professionals are often duplicated activities during triage and treatment in both urgent care centers (UCCs) and EDs [10].

Redundancy in data capture has been reported to reduce the quality of patient care [11], and a resultant practice recommendation was to take steps to resolve this issue. Similarly, a clinical study of randomly selected practices found that repetitive clinical notes can hinder coordination of patient care between health care professionals [12].

In the ED setting, a waiting room–based patient self-symptom and history-taking app (Ada Health, Germany) facilitated patient-to-health care professional communication, and triage nurses perceived this app as also having a workflow benefit by saving time [13]. The tool uses a probabilistic reasoning engine to collect demographic information, medical history, and symptoms. In a previous usability study, 97.8% (511/522) of patients found the symptom-taking system easy to use in the primary care waiting room [14]. A clinical vignette study showed that the system’s reasoning engine has similar levels of coverage, accuracy, and safety as human general practitioners [15], which is important for gathering comprehensive primary care histories. Symptom-taking and assessment tools from other providers have been judged by patients to provide useful diagnostic advice and to be easy to use [16,17]. However, the potential workflow benefits that might be experienced by using this tool in a more urgent setting remain unclear.

The term UCC can refer to several types of services, including walk-in centers, minor injury units, and urgent treatment centers, all with different levels of service [18-21]. As modeled in this study, a typical UCC is led by a physician (general practitioner), is open every day of the week, and is equipped to diagnose and treat common ailments. In the United Kingdom, this type of unit is known as an “urgent treatment center” [22]. Most prior research on triage, waiting, and consultation time distributions

has been carried out in primary health clinics [23-25] or the ED [26-29]; thus, relatively little, with the exception of one study [30], has been reported for UCCs.

System Simulation for Workflow Efficiency

We used a system simulation approach to understand the potential UCC flow and efficiency effects of a patient self-symptom and history-taking app. We specifically tested the hypothesis that waiting times and crowding in a UCC could be reduced through the introduction of a digital history-taking tool, and that system efficiency would be greater with use of the tool than through the addition of staff.

Methods

Simulation Development

We compared a scenario in which there was no patient self-system and history-taking tool to a scenario in which every patient entering the UCC waiting room had used the tool. Patient usage of the tool could be either: (i) at home (using a webpage or phone app); (ii) using check-in kiosks in a colocated ED waiting room, before fast-track redirection to the associated UCC; or (iii) using check-in kiosks at the UCC. In each case, it was modeled that the assessment report’s questions, answers, demographics, and symptoms would be transferred to the UCC’s electronic medical record system.

Parameter Development: Clinical Setting

We simulated a typical UCC in the first 4 hours of its opening. At the start of the patient journey (see Figure 1), a triage nurse assesses the symptoms of the patient. The patient then visits the doctor and then either visits the examination/treatment room (with probability λ) or is discharged (with probability $1 - \lambda$). If a patient visits the examination/treatment room, they are either redirected to the doctor for further investigations (with probability ω) or discharged (with probability $1 - \omega$). Triage duration, consultation duration, number of staff in service, and arrival rate of the patients all affect the patient flow in the UCC. The baseline scenario of staffing of the UCC was taken from previous reports [18-21], and professional experience of one author (SU) and another colleague (Adel Baluch, Medical Director, Ada Health GmbH) who have each worked for over 5 years in National Health Service general practices, UCCs, and EDs. We assumed that there were two triage nurses, two doctors, one nurse for examination/treatment, and one administrator responsible for discharge (Table 1). We simulated the effects of the history-taking tool on queue size, waiting time for triage nurses, idle time, and utilization of triage nurses and doctors. Waiting times were modeled based on data collected from the ED setting [26,27].

Figure 1. Illustration of the urgent care center, where patients arrive without any planned appointment. In the first step, a triage nurse runs a symptom assessment, and then patients are directed to the doctor. Depending on their situation, they may be examined/treated by another nurse and then discharged, sent back to a doctor, or discharged immediately by administrative staff.

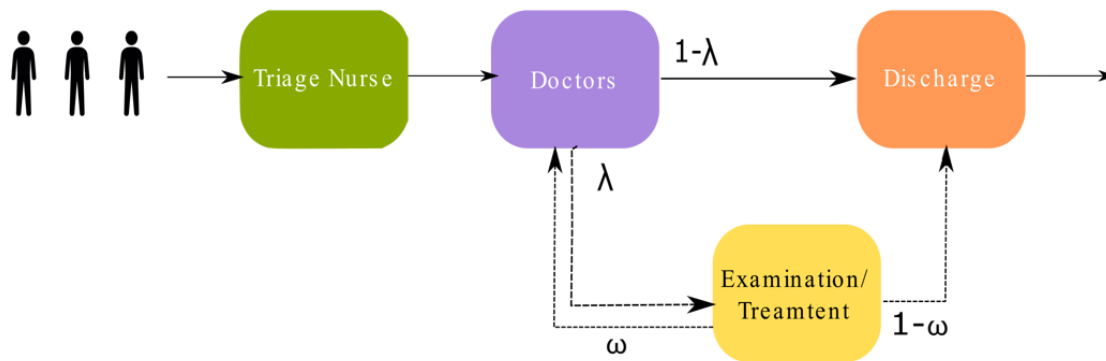


Table 1. Baseline settings for simulation of the urgent care center.

Baseline setting	Triage nurse	Doctor	Examination/treatment nurse	Administration staff
Average duration of interaction with patient (minutes)	15	20	15	5
Number in service	2	2	1	1

Parameter Development: Time Savings

Our model required a parameter for how much time could be saved through digital history-taking. A 2017 pilot implementation of a symptom app assessment in a busy (10,000 patients) UK primary care practice saved an estimated 1.9 minutes, as reported by doctors from over 300 primary care consultations (personal communication of unpublished pilot report, Dr. Vishal Virani, Business Development & Client Success Director, Ada Health GmbH). A 2019 pilot involving structured interviews with 5 ED clinicians who viewed the handover report produced by the app estimated a time saving between 4 and 6 minutes (personal communication by email of unpublished product development data, Joseph Wolanski, Ada Health GmbH). Finally, in an observational study, time savings in the ED were estimated in the range of 2.5-5 minutes by triage nurses and physicians [13]. Based on these data, a range of potential time savings were modeled in this study.

Setting Model Parameters: Crowding

First, we simulated the flow with different arrival rates to cause crowding, defined as more than 5 patients waiting for staff. We simulated the patient trajectory starting with an arrival rate of 0.1 patients per minute. Solutions were found to reach stability after 5000 simulations.

An arrival rate of 0.2 patients per minute (ie, one new patient every 5 minutes) was used, as described in further detail in [Multimedia Appendix 1](#).

To explore our hypothesis that crowding can be reduced through the addition of a digital tool, we simulated *what-if* scenarios. Alongside this, we varied staffing from the baseline settings, as our hypothesis recognized that crowding can also likely be reduced by provision of more staff. We measured queue status, waiting time for the triage nurse, idle time, and utilization of

triage nurses and doctors. Waiting time was defined as the interval between patient readiness for nurse triage and the end of the triage consultation, excluding the consultation duration. Idle time was defined as the period when one or more health care professionals is unoccupied. Utilization was defined as the ratio of the time the health care professionals are occupied to the total simulated time. Based on the previous study [13], we used a range of time savings by the app for triage (2.5, 3, 3.5, 4, 4.5, and 5 minutes) and for consultation (1.5, 2, 2.5, 3, and 3.5 minutes) to parameterize the model. We simulated 33 scenarios, including the baseline setting.

Statistical Analysis

We used the package *Simmer* (version 4.4.0) [31], a process-oriented and trajectory-based discrete-event simulation (DES) package for R. Measures are reported after 5000 simulation runs as the overall mean and 95% CI, with the exception of utilization, which is reported as the median and IQR (as is standard in the DES package). The baseline case scenario was a UCC staffed with two triage nurses, two doctors, one treatment nurse, and one administrator responsible for discharge. We assumed that patient arrivals, triage, consultation, and discharge (all events in the patient flow through the UCC) follow a Poisson distribution and therefore the time interval distribution between all events follow exponential distributions. This approach stochastically models the variable duration of each event, including the variable patient-to-patient time-saving potential of the symptom and history report.

Results

Effect of Additional Staff

[Table 2](#) shows how different staffing scenarios and use of the symptom and history-taking app could alter crowding.

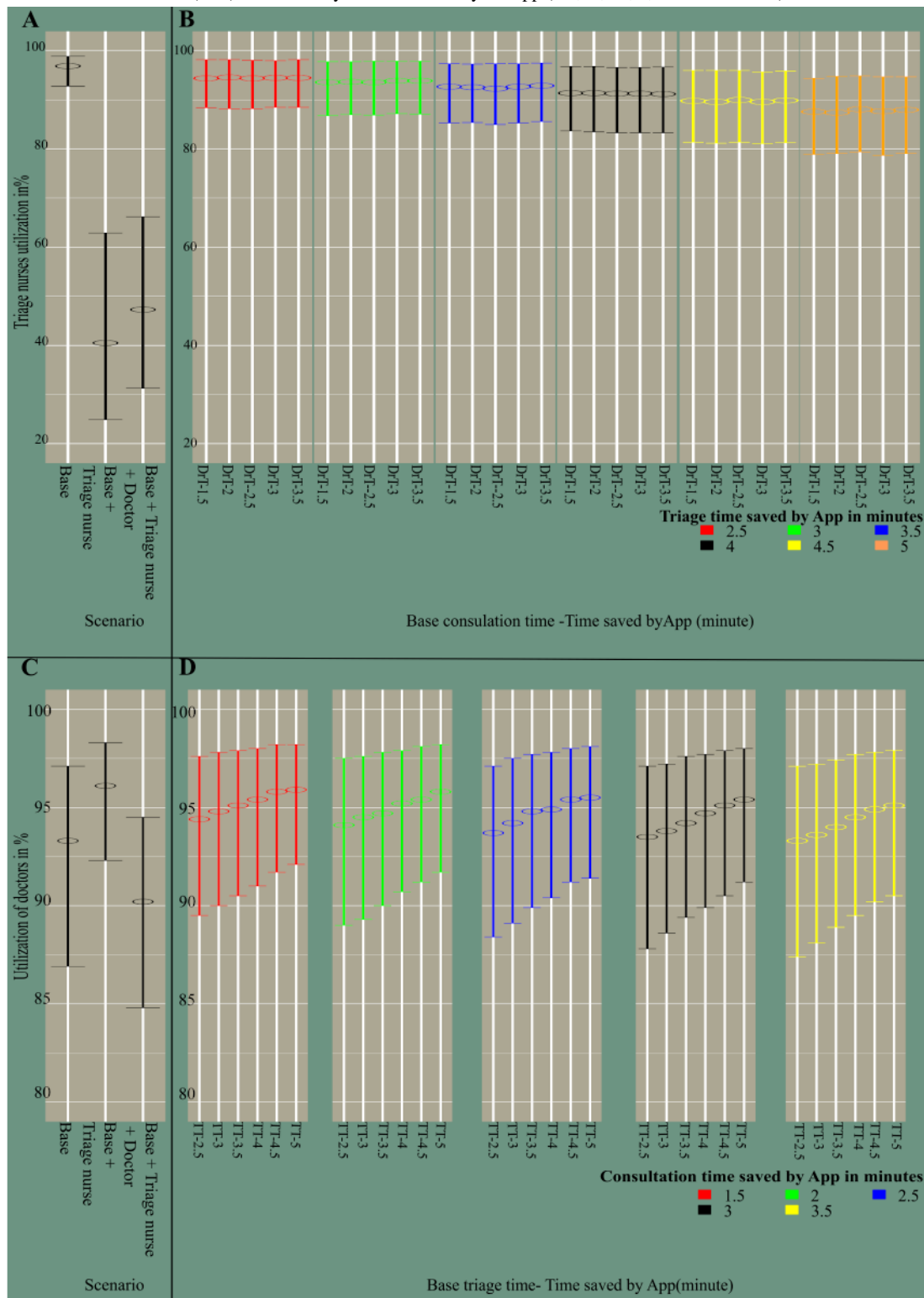
Table 2. Effect of adding extra staff or using a digital symptom and history-taking app on queue sizes, idle time, and utilization of staff members, and patient waiting time for the triage nurse.

Scenario	Queue size for triage nurses (number of patients), mean (95% CI)	Queue size for doctors (number of patients), mean (95% CI)	Idle time of triage nurses (minutes), mean (95% CI)	Idle time of doctors (minutes), mean (95% CI)	Utilization of triage nurses (%), median (IQR)	Utilization of doctors (%), median (IQR)	Waiting time for triage nurses (minutes), mean (95% CI)
Baseline	8.47 (8.44-8.49)	5.44 (5.42-5.46)	13.99 (13.50-4.50)	24.10 (23.40-24.80)	96.9 (92.8-98.9)	93.3 (86.9-97.1)	34.05 (33.90-34.21)
Baseline + triage nurse	3.4 (3.37-3.39)	9.53 (9.51-9.56)	61.04 (59.86-62.22)	13.43 (13.00-13.86)	40.5 (24.9-62.9)	96.1 (93.3-98.3)	13.2 (13.13-13.28)
Baseline + triage nurse + doctor	3.47 (3.46-3.48)	5.57 (5.55-5.58)	59.78 (58.63-60.94)	40.34 (39.55-41.13)	47.3 (31.3-66.2)	90.2 (84.8-94.5)	13.54 (13.46-13.62)
Baseline + digital tool (assuming minimum time saving)	6.29 (6.27-6.31)	6.82 (6.80-6.84)	22.74 (22.05-23.43)	19.32 (18.72-19.92)	94.4 (88.4-98.2)	94.4 (89.5-97.6)	25.44 (25.32-25.56)
Baseline + digital tool (assuming maximum time saving)	3.84 (3.82-3.85)	8.2 (8.17-8.22)	41.73 (40.80-42.6)	17.04 (16.54-17.53)	88.0 (79.1-94.7)	95.1 (90.5-97.9)	15.5 (15.48-15.64)

Addition of an extra nurse reduced the queue length for triage nurses by around 60% but led to an approximately 75% increase in the queue length for the doctors. Providing one additional doctor could reduce the number of patients waiting for doctors to a similar situation as the baseline case (see Figure S1 in [Multimedia Appendix 1](#)). Adding one extra triage nurse resulted in a 336% increase of triage nurses' idle time and a 44% decrease of the doctor's idle time as more patients would be transferred to consultation in a shorter time (see Figure S2 in

[Multimedia Appendix 1](#)). Adding one extra doctor led to a 67% increase of the mean idle time of doctors (Figure S2 in [Multimedia Appendix 1](#)). In the baseline case, the median triage nurses' utilization was 96.9% and adding one extra triage nurse reduced this value to 40.5% (Figure 2). The median utilization of doctors was consistently maintained at a level of 90% or above. Adding a triage nurse led to a 61.23% reduction in the average waiting time for triage nurses (see Figure S3 in [Multimedia Appendix 1](#)).

Figure 2. Utilization of triage nurses (A and B) and doctors (C and D) in percentage. A and C represent the idle time of the staff in scenarios where no app is used, which is the baseline case setting, and in scenarios with extra staff. B and D indicate the idle time of triage nurses and doctors in 30 different scenarios combining different time savings by the app in the triage and consultation processes. In B, the X-axis labels show triage time as the baseline triage time (TT) subtracted by time saved by the app (2.5, 3, 3.5, 4, 4.5, and 5 minutes). In D, the X-axis labels show the consultation time as the baseline doctor’s consultation time (DrT) subtracted by the time saved by the app (1.5, 2, 2.5, 3, and 3.5 minutes).



Effect of the Symptom and History-Taking App

Figure S1 in Multimedia Appendix 1 shows the impact of applying the symptom and history-taking app on the queue sizes of triage nurses and doctors in comparison to the baseline setting and the addition of extra staff. For the scenario where the time

saved per patient by the app was modeled as 5 minutes, the time-saving impact of the app was equivalent to adding one triage nurse. Even when the app consultation time saving was modeled as 2.5 minutes per patient, this reduced patient queue length for triage by 25.73%. However, when nurse-led triage took less time per patient, the rate of flow to doctors increased,

with a consequent increase in doctors' queue size (see Table S1 in [Multimedia Appendix 2](#)).

Idle and Utilization Time of Triage Nurses and Doctors

Longer triage times and shorter consultation times led to longer idle times for the doctors (Figure S2 in [Multimedia Appendix 1](#)). Assuming maximum app time saving, the triage nurses' idle time almost doubled, whereas the doctors' idle time was reduced by 30%. By contrast, assuming minimum app time saving, the average triage nurses' idle time increased by about 62% and the average doctors' idle time reduced by less than 20% (Table 2 and Table S2 in [Multimedia Appendix 2](#)). Median triage nurses' utilization dropped only moderately, by about 9%, when modeling maximum app time savings. For minimum modeled app time saving, this drop in utilization was only 2.5%. Conversely, the median utilization of doctors was less dependent on the amount of time saved and was always above 93% (Figure 2, Table 2, Table S3 in [Multimedia Appendix 2](#)).

Waiting Time for Triage Nurses

The more time saved by the app, the less time the patient needed to wait for a triage nurse (Figure S3 in [Multimedia Appendix 1](#)). When maximum app time saving was modeled, waiting time for a triage nurse dropped by 54.88%. (Table 2 and Table S4 in [Multimedia Appendix 2](#)). When minimal app time saving was modeled, the waiting time for triage dropped by 25.28%.

Discussion

Principal Findings

We simulated patient flow of a UCC in three conditions: (a) baseline, (b) with extra staff, and (c) with a digital symptom-taking tool. The shortest queue size and waiting time for triage nurses were achieved with the provision of one extra triage nurse (ie, a total of three triage nurses) and one additional doctor (ie, a total of two doctors). However, this approach may not be feasible due to limitations of available staff and high costs. Therefore, we hypothesized that use of a symptom and history-taking app before the triage process could be another possible solution. These apps have the potential to improve the patient flow in health care facilities such as hospitals, primary clinics, EDs, and UCCs [7], where a long queue of patients not only places substantial pressure on the health care workers but also on patients.

Our results suggest that for all measured variables, the amount of time saved by the app is an important determinant of the patient waiting time and system efficiency improvement. We found an amplification of time efficiency, through which relatively modest time savings per patient consultation accrued into substantial reduction in queuing time overall. The shortest modeled time saving from the app (2.5 minutes per patient) reduced the patient waiting time for triage by 25.28% and the longest time modeled from the app (5 minutes per patient) led to a 54.88% reduction in patient waiting time for a triage nurse.

Although crowding can also be resolved by additional staff, the simulation suggested that simply adding triage nurses may be inefficient as additional staff are only required at the busiest times. A digital symptom tool that could save 5 minutes per

patient led to a reduction in waiting time equivalent to employing one extra triage nurse. Adding a triage nurse would have lowered staff utilization from 88% to 40%.

Simulation for Improving Health System Efficiency

Simulation is an accepted and powerful method for hypothesis generation for the effects of new health care interventions on overall system efficiency. Simulation methods such as system dynamics, agent-based simulation, and discrete event simulation have gained substantial attention as helpful methods to tackle the complexities of analysis of patient flow in different areas. These applications include: (i) the detection of bottlenecks of the patient flow in health care facilities, (ii) optimizing flow management strategies such as scheduling and resource allocation rules, and (iii) estimating treatment cost in terms of the lengths of stay of patients [9,32,33]. Results of many simulation-based studies have already been implemented in real-world settings for better management of patient flow. One example evaluated scheduling, process flow, and resource levels in an oncology center [34], where the implementation of the changes proposed by the simulations resulted in improvement of the center's system-wide performance. Another example applied the techniques explored here to a military outpatient primary care clinic. Simulation revealed a hybrid appointment/walk-in model for improving patient flow and optimized care provider utilization [35]. A final example applied a simulation model to identify factors contributing to flow blockage in an outpatient clinic, and its application led to significant improvements in real-life patient waiting time and physician utilization [36].

Comparisons to the Wider Literature

One of the principal reasons that patients choose to go to a UCC is that they perceive waiting times to be lower than those experienced in general practitioner clinics or in the ED [37]. However, we were unable to identify any time-series studies that reported waiting times or other clinical processes in UCCs, and there has been little systematic data gathering on UCC clinical efficiency [18]. There is more substantial health service delivery and clinical efficiency research for the ED setting [38]; although time-series studies have been performed, the length of recording clinical history and symptoms, and how much time can be saved through digital history-taking tools have not been reported with certainty. We found no studies investigating the benefit or performance of self-assessment with a digital assessment tool in the UCC; however, some studies have reported the potential of self-triage for optimizing flow in subsections of EDs or in primary care units. Investigation of a bilingual self-triage kiosk in a pediatric ED showed that the system enabled parents to provide symptoms and history faster and more accurately than routine nurse-led triage [39].

Unanswered Questions and Future Research

It is widely recognized that many promising digital innovations in health care are ultimately not adopted in practice, or are abandoned soon after limited local pilot utilization [40]. Often, it is not the limitations of the technology or difficulties in implementation that ultimately determine the success of the pilots and wider adoption, but rather the dynamic interactions

between many of these factors [41]. This study explored the potential effects of a patient digital symptom and history-taking tool on patient flow and queuing, but did not explore the wider implications of the technology for the quality-of-care delivery, patient experience, patient safety, or the working experience of health care staff. These interlinked phenomena will be addressed in future studies.

ED crowding is mainly caused by patients who do not require urgent treatment [5] but whose medical history must be documented, accounting for approximately 41% of ED doctors' time [10]. Crowding also leads to interruptions, which impair history-taking and documentation, particularly for inexperienced junior physicians who are overstretched [5].

Future research (including simulation studies, clinical investigations, and technology rollouts) should seek to understand the potential of such tools in reducing documentation burden, facilitating fast tracking, increasing patient safety, improving documentation accuracy, and ultimately reducing crowding.

Strengths and Limitations

We used DES to simulate a queue of events. The choice of modeling technique, model structure, and parameter values limited the generalizability of the results as the nature of UCCs varies substantially [18]. In our model, we only considered a UCC without any planned appointments. We also assumed a first-in-first-out flow, irrespective of the urgency of treatment of individual patients. Patients and staff were all treated as passive, and we did not consider any ongoing learning that can influence patient and health worker interactions. We also assumed that there were enough digital devices available such that digital symptom assessment would not itself lead to another queue. We simulated 33 UCC setups, which were modified from a representative UCC baseline scenario (taken from experience and literature descriptions [18-21]) and cover a wide range of realistic UCC staffing scenarios. These 33 scenarios

provide a balance between the range of real-world scenario coverage and study complexity, and also provide a reasonable basis of extrapolation of results. As UCCs vary substantially in terms of their staffing, resources, business, and layout, future studies should build on our results and simulate several real UCCs and include actual time measurements to substantiate the parameters.

We modeled under the assumption that the time spent taking history leads to a time saving for both the triage nurse and for the treating physician. One example from the literature highlights the level of duplication in a typical ED setting [10], where a history was taken by: (i) the triage nurse, (ii) clerking (student) physician, (iii) second clerking on transfer to the acute medical unit, (iv) at history review in the general ward round, and (v) at history retaking on admission to a specialty ward. In the ED setting, the retaking of clinical history provides no clinical benefit, with the history often recorded nearly verbatim to the previous history, as part of a recognized "futile clinical cycle" [10]. However, we acknowledge that in some cases, the histories taken by the triage nurse and the treating doctor have different purposes. We assumed that the information queried and the time spent in both cases overlapped to a large degree. Finally, we assumed that nurses and doctors could assess the recorded symptoms within their standard workflow.

Conclusions

This simulation showed that even a small reduction in the time taken to assess symptoms can lead to a substantial reduction in the time patients wait for triage nurses, which could in turn lead to reduced patient anxiety, lower staff anxiety, and improved patient care. Compared to baseline, the use of a digital symptom-taking tool shortened the average patient waiting time to the same extent as adding an additional triage nurse to the UCC, with the additional advantages of higher staff efficiency. Such approaches have the potential to streamline service provision and accelerate approaches to digitalization in urgent care settings.

Acknowledgments

We are grateful to Adel Baluch (Medical Director, Ada Health GmbH) for advice on the normal structure, layout, and working practices of UCCs.

Authors' Contributions

MM, JM, CN, SU, and SG contributed to the planning (study conception and protocol development). MM carried out the simulation setup and simulations. MM, JM, and SG contributed to the data analysis and interpretation. MM, JM, CN, SU, PW, and SG contributed to manuscript writing. All authors contributed to commenting on drafts of the manuscript. SG is the guarantor and corresponding author of this work, and attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

Conflicts of Interest

MM, JM, SU, CN, and SG are employees or company directors of Ada Health GmbH, and some of these authors hold stock options in the company. PW has a consultancy contract with Ada Health GmbH. The Ada Health GmbH research team has received research grant funding from Fondation Botnar and the Bill & Melinda Gates Foundation. PW is an associate editor at Journal of Medical Internet Research, and is on the editorial advisory boards of BMJ, BMC Medicine, The Patient, and Digital Biomarkers. PW is employed by Wicks Digital Health Ltd, which has received funding from Ada Health, AstraZeneca, Baillie

Gifford, Bold Health, Camoni, Compass Pathways, Coronna, EIT, Happify, HealthUnlocked, Inbeoo, Kheiron Medical, Sano Genetics, Self Care Catalysts, The Learning Corp, The Wellcome Trust, VeraSci, and Woebot.

Multimedia Appendix 1

Supplementary methods and results.

[[DOCX File, 526 KB - formative_v5i5e26402_app1.docx](#)]

Multimedia Appendix 2

Triage nurses' and doctors' queue sizes in 33 simulated scenarios (Table S1), triage nurses' and doctors' idle time in 33 simulated scenarios (Table S2), utilization of triage nurses and doctors in % (Table S3), and waiting time for the nurses in minutes (Table S4).

[[XLSX File \(Microsoft Excel File\), 20 KB - formative_v5i5e26402_app2.xlsx](#)]

References

1. Derlet RW, Richards JR. Ten solutions for emergency department crowding. *West J Emerg Med* 2008 Jan;9(1):24-27 [[FREE Full text](#)] [Medline: [19561699](#)]
2. Zhao L, Lie B. Modeling and simulation of patient flow in hospitals for resource utilization. *Simul Notes Eur* 2010 Aug;20(2):41-50. [doi: [10.11128/sne.20.tn.09976](#)]
3. Epstein SK, Huckins DS, Liu SW, Pallin DJ, Sullivan AF, Lipton RI, et al. Emergency department crowding and risk of preventable medical errors. *Intern Emerg Med* 2012 Apr;7(2):173-180. [doi: [10.1007/s11739-011-0702-8](#)] [Medline: [22009553](#)]
4. Pines JM, Hilton JA, Weber EJ, Alkemade AJ, Al Shabanah H, Anderson PD, et al. International perspectives on emergency department crowding. *Acad Emerg Med* 2011 Dec;18(12):1358-1370. [doi: [10.1111/j.1553-2712.2011.01235.x](#)] [Medline: [22168200](#)]
5. Schmiedhofer M, Möckel M, Slagman A, Frick J, Ruhla S, Searle J. Patient motives behind low-acuity visits to the emergency department in Germany: a qualitative study comparing urban and rural sites. *BMJ Open* 2016 Nov 16;6(11):e013323 [[FREE Full text](#)] [doi: [10.1136/bmjopen-2016-013323](#)] [Medline: [27852722](#)]
6. Tekwani K, Kerem Y, Mistry C, Sayger B, Kulstad E. Emergency department crowding is associated with reduced satisfaction scores in patients discharged from the emergency department. *West J Emerg Med* 2013 Feb;14(1):11-15 [[FREE Full text](#)] [doi: [10.5811/westjem.2011.11.11456](#)] [Medline: [23447751](#)]
7. Lorenzetti DL, Quan H, Lucyk K, Cunningham C, Hennessy D, Jiang J, et al. Strategies for improving physician documentation in the emergency department: a systematic review. *BMC Emerg Med* 2018 Oct 25;18(1):36 [[FREE Full text](#)] [doi: [10.1186/s12873-018-0188-z](#)] [Medline: [30558573](#)]
8. Ashour OM, Okudan Kremer GE. Dynamic patient grouping and prioritization: a new approach to emergency department flow improvement. *Health Care Manag Sci* 2016 Jun;19(2):192-205. [doi: [10.1007/s10729-014-9311-1](#)] [Medline: [25487711](#)]
9. Bhattacharjee P, Ray PK. Patient flow modelling and performance analysis of healthcare delivery processes in hospitals: A review and reflections. *Comput Ind Eng* 2014 Dec;78:299-312. [doi: [10.1016/j.cie.2014.04.016](#)]
10. Sabin J, Khan W, Subbe CP, Franklin M, Abulela I, Khan A, et al. 'The time it takes...' How doctors spend their time admitting a patient during the acute medical take. *Clin Med (Lond)* 2016 Aug;16(4):320-324 [[FREE Full text](#)] [doi: [10.7861/clinmedicine.16-4-320](#)] [Medline: [27481373](#)]
11. Siegler JE, Patel NN, Dine CJ. Prioritizing paperwork over patient care: why can't we do both? *J Grad Med Educ* 2015 Mar;7(1):16-18 [[FREE Full text](#)] [doi: [10.4300/JGME-D-14-00494.1](#)] [Medline: [26217415](#)]
12. O'Malley AS, Grossman JM, Cohen GR, Kemper NM, Pham HH. Are electronic medical records helpful for care coordination? Experiences of physician practices. *J Gen Intern Med* 2010 Mar;25(3):177-185 [[FREE Full text](#)] [doi: [10.1007/s11606-009-1195-2](#)] [Medline: [20033621](#)]
13. Scheder-Bieschin J, Bluemke B, Buijzer ED, Echterdiek F, Nacsa J, Ondresik M, et al. Improving emergency department patient-doctor conversation through an artificial intelligence symptom taking tool: an action-oriented design pilot study. *medRxiv*. 2020 Nov 16. URL: <https://www.medrxiv.org/content/10.1101/2020.11.13.20230953v1.full> [accessed 2021-05-16]
14. Miller S, Gilbert S, Virani V, Wicks P. Patients' utilization and perception of an artificial intelligence-based symptom assessment and advice technology in a British primary care waiting room: exploratory pilot study. *JMIR Hum Factors* 2020 Jul 10;7(3):e19713 [[FREE Full text](#)] [doi: [10.2196/19713](#)] [Medline: [32540836](#)]
15. Gilbert S, Mehl A, Baluch A, Cawley C, Challiner J, Fraser H, et al. How accurate are digital symptom assessment apps for suggesting conditions and urgency advice? A clinical vignettes comparison to GPs. *BMJ Open* 2020 Dec 16;10(12):e040269 [[FREE Full text](#)] [doi: [10.1136/bmjopen-2020-040269](#)] [Medline: [33328258](#)]
16. Meyer AND, Giardina TD, Spitzmueller C, Shahid U, Scott TMT, Singh H. Patient perspectives on the usefulness of an artificial intelligence-assisted symptom checker: cross-sectional survey study. *J Med Internet Res* 2020 Jan 30;22(1):e14679 [[FREE Full text](#)] [doi: [10.2196/14679](#)] [Medline: [32012052](#)]

17. Moreno Barriga E, Pueyo Ferrer I, Sánchez Sánchez M, Martín Baranera M, Masip Utset J. A new artificial intelligence tool for assessing symptoms in patients seeking emergency department care: the Mediktör application. *Emergencias* 2017;29(6):391-396 [FREE Full text] [Medline: 29188913]
18. Chen CE, Chen CT, Hu J, Mehrotra A. Walk-in clinics versus physician offices and emergency rooms for urgent care and chronic disease management. *Cochrane Database Syst Rev* 2017 Feb 17;2:CD011774 [FREE Full text] [doi: [10.1002/14651858.CD011774.pub2](https://doi.org/10.1002/14651858.CD011774.pub2)] [Medline: 28211045]
19. Urgent treatment centres - FAQs to support implementation. NHS England and NHS Improvement. 2019. URL: <https://www.england.nhs.uk/wp-content/uploads/2017/07/urgent-treatment-centres-faqs-v2.0.pdf> [accessed 2021-05-16]
20. Quick Guide: improving access to urgent treatment centres using the Directory of Services. NHS England and NHS Improvement. 2019. URL: <https://www.england.nhs.uk/wp-content/uploads/2019/08/quick-guide-improving-access-to-utc-using-dos.pdf> [accessed 2021-05-16]
21. Urgent and Emergency Care Review Team. Urgent Treatment Centres - Principles and Standards. NHS England Guidance. 2017 Jul. URL: <https://www.england.nhs.uk/wp-content/uploads/2017/07/urgent-treatment-centres%E2%80%93Principles-standards.pdf> [accessed 2021-03-19]
22. When to visit an urgent treatment centre (walk-in centre or minor injury unit). National Health Service. URL: <https://www.nhs.uk/using-the-nhs/nhs-services/urgent-and-emergency-care/when-to-visit-an-urgent-care-centre/> [accessed 2020-03-09]
23. Ahmad BA, Khairatul K, Farnaza A. An assessment of patient waiting and consultation time in a primary healthcare clinic. *Malays Fam Physician* 2017;12(1):14-21 [FREE Full text] [Medline: 28503269]
24. Irving G, Neves AL, Dambha-Miller H, Oishi A, Tagashira H, Verho A, et al. International variations in primary care physician consultation time: a systematic review of 67 countries. *BMJ Open* 2017 Nov 08;7(10):e017902 [FREE Full text] [doi: [10.1136/bmjopen-2017-017902](https://doi.org/10.1136/bmjopen-2017-017902)] [Medline: 29118053]
25. Potisek NM, Malone RM, Shilliday BB, Ives TJ, Chelminski PR, DeWalt DA, et al. Use of patient flow analysis to improve patient visit efficiency by decreasing wait time in a primary care-based disease management programs for anticoagulation and chronic pain: a quality improvement study. *BMC Health Serv Res* 2007 Jan 15;7:8 [FREE Full text] [doi: [10.1186/1472-6963-7-8](https://doi.org/10.1186/1472-6963-7-8)] [Medline: 17224069]
26. Ashour OM, Okudan Kremer GE. A simulation analysis of the impact of FAHP–MAUT triage algorithm on the emergency department performance measures. *Expert Syst Appl* 2013 Jan;40(1):177-187. [doi: [10.1016/j.eswa.2012.07.024](https://doi.org/10.1016/j.eswa.2012.07.024)]
27. Maull RS, Smart PA, Harris A, Karasneh AA. *Serv Indust J* 2009 Jul 14;29(7):923-941. [doi: [10.1080/02642060902749534](https://doi.org/10.1080/02642060902749534)]
28. Spechbach H, Rochat J, Gaspoz J, Lovis C, Ehrlé F. Patients' time perception in the waiting room of an ambulatory emergency unit: a cross-sectional study. *BMC Emerg Med* 2019 Aug 01;19(1):41 [FREE Full text] [doi: [10.1186/s12873-019-0254-1](https://doi.org/10.1186/s12873-019-0254-1)] [Medline: 31370794]
29. Wiler JL, Bolandifar E, Griffey RT, Poirier RF, Olsen T. An emergency department patient flow model based on queueing theory principles. *Acad Emerg Med* 2013 Sep;20(9):939-946. [doi: [10.1111/acem.12215](https://doi.org/10.1111/acem.12215)] [Medline: 24050801]
30. Patwardhan A, Davis J, Murphy P, Ryan SF. Comparison of waiting and consultation times in convenient care clinics and physician offices: a cross-sectional study. *J Prim Care Community Health* 2013 Apr 01;4(2):124-128 [FREE Full text] [doi: [10.1177/2150131912450030](https://doi.org/10.1177/2150131912450030)] [Medline: 23799720]
31. Ucar I, Smeets B, Azcorra A. Discrete-event simulation for R. *J Stat Soft* 2019;90(2):1-30. [doi: [10.18637/jss.v090.i02](https://doi.org/10.18637/jss.v090.i02)]
32. Cassidy R, Singh NS, Schiratti P, Semwanga A, Binyaruka P, Sachingongu N, et al. Mathematical modelling for health systems research: a systematic review of system dynamics and agent-based models. *BMC Health Serv Res* 2019 Nov 19;19(1):845 [FREE Full text] [doi: [10.1186/s12913-019-4627-7](https://doi.org/10.1186/s12913-019-4627-7)] [Medline: 31739783]
33. Zhang X. Application of discrete event simulation in health care: a systematic review. *BMC Health Serv Res* 2018 Sep 04;18(1):687 [FREE Full text] [doi: [10.1186/s12913-018-3456-4](https://doi.org/10.1186/s12913-018-3456-4)] [Medline: 30180848]
34. Matta ME, Patterson SS. Evaluating multiple performance measures across several dimensions at a multi-facility outpatient center. *Health Care Manag Sci* 2007 Jun;10(2):173-194. [doi: [10.1007/s10729-007-9010-2](https://doi.org/10.1007/s10729-007-9010-2)] [Medline: 17608058]
35. Findlay M, Grant H. An application of discrete-event simulation to an outpatient healthcare clinic with batch arrivals. 2011 Presented at: 2011 Winter Simulation Conference (WSC); 2011; Phoenix, Arizona p. 1166. [doi: [10.1109/wsc.2011.6147839](https://doi.org/10.1109/wsc.2011.6147839)]
36. Chand S, Moskowitz H, Norris JB, Shade S, Willis DR. Improving patient flow at an outpatient clinic: study of sources of variability and improvement factors. *Health Care Manag Sci* 2009 Sep;12(3):325-340. [doi: [10.1007/s10729-008-9094-3](https://doi.org/10.1007/s10729-008-9094-3)] [Medline: 19739363]
37. Coster JE, Turner JK, Bradbury D, Cantrell A. Why do people choose emergency and urgent care services? a rapid review utilizing a systematic literature search and narrative synthesis. *Acad Emerg Med* 2017 Sep;24(9):1137-1149. [doi: [10.1111/acem.13220](https://doi.org/10.1111/acem.13220)] [Medline: 28493626]
38. Rui P, Kang K. National Hospital Ambulatory Medical Care Survey: emergency department summary tables. US Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Health Statistics., 2017. URL: https://www.cdc.gov/nchs/data/nhamcs/web_tables/2017_ed_web_tables-508.pdf [accessed 2020-04-10]
39. Sinha M, Khor K, Amresh A, Drachman D, Frechette A. The use of a kiosk-model bilingual self-triage system in the pediatric emergency department. *Pediatr Emerg Care* 2014 Jan;30(1):63-68. [doi: [10.1097/PEC.0000000000000037](https://doi.org/10.1097/PEC.0000000000000037)] [Medline: 24378865]

40. Greenhalgh T, Wherton J, Papoutsis 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](https://doi.org/10.1186/s12916-018-1050-6)] [Medline: [29754584](https://pubmed.ncbi.nlm.nih.gov/29754584/)]
41. Cresswell K, Sheikh A. Organizational issues in the implementation and adoption of health information technology innovations: an interpretative review. *Int J Med Inform* 2013 May;82(5):e73-e86. [doi: [10.1016/j.ijmedinf.2012.10.007](https://doi.org/10.1016/j.ijmedinf.2012.10.007)] [Medline: [23146626](https://pubmed.ncbi.nlm.nih.gov/23146626/)]

Abbreviations

DES: discrete-event simulation

ED: emergency department

UCC: urgent care centre

Edited by C Lovis; submitted 10.12.20; peer-reviewed by A Maier, S Fetzler, A Landmark; comments to author 20.02.21; revised version received 19.03.21; accepted 14.04.21; published 21.05.21.

Please cite as:

Montazeri M, Multmeier J, Novorol C, Upadhyay S, Wicks P, Gilbert S

Optimization of Patient Flow in Urgent Care Centers Using a Digital Tool for Recording Patient Symptoms and History: Simulation Study

JMIR Form Res 2021;5(5):e26402

URL: <https://formative.jmir.org/2021/5/e26402>

doi: [10.2196/26402](https://doi.org/10.2196/26402)

PMID: [34018963](https://pubmed.ncbi.nlm.nih.gov/34018963/)

©Maryam Montazeri, Jan Multmeier, Claire Novorol, Shubhanan Upadhyay, Paul Wicks, Stephen Gilbert. Originally published in *JMIR Formative Research* (<https://formative.jmir.org>), 21.05.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Formative Research*, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

A Provider-Facing eHealth Tool for Transitioning Youth With Special Health Care Needs From Pediatric to Adult Care: Mixed Methods, User-Engaged Usability Study

Darcy Jones McMaughan^{1,2}, MA, PhD; Sherry Lin¹, MPH, PhD; Jennifer Ozmetin¹, BSc; Judith Gayle Beverly¹, MPA; Joshua Brog¹, MPH; Emily Naiser³, MPH

¹Department of Health Policy and Management, School of Public Health, Texas A&M University, College Station, TX, United States

²Oklahoma State University, Stillwater, OK, United States

³Public Policy Research Institute, Texas A&M University, College Station, TX, United States

Corresponding Author:

Darcy Jones McMaughan, MA, PhD

Oklahoma State University

423 Willard

Stillwater, OK

United States

Phone: 1 979 204 3983

Email: darcy.mcmaughan@okstate.edu

Abstract

Background: There is a need for medical education on health care transitions for youth with special health care needs. The Texas Transition Toolkit (the tool) supports providers through a *one-stop shop* for researching literature on care transitions, a catalog of care transition tools, and guides for developing care transition programs.

Objective: This study aims to assess the functionality and usability of the tool with providers working with transition-aged children and youth with special health care needs (representative users).

Methods: The tool was evaluated using a triangulated mixed methods case study approach consisting of a concurrent think-aloud phase, a satisfaction survey, and a survey of problem relevance and task performance to operationalize and capture functionality and usability. Our mixed methods deep dive into the functionality and usability of the tool focused on 10 representative users from one medical home in Texas and 5 website design experts.

Results: Representative users found the tool to be highly relevant, as demonstrated by the satisfaction score for *relevance* (138/150, 92%). According to the users, the tool provided comprehensive information related to health care transitions for youth with special health care needs, with a satisfaction score of 87.3% (131/150) for *comprehensive*. Overall satisfaction with the tool was high at 81.92% (1065/1300) with a cutoff score of 73.33% (953.4/1300) indicating high satisfaction, but users reported relatively lower satisfaction with *search* (114/150, 76%) and *navigation (ease of use)*: 114/150, 76%; hyperlinks: 163/200, 81.5%; structure: 159/200, 79.5%). They experienced search- and navigation-related problems (total problems detected: 21/31, 68%) and, based on quality checks, had a relatively low task completion rate for tasks involving finding information (60/80, 75%), which required searching and navigation. The problems identified around search and navigation functionality were relevant (relevance scores ranging from 14.5 to 22, with a cutoff score of 11.7 indicating relevance).

Conclusions: The tool may help bridge the gaps in training on health care transitions for youth with special health care needs in US medical education. The tool can be used to create structured protocols to help improve provider knowledge, collaboration across pediatric and adult care providers, and the continuity of care as youth with special health care needs transition from pediatric to adult care. The results provided a road map for optimizing the tool and highlighted the importance of evaluating eHealth technologies with representative users.

(JMIR Form Res 2021;5(5):e22915) doi:[10.2196/22915](https://doi.org/10.2196/22915)

KEYWORDS

youth with special health care needs; health care transitions; eHealth; usability; concurrent think aloud method

Introduction

Background

Pediatric to adult health care transitions are a health promotion priority for youth with special health care needs in the United States [1]. Youths with special health care needs have chronic physical, developmental, behavioral, or emotional conditions that require health care and other services beyond typical use [2]. These youths need a higher level of well-coordinated health care, often across multiple providers and specialty types, to realize optimal health and wellness [3,4]. Common conditions among youth with special health care needs include a range of diagnoses that necessitate an even broader range of types and intensities of health care transition support. Diabetes, sickle cell disease, autism, and spina bifida are examples of the diversity of diagnoses among youth with special health care needs. They may need differing care transition support for a variety of aspects of their health care, such as medication management, health status monitoring, physical therapies, and behavioral therapies.

An estimated 5 million youths with special health care needs in the United States are between 12 and 17 years old—the *transition age*, during which health care transition planning should begin [5]. Seamlessly moving from pediatric to adult health care is a critical aspect of transitioning from childhood to adulthood for youth with special health care needs, and structured pediatric to adult care transition programs are associated with positive outcomes [6-8]. Without appropriate health care transition services, youths with special health care needs are vulnerable to adverse outcomes, such as loss of continuity of care [9,10], preventable morbidity, emergency department visits, and hospital visits [11,12].

Pediatric to adult health care transitions should be characterized by having an organizational policy in place that supports care transitions, developing tracking and monitoring processes that follow youths with special health care needs as they transition, assessing youth transition readiness, and transferring care [13]. According to the care transitions consensus statement by the American Academy of Pediatrics, the American Academy of Family Physicians, and the American College of Physicians, providers who care for youth with special health care needs should “(1) understand the rationale for transition from child-oriented to adult-oriented care; (2) have the knowledge and skills to facilitate that process; and (3) know if, how, and when transfer of care is indicated” [14]. It is also vitally important for health care institutions to have organizational policies that support care transitions at the pediatric and adult care levels. The adult team is crucial in supporting care engagement post transfer of care, but much of the emphasis is placed on the pediatric team.

Most youths with special health care needs do not receive adequate care transition services. Across the United States, only 17% of youths with special health care needs received appropriate care transitioning support [15], and access to health services decreased as the youths aged into adulthood [16]. Lack of adequate care transition support may be exacerbated by barriers to care. Although many barriers to care transitions exist

for youth with special health care needs, providers also experience barriers to providing care transition services. These barriers can include lack of communication and coordination between the pediatric and adult sides of care, the perceived difficulty of working with youth with special health care needs, and provider hesitation to be involved in the care transition process [17,18]. Provider hesitation may be because of discomfort in providing care transition services, perceived lack of knowledge, lack of training on providing care for youth with special health care needs, and lack of resources (time and low or no reimbursement) associated with providing care transition services [17,19-22].

To assist in addressing issues related to lack of information, communication, and limited resources surrounding care transitions for youth with special health care needs, we tested the functionality and usability of an internet communication technology developed to assist health care providers in transitioning youths from pediatric to adult care. Funded by the Texas Department of State Health Services under the auspices of the federal Maternal and Child Health Bureau, this particular internet communication technology, the Texas Transition Toolkit [23], is an eHealth initiative providing three main services: a *one-stop shop* to research literature on transition care, a catalog of relevant tools to assist providers in creating care transition plans, and steps for how to develop and communicate a transition program. This tool is a state-level initiative that complements the federal Got Transition program [24] and is in line with the Health Resources and Services Administration Maternal and Child Health Bureau Title V Services Block Grant Program’s transition performance measure, specifically advocating for increasing the number of youths receiving appropriate care transition services [25].

Internet communication technologies such as this tool have the potential to provide rapid access to health and medical knowledge. Under one of the many definitions of *eHealth*, these health care system-oriented applications can improve access to quality health care at lower costs. Through eHealth applications, health-related information and knowledge can reach rural, isolated, low-population, or impoverished communities with very little marginal cost to the community or health service provider [26]. As these benefits of eHealth internet communication technology are mitigated by the quality of the application, it is important to evaluate the functionality and usability of eHealth internet communication technology with end users. Functionality assesses whether the technology works the way it is intended to work and delivers the expected results, whereas usability assesses users’ reactions to and interactions with the technology (including perceived utility and satisfaction) [27].

Objective

This paper describes the formative evaluation of the functionality and usability of this eHealth internet communication tool using a mixed methods usability test. As pragmatic researchers, we determined our research method through our research question: do representative users find the tool functional and usable? [28]. The project was exempted

from the institutional review board of the Texas A&M University.

Methods

Study Design

The usability test used a triangulated mixed methods case study approach consisting of a concurrent think-aloud phase, a satisfaction survey, and a survey of problem relevance and task performance to operationalize and capture functionality and usability [29,30]. The triangulation mixed methods design uses quantitative and qualitative methods to study an issue [29]. Triangulation considers the qualitative and quantitative aspects of a single study separately during the discovery and analyses stages but integrates the findings of each paradigm to develop new knowledge [31]. This approach examines multiple cases, in this case, multiple representative users, to more fully understand the phenomenon of interest—the functionality and usability of the tool [32]. The evaluation of the functionality and usability of the tool involved assessing for two aspects of the experiences of representative users: a qualitative exploration of problems encountered and a quantitative assessment of satisfaction. To confirm the outcomes of the usability test and further measure the functionality of the tool, the research team explored the relevance of the detected problems and representative users' task performance. If the number and type of tasks and associated quality checks successfully completed or failed aligns with the concurrent think-aloud and the satisfaction analyses, it supports both the robustness of the results and indicates functionality and usability issues for representative users. Similarly, the presence of *high-relevance* problems in the concurrent think-aloud and satisfaction analyses adds further support to the robustness of the results in detecting important problems and where the functionality and usability of the tool possibly fails. Relevance, in this context, is how connected the type of problem is to the functionality and usability of the tool. It consists of the likelihood of the problem occurring combined with the impact the problem has on the functionality and usability of the tool for representative users, should the problem occur.

Participants

Two groups of participants were recruited for the usability test. One group consisted of representative users—providers working with youths with special health care needs as they transitioned from pediatric to adult care. This group participated in the concurrent think-aloud phase of the usability test and completed the satisfaction survey, thus providing information on problems encountered when using the tool and end user satisfaction with the tool. The second group consisted of website design professionals with experience in designing informational websites for governmental organizations. This group provided information on the relevance of the problems encountered by the representative users.

On the basis of the criterion sampling, 10 representative users were recruited from a medical home in Texas [33]. As the think-aloud method provides a rich source of data, a small sample of subjects generally suffice to discover individual (rather than population) knowledge and experiences [34]. The

medical home was an opportunity sample that met the usability trial criterion established a priori, that is, the medical home expressed an interest in a care transition program for youth with special health care needs, had the ability to implement a care transition program, provided care for at least 25 youths with special health care needs, and was able to suggest a champion of care transitions. Through this champion, we were able to recruit a full sample of representative users during a professional conference on care transitions held at the medical home. To be eligible to participate as representative users, the providers must have experience working with youth with special health care needs on the pediatric or adult side of care.

Five website design experts from a research organization affiliated with higher education and a robust history of government website design were recruited to provide expert assessment of the relevance of each problem encountered by representative users during the concurrent think-aloud phase of the usability test. The experts were recruited via snowball recruitment, starting with a government website design consultant familiar with internet communication technology. These experts included software application developers (n=2), web application developers (n=2), and a web and information designer (n=1) with experience in the field ranging from 11 to 17 years.

Qualitative and Quantitative Assessments

Qualitative: Exploring Problems Encountered Through the Concurrent Think-Aloud Method

eHealth internet communication tools can be evaluated using a variety of techniques; most originate from the fields of human-computer interaction and media design, which can be broadly grouped into two modes: expert-focused methods and user-focused methods. One example of a user-focused method is the think-aloud usability testing method [35]. This approach, which originates in the evaluation of physical tasks and then used primarily for software evaluations, is also frequently used for evaluating websites [36,37]. A form of the think-aloud approach, the concurrent think-aloud method, is an evaluation method that involves representative users completing tasks using the tool and simultaneously verbalizing their thoughts. During the recorded sessions, representative users completed tasks according to a predetermined scenario (a vignette) while verbalizing their experiences. Analyses of these verbal reports provided detailed insights into the functionality and usability problems encountered by representative users.

To evaluate the functionality and usability of the tool using the concurrent think-aloud method, the research team developed a set of tasks related to the intended purpose of the tool. For example, representative users were asked to *find a peer-reviewed article focusing on care transitions for youth with type 1 diabetes* and to denote the completion of the task (yes or no). All tasks could be carried out independently to minimize dependency bias and reduce the risk that users would stall after one or two tasks. When applicable, a task contained an associated quality check. The quality check served to determine if the user actually completed the task successfully, in addition to perceived successful completion. These quality checks

consisted of asking representative users to write down and speak aloud a few words from the completed subtasks—for example, when asked to find a peer-reviewed article on care transitions for youth with type 1 diabetes, users were subsequently asked to write and speak aloud the first three words of the article title. The tasks were introduced by a vignette, which explained the context and provided the details necessary to perform the tasks. The vignette described a youth with type 1 diabetes poised to graduate high school, transition to attending university, and transition to adult health care without the support of family. Each task represented an action that representative users were likely to perform when using the tool to provide care transition assistance for the youth in the vignette: navigating to pages containing information on care transitions, finding and downloading articles from the evidence base, locating care transition tools, and gathering information on key aspects of developing a care transition program ([Multimedia Appendix 1](#)).

During the conference intermission, representative users received concurrent think-aloud tasks and oral instructions on how to carry them out. These instructions, which the facilitator read out loud from a script for consistency, instructed each user to “think aloud while performing your tasks, and pretend that the facilitator is not there. Do not turn to them for assistance. If you fall silent for a while, the facilitator will remind you to keep talking aloud. Finally, remember that it is the T3 Website, and not you, that is being tested.” All interactions between the representative users and the tool were recorded.

Quantitative: Assessing Satisfaction With the Website Evaluation Questionnaire

To assess satisfaction with the tool, representative users completed the Website Evaluation Questionnaire upon finishing the concurrent think-aloud phase of the usability test. The Website Evaluation Questionnaire is a 26-item Likert-type scale ([Multimedia Appendix 2](#)). Items from the questionnaire load onto one of eight dimensions: ease of use, hyperlinks, structure, relevance, comprehension, completeness, layouts, and search options. For each item, participants circled the response that best characterized their attitude on the item statement: strongly disagree, disagree, neither agree nor disagree, agree, and strongly agree. Each dimension also contained one reverse-coded item. Responses to these items thus represented a disagreement, meaning disagreement with a negative statement about the tool (in other words, satisfaction with the tool along that measure; see Elling et al [35] for the structure of the Website Evaluation Questionnaire).

The Website Evaluation Questionnaire was specifically designed for the evaluation of government websites and is based on user attitudes toward three global factors: the interaction process, the outcomes of the interaction process, and the esthetics of the website [38]. Each of these global factors provided the conceptual foundation for the four main quality factors assessed by the Website Evaluation Questionnaire. The first factor is navigation, which assesses user satisfaction regarding the process of seeking information via the tool. This is particularly relevant to measure the functionality and usability of the tool, given that the purpose of the tool is to assist representative users

in finding information on care transitions for youth with special health care needs. Navigation is assessed through the dimensions of ease of use, hyperlinks, and structure, each of which was measured by three (ease of use) or four (hyperlinks and structure) items. The items interrogate users on, for example, whether the website is user friendly (“I consider this website user friendly”), whether they could find the information they needed via the hyperlinks (“under the hyperlinks, I found the information I expected to find there”), and whether the structure of the website supports information-seeking (“the convenient set-up of the website helps me find the information I am looking for”). The second dimension, content, measured the outcomes of the process of seeking information and contained the construct relevance, completeness, and comprehension. Each of these constructs was measured by three items per construct that focused on the perceived utility of the website (ie, “this website offers information that I find useful”), how easy it is to understand the website (“I find the information in this website easy to understand”), and whether the website provides enough information (“this website provides me with sufficient information”). The third dimension was layout, which assesses the esthetics of the website through three items, which focused on the appeal and attractiveness of the website (ie, “I like the way this website looks”). The fourth and final dimension, search options, contained three items assessing the usefulness of information retrieved through the search process (ie, “the search option on this website gives me useful results”). The Website Evaluation Questionnaire was tested in controlled and real-life settings and found to be valid and reliable [35].

Quantitative: Assessing the Relevance of Problems Encountered and Task Performance

To assess the relevance of problems encountered with the tool, the web design experts participating in this aspect of the study were asked to evaluate the detected problems in terms of likelihood and impact. Our research team explained to the web design experts that we analyzed recordings of a usability test of the tool and identified verbal indicators of problems experienced by representative users, such as expressions of doubt, task difficulty, incomprehensibility, or annoyance related to the use of the tool. We further explained that to measure the relevance of the problems detected, we asked them, as experts in web design, to rate each problem on a scale of 1 to 5 using two Likert-type scales—one for likelihood and one for impact. Likelihood was defined for the experts as how likely a typical user would experience the problem detected. Impact was defined as how much impact the detected problem would have on the functionality and usability of the website, should that problem occur. We included quotes from the representative users as examples of the detected problems and provided the tool website address for reference ([Multimedia Appendix 3](#)).

Analysis

Once the usability test was completed, the research team transcribed the verbalizations, abstracted data from the surveys, and charted the participants’ navigation through the tool. The functionality and usability of the tool were understood through a thematic analysis of the concurrent think-aloud data using NVivo 12 (QSR International) [39], and descriptive analyses

of the Website Evaluation Questionnaire, task performance, and relevance data using Microsoft Excel [40] and Stata 15 (StataCorp) [41].

Qualitative: Concurrent Think-Aloud Method

This analysis identified the number and types of problems detected through deep readings of recordings of the usability test sessions. To analyze the data obtained from the concurrent think-aloud aspect of the usability test, the researchers (DJM and JO) used a summative content analysis to identify and quantify the content of the data and understand the contextual use of words and phrases within the content [42]. These words and phrases were synthesized into salient categories (problem types). During this analysis, the researchers interpreted and assigned meanings to the synthesized categories [42]. Problems identified were confirmed through team discussions.

Quantitative: Satisfaction With the Tool

Data from the Website Evaluation Questionnaire were coded based on the range of numerical values of each Likert-type item (1-5). Each numerical value in the range weighted the responses to produce a weighted score with *strongly disagree* weighted with 1 and *strongly agree* weighted with 5 and with negative items reverse coded. The weighted score was used to generate a satisfaction ratio based on the sum of the 26 items through a ratio of the total score to the highest possible score, with the highest possible score representing unanimous strong agreement across all representative users. The ratio was then multiplied by 100 to generate the overall satisfaction score. A dimension satisfaction score was similarly generated for each of the eight dimensions.

Cutoff points were determined by subtracting the lowest score on the Likert scale (260) from the highest score on the scale (1300) and dividing by the number of levels of satisfaction (3—for low, medium, and high satisfaction), thus creating an interval value of 346.6. This value was added to each score to create three categories of satisfaction: low (260-606.6), medium (606.7-953.3), and high (953.4-1300). By applying these cutoff scores to the ratio of the total score to the highest possible score, we determined that a satisfaction score of 73.33% (953.4/1300) is the cutoff for high satisfaction. The same process was repeated across dimensions for satisfaction cutoff scores of low (30-70), medium (71-111), and high (112-150) for three item dimensions and low (40-93.3), medium (93.4-146.7), and high (146.8-200) for four item dimensions. The resulting high satisfaction cutoff score was 74.6% (112/150) for the three item dimensions and 73.4% (146.8/200) for the four item dimensions. Sensitivity analyses were conducted using cutoff scores based on mean and mode with no changes to the results.

Quantitative: Relevance of the Problems Encountered and Task Performance

The relevance score for each problem was created through the square root of the multiplied likelihood and impact scores [43]. Likelihood and impact scores were created by weighting each numerical value in the range of the responses to produce a weighted score with *unlikely* or *no impact* weighted with 1 and

highly likely or *high impact* weighted with 5. The weighted score was used to generate a likelihood or impact score for each problem through a ratio of the total score to the highest possible score, with the highest possible score being 25. Cutoff points were determined by creating a three-level (low, medium, and high) interval value. A score of 5 to 11.6 indicated low likelihood, low impact, or low relevance. A score of 11.7 to 18.3 was indicative of likely, impactful, and relevant. A score of 18.4 to 25 signified that the problem type was highly likely, had high impact, or was highly relevant. Task performance was understood through the completion rate for each task and the associated quality checks.

Results

Representative Users

Given the small sample size and potential for identification, demographic data of representative users were fuzzed. Approximately 90% of the representative users who participated in the usability test were White, non-Hispanic women. Every user worked with at least one youth with special health care needs, and approximately 90% had a professional caseload where more than half of the patients were youths with special health care needs. More than one-third of the users worked solely with youths with special health care needs. Almost all (approximately 80%) of the representative users transitioned youths from pediatric to adult care. Half of the users worked with youths with special health care needs on both ends of the continuum—pediatrics and adult care. Approximately one-fourth of the users were physicians and approximately another quarter were nurse practitioners. The remaining representative users were social workers (approximately one-fourth), nurses, and transition specialists (clinical team coordinators, youth service specialists, and transition coaches). Some users served multiple roles (eg, nurses and certified educators for a particular chronic condition).

Qualitative: Problems Encountered Through the Concurrent Think-Aloud Method

Table 1 displays the number of problems detected, the different types of problems detected, and the frequency of each type of problem. A total of 31 problems were detected through the analysis of the concurrent think-aloud transcripts. These problems were aggregated into 10 categories or type of problem, most of which (29/31, 93%) focused on issues associated with finding information on health care transitions, such as the utility of search criteria, finding the search bar, finding disease-specific resources, and finding resources blocked by a firewire. Among these problems, approximately one-fourth (8/31, 26%) centered on the utility of search criteria, meaning representative users found it difficult to create search terms that would return resources on a particular topic in health care transitions:

Still looking...self-care management, maybe I'm looking in the wrong place. Care transition, EPIC transition planning tool...uh I'm not finding the self-care. Okay, I'll probably give up on that.

Table 1. Frequency of detected T3 problems by type (N=31).

Problem type	Examples	Frequency, n (%)
Utility of search criteria	“Okay so if I type in the wrong thing it makes it more difficult. Still looking...self-care management, maybe I’m looking in the wrong place. Care transition, EPIC transition planning tool...uh I’m not finding the self-care. Okay, I’ll probably give up on that.”	8 (26)
Finding the search bar is difficult	“I’m looking for that, scrolling. Hmm, is there a search bar? That would have been easy, oh here we go, I found it.”	4 (13)
Finding disease-specific resources	“So, I still think that it would be better to organize this page by general versus...and then also you could have some general articles and then you could have some disease-specific articles and the disease-specific articles could be in alphabetical order to make it easier to find because I kinda gave up on that one.”	3 (10)
Email of website contact opens email software or application	“Oh okay, so you have to add an account, so you have to actually put your email in? Okay can I close that?”	3 (10)
Cannot locate the email of the website contact person	“Copy any contact email address. Let’s see. Okay so I’m going to the wrong place and I’m going back to look. Okay, so it has a contact person but no email address.”	3 (10)
Need clarity on who to contact for an article behind a firewall	“I can’t really find who to talk to about getting this article.”	3 (10)
Some windows blocked by a firewall or files won’t open	“There were, I’m going to score it a four because there were some windows blocked by a firewall.”	3 (10)
Search bar is missing	“Is there a search bar here? That would be helpful under the tools, search for articles and tools.”	2 (6)
Clicking on aspects of the webpage results in no action	“I should not be doing this because when I click on it, it doesn’t work.”	1 (3)
Difficulty returning to a previous page (<i>going back</i>)	“Do I hit back-arrow or close?”	1 (3)

In total, 19% (6/31) of the problems resulted from the difficulty in finding the search bar. The representative users found the placement of the search bar confusing to the point that some thought the search bar was missing:

Is there a search bar here? That would be helpful under the tools, search for articles and tools.

Approximately 10% (3/31) of the problems were associated with users encountering difficulties finding disease-specific resources:

So, I still think that it would be better to organize this page by general versus...and then also you could have some general articles and then you could have some disease-specific articles and the disease-specific articles could be in alphabetical order to make it easier to find because I kinda gave up on that one.

When attempting to access resources behind a firewall, users encountered a number of problems, including blocked windows, software application openings, and an inability to access the contact information of the website manager:

I can’t really find who to talk to about getting this article.

Quantitative: Satisfaction With the Tool

The overall satisfaction score for the tool was 81.92% (1065/1300), indicating high satisfaction based on the overall satisfaction cutoff score of 73.33% (953.4/1300). The tool scored highest in relevance (138/150, 92%), followed by comprehension and layout (both 131/150, 87.3%), and the lowest in search options (114/150, 76%) and ease of use (114/150, 76%; [Table 2 \[35\]](#)). Satisfaction was high for each dimension.

Table 2. Representative user satisfaction with the tool based on the Website Evaluation Questionnaire.

Dimension and item ^a	Satisfaction scores	
	Dimension score, n (%)	Item score, n (%)
Relevance	138 (92)	
I find the information in website helpful.		46 (92)
Website offers information I find useful.		47 (94)
Information in this website is of little use to me. ^b		45 (90)
Comprehension	131 (87.3)	
Language used in website is clear to me.		46 (92)
I find the information in website easy to understand.		42 (84)
I find many words in website difficult to understand. ^b		43 (86)
Layout	131 (87.3)	
I like the way this website look.		42 (84)
I find the design of this website appealing.		43 (86)
I think this website looks unattractive. ^b		46 (92)
Structure	159 (79.5)	
I know where to find information I need on this website.		38 (76)
I find the structure of this website clear.		42 (84)
I was constantly redirected on this website. ^b		39 (78)
The convenient set-up of the website helps me find the information I am looking for		40 (80)
Hyperlinks	163 (81.5)	
Homepage clearly directs me towards information I need.		39 (78)
Homepage immediately points me to information I need.		40 (80)
Under hyperlinks, I found information I expected to find.		43 (86)
It is unclear which hyperlink leads to information I need. ^b		41 (82)
Completeness	115 (76.6)	
This website provides me with sufficient information.		41 (82)
I find the information in this website precise.		37 (74)
I find the information in this website incomplete. ^b		37 (74)
Search options	114 (76)	
Search option helps me find the right information quickly.		38 (76)
Search option gives me useful results.		40 (80)
Search option gives me too many irrelevant results. ^b		36 (72)
Ease of use	114 (76)	
I find this website easy to use.		38 (76)
I consider this website user friendly.		38 (76)
I had difficulty using this website. ^b		38 (76)

^aItem wording truncated for parsimony. Please see the Website Evaluation Questionnaire for complete item wording.

^bReverse-coded items. The score represents a disagreement score, meaning disagreement with a negative statement about the tool (thus satisfaction with the tool along that measure).

Quantitative: Relevance of the Problems Encountered and Task Performance

Table 3 shows task performance. This is the number of tasks and quality checks successfully completed. Overall, 89%

(89/100) of all the tasks and 75% (60/80) of the quality checks were completed successfully. Participants were least likely to find a specific tool for early adolescence, with a 70% (7/10) completion and a 70% (7/10) quality check success rate.

Table 3. Participant (n=10) task performance and associated quality checks.^a

Task and check ^b	Completion, n (%)
Navigate to page contains a database of peer-reviewed articles.	9 (90)
Write the first 3 words of the database page title.	8 (80)
Find peer-reviewed article on care transition for T1 diabetes.	10 (100)
Write the first 3 words of the article title.	8 (80)
Download the article and view the abstract.	10 (100)
Navigate back to the homepage.	9 (90)
Find the contact information of someone from the T3.	8 (80)
Copy the contact's email here.	5 (50)
A tool by Parent to Parent provides a timeline and action items for parents of youth with special health care needs. Find and view this tool.	9 (90)
Write the first 3 words of the tool title.	7 (70)
Find a tool to discuss level of distress around chronic disease self- management.	9 (90)
Write the first 3 words of the tool title.	9 (90)
You are charged with improving care transition for youth with special health care needs for your institution. Find the page that would be the most useful.	9 (90)
Write the page you found this on.	8 (80)
According to the T3, is a champion important?	9 (90)
Write the page you found this on.	8 (80)
Find a tool specific for early adolescence.	7 (70)
Write the first 3 words of the tool title.	7 (70)

^aAggregating by tasks and checks, we found an 89% (89/100) and 75% (60/80) completion rate, respectively.

^bItem wording truncated for parsimony.

Table 4 provides the likelihood, impact, and relevance scores of each problem type. Focusing on relevance, every problem type was found to be relevant based on a cutoff score of 11.7 for relevance, and 60% (6/10) of the problems were found to be highly relevant to the functioning of the tool based on a cutoff score of 18.4 for highly relevant. Problems that interrupted the ability of end users to locate information because of the layout or mechanical workings of the tool were rated as having the highest relevance to the functioning of the tool. These involved

the tags programmed on search terms (*trouble finding search criteria that result in what user wants*: relevance=18.4), organization of information (hard to find articles by disease: relevance=21.5), and inability to access information (*email of website contact opens email software*: relevance=20.4; *need clarity on who to contact for article behind a firewall*: relevance=21; *some windows blocked by a firewall or files will not open*: relevance=22).

Table 4. Relevance of each problem type (n=10).

Problem type	Relevance ^a $\sqrt{\text{Likelihood} \times \text{Impact}}$, weighted scores		
	Likelihood	Impact	Relevance
Trouble finding search criteria result in what user wants	20	17	18.4
Finding the search bar is difficult	18	17	17.5
Hard to find articles by disease	21	22	21.5
Email of website contact opens email software or application	19	22	20.4
Cannot locate the email of the website contact person	17	17	17
Need clarity on who to contact for article behind a firewall	20	22	21
Some windows blocked by a firewall or files won't open	22	22	22
Search bar is missing	20	24	21.9
Clicking on aspects of the webpage results in no action	14	18	15.9
Difficulty returning to a previous page ("going back")	15	14	14.5

^aRelevance = $\sqrt{\text{Likelihood} \times \text{Impact}}$ or the square root of the likelihood score \times the impact score (Van den Haak et al [43]). Cutoff scores were 5-11.66 for low relevance, 11.67-18.33 for relevant, and 18.34-25 for highly relevant.

Discussion

Principal Findings

The functionality and usability testing of the tool with representative users highlighted the usefulness of this eHealth internet communication technology among providers working with youth with special health care needs. Representative users found the tool to be highly relevant, as demonstrated by the satisfaction score for the dimension *relevance* (138/150, 92.0%). According to the users, the tool provided comprehensive information related to health care transitions for youth with special health care needs, with a satisfaction score of 87.3% (131/150) on the comprehensive dimension. This is important, as it suggests that the tool may help bridge the gaps in training on health care transitions for youth with special health care needs in US medical education [44]. Gaps in education on health care transitions exist at all levels of medical education [44], and providers often express concerns about the lack of knowledge on transitioning care for youth with special health care needs [45,46]. Furthermore, youths transitioning from pediatric to adult care also want access to medical education on health care transitions [47], which can be provided through the tool. There is also a need to provide information on health care transitions among youth with special health care needs to adult providers, who may lack knowledge of the process [48]. An eHealth internet communication technology such as the tool, which includes evidence-based literature on health care transitions and templates for transition protocols, can be used to create a structured tool to help improve collaboration across pediatric and adult care providers and continuity of care [49]. The tool does not rely on the ability of a health care organization to integrate the tool into their system; it can be implemented within an organization or external to an organization. Although integration within the systems of care is ideal [50], in other studies of educational transition tools, health care providers expressed concern over the ability to integrate tools, given the characteristics and deficits of health care systems [51].

Although the tool was well received by representative users, the usability test results identified areas of concern regarding functionality and usability. Users reported the most difficulty in two areas of functionality and usability: search and navigation. This was reflected in both the concurrent think-aloud and satisfaction survey results and was supported by the task performance and relevance analyses. Representative users reported lower satisfaction with search and navigation dimensions, relatively high number of search- and navigation-related problems and low task completion for tasks involving finding tools that require searching and navigation. The problems identified around the search and navigation functionality were also found to be relevant by web design experts. Each of these areas of analyses triangulates on search and navigation issues, suggesting the robustness of the results and allowing researchers to fine-tune the tool to optimize performance.

Uncovering these inefficiencies and subsequently optimizing the tool improves its functionality and usability, which may drive the future use of the technology for its specified purpose [52]. Studies of commercial technologies show that end users are more likely to utilize a technology if it is easy to use [53-55]. This is especially important for eHealth internet communication technology, such as the tool, which is intended to improve health outcomes in the long term through the dissemination of information and best practices. Although the results from the usability test provide a clear road map for optimizing the tool, they also highlight the importance of evaluating eHealth and internet communication technology with representative users. Often, internet communication technology evaluations focus too intensely on the technical aspects of the technology rather than the needs and expectations of end users [56]. Conducting functionality and usability assessments before widespread implementation is essential for forecasting usage and ensuring that internet communication technology has the intended impact [57].

Limitations and Future Research

Although the concurrent think-aloud methodology and the satisfaction survey provided a robust way to ascertain the functionality and usability of this eHealth internet communication technology, there are certainly other acceptable methods. In fact, even with the growing interest in website functionality and usability and the growing use of internet communication technology for eHealth applications, there is no consensus on the definition of usability. There exists a multitude of possible dimensions and measurement techniques in the field of internet communication technology research. Choosing between these dimensions and techniques to evaluate the technology involves a certain amount of bias, which could result in the website performing better in the laboratory than in a real-life setting. However, this usability test followed quality measures for usability studies of eHealth applications [58]. We used a valid and reliable tool, chose our study design based on the objectives of the study, used across-method triangulation, and included both representative users and experts in assessing usability.

Few evaluated measures exist that can be used to understand the impact of eHealth technology on health outcomes [59], which has potentially contributed to the limited evaluation of eHealth [60]. This is a particularly important limitation when the website is eHealth internet communication technology and is expected to affect health outcomes. Thus, further research is needed on the effect of the tool on care transition outcomes, particularly in isolated or resource-poor communities. Evaluation research on holistic care transition tools and programs is limited, with the majority of outcomes-focused evaluations targeting narrow, disease-specific populations of youth with special health care needs [6]. Although disease-specific instruments fall under the triple aim domain of population health [61], considering that most youths with special health care needs experience multiple comorbid conditions across their lifetime, more empirical evidence is needed on tools dedicated to broader care transitions. This tool, with its broad focus, answers calls to assist care providers in transitioning youth with special health care needs using a biopsychosocial model [62].

Conflicts of Interest

None declared.

Multimedia Appendix 1

Texas Transition Toolkit think-aloud usability testing: vignette, tasks, and quality checks.

[\[DOCX File , 26 KB - formative_v5i5e22915_app1.docx \]](#)

Multimedia Appendix 2

The website evaluation questionnaire.

[\[DOCX File , 19 KB - formative_v5i5e22915_app2.docx \]](#)

Multimedia Appendix 3

Texas Transition Toolkit usability trial: website problem relevance.

[\[DOCX File , 22 KB - formative_v5i5e22915_app3.docx \]](#)

References

1. Coller RJ, Berry JG, Kuo DZ, Kuhlthau K, Chung PJ, Perrin JM, et al. Health system research priorities for children and youth with special health care needs. *Pediatrics* 2020 Mar;145(3):e20190673 [[FREE Full text](#)] [doi: [10.1542/peds.2019-0673](https://doi.org/10.1542/peds.2019-0673)] [Medline: [32024751](https://pubmed.ncbi.nlm.nih.gov/32024751/)]
2. McPherson M, Arango P, Fox H, Lauver C, McManus M, Newacheck PW, et al. A new definition of children with special health care needs. *Pediatrics* 1998 Jul;102(1 Pt 1):137-140. [doi: [10.1542/peds.102.1.137](https://doi.org/10.1542/peds.102.1.137)] [Medline: [9714637](https://pubmed.ncbi.nlm.nih.gov/9714637/)]
3. Council on Children with Disabilities Medical Home Implementation Project Advisory Committee. Patient- and family-centered care coordination: a framework for integrating care for children and youth across multiple systems. *Pediatrics* 2014 May;133(5):1451-1460. [doi: [10.1542/peds.2014-0318](https://doi.org/10.1542/peds.2014-0318)] [Medline: [24777209](https://pubmed.ncbi.nlm.nih.gov/24777209/)]
4. Boulet S, Boyle CA, Schieve LA. Health care use and health and functional impact of developmental disabilities among US children, 1997-2005. *Arch Pediatr Adolesc Med* 2009 Jan;163(1):19-26 [[FREE Full text](#)] [doi: [10.1001/archpediatrics.2008.506](https://doi.org/10.1001/archpediatrics.2008.506)] [Medline: [19124699](https://pubmed.ncbi.nlm.nih.gov/19124699/)]
5. National Survey of Children's Health (NSCH). United States Census Bureau. URL: <https://www.census.gov/programs-surveys/nsch/data.html> [accessed 2019-05-04]
6. Schmidt A, Ilango SM, McManus MA, Rogers KK, White PH. Outcomes of pediatric to adult health care transition interventions: an updated systematic review. *J Pediatr Nurs* 2020 Mar;51:92-107 [[FREE Full text](#)] [doi: [10.1016/j.pedn.2020.01.002](https://doi.org/10.1016/j.pedn.2020.01.002)] [Medline: [31981969](https://pubmed.ncbi.nlm.nih.gov/31981969/)]
7. Gabriel P, McManus M, Rogers K, White P. Outcome evidence for structured pediatric to adult health care transition interventions: a systematic review. *J Pediatr* 2017 Sep;188:263-269 [[FREE Full text](#)] [doi: [10.1016/j.jpeds.2017.05.066](https://doi.org/10.1016/j.jpeds.2017.05.066)] [Medline: [28668449](https://pubmed.ncbi.nlm.nih.gov/28668449/)]

8. Crowley R, Wolfe I, Lock K, McKee M. Improving the transition between paediatric and adult healthcare: a systematic review. *Arch Dis Child* 2011 Jun;96(6):548-553 [FREE Full text] [doi: [10.1136/adc.2010.202473](https://doi.org/10.1136/adc.2010.202473)] [Medline: [21388969](https://pubmed.ncbi.nlm.nih.gov/21388969/)]
9. Refaeli T, Mangold K, Zeira A, Königeter S. Continuity and discontinuity in the transition from care to adulthood. *Br J Soc Work* 2016 Mar 21;325-342 [FREE Full text] [doi: [10.1093/bjsw/bcw016](https://doi.org/10.1093/bjsw/bcw016)]
10. Montano CB, Young J. Discontinuity in the transition from pediatric to adult health care for patients with attention-deficit/hyperactivity disorder. *Postgrad Med* 2012 Sep;124(5):23-32. [doi: [10.3810/pgm.2012.09.2591](https://doi.org/10.3810/pgm.2012.09.2591)] [Medline: [23095423](https://pubmed.ncbi.nlm.nih.gov/23095423/)]
11. Jordan L, Swerdlow P, Coates T. Systematic review of transition from adolescent to adult care in patients with sickle cell disease. *J Pediatr Hematol Oncol* 2013 Apr;35(3):165-169 [FREE Full text] [doi: [10.1097/MPH.0b013e3182847483](https://doi.org/10.1097/MPH.0b013e3182847483)] [Medline: [23511487](https://pubmed.ncbi.nlm.nih.gov/23511487/)]
12. Shepard C, Doerge E, Eickmeyer A, Kraft K, Wan J, Stoffel J. Ambulatory care use among patients with spina bifida: change in care from childhood to adulthood. *J Urol* 2018 Apr;199(4):1050-1055 [FREE Full text] [doi: [10.1016/j.juro.2017.10.040](https://doi.org/10.1016/j.juro.2017.10.040)] [Medline: [29113842](https://pubmed.ncbi.nlm.nih.gov/29113842/)]
13. Got Transition. URL: <https://gottransition.org> [accessed 2020-06-28]
14. American Academy of Pediatrics, American Academy of Family Physicians, American College of Physicians-American Society of Internal Medicine. A consensus statement on health care transitions for young adults with special health care needs. *Pediatrics* 2002 Dec;110(6 Pt 2):1304-1306. [Medline: [12456949](https://pubmed.ncbi.nlm.nih.gov/12456949/)]
15. Lebrun-Harris LA, McManus MA, Ilango SM, Cyr M, McLellan S, Mann MY, et al. Transition planning among US youth with and without special health care needs. *Pediatrics* 2018 Oct;142(4):e20180194 [FREE Full text] [doi: [10.1542/peds.2018-0194](https://doi.org/10.1542/peds.2018-0194)] [Medline: [30224366](https://pubmed.ncbi.nlm.nih.gov/30224366/)]
16. de Silva PS, Fishman LN. The transition of the gastrointestinal patient from pediatric to adult care. *Pediatr Clin North Am* 2017 Jun;64(3):707-720. [doi: [10.1016/j.pcl.2017.02.001](https://doi.org/10.1016/j.pcl.2017.02.001)] [Medline: [28502447](https://pubmed.ncbi.nlm.nih.gov/28502447/)]
17. Okumura M, Heisler M, Davis M, Cabana M, Demonner S, Kerr E. Comfort of general internists and general pediatricians in providing care for young adults with chronic illnesses of childhood. *J Gen Intern Med* 2008 Oct;23(10):1621-1627 [FREE Full text] [doi: [10.1007/s11606-008-0716-8](https://doi.org/10.1007/s11606-008-0716-8)] [Medline: [18661191](https://pubmed.ncbi.nlm.nih.gov/18661191/)]
18. White PH, Cooley WC, American Academy of Pediatrics, American Academy of Family Physicians. Supporting the health care transition from adolescence to adulthood in the medical home. *Pediatrics* 2018 Nov;142(5):e20182587 [FREE Full text] [doi: [10.1542/peds.2018-2587](https://doi.org/10.1542/peds.2018-2587)] [Medline: [30348754](https://pubmed.ncbi.nlm.nih.gov/30348754/)]
19. Hergenroeder A, Wiemann C, Cohen M. Current issues in transitioning from pediatric to adult-based care for youth with chronic health care needs. *J Pediatr* 2015 Dec;167(6):1196-1201 [FREE Full text] [doi: [10.1016/j.jpeds.2015.08.005](https://doi.org/10.1016/j.jpeds.2015.08.005)] [Medline: [26340879](https://pubmed.ncbi.nlm.nih.gov/26340879/)]
20. Oswald D, Gilles D, Cannady M, Wenzel D, Willis J, Bodurtha J. Youth with special health care needs: transition to adult health care services. *Matern Child Health J* 2013 Dec;17(10):1744-1752 [FREE Full text] [doi: [10.1007/s10995-012-1192-7](https://doi.org/10.1007/s10995-012-1192-7)] [Medline: [23160763](https://pubmed.ncbi.nlm.nih.gov/23160763/)]
21. Treadwell M, Telfair J, Gibson RW, Johnson S, Osunkwo I. Transition from pediatric to adult care in sickle cell disease: establishing evidence-based practice and directions for research. *Am J Hematol* 2011 Jan 08;86(1):116-120 [FREE Full text] [doi: [10.1002/ajh.21880](https://doi.org/10.1002/ajh.21880)] [Medline: [21061308](https://pubmed.ncbi.nlm.nih.gov/21061308/)]
22. Peter NG, Forke CM, Ginsburg KR, Schwarz DF. Transition from pediatric to adult care: internists' perspectives. *Pediatrics* 2009 Feb;123(2):417-423. [doi: [10.1542/peds.2008-0740](https://doi.org/10.1542/peds.2008-0740)] [Medline: [19171604](https://pubmed.ncbi.nlm.nih.gov/19171604/)]
23. Texas transitions toolkit. 2019. URL: <https://www.texastransitiontoolkit.org/> [accessed 2019-08-10]
24. Got transition (2014-2020). The National Alliance to Advance Adolescent Health. URL: <https://gottransition.org> [accessed 2019-08-31]
25. Lu M, Lauver C, Dykton C, Kogan MD, Lawler MH, Raskin-Ramos L, et al. Transformation of the title V maternal and child health services block grant. *Matern Child Health J* 2015 May;19(5):927-931 [FREE Full text] [doi: [10.1007/s10995-015-1736-8](https://doi.org/10.1007/s10995-015-1736-8)] [Medline: [25854797](https://pubmed.ncbi.nlm.nih.gov/25854797/)]
26. Taylor L, Waller M, Portnoy JM. Telemedicine for allergy services to rural communities. *J Allergy Clin Immunol Pract* 2019;7(8):2554-2559 [FREE Full text] [doi: [10.1016/j.jaip.2019.06.012](https://doi.org/10.1016/j.jaip.2019.06.012)] [Medline: [31238163](https://pubmed.ncbi.nlm.nih.gov/31238163/)]
27. Bertot J, Jaeger PT. User-centered e-government: challenges and benefits for government web sites. *Gov Inf Q* 2006 Jan;23(2):163-168 [FREE Full text] [doi: [10.1016/j.giq.2006.02.001](https://doi.org/10.1016/j.giq.2006.02.001)]
28. Onwuegbuzie A, Leech N. On becoming a pragmatic researcher: the importance of combining quantitative and qualitative research methodologies. *Int J Soc Res Methodol* 2005 Dec;8(5):375-387 [FREE Full text] [doi: [10.1080/13645570500402447](https://doi.org/10.1080/13645570500402447)]
29. Creswell JW, Clark VL, Gutmann ML, Hanson WE. An expanded typology for classifying mixed methods research into designs. In: Taskakori AM, Teddlie CB, editors. *Handbook of Mixed Methods in Social & Behavioral Research*. Thousand Oaks, California, United States: Sage Publications Inc; 2002:209-240.
30. Merriam S, Tisdell E. *Qualitative Research: A Guide to Design and Implementation*. Hoboken, NJ: John Wiley & Sons; 2015.
31. Whitehead D, LoBiondo-Wood G, Haber J. *Nursing and Midwifery Research: Methods and Appraisal for Evidence Based Practice*. Amsterdam, NL: Elsevier; 2016:1-450.
32. Stake RE. *Multiple Case Study Analysis*. New York City, NY: Guilford press; 2013:1-978.

33. Gall MD, Gall JP, Borg WR. Educational Research: An Introduction, 8th Edition. New York City, NY: Pearson; 2007:1-712.
34. Fonteyn M, Kuipers B, Grobe SJ. A description of think aloud method and protocol analysis. *Qual Health Res* 2016 Jul 01;3(4):430-441 [FREE Full text] [doi: [10.1177/104973239300300403](https://doi.org/10.1177/104973239300300403)]
35. Elling S, Lentz L, de Jong M, van den Bergh H. Measuring the quality of governmental websites in a controlled versus an online setting with the 'Website Evaluation Questionnaire'. *Gov Inf Q* 2012 Jul;29(3):383-393 [FREE Full text] [doi: [10.1016/j.giq.2011.11.004](https://doi.org/10.1016/j.giq.2011.11.004)]
36. Hoc J, Leplat J. Evaluation of different modalities of verbalization in a sorting task. *Int J Man Mach Stud* 1983 Mar;18(3):283-306 [FREE Full text] [doi: [10.1016/s0020-7373\(83\)80011-x](https://doi.org/10.1016/s0020-7373(83)80011-x)]
37. Sienot M. Pretesting web sites: a comparison between the plus-minus method and the think-aloud method for the World Wide Web. *J Bus Tech Commun* 2016 Jul 25;11(4):469-482 [FREE Full text] [doi: [10.1177/1050651997011004006](https://doi.org/10.1177/1050651997011004006)]
38. Elling S, Lentz L, De Jong M. Website evaluation questionnaire: development of a research-based tool for evaluating informational websites. In: *Electronic Government*. Berlin, Heidelberg: Springer; 2007:293-304.
39. NVivo Qualitative Data Analysis Software. QRS International. URL: <https://qsrinternational.com/nvivo/nvivo-products/> [accessed 2019-08-10]
40. Microsoft Excel. Microsoft Corporation. URL: <https://office.microsoft.com/excel> [accessed 2019-12-30]
41. Stata Statistical Software: Release 15. StataCorp. 2017. URL: www.stata.com [accessed 2018-05-16]
42. Hsieh H, Shannon SE. Three approaches to qualitative content analysis. *Qual Health Res* 2005 Nov;15(9):1277-1288 [FREE Full text] [doi: [10.1177/1049732305276687](https://doi.org/10.1177/1049732305276687)] [Medline: [16204405](https://pubmed.ncbi.nlm.nih.gov/16204405/)]
43. van den Haak M, De Jong M, Schellens P. Evaluation of an informational web site: three variants of the think-aloud method compared. *Tech Comm* 2007 Feb 01;54(1):58-71.
44. Sharma N, O'Hare K, Antonelli RC, Sawicki GS. Transition care: future directions in education, health policy, and outcomes research. *Acad Pediatr* 2014;14(2):120-127 [FREE Full text] [doi: [10.1016/j.acap.2013.11.007](https://doi.org/10.1016/j.acap.2013.11.007)] [Medline: [24602574](https://pubmed.ncbi.nlm.nih.gov/24602574/)]
45. Dressler P, Nguyen TK, Moody EJ, Friedman SL, Pickler L. Use of transition resources by primary care providers for youth with intellectual and developmental disabilities. *Intellect Dev Disabil* 2018 Feb;56(1):56-68 [FREE Full text] [doi: [10.1352/1934-9556-56.1.56](https://doi.org/10.1352/1934-9556-56.1.56)] [Medline: [29389256](https://pubmed.ncbi.nlm.nih.gov/29389256/)]
46. Kuhlthau K, Warfield ME, Hurson J, Delahaye J, Crossman MK. Pediatric provider's perspectives on the transition to adult health care for youth with autism spectrum disorder: current strategies and promising new directions. *Autism* 2015 Apr;19(3):262-271 [FREE Full text] [doi: [10.1177/1362361313518125](https://doi.org/10.1177/1362361313518125)] [Medline: [24497626](https://pubmed.ncbi.nlm.nih.gov/24497626/)]
47. Sezgin E, Weiler M, Weiler A, Lin S. Proposing an ecosystem of digital health solutions for teens with chronic conditions transitioning to self-management and independence: exploratory qualitative study. *J Med Internet Res* 2018 Sep 06;20(9):e10285 [FREE Full text] [doi: [10.2196/10285](https://doi.org/10.2196/10285)] [Medline: [30190253](https://pubmed.ncbi.nlm.nih.gov/30190253/)]
48. Michaud S, Dasgupta K, Bell L, Yale JF, Anjachak N, Wafa S, et al. Adult care providers' perspectives on the transition to adult care for emerging adults with Type 1 diabetes: a cross-sectional survey. *Diabet Med* 2018 Jul;35(7):846-854 [FREE Full text] [doi: [10.1111/dme.13627](https://doi.org/10.1111/dme.13627)] [Medline: [29577410](https://pubmed.ncbi.nlm.nih.gov/29577410/)]
49. Tanner A, Philbin MM, Ma A, Chambers BD, Nichols S, Lee S, Adolescent Trials Network for HIV/AIDS Interventions. Adolescent to adult HIV health care transition from the perspective of adult providers in the United States. *J Adolesc Health* 2017 Oct;61(4):434-439 [FREE Full text] [doi: [10.1016/j.jadohealth.2017.05.011](https://doi.org/10.1016/j.jadohealth.2017.05.011)] [Medline: [28754584](https://pubmed.ncbi.nlm.nih.gov/28754584/)]
50. Bond J, Shanske S, Hoffman R, Ross AM. Piloting a structured developmental tool to assess transition readiness for youth with special health-care needs: a mixed-methods exploration of health-care provider experiences. *J Child Health Care* 2020 Mar 17;24(1):92-105 [FREE Full text] [doi: [10.1177/1367493518823899](https://doi.org/10.1177/1367493518823899)] [Medline: [30773898](https://pubmed.ncbi.nlm.nih.gov/30773898/)]
51. Morsa M, Gagnayre R, Pomey M, Deccache C, Lombraill P. Developmentally appropriate patient education during transition: a study of healthcare providers' and parents' perspective. *Health Educ J* 2019 Nov 18;79(4):377-389 [FREE Full text] [doi: [10.1177/0017896919888559](https://doi.org/10.1177/0017896919888559)]
52. Venkatesh V, Hoehle H, Aljafari R. A usability evaluation of the Obamacare website. *Gov Inf Q* 2014 Oct;31(4):669-680 [FREE Full text] [doi: [10.1016/j.giq.2014.07.003](https://doi.org/10.1016/j.giq.2014.07.003)]
53. Monsuwé TP, Dellaert BG, de Ruyter K. What drives consumers to shop online? A literature review. *Int J of Service Industry Mgmt* 2004 Feb;15(1):102-121 [FREE Full text] [doi: [10.1108/09564230410523358](https://doi.org/10.1108/09564230410523358)]
54. Rose GM, Straub DW. The effect of download time on consumer attitude toward the e-service retailer. *e-Serv J* 2001;1(1):55 [FREE Full text] [doi: [10.2979/esj.2001.1.1.55](https://doi.org/10.2979/esj.2001.1.1.55)]
55. Huang Z. The impact of usability, functionality and sociability factors on user shopping behavior in social commerce design. In: *HCI in Business, Government, and Organizations*. Switzerland: Springer; 2018:303-312.
56. Verdegem P, Verleye G. User-centered e-Government in practice: a comprehensive model for measuring user satisfaction. *Gov Inf Q* 2009 Jul;26(3):487-497 [FREE Full text] [doi: [10.1016/j.giq.2009.03.005](https://doi.org/10.1016/j.giq.2009.03.005)]
57. Tomlin C. The top 7 healthcare.gov UX failures caused by poor forecasting of usage. *Useful Usability*. 2013. URL: <http://www.usefulness.com/top-7-healthcare-gov-ux-failures> [accessed 2018-08-31]
58. Silva AG, Simões P, Santos R, Queirós A, Rocha NP, Rodrigues M. A scale to assess the methodological quality of studies assessing usability of electronic health products and services: Delphi study followed by validity and reliability testing. *J Med Internet Res* 2019 Nov 15;21(11):e14829 [FREE Full text] [doi: [10.2196/14829](https://doi.org/10.2196/14829)] [Medline: [31730036](https://pubmed.ncbi.nlm.nih.gov/31730036/)]

59. Wakefield BJ, Turvey CL, Nazi KM, Holman JE, Hogan TP, Shimada SL, et al. Psychometric properties of patient-facing ehealth evaluation measures: systematic review and analysis. *J Med Internet Res* 2017 Oct 11;19(10):e346 [[FREE Full text](#)] [doi: [10.2196/jmir.7638](https://doi.org/10.2196/jmir.7638)] [Medline: [29021128](https://pubmed.ncbi.nlm.nih.gov/29021128/)]
60. Enam A, Torres-Bonilla J, Eriksson H. Evidence-based evaluation of ehealth interventions: systematic literature review. *J Med Internet Res* 2018 Nov 23;20(11):e10971 [[FREE Full text](#)] [doi: [10.2196/10971](https://doi.org/10.2196/10971)] [Medline: [30470678](https://pubmed.ncbi.nlm.nih.gov/30470678/)]
61. Prior M, McManus M, White P, Davidson L. Measuring the "triple aim" in transition care: a systematic review. *Pediatrics* 2014 Dec;134(6):1648-1661. [doi: [10.1542/peds.2014-1704](https://doi.org/10.1542/peds.2014-1704)] [Medline: [25422015](https://pubmed.ncbi.nlm.nih.gov/25422015/)]
62. Kuo AA, Ciccarelli MR, Sharma N, Lotstein DS. A health care transition curriculum for primary care residents: identifying goals and objectives. *Pediatrics* 2018 Apr;141(Suppl 4):346-354 [[FREE Full text](#)] [doi: [10.1542/peds.2016-4300L](https://doi.org/10.1542/peds.2016-4300L)] [Medline: [29610416](https://pubmed.ncbi.nlm.nih.gov/29610416/)]

Edited by G Eysenbach; submitted 03.08.20; peer-reviewed by A Abubakar, A Tanner; comments to author 02.12.20; revised version received 27.03.21; accepted 13.04.21; published 25.05.21.

Please cite as:

McMaughan DJ, Lin S, Ozmetin J, Beverly JG, Brog J, Naiser E

A Provider-Facing eHealth Tool for Transitioning Youth With Special Health Care Needs From Pediatric to Adult Care: Mixed Methods, User-Engaged Usability Study

JMIR Form Res 2021;5(5):e22915

URL: <https://formative.jmir.org/2021/5/e22915>

doi: [10.2196/22915](https://doi.org/10.2196/22915)

PMID: [34032579](https://pubmed.ncbi.nlm.nih.gov/34032579/)

©Darcy Jones McMaughan, Sherry Lin, Jennifer Ozmetin, Judith Gayle Beverly, Joshua Brog, Emily Naiser. Originally published in *JMIR Formative Research* (<https://formative.jmir.org>), 25.05.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Formative Research*, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Remote Rating of Atopic Dermatitis Severity Using Photo-Based Assessments: Proof-of-Concept and Reliability Evaluation

Zarqa Ali¹, MD, PhD; Kristina Melbardis Joergensen², MD; Anders Daniel Andersen², PhD; Andrei Chiriac³, MD; Theis Bjerre-Christensen², MD; Ionela Manole², MD; Ana-Maria Dutei², MD; Irina Deaconescu², MD; Alina Suru², MD; Adina Serban², MD; Ari Pall Isberg², MSc; Priyanka Dahiya², MSc; Simon Francis Thomsen^{1,4}, MD, PhD, DMSc; John Robert Zibert², PhD

¹Department of Dermato-Venereology and Wound Healing Centre, Bispebjerg Hospital, Copenhagen, Denmark

²Studies&Me, Copenhagen, Denmark

³Omhu A/S, Copenhagen, Denmark

⁴Department of Biomedical Sciences, University of Copenhagen, Copenhagen, Denmark

Corresponding Author:

Zarqa Ali, MD, PhD

Department of Dermato-Venereology and Wound Healing Centre

Bispebjerg Hospital

Bispebjerg Bakke 23

Copenhagen, 2400

Denmark

Phone: 45 60122596

Email: zarqa_ali@hotmail.com

Abstract

Background: Digital imaging of dermatological patients is a novel approach to remote assessment and has recently become more relevant since telehealth and remote decentralized clinical trials are gaining ground.

Objective: We aimed to investigate whether photographs taken by a smartphone are of adequate quality to allow severity assessments to be made and to explore the usefulness of an established atopic dermatitis severity assessment instrument on photograph evaluation.

Methods: During scheduled visits in a previously published study, the investigating doctor evaluated the severity of atopic dermatitis using the Scoring AD (SCORAD) index and took photographs of the most representative lesions (target lesions) with both a smartphone and a digital single-lens reflex camera (DSLR). The photographs were then assessed by 5 dermatologists using the intensity items of the SCORAD (iSCORAD), which consists of erythema, oedema/papulation, excoriations, lichenification, oozing/crusts, and dryness (scale 0-3, maximum score 18). The mean iSCORAD of the photographs was calculated and compared with in-person assessments using Pearson correlation and Bland-Altman plots. Intraclass correlation coefficients were used for interrater reliability.

Results: A total of 942 photographs from 95 patients were assessed. The iSCORAD based on smartphone photographs correlated strongly with the evaluations performed in person (iSCORAD: $r=0.78$, $P<.001$; objective SCORAD: $r=0.81$, $P<.001$; and total SCORAD: $r=0.78$, $P<.001$). For iSCORAD specifically, a Bland-Altman plot showed a difference in mean score of 1.31 for in-person and remote iSCORAD. In addition, the interrater agreement between the 5 rating dermatologists was 0.93 (95% CI 0.911-0.939). A total of 170 lesions were photographed, and the difference in mean scores was 1.32, 1.13, and 1.43 between in-person and remote evaluations based on photographs taken by a DSLR camera, a smartphone without flash, and a smartphone with flash, respectively.

Conclusions: In terms of quality, remote atopic dermatitis severity assessments based on photographs are comparable to in-person assessments, and smartphone photos can be used to assess atopic dermatitis severity to a similar degree as photographs from a DSLR camera. Further, the variation in how the dermatologists in this study rated the iSCORAD based on the photographs was very low.

(JMIR Form Res 2021;5(5):e24766) doi:[10.2196/24766](https://doi.org/10.2196/24766)

KEYWORDS

atopic dermatitis; eczema; remote assessment; photo; photographs; EASI; SCORAD; severity; assessment; agreement

Introduction

Digital imaging of dermatological patients is a novel approach to remote assessment and has recently become more relevant since telehealth and remote clinical trials are gaining ground.

Clinical trials are a cornerstone of drug development and provide scientific evidence on safety and efficacy of a new pharmaceutical drug. However, traditional clinical trials take a long time to complete and are expensive and inefficient in terms of high dropout rates [1]. Fully decentralized virtual clinical trials (VCTs) that incorporate remote outcome assessments may accelerate clinical trials, increase adherence, reduce dropout rates, and bring new treatments to the market faster [2]. Teledermatology has grown over the last two decades, and the visual nature of dermatology makes it ideal for the practice of telemedicine. Teledermatology is cost-effective [3], effective in managing dermatologic diseases [4], has better diagnostic accuracy [5], and is satisfying for both patients and providers [6].

The foundation for both VCTs in dermatology and teledermatology is remote assessment, including digital assessment of photographs of skin conditions. However, little is known about remote assessment of many dermatological diseases including atopic dermatitis (AD).

Several assessment tools have been developed to grade AD severity in the clinic. Although many of these tools have been validated when used in in-person settings, it is unknown to what extent they can be applied to assess photographs remotely. To our knowledge, only one study by Hughes et al [7] has investigated the concordance between assessment of AD in person compared to a standardized set of full-body digital photographs captured by a clinical research coordinator. They reported an excellent agreement between in-person assessment and remote assessment of photographs with respect to body surface area, Eczema Area and Severity Index, and Scoring AD (SCORAD) scales. However, to better accommodate the promises of VCT in which most of the study tasks are conveniently performed on participants' smartphones from the comfort of their own home, the number of photographs required from participants in dermatological trials should be minimized.

The primary objective of this study was to investigate whether photographs taken by a smartphone are of adequate quality to allow severity assessment to be made. The secondary purpose was to determine whether SCORAD can be applied to the evaluation of photographs.

Methods

Data Collection

The data used in the present study were from a previously published study (Atopisk Dermatit Eksem Studie [ADES]) [5]. The study was originally designed to investigate adherence to treatment using a memory button with an associated smartphone

app (Klikkit, The HabLab ApS) among patients with AD. Although not originally designed for this purpose, data from in-person severity assessments together with digital photographs of lesions taken by the doctor have been evaluated for further analysis. A medical doctor trained in AD assessment by a certified dermatologist evaluated AD severity using SCORAD [8,9] during 2 scheduled in-person visits in the clinic. The doctor took digital photographs of AD lesions using both a smartphone (with and without flash) and a digital single-lens reflex camera (DSLR). These photographs were used for severity assessment by 5 blinded dermatologists for the purpose of this analysis.

The number of target lesions from each patient was selected based on the overall number of active lesions present as determined by the investigator during the first ADES visit. If a patient had 2 active lesions, both of these were photographed. In cases of >2 active lesions, the investigator made an overall judgment of which ones to photograph based on lesion size and the presence of the following clinical signs: (1) excoriation, (2) oozing, (3) erythema, (4) lichenification, (5) dryness, and (6) swelling.

Photograph-Based Severity Assessment

Elements from SCORAD were used to determine the severity of AD from all of the individual photographs. SCORAD is a clinical scoring tool composed of both a subjective (itch and sleep quality) and an objective part (objSCORAD) [10]. ObjSCORAD consists of evaluations of both disease intensity and extent. The intensity part of the SCORAD (iSCORAD) is based on the rating of the following 6 items: erythema, oedema/papulation, excoriations, lichenification, oozing/crusts, and dryness. These items were used to assess all the photographs to obtain a remote iSCORAD. Each item can be graded on a scale of 0 to 3, and the overall intensity score can therefore vary from 0 to 18. Each photograph was presented independently and in a random order to 5 blinded dermatologists on an iPad, on which the dermatologists would rate each of the 6 items on a scale from 0 to 3 and thereby assign a remote iSCORAD. In cases when one item could not be rated from the photograph, the dermatologist would choose "not applicable" for that specific item and the entire photo was consequently discarded.

Statistical Analysis

To calculate one single iSCORAD per patient, the mean of all available photographs for a patient was calculated. This iSCORAD was compared with the clinical assessment performed in person. To investigate the concordance between in-person and photo-based severity assessments, Pearson correlation and Bland Altman plots were performed. To examine the relationship between photo-based assessments and the total severity scoring performed in person, Pearson correlations were used to compare photo-based iSCORAD vs in-person iSCORAD, objSCORAD, and total SCORAD, respectively. Bland-Altman plots were constructed to calculate the average bias and limits of agreement between the methods. To investigate the interrater reliability of the photo-based severity ratings, the intraclass correlation coefficient (ICC) was used. This was performed using the *icc*

function from the *irr* package in R (The R Foundation) [11] and included 95% CI values. For interrater reliability, the ICC estimates were based on two-way random-effects models, absolute agreement, and average measure. An ICC >0.90, 0.75-0.90, 0.50-0.75, and <0.50 indicate an excellent, good, moderate, and poor agreement, respectively [12]. Statistical analyses were performed using the computing environment R (R Core Team) and RStudio (RStudio, PBC).

Results

Of the 95 participants who were assessed by the investigator using the SCORAD in clinic, 50 (52%) were categorized as having mild AD, 36 (38%) as having moderate AD, and 10 (10%) as having severe AD. From these, a total of 942 photos

were evaluated by all 5 dermatologists. The median number of photographed lesions per patient was 3 (range 2-4).

In-Person Assessment vs Remote Assessment

The smartphone-based iSCORAD correlated strongly with the iSCORAD rated in person ($r=0.78$, $P<.001$). In addition, the remote iSCORAD correlated strongly with the objSCORAD ($r=0.81$, $P<.001$), and total SCORAD ($r=0.78$, $P<.001$) obtained in person (Figure 1).

The difference in mean scores for the Bland-Altman plot for the comparison between in-person and remote iSCORAD was 1.31 (Figure 2).

The interrater agreement between the 5 dermatologists assessing the photographs remotely was 0.93 (95% CI 0.911-0.939) (Table 1).

Figure 1. Correlation between the intensity items of Scoring AD (iSCORAD; 0-18) assessed from photos and in-person assessments based on (A) the intensity items of SCORAD, (B) the objective SCORAD (objSCORAD), and (C) total SCORAD.

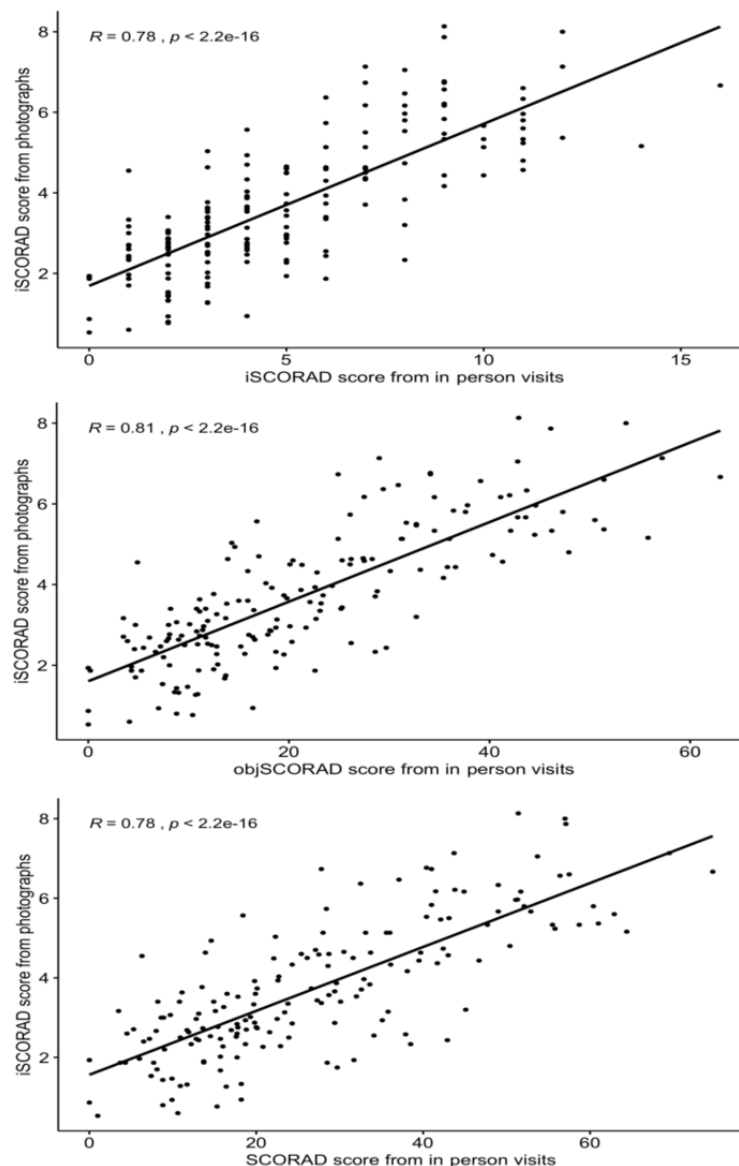


Figure 2. A Bland-Altman plot analyzing the difference between the intensity items of Scoring AD (iSCORAD) assessed in person and from photographs taken via smartphone. The solid line represents the mean difference, the broken line the 1.96 SD, and the dotted line represents zero.

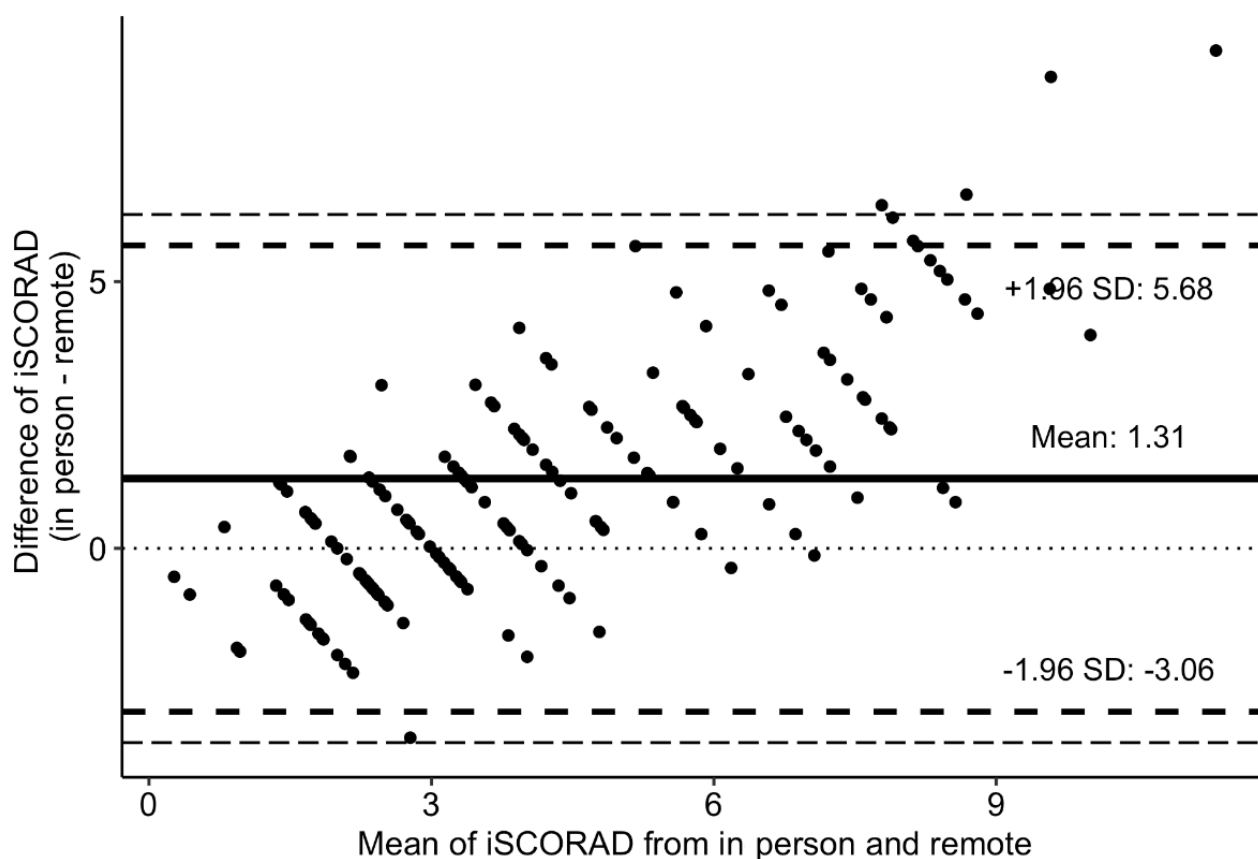


Table 1. Interrater agreement for the remote assessments done by 5 dermatologists for the intensity items of Scoring AD (iSCORAD) for different camera types.

Photographs	ICC ^a (95% CI)
All photographs	0.926 (0.911-0.939)
DSLR ^b	0.932 (0.913-0.947)
Smartphone without flash	0.919 (0.894-0.938)
Smartphone with flash	0.926 (0.908-0.941)

^aICC: intraclass correlation coefficient.

^bDSLR: digital single-lens reflex.

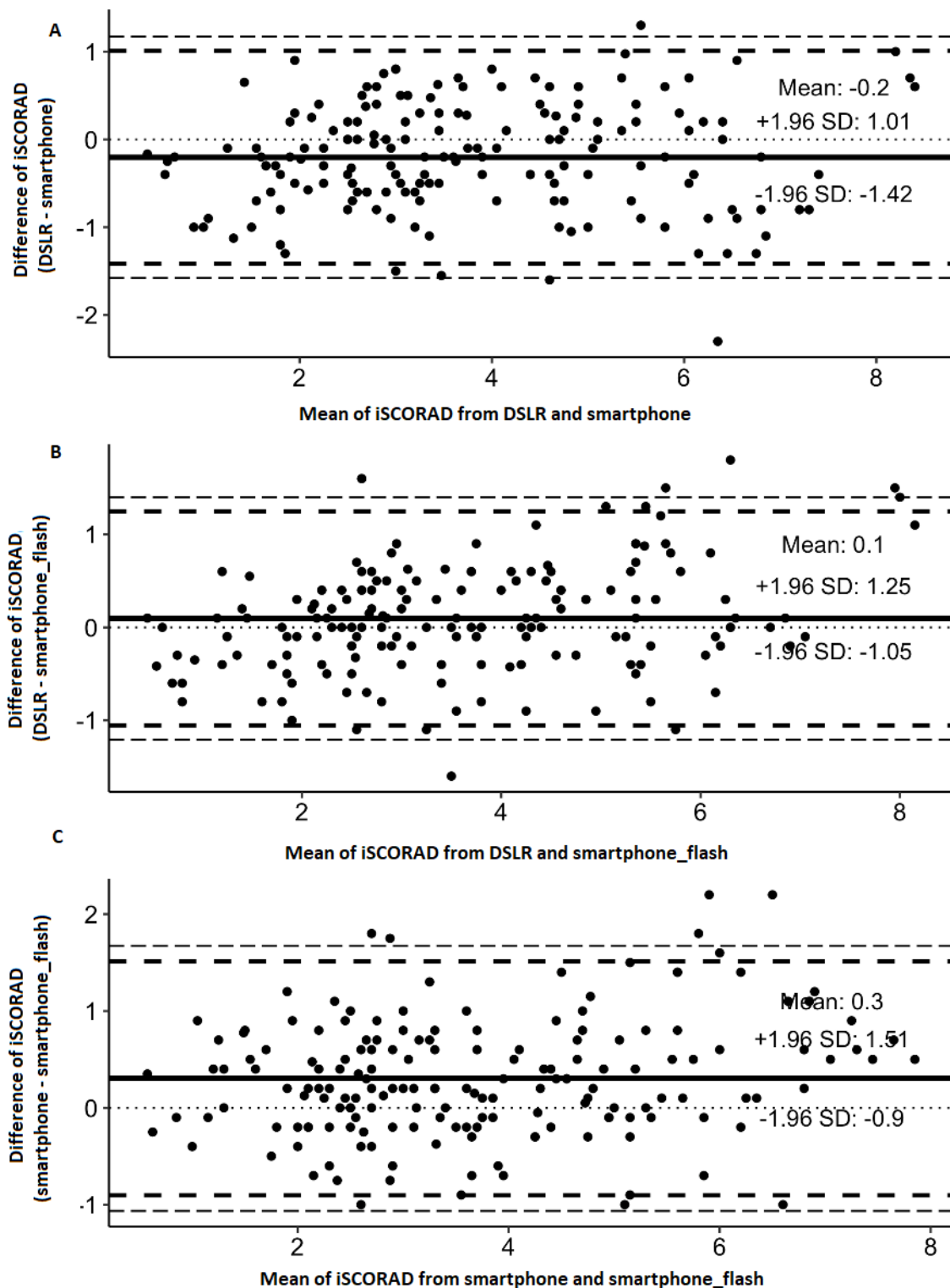
Comparison of the Different Camera Types

In total, 170 lesions were photographed with all 3 camera types (ie, DSLR camera, smartphone without flash, and smartphone with flash). The difference in mean scores was 1.32 (95% CI -3.08 to 5.71), 1.13 (95% CI -3.27 to 5.53), and 1.43 (95% CI -3.05 to 5.92) between in-person evaluation and remote evaluation based on photographs taken by a DSLR camera, a

smartphone without flash, and a smartphone with flash, respectively.

The difference in mean scores for the Bland-Altman plot for the comparison between remote evaluations based on the different camera types was as follows: -0.2 for the DSLR camera and the smartphone without flash, 0.1 for the DSLR camera and the smartphone with flash, and 0.3 for the smartphone with and without flash (Figure 3).

Figure 3. A Bland-Altman plot of the difference between the intensity items of Scoring AD (iSCORAD) remotely assessed based on (A) digital single-lens reflex camera (DSLR) camera and smartphone without flash, (B) DSLR camera and smartphone with flash, and (C) smartphone with and without flash. The solid line represents the mean difference, the broken line the 1.96 SD, and the dotted line represents zero.



Discussion

Principal Findings

In a setup where clinical assessments are conducted remotely, it is important to be certain that the assessments and clinical decisions made are similar to conventional clinical practice (in-person assessment). In this study, we showed that

smartphone-based severity assessments are strongly correlated with in-person assessments. Further, photographs taken with a smartphone are similar to DSLR photographs in the assessment of AD severity using iSCORAD.

It has been demonstrated that the Psoriasis Area and Severity Index can be determined with moderate to good accuracy by dermatologists using standardized digital photos to assess the

severity of psoriasis [13]. In patients with acne, Total Inflammatory Lesion Count was found to be the most reliable way to remotely track progress over time [14], whereas the Leeds technique and the Investigator's Global Assessment designed to grade acne during in-person visits were not reliable in the assessment of digital photos of acne [14]. Further, a pilot study showed that a clinician viewing 3D photos could accurately measure and assess a diabetic foot ulcer remotely [15].

The assessment of AD severity relies on the assessment of clinical manifestations and subjective symptoms, as there is no specific and adequate serological or laboratory test to diagnose or monitor AD. A systematic review performed by Hill et al [16] found 62 different AD severity scales used in clinical trials, of which SCORAD was among the most commonly used. The level of agreement between different raters to give a consistent assessment of AD severity for the same patient has been investigated previously for SCORAD. Bozek et al [17] reported an ICC value of 0.66 for the intrarater reliability for objSCORAD with 10 trained dermatologists assessing 10 adult patients with AD. Zhao et al [18] also investigated in-person reliability for objSCORAD where 12 patients with AD were assessed by 5 trained dermatologists. In that study, an ICC of 0.498 (95% CI 0.234-0.785) and 0.446 (95% CI 0.037-0.730) for interrater and intrarater reliability, respectively, was reported. In another study with full-body photographs of 20 patients with AD of different skin colors assessed by 5 assessors showed that the interrater ICC for objSCORAD was -0.089 for highly pigmented patients, 0.588 for mildly pigmented patients, and 0.586 for nonpigmented patients [18].

In our study, there was a strong and significant correlation between in-person severity assessment and the 5 dermatologists' remote assessments of photographs. Further, the degree of severity assessed remotely based on smartphone photographs was similar to those based on DSLR photographs. The widespread use and ownership of smartphones in the general public may suggest that, with the right training, patients may be able to use their own devices in clinical trials to photograph lesions without compromising the clinical evaluations.

Strengths and Limitations

Our study has both important strengths and limitations that need to be addressed. It is a large study with 5 dermatologists rating hundreds of photographs remotely. The extent to which different

camera types influence severity assessments based on photos has been investigated for the first time. In real-life settings, the photographs will often be taken by a smartphone and not a DSLR camera due to the ubiquity of smartphones in today's society. In VCTs and teledermatology, the photographs will often be taken by the patients themselves and not by the clinician. Therefore, it is important to demonstrate that smartphones are valid tools to collect photographs that can be used to assess severity to the same degree as photographs taken by a DSLR camera. An important limitation is that the photographs used in this study are from a previously conducted study and therefore not collected for the purpose of this research. This explains why only iSCORAD was assessed remotely in our study, since information on disease extent, itch, and sleep quality was not available for remote assessments. Lastly, on average, the in-person intensity ratings were 1 point higher than the ones based on photographs. The trend appears to be linear, meaning that patients with greater severity are increasingly not being scored as "severe" in the remote assessments as they are in person. This could be due to lack of experience by the clinician rating the patients in person since the physician was not a trained dermatologist and the remote assessors were certified dermatologists with at least 5 years of experience in the field. Another explanation could be that the global overview the in-person physician has is lacking when remote assessments are done based on photographs. Future studies should therefore investigate interrater and intrarater reliability between in-person assessment and smartphone photographs taken by the patient at home to investigate the real-world scenario of future virtual trials and use in teledermatology.

Conclusion

In conclusion, this large study based on 5 dermatologists' assessments of hundreds of photographs showed that remote severity assessments are strongly associated with in-person assessments. We also found that smartphones are valid tools to collect photographs and can be used to assess AD severity to the same degree as photographs from a DSLR camera. Further, variation in how the dermatologists rated the iSCORAD based on the photographs was very low. Although this study clearly demonstrates the potential for remote severity assessment of AD, the validity and reliability of the photograph-based methodology should be investigated in a properly designed method-comparison study before implementation.

Conflicts of Interest

ZA and SFT have no conflicts of interest. KMJ, ADA, AC, TB-C, IM, A-MD, ID, A Suru, A Serban, API, PD, and JRZ are employed by Studies&Me, which sponsored this study.

References

1. Fogel DB. Factors associated with clinical trials that fail and opportunities for improving the likelihood of success: A review. *Contemp Clin Trials Commun* 2018 Sep;11:156-164 [FREE Full text] [doi: [10.1016/j.conctc.2018.08.001](https://doi.org/10.1016/j.conctc.2018.08.001)] [Medline: [30112460](https://pubmed.ncbi.nlm.nih.gov/30112460/)]
2. Shore C, Khandekar E, Alper J, editors. *Virtual Clinical Trials Challenges and Opportunities: Proceedings of a Workshop*. Washington, DC: The National Academies Press; 2019.

3. Yang X, Barbieri JS, Kovarik CL. Cost analysis of a store-and-forward teledermatology consult system in Philadelphia. *J Am Acad Dermatol* 2019 Sep;81(3):758-764. [doi: [10.1016/j.jaad.2018.09.036](https://doi.org/10.1016/j.jaad.2018.09.036)] [Medline: [30287316](https://pubmed.ncbi.nlm.nih.gov/30287316/)]
4. Armstrong AW, Ford AR, Chambers CJ, Maverakis E, Dunnick CA, Chren M, et al. Online Care Versus In-Person Care for Improving Quality of Life in Psoriasis: A Randomized Controlled Equivalency Trial. *J Invest Dermatol* 2019 May;139(5):1037-1044 [FREE Full text] [doi: [10.1016/j.jid.2018.09.039](https://doi.org/10.1016/j.jid.2018.09.039)] [Medline: [30481495](https://pubmed.ncbi.nlm.nih.gov/30481495/)]
5. Warshaw EM, Hillman YJ, Greer NL, Hagel EM, MacDonald R, Rutks IR, et al. Teledermatology for diagnosis and management of skin conditions: a systematic review. *J Am Acad Dermatol* 2011 Apr;64(4):759-772. [doi: [10.1016/j.jaad.2010.08.026](https://doi.org/10.1016/j.jaad.2010.08.026)] [Medline: [21036419](https://pubmed.ncbi.nlm.nih.gov/21036419/)]
6. Mounessa JS, Chapman S, Braunberger T, Qin R, Lipoff JB, Dellavalle RP, et al. A systematic review of satisfaction with teledermatology. *J Telemed Telecare* 2017 Mar 28;24(4):263-270. [doi: [10.1177/1357633x17696587](https://doi.org/10.1177/1357633x17696587)]
7. Hughes ME, Aralis H, Bruhn KW, Cotliar J, Craft N, DeLuca IJ, et al. "A Reliability Study Using Network-Oriented Research Assistant (NORA) to Evaluate the Use of Digital Photographs in the Assessment of Atopic Dermatitis". *J Am Acad Dermatol* 2019 Jan 28. [doi: [10.1016/j.jaad.2019.01.043](https://doi.org/10.1016/j.jaad.2019.01.043)] [Medline: [30703454](https://pubmed.ncbi.nlm.nih.gov/30703454/)]
8. Kunz B, Oranje A, Labrèze L, Stalder J, Ring J, Täieb A. Clinical validation and guidelines for the SCORAD index: consensus report of the European Task Force on Atopic Dermatitis. *Dermatology* 1997 Oct 6;195(1):10-19 [FREE Full text] [doi: [10.1159/000245677](https://doi.org/10.1159/000245677)] [Medline: [9267730](https://pubmed.ncbi.nlm.nih.gov/9267730/)]
9. Secretariat of the European Task Force on Atopic, Dermatitis. Severity scoring of atopic dermatitis: the SCORAD index. *Dermatology* 1993 Oct 8;186(1):23-31 [FREE Full text] [doi: [10.1159/000247298](https://doi.org/10.1159/000247298)] [Medline: [8435513](https://pubmed.ncbi.nlm.nih.gov/8435513/)]
10. Chopra R, Vakharia PP, Sacotte R, Patel N, Immaneni S, White T, et al. Relationship between EASI and SCORAD severity assessments for atopic dermatitis. *J Allergy Clin Immunol* 2017 Dec;140(6):1708-1710.e1 [FREE Full text] [doi: [10.1016/j.jaci.2017.04.052](https://doi.org/10.1016/j.jaci.2017.04.052)] [Medline: [28625808](https://pubmed.ncbi.nlm.nih.gov/28625808/)]
11. Gamer ML, Sing I. irr: Various Coefficients of Interrater Reliability and Agreement. CRAN. 2019. URL: <https://CRAN.R-project.org/package=irr> [accessed 2021-04-28]
12. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of Chiropractic Medicine* 2016 Jun;15(2):155-163. [doi: [10.1016/j.jcm.2016.02.012](https://doi.org/10.1016/j.jcm.2016.02.012)]
13. Singh P, Soyer H, Wu J, Salmhofer W, Gilmore S. Tele-assessment of Psoriasis Area and Severity Index: a study of the accuracy of digital image capture. *Australas J Dermatol* 2011 Nov;52(4):259-263. [doi: [10.1111/j.1440-0960.2011.00800.x](https://doi.org/10.1111/j.1440-0960.2011.00800.x)] [Medline: [22070699](https://pubmed.ncbi.nlm.nih.gov/22070699/)]
14. Bergman H, Tsai KY, Seo S, Kvedar JC, Watson AJ. Remote assessment of acne: the use of acne grading tools to evaluate digital skin images. *Telemed J E Health* 2009 Jun;15(5):426-430. [doi: [10.1089/tmj.2008.0128](https://doi.org/10.1089/tmj.2008.0128)] [Medline: [19548822](https://pubmed.ncbi.nlm.nih.gov/19548822/)]
15. Bowling F, King L, Paterson J, Hu J, Lipsky B, Matthews D, et al. Remote assessment of diabetic foot ulcers using a novel wound imaging system. *Wound Repair Regen* 2011;19(1):25-30. [doi: [10.1111/j.1524-475X.2010.00645.x](https://doi.org/10.1111/j.1524-475X.2010.00645.x)] [Medline: [21134035](https://pubmed.ncbi.nlm.nih.gov/21134035/)]
16. Hill MK, Kheirandish Pishkenari A, Braunberger TL, Armstrong AW, Dunnick CA. Recent trends in disease severity and quality of life instruments for patients with atopic dermatitis: A systematic review. *J Am Acad Dermatol* 2016 Nov;75(5):906-917. [doi: [10.1016/j.jaad.2016.07.002](https://doi.org/10.1016/j.jaad.2016.07.002)] [Medline: [27615798](https://pubmed.ncbi.nlm.nih.gov/27615798/)]
17. Božek A, Reich A. Assessment of Intra- and Inter-Rater Reliability of Three Methods for Measuring Atopic Dermatitis Severity: EASI, Objective SCORAD, and IGA. *Dermatology* 2017 May 12;233(1):16-22. [doi: [10.1159/000472711](https://doi.org/10.1159/000472711)] [Medline: [28494438](https://pubmed.ncbi.nlm.nih.gov/28494438/)]
18. Zhao C, Wijayanti A, Doria M, Harris A, Jain S, Legaspi K, et al. The reliability and validity of outcome measures for atopic dermatitis in patients with pigmented skin: A grey area. *Int J Womens Dermatol* 2015 Aug;1(3):150-154 [FREE Full text] [doi: [10.1016/j.ijwd.2015.05.002](https://doi.org/10.1016/j.ijwd.2015.05.002)] [Medline: [28491979](https://pubmed.ncbi.nlm.nih.gov/28491979/)]

Abbreviations

- AD:** atopic dermatitis
- ADES:** Atopisk Dermatit Eksem Studie
- DSLR:** digital single-lens reflex camera
- ICC:** intraclass correlation coefficient
- iSCORAD:** SCORAD-intensity
- objSCORAD:** objective SCORAD
- SCORAD:** Scoring AD
- VCT:** virtual clinical trial

Edited by G Eysenbach; submitted 04.10.20; peer-reviewed by T Schopf, A Dyer, AUR Bacha; comments to author 17.11.20; revised version received 08.12.20; accepted 23.04.21; published 25.05.21.

Please cite as:

Ali Z, Joergensen KM, Andersen AD, Chiriac A, Bjerre-Christensen T, Manole I, Dutei AM, Deaconescu I, Suru A, Serban A, Isberg AP, Dahiya P, Thomsen SF, Zibert JR

Remote Rating of Atopic Dermatitis Severity Using Photo-Based Assessments: Proof-of-Concept and Reliability Evaluation

JMIR Form Res 2021;5(5):e24766

URL: <https://formative.jmir.org/2021/5/e24766>

doi: [10.2196/24766](https://doi.org/10.2196/24766)

PMID: [34032580](https://pubmed.ncbi.nlm.nih.gov/34032580/)

©Zarqa Ali, Kristina Melbardis Joergensen, Anders Daniel Andersen, Andrei Chiriac, Theis Bjerre-Christensen, Ionela Manole, Ana-Maria Dutei, Irina Deaconescu, Alina Suru, Adina Serban, Ari Pall Isberg, Priyanka Dahiya, Simon Francis Thomsen, John Robert Zibert. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 25.05.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Changes in Perceived Stress Following a 10-Week Digital Mindfulness-Based Stress Reduction Program: Retrospective Study

Aarathi Venkatesan¹, PhD; Holly Krymis¹, MSW; Jenny Scharff¹, MSc; Art Waber¹, MFA, NBC-HWC

Vida Health, San Francisco, CA, United States

Corresponding Author:

Aarathi Venkatesan, PhD

Vida Health

100 Montgomery St #750

San Francisco, CA, 94104

United States

Phone: 1 (415) 989 1017

Email: aarathi.venkatesan@vida.com

Abstract

Background: As the need for effective scalable interventions for mental health conditions such as depression, anxiety, and stress has grown, the digital delivery of mindfulness-based stress reduction (MBSR) has gained interest as a promising intervention in this domain.

Objective: This study aims to evaluate the changes in perceived stress following a 10-week digital MBSR program that combined an app-based digital program with weekly one-on-one remote sessions with a health coach.

Methods: This study used a retrospective, observational design. A total of 229 participants with moderate-to-high perceived stress scores as assessed by the Perceived Stress Scale (PSS)-10 enrolled in the 10-week Vida Health MBSR program. The program included weekly remote sessions with a certified health coach and digital content based on concepts fundamental to mindfulness practice. The PSS-10 was used to evaluate perceived stress. Of the 229 participants, 131 (57.2%) were considered program completers and provided at least one follow-up PSS-10. A secondary analysis examined the changes in stress scores at 6 months. This analysis was restricted to participants who had been enrolled in the program for at least 6 months ($n=121$). To account for random and fixed effects, linear mixed effects modeling was used to assess changes in stress scores over time. An intention-to-treat approach was used to evaluate the changes in perceived stress across the entire study cohort, including those who were lost to follow-up. In addition, a reliable change index was computed to evaluate the changes in scores from the baseline.

Results: The findings revealed a significant positive association between program time and stress reduction ($B=-0.365$; $P<.001$) at 12 weeks. We observed an average reduction in stress scores of 3.17 points (95% CI -3.93 to -2.44) by program week 6 and 4.86 points (95% CI -5.86 to -3.85) by program week 12. Overall, 83.2% (109/131) of participants showed a reduction in stress scores by week 12, with 40.5% (53/131) of participants showing reliable improvement at 12 weeks and 47.8% (56/131) of participants showing a shift to a lower stress level category (ie, moderate-to-low stress). The intention-to-treat analysis revealed a significant, although attenuated, reduction in stress scores at 12 weeks ($B=-0.23$; $P<.001$). Participants who completed more lessons had an increased likelihood of moving down at least one stress level category (odds ratio 1.512, 95% CI 1.056 to 2.166; $P=.02$). In assessing medium-term outcomes, among participants who had completed at least 6 months in the program, 48.8% (59/121) of members provided a 6-month assessment. We observed a significant reduction in stress scores at 6 months ($t_{58}=10.24$; $P<.001$), with 61% (36/59) of participants showing reliable improvement.

Conclusions: The findings of this retrospective, observational study suggest that a blended, digital mindfulness-based intervention may support program uptake and meaningful, sustained reduction in stress outcomes.

(JMIR Form Res 2021;5(5):e25078) doi:[10.2196/25078](https://doi.org/10.2196/25078)

KEYWORDS

perceived stress; health coaching; digital mental health intervention; digital therapeutics; mobile phone

Introduction

Background

It is without question that the demand for mental health services has substantially increased in recent times, given the health, social, and economic ramifications of the COVID-19 pandemic [1,2]. Estimates suggest that nearly 40% of the US population will need treatment during their lifespan for anxiety, depression, or other common mental health conditions [3,4]. Most mental health concerns are psychiatric disorders, including anxiety, depression, and other stress-related conditions, that are treatable and, in some cases, preventable [5-7].

Barriers to seek care that are unique to mental health have been well documented in the literature. Digital mental health interventions (DMHIs) delivered via apps, web-based platforms, or text messaging have shown promise in addressing these concerns. They appear to reduce barriers to access conventional forms of mental health services, including cost, mental health stigma, and accessibility [8-10]. In addition, although DMHIs can be heterogeneous in terms of their approach, area of focus, and method of delivery, they appear to be as effective as traditional forms of in-person treatment interventions [11-13].

More recently, there has been a growing interest in the clinical utility of mindfulness-based interventions (MBIs) in improving mental health, stress management, and well-being. Mindfulness meditation is the act of purposefully paying attention to the present moment and being aware of mental states and processes with a sense of openheartedness, curiosity, and kindness in a nonjudgmental manner [14,15]. It has been proposed that increased awareness and nonjudgmental acceptance of experience facilitate emotional regulation and overall well-being [16,17]. Mindfulness-based stress reduction (MBSR) was originally developed as a treatment intervention for reducing continual stress that accompanies chronic pain [14,18]. However, growing evidence suggests that MBIs are as effective in improving both physical and mental health outcomes [16,19,20].

Research on MBIs and mental health outcomes has spanned two decades, and the general consensus has been that MBIs have a significant effect on improving stress and anxiety [21]. However, research focused on the efficacy of app-based or digitally delivered MBIs is still nascent. Recently, in a randomized controlled trial, Flett et al [22] reported short-term improvements in depressive symptoms among college students assigned to a digital mindfulness intervention compared with a control group. Despite the abundance of mindfulness-based DMHIs, Mani et al [23] noted that part of the challenge in evaluating these programs is that they differ substantially in content, particularly with respect to supporting core mindfulness practices, such as acknowledgment of thoughts and emotions, guided meditations, breath awareness techniques, body scans, and yoga movements. This content variability limits efforts to systematically evaluate the efficacy of mindfulness-based DMHIs. Indeed, a recent study by Cavanagh et al [24] observed improvements in perceived stress, anxiety, and depression among university staff and students who participated in a 2-week self-guided MBI compared with a wait-list control. However,

the study found no differences in mental health outcomes between the type of treatment intervention: the intervention arm that incorporated guided mindfulness practice and psychoeducation was no more effective than the psychoeducation-only treatment group. These findings suggest that digital MBIs may differ in their effectiveness from their traditional counterparts because they use different subsets of activities and practices. Despite these inconsistencies, MBIs appear to be effective in stress management and well-being.

In their meta-analytic review of web-based MBIs, Spijkerman et al [25] observed a moderate effect of MBIs on stress ($g=0.51$), suggesting that digital MBIs can be effective in reducing perceived stress. However, the review also noted that most digital MBI studies include brief interventions, ranging from 2 to 8 sessions, with adherence rates between 35% and 92%. A subgroup analysis observed a larger effect size when comparing provider-supported interventions ($g=0.89$) with those lacking therapist guidance ($g=0.19$). These findings suggest that digital interventions that incorporate support by a therapist or coaching may facilitate program efficacy and adherence.

Health coaching has recently emerged as a promising behavioral intervention for improving health outcomes and adherence to mobile app platforms [26,27]. Defined as “a patient-centred process that is based upon behavior change theory and is delivered by health professionals with diverse backgrounds” [26], health coaching is a supportive model in teaching evidence-based interventions, which improve health outcomes by providing individuals with the knowledge, skills, and confidence to manage their health conditions [28-31]. Several frameworks are used by health coaches in their intervention approach, frequently including motivational interviewing and solution-focused goal setting [32-34].

In summary, although there exists an established body of research evaluating the effectiveness of digital MBIs, there has been less focus on the emerging trend of exclusively remote individualized health coaching combined with digital mindfulness tools. Furthermore, there is considerable variability in the structure of digital mindfulness interventions (eg, guided practice vs psychoeducation only and intervention duration). In this study, we evaluated the Vida Health MBSR digital intervention for moderate-to-high perceived stress. Vida Health is a commercially available app that offers tailored digital health programs paired with one-on-one coaching with health education specialists or licensed therapists. The Vida Health app is available in all 50 states with program offerings, including MBSR, cognitive behavioral therapy for depression and anxiety, and chronic disease management.

Objective

The purpose of this study is to evaluate the effect of a mindfulness-based DMHI program delivered via a smartphone on perceived stress. The program pairs individualized remote health coaching, with tailored lessons and tools to introduce and facilitate the practice of mindfulness. Our primary hypothesis was that participants enrolled in the Vida Health MBSR program would show improvements in perceived stress scores at the end of the program. In addition, we predicted that program engagement, as measured by the extent of coach

interaction or program content completed, would be positively associated with reductions in perceived stress scores.

Methods

Study Design

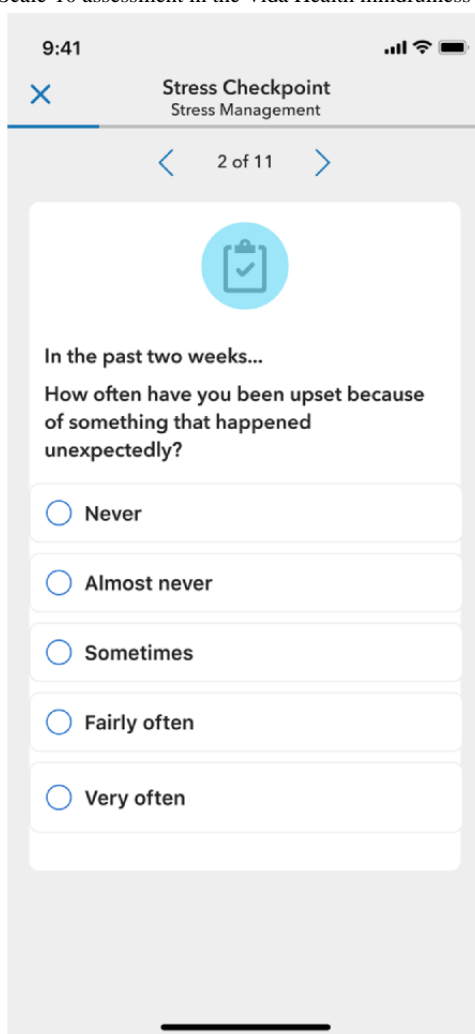
This study used an observational, retrospective design to evaluate changes in perceived stress among participants who completed Vida Health app-based and coach-supported MBSR program. Individuals from across the United States were invited to use the Vida platform, paid for by their employers. Participants joined the Vida MBSR program between March 2018 and May 2020. The institutional review board (Western IRB) waived the requirement for informed consent because the study was determined to have minimal risk, and the data were fully anonymized before analysis.

Outcome Measure

Stress was assessed using the clinically validated 10-item Perceived Stress Scale (PSS-10) [35]. The PSS-10 is an industry-standard assessment instrument designed to measure the perception of stress and how a variety of life situations may occur to them as uncertain, unmanageable, or overburdening. PSS-10 also tracks clients' feelings and thoughts during the intervening periods between assessments [36].

The delivery of the PSS assessment occurred via the Vida app (an example of the assessment is shown in Figure 1). The PSS-10 was sent every 2 weeks during the 10-week program intervention and every 3 months in the postintervention phase. Although participants were encouraged to complete the survey on the day on which it was received, they had the option to complete the assessment up to 2 weeks after receipt. After that point, the next scheduled assessment became available in the app.

Figure 1. A screenshot of the Perceived Stress Scale-10 assessment in the Vida Health mindfulness-based stress reduction digital program.



9:41

Stress Checkpoint
Stress Management

< 2 of 11 >

Never

Almost never

Sometimes

Fairly often

Very often

Study Population and Recruitment

The study included adults aged ≥ 18 years, who owned a smartphone or tablet, were fluent in English, and had a score indicative of moderate perceived stress or higher at program intake, as measured by the PSS (PSS-10 ≥ 14). Participants were recruited between March 2018 and May 2020 from employers

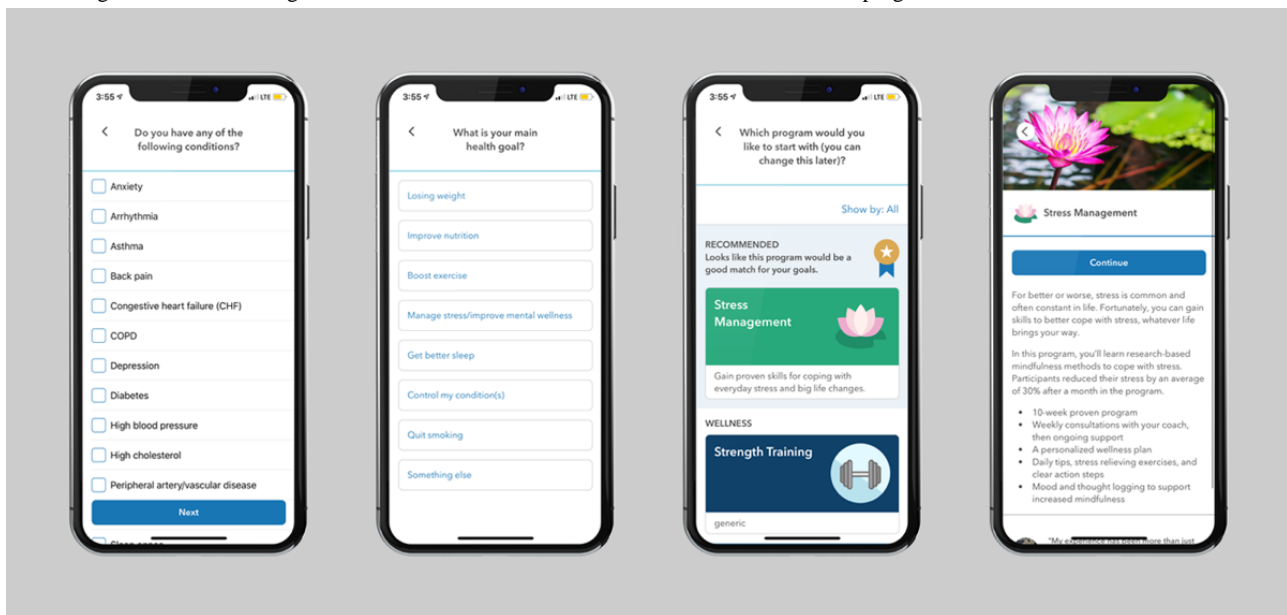
that offered the Vida Health MBSR program to employees and spouses as part of their health plan. They were recruited through a combination of email announcements, paper flyers, and onsite events at their employer and were directed to download the Vida Health app from the App Store (iOS version) or Google Play Store (Android version). On downloading the app, participants enrolled in the program by completing a brief set of registration

questions that included name, contact information, basic demographics (age, gender, height, and weight), existing health conditions, and main health goals (flow shown in Figure 2).

After enrolling in the program, study participants were paired with a certified health coach. Coaches were required to have professional health and wellness coach training from a National Board for Health and Wellness Coaching (NBHWC) accredited program. Coaches performed a video call–based intake in which they determined whether the client was still a good fit for the program. Participants excluded from the study were referred to Vida Health Care Navigators, licensed mental health professionals, who then performed a psychosocial assessment, based on which the participants were referred either internally

to the Vida Health Therapy program or external local resources. Conditions that led to exclusion criteria included self-reported moderate-to-severe depressive or anxiety symptoms, reports of suicidality, homicidality or presenting psychosis, active posttraumatic stress disorder, addiction to substances or alcohol, or significant health problems exacerbated by breathing exercises or yoga. Participants scoring in the high perceived stress range (PSS-10 >27) were also offered the option of completing an assessment with a licensed mental health professional to discuss treatment options (remain in the MBSR program, transfer to the Vida cognitive behavioral therapy program for anxiety or connect with external resources to mental health services).

Figure 2. Program screens showing flow into the Vida Health mindfulness-based stress reduction program.



Study Intervention

Following enrollment, participants received the Vida Health MBSR program, a 10-week digital therapeutic intervention for moderate perceived stress. The intervention was delivered via a mobile app. The evaluated program had three core components: (1) remote personal health coaching with licensed providers, (2) educational lessons and meditation or mindfulness practices, and (3) progress tracking.

At the start of the MBSR program, participants were asked to select a health coach with whom they would have 30-minute weekly video call sessions throughout the course of the 10-week intervention. During the video call sessions, the coach and participant routinely discussed progress, set goals, and reviewed strategies for applying learned concepts and skills to the participants' daily lives. In addition to weekly video calls, participants had unlimited app-based messaging support from their coaches. Digital content, which was based on MBSR techniques and could be audio-, video-, or text-based, was automatically sent to participants' mobile apps every few days. The materials reviewed core concepts of MBSR, including the mind-body connection, meditations, body and breath awareness, and yoga and mindfulness techniques. Health coaches could also send additional content depending on the participants'

needs. Participants were asked to engage with the lessons and practices as they came, but they had access to all of the content throughout the program. Finally, participants were asked to track their stress levels every 2 weeks via the PSS-10 survey. Examples of how the participants experienced these core program components are shown in Figure 3.

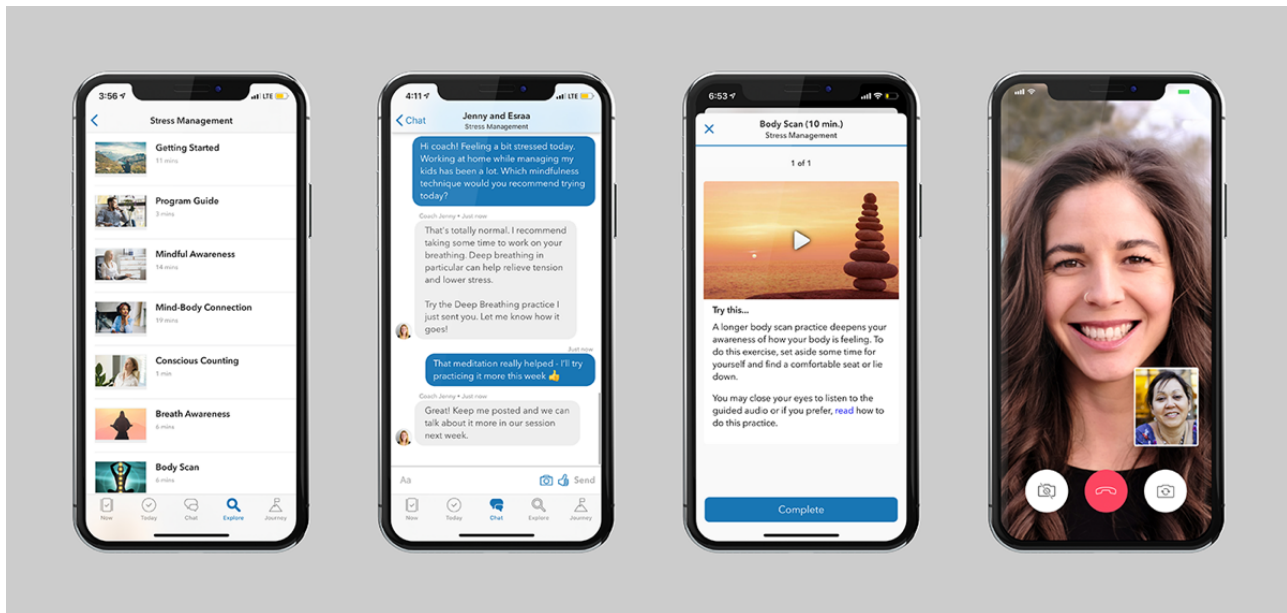
The core intervention involved lessons and tools built on concepts fundamental to integrating daily mindfulness-based approaches. These lessons and mindful practices focused on building mindful awareness, establishing healthy relationships with existing stress or stressors, and integrating skills to increase a participant's autonomy of control over managing and responding to stressors. The lessons and practices were designed for consumption in short intervals, and participants had the option of completing the same lesson or tool multiple times. Were there lessons that were not part of the core?

In addition to the core components of the program, participants received access to various app features to support them in reducing their stress. This included graphs detailing the participant's progress over time and a home screen that used machine learning models to recommend helpful actions for the participant to take or useful information for the participant to know. These models could prompt the participant to read a

specific piece of content, send a message to their coach, log their stress, or view their progress in relation to other members using the Vida app. Participants could also generate habits, or

actions to take in real life, that the app reminded them to complete. Participants were able to access and engage with all the functionality of the Vida app during their leisure time.

Figure 3. Features of the mobile app available to participants completing the mindfulness-based stress reduction program.



Statistical Analyses

All data preparation and analyses were performed using Python version 3.7.9 (Python Software Foundation) and STATA version 16.1 (StataCorp LLC). Baseline was defined as the PSS-10 score completed on the program intake. The week of follow-up assessment completion was computed as the difference in weeks between the survey completion date and the program start date. Baseline differences in stress scores between program completers and noncompleters were assessed using a two-tailed *t* test. As participants were drawn from multiple employers, a Kruskal-Wallis test was performed to assess any potential differences between organizations.

The assessment schedule allowed the participants to complete the PSS-10 survey at any time in a 2-week period. In other words, a participant could complete the survey on the day of receipt or 10 days thereafter. Owing to this inconsistent cadence of completion, only a subset of participants completed the survey in any given week. We used curve fitting, a previously described data imputation technique [37], which can be used to address data sparsity. In brief, we first compiled the time series of PSS-10 scores for each participant. Using the `curve_fit` function from the SciPy library for Python version 3.7.6 [38], 3 different functional forms (linear, quadratic, and sigmoidal) were fit to each participant's data. The fit that yielded the lowest root mean squared error was selected for that participant. This procedure was applied to all participants, and the resulting curves yielded data for all weeks that the participant was in the program.

Program engagement was operationalized as two variables: the number of coach consultations and the number of core lessons completed during the intervention. The program engagement factors were each scaled. As described in an earlier paper [39], to adjust for a significant right skew, the number of completed lessons and the number of messages sent were right Winsorized

at the 99th percentile. This method retains superusers who may complete many lessons during the program but prevents any one participant from excessively influencing the analyses.

To determine the effect of program time and engagement on changes in perceived stress scores, we used a linear mixed effects model (MLM). Linear MLMs address potential heterogeneity in the data because of possible differences across organizations and provider effectiveness. In this analysis, employer organization and health coaches were entered into the model as random effects. Change in PSS-10 scores from the baseline was entered into the model as the outcome variable. Fixed factors included baseline PSS-10 scores, program time in weeks, gender, number of lessons completed, and number of coach consultations. The reliable change index was computed to estimate the proportion of participants who experienced a reliable improvement in stress outcomes at the end of the intervention and at 3 months post intervention [40]. All analyses were performed using the StatsModels module of Python version 3.7.6 [41] and STATA version 16.1.

Results

Overview

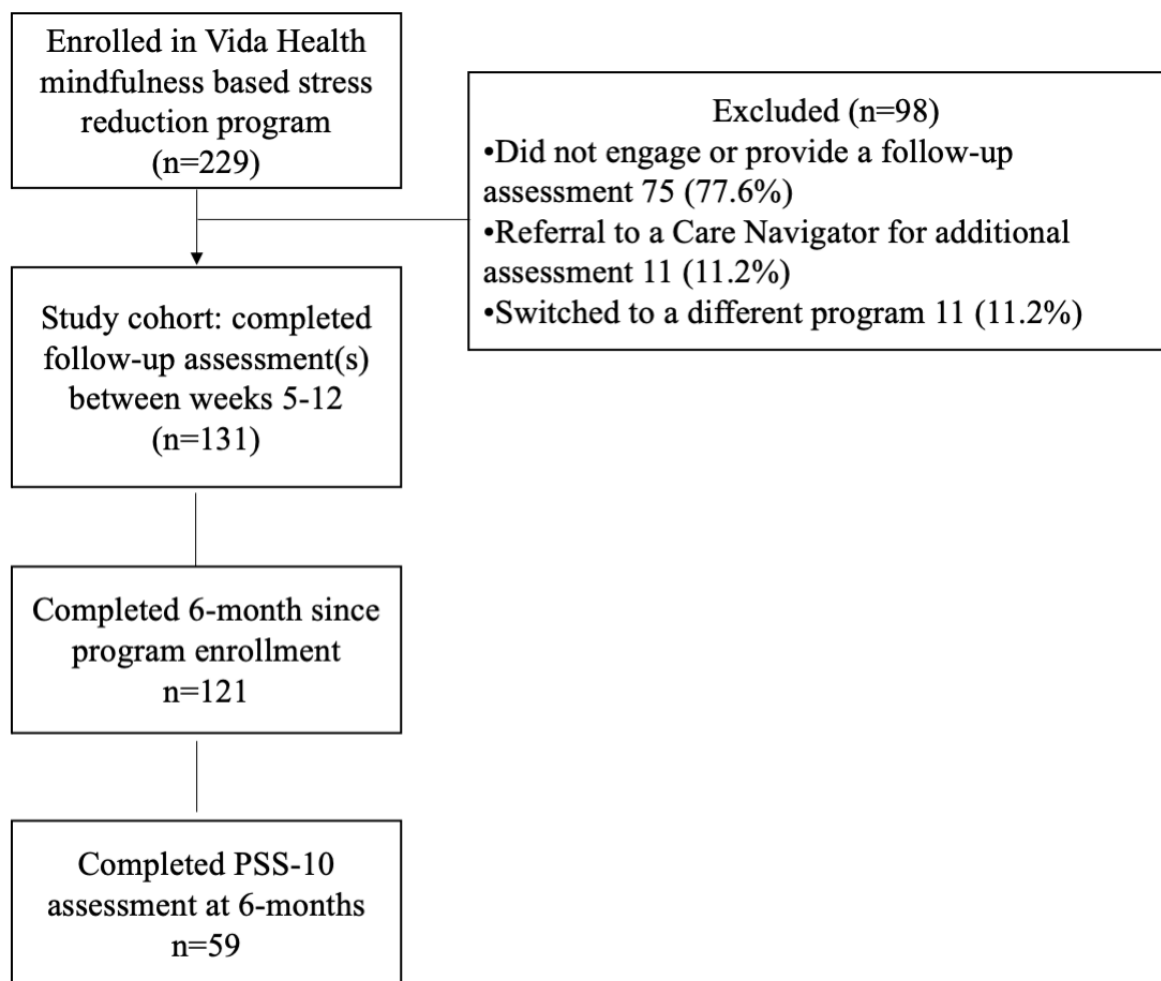
A total of 229 participants with a baseline PSS score ≥ 14 (moderate stress or higher) were enrolled in the Vida MBSR program between March 2018 and May 2020. Study enrollment was restricted to participants who completed at least one in-app program lesson or had at least one coach interaction (video consultation or text message to coach) during the study period. The PSS-10 score reported at program intake was used as the baseline PSS-10 score. A schematic of the participant flow is shown in Figure 4. Of the enrolled participants, 42.7% (98/229) failed to complete at least one follow-up PSS-10 assessment after the first month of the program. In the absence of follow-up

assessment, these participants were considered program noncompleters and were excluded from all subsequent analyses. Unless otherwise noted, analysis was restricted to the treatment cohort, defined as participants who completed at least one follow-up PSS assessment within 5 to 12 weeks from the start of the program. Overall, 57.2% (131/229) of the participants met the inclusion criteria in the treatment cohort.

To address potential systematic baseline differences between program noncompleters and the treatment cohort, we performed a two-tailed *t* test that showed a nonsignificant difference in baseline PSS-10 scores between groups ($P=.10$). However, a

two-tailed chi-square test suggested that more program noncompleters scored in the higher perceived stress range (PSS-10 ≥ 27) than participants in the treatment cohort ($\chi^2_{1}=5.1$; $P=.02$). This is expected, as participants scoring in the higher severity range were typically referred to external services or Vida Health Therapy rather than remaining in the MBSR program. There were no significant gender ($P=.20$) or age-based baseline differences between groups ($P=.45$). As noted earlier, participants were drawn from employer-based organizations. The results of the Kruskal-Wallis test showed no significant differences in baseline PSS-10 scores between organizations ($P=.38$).

Figure 4. A schematic of participant flow. PSS: Perceived Stress Scale.



Baseline Characteristics

The treatment cohort comprised 131 participants. Baseline characteristics are presented in [Table 1](#). Of 131 participants, 121 (92.4%) reported experiencing moderate stress at baseline, with 10 (7.6%) reporting high perceived stress at baseline. The study included more women than men. There was a marginally

significant trend, suggesting that women reported higher levels of perceived stress than men at baseline ($t_{129}=1.93$; $P=.06$). All participants engaged with their coach either via text messages or remote consultations. Of 131 participants, 96 (73.3%) participants had completed at least one consultation with their coach during the intervention period and 130 (99.2%)

participants had messaged their coach during the program in [Table 2](#).
intervention. A summary of program engagement is presented

Table 1. Demographic characteristics of treatment cohort (N=131).

Characteristics	Values
Gender, n (%)	
Female	82 (62.6)
Male	49 (37.4)
Age (years), mean (SD)	
Female	38.02 (10.92)
Male	40.75 (10.72)
Baseline Perceived Stress Scale -10, mean (SD)	
Female	38.02 (10.92)
Male	40.75 (10.72)

Table 2. Summary statistics for program engagement (N=131).

Variables	Consultations (mean 3.87, SD 3.29)		Messages (mean 28.57, SD 5.35)		Lesson completions (mean 10.56, SD 8.21)	
	Correlation coefficients	P value	Correlation coefficients	P value	Correlation coefficients	P value
Consultations	1.0	N/A ^a	0.040	.64	0.094	.28
Messages	0.040	.64	1.0	N/A	0.353	<.001

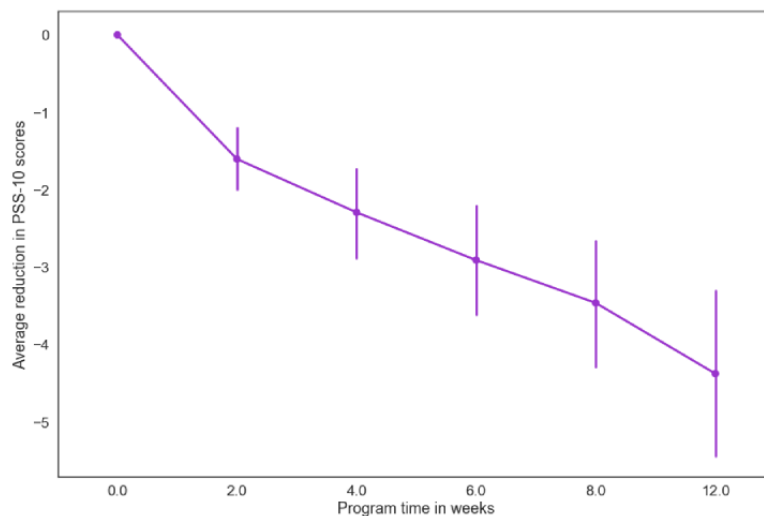
^aN/A: not applicable.

Perceived Stress Outcomes

Overall, 83.2% (109/131) of participants experienced a reduction in PSS-10 scores from baseline by program week 12. Of 131 participants, 56 (47.8%) moved down at least one perceived stress level (ie, moderate-to-low stress) and 53 (40.5%) had a reliable improvement in perceived stress scores. There was a significant effect of program week on reduction in PSS-10 scores relative to baseline (B=-0.365; P<.001) such that increased program time was associated with greater perceived stress

reduction ([Figure 5](#)). We observed an average reduction of 3.17 points (95% CI -3.93 to -2.44) by program week 6 and a reduction of 4.86 points by week 12 (95% CI -5.86 to -3.85). In addition, the analysis revealed a significant inverse association between baseline PSS-10 scores and average reduction (B=-0.402; P<.001). Higher baseline scores were associated with greater reductions in PSS-10 scores by program week 12. We observed a nonsignificant trend, suggesting that women showed a greater reduction in perceived stress than men (B=-0.362; P=.09).

Figure 5. Estimated marginal means of changes in 10-item Perceived Stress Scale as a function of program time. PSS: Perceived Stress Scale.



Engagement-Based Outcomes

The number of coach consultations had no significant effect on the perceived stress scores ($P=.69$). However, there was a significant effect between core lesson completion and reduction in stress scores by program week 12 ($B=-1.420$; $P<.001$). Participants who completed a higher number of core lessons had an increased likelihood of achieving at least one-level reduction in perceived stress scores (odds ratio 1.512, 95% CI 1.056-2.166; $P=.02$).

Medium-Term Perceived Stress Outcomes

In addition, we examined the three-month postintervention outcomes. A total of 92.4% (121/131) of participants had completed at least 6 months from the start of program enrollment. Of this cohort, 48.8% (59/121) provided a follow-up PSS-10 assessment at month 6. A paired t test revealed a significant reduction in PSS-10 scores from baseline ($t_{58}=10.24$; $P<.001$). By 6 months, participants showed an average reduction of 6.77 points (95% CI 5.45 to 8.09) in perceived stress scores relative to baseline. A reduction of at least one stress level was observed in 57% (34/59) of participants, and 61% (36/59) of the participants showed reliable improvement in stress outcomes relative to baseline. A Hedges g calculation suggested a large effect size ($g=1.37$; 95% CI 0.97 to 1.77).

Discussion

Principal Findings

The Vida Health digital MBSR program is an app-based mental health intervention that combines one-on-one weekly remote video sessions with a coach along with tailored digital content based on core concepts of mindfulness practice. The goal of this study is to assess changes in stress outcomes following this MBSR intervention among participants with moderate-to-high perceived stress scores at baseline. The results showed a significant and reliable reduction in perceived stress scores within 12 weeks, which seemed to be maintained at month 6. Higher program engagement, as measured by the completion of core lesson content, was associated with an increased likelihood of a shift to a lower stress-level category (ie, moderate-to-low stress). Although our findings suggest that the Vida Health digital MBSR intervention is associated with improvements in perceived stress, the study design (ie, lack of a control or comparison group) limits our ability to draw causal inferences from these results.

The results of this study are consistent with previous research [12,13,21] that has observed improvements in mental health metrics following digital, app-based interventions. In their meta-analysis, Spijkerman et al [25] observed a stronger effect size for mindfulness interventions that included therapist guidance compared with self-guided interventions. In this study, the frequency of coach consultations was not associated with changes in perceived stress. However, program content completion showed a significant, positive relationship with stress reduction such that participants who completed more core lessons had a greater likelihood of a one-level reduction in stress scores. Coach consults were not correlated with lesson completion; in other words, participants who had more consults

did not appear to complete more lessons. We observed a modest, positive association between the number of messages sent to the coach and lessons completed. Although the Vida MBSR program can be self-guided and completed independently of a coach, it is important to note that all participants included in the study had engaged with their coach (either via text messaging or video consultation) during the intervention period, with 73.3% (96/131) having completed at least one consultation. Although not directly evaluated in this study, it is possible that the benefit of coach guidance was in facilitating program uptake and adherence. It has been noted in earlier research that program adherence or retention is a commonly observed challenge of DMHIs, with uptake rates for DMHIs ranging from 0.5% to 28.6% [40]. In this study, we observed that participants who had more coach consultations were more likely to complete postintervention PSS-10 assessments (odds ratio 1.88, 95% CI 1.39 to 2.54; $P<.001$), suggesting that perhaps the coaches serve as a program anchor. In summary, this study observed a significant reduction in perceived stress scores following a 10-week digital MBSR intervention. In addition, the study provides preliminary insight into the role of program content engagement as a possible moderator of this effect. Together, they lend support to the utility and possible efficacy of digital DMHIs that incorporate sound-validated mental health interventions, adding MBSR to the arsenal of options.

Limitations

This study used a retrospective observational design that lacked a comparison group. The possibility of self-selection bias and the lack of randomization limit the generalizability of our findings and the ability to draw causal inferences regarding the effect of the digital MBSR intervention on stress reduction. As noted earlier, 42.7% (98/229) of the enrolled participants failed to provide an assessment after their initial intake. Although we did not observe any significant systematic demographic differences across program nonstarters and those who remained in the program, it is possible that the groups differed on factors not assessed in this study, such as the presence of other comorbid mental health conditions. We did note that more program nonstarters scored in the high stress range than participants who remained in the program. Although participants with high stress scores were eligible to enroll in this study, participants scoring in the high perceived stress range (PSS-10 >27) were offered the services of a Vida Health Care Navigator to assess the suitability of the MBSR program in addressing their mental health concerns. Indeed, as shown in Figure 4, 22% (22/98) of the program noncompleters had a care navigation consult for external resources to care or switched to a different program. Nevertheless, over three-fourth of program nonstarters failed to engage in the program and provide follow-up assessments. Overall, the observed retention in this study is consistent with previous research that reported an adherence rate ranging from 35% to 92% [25].

Perceived stress was the key outcome metric assessed in this study. However, MBSRs have been associated with improvements in additional measures of mental health, such as depression, anxiety, mood, and well-being [21,22,24]. Additional research is warranted to better define and measure the impact of DMHIs on the treatment of stress management.

Future research should incorporate more comprehensive measures of mental health and well-being to better evaluate the possible benefits of digital MBSR interventions. Although we observed a significant positive association between lesson completion and reduction in perceived stress scores, it is possible that participants who experienced improvement were more motivated to complete additional lessons. Moreover, factors not assessed in this study, such as frequency of mindful meditation practice, may account for the observed association between lesson completion and stress scores. In addition, the lessons incorporated psychoeducation and guided practice. It has been suggested that MBSR interventions that incorporate guided lessons may be no more effective than psychoeducation alone [24]. Although research suggests that therapist-supported DMHIs can be as effective as conventional in-person forms of therapy [12,13], further clarity is needed on the role of certified health coaches in MBI programs. Additional research unpacking patterns of engagement in digital interventions, consumption of program content, and their association with mental health outcomes is warranted.

DMHIs through the integration of mindful awareness lessons, practices, and health coaching can be effective in improving mental health care accessibility, cost-effectiveness, and increasing support services to a larger demographic. Future research should involve equivalence trials comparing DMHIs and in-person behavioral health interventions on MBIs for stress management, further examining the importance of the role of health coaches in DMHIs.

Conclusions

The recent growth and accessibility of smartphones has facilitated the continual development and deployment of mobile-based apps, making it practical for individuals to access the DMHI. Mobile phones facilitate the ability for interventions to enter into the daily lives of individuals, allowing unobtrusive monitoring of activities and contexts, and promote the possibility for interventions at opportune moments, that is, when most needed or desired [42]. Mobile phones are particularly beneficial for mental health care accessibility, as their ownership is largely unrestricted by socioeconomic or demographic status. In addition, they are the preferred form of communication among younger populations, the age group with a decreased likelihood of seeking treatment or support services when affected by mental health conditions [43,44]. MBSR interventions have been shown to be effective in improving mental health outcomes. In this study, adults with moderate-to-high perceived stress completed a 10-week digital MBSR intervention. The intervention paired one-on-one coaching sessions with tailored, guided digital content based on the core concepts of mindfulness practice. We observed significant and reliable postintervention reductions in perceived stress at 12 weeks and 6 months. Although the nonrandomized study design, participant attrition, and the lack of a control group are study limitations, the findings of the study suggest that mindfulness-based digital intervention may be effective in the treatment and management of mental health.

Acknowledgments

The authors are grateful to the outstanding team of Vida Health coaches and their unfailing support from their clients. The data sets analyzed for this study are available from the corresponding author on request.

Authors' Contributions

HK and AW developed program content and supported coach training. Program operations and content delivery were managed by JS. AV developed the concept and design of the study and performed the statistical analyses and data interpretation. Data acquisition was supervised by AV and JS. The manuscript was prepared by AV, JS, and HK.

Conflicts of Interest

AV, HK, JS, and AW are employees of Vida Health and receive compensation through salaries and/or equity.

References

1. 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. [doi: [10.2196/19347](https://doi.org/10.2196/19347)] [Medline: [32459186](https://pubmed.ncbi.nlm.nih.gov/32459186/)]
2. Pfefferbaum B, North CS. Mental health and the COVID-19 pandemic. *N Engl J Med* 2020 Aug 06;383(6):510-512. [doi: [10.1056/NEJMp2008017](https://doi.org/10.1056/NEJMp2008017)] [Medline: [32283003](https://pubmed.ncbi.nlm.nih.gov/32283003/)]
3. Wittchen HU, Jacobi F, Rehm J, Gustavsson A, Svensson M, Jönsson B, et al. The size and burden of mental disorders and other disorders of the brain in Europe 2010. *Eur Neuropsychopharmacol* 2011 Sep;21(9):655-679. [doi: [10.1016/j.euroneuro.2011.07.018](https://doi.org/10.1016/j.euroneuro.2011.07.018)] [Medline: [21896369](https://pubmed.ncbi.nlm.nih.gov/21896369/)]
4. Kessler RC, Chiu WT, Demler O, Merikangas KR, Walters EE. Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. *Arch Gen Psychiatry* 2005 Jun;62(6):617-627 [FREE Full text] [doi: [10.1001/archpsyc.62.6.617](https://doi.org/10.1001/archpsyc.62.6.617)] [Medline: [15939839](https://pubmed.ncbi.nlm.nih.gov/15939839/)]
5. Dragano N, He Y, Moebus S, Jöckel K, Erbel R, Siegrist J, Heinz Nixdorf Recall Study. Two models of job stress and depressive symptoms. Results from a population-based study. *Soc Psychiatry Psychiatr Epidemiol* 2008 Jan;43(1):72-78. [doi: [10.1007/s00127-007-0267-z](https://doi.org/10.1007/s00127-007-0267-z)] [Medline: [17960320](https://pubmed.ncbi.nlm.nih.gov/17960320/)]

6. Harvey SB, Modini M, Joyce S, Milligan-Saville JS, Tan L, Mykletun A, et al. Can work make you mentally ill? A systematic meta-review of work-related risk factors for common mental health problems. *Occup Environ Med* 2017 Mar;74(4):301-310. [doi: [10.1136/oemed-2016-104015](https://doi.org/10.1136/oemed-2016-104015)] [Medline: [28108676](https://pubmed.ncbi.nlm.nih.gov/28108676/)]
7. Knudsen AK, Harvey SB, Mykletun A, Øverland S. Common mental disorders and long-term sickness absence in a general working population. The Hordaland Health Study. *Acta Psychiatr Scand* 2013 Apr;127(4):287-297. [doi: [10.1111/j.1600-0447.2012.01902.x](https://doi.org/10.1111/j.1600-0447.2012.01902.x)] [Medline: [22775341](https://pubmed.ncbi.nlm.nih.gov/22775341/)]
8. Renton T, Tang H, Ennis N, Cusimano MD, Bhalerao S, Schweizer TA, et al. Web-based intervention programs for depression: a scoping review and evaluation. *J Med Internet Res* 2014;16(9):e209 [FREE Full text] [doi: [10.2196/jmir.3147](https://doi.org/10.2196/jmir.3147)] [Medline: [25249003](https://pubmed.ncbi.nlm.nih.gov/25249003/)]
9. Rai M, Vigod SN, Hensel JM. Barriers to office-based mental health care and interest in e-communication with providers: a survey study. *JMIR Ment Health* 2016 Aug 01;3(3):e35 [FREE Full text] [doi: [10.2196/mental.6068](https://doi.org/10.2196/mental.6068)] [Medline: [27480108](https://pubmed.ncbi.nlm.nih.gov/27480108/)]
10. Andrews G, Cuijpers P, Craske MG, McEvoy P, Titov N. Computer therapy for the anxiety and depressive disorders is effective, acceptable and practical health care: a meta-analysis. *PLoS One* 2010;5(10):e13196 [FREE Full text] [doi: [10.1371/journal.pone.0013196](https://doi.org/10.1371/journal.pone.0013196)] [Medline: [20967242](https://pubmed.ncbi.nlm.nih.gov/20967242/)]
11. Karyotaki E, Riper H, Twisk J, Hoogendoorn A, Kleiboer A, Mira A, et al. Efficacy of self-guided internet-based cognitive behavioral therapy in the treatment of depressive symptoms: a meta-analysis of individual participant data. *JAMA Psychiatry* 2017 Apr 01;74(4):351-359. [doi: [10.1001/jamapsychiatry.2017.0044](https://doi.org/10.1001/jamapsychiatry.2017.0044)] [Medline: [28241179](https://pubmed.ncbi.nlm.nih.gov/28241179/)]
12. 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* 2015 Sep 22;448-455. [doi: [10.1093/schbul/sbv132](https://doi.org/10.1093/schbul/sbv132)] [Medline: [26400871](https://pubmed.ncbi.nlm.nih.gov/26400871/)]
13. Firth J, Torous J, Nicholas J, Carney R, Prata A, Rosenbaum S, et al. The efficacy of smartphone-based mental health interventions for depressive symptoms: a meta-analysis of randomized controlled trials. *World Psychiatry* 2017 Oct;16(3):287-298 [FREE Full text] [doi: [10.1002/wps.20472](https://doi.org/10.1002/wps.20472)] [Medline: [28941113](https://pubmed.ncbi.nlm.nih.gov/28941113/)]
14. 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](https://doi.org/10.1037/0022-3514.84.4.822)] [Medline: [12703651](https://pubmed.ncbi.nlm.nih.gov/12703651/)]
15. Kabat-Zinn J. Mindfulness-based interventions in context: past, present, and future. *Clin Psychol Sci Pract* 2003;10(2):144-156. [doi: [10.1093/clipsy.bpg016](https://doi.org/10.1093/clipsy.bpg016)]
16. Baer. Mindfulness training as a clinical intervention: a conceptual and empirical review. *Clin Psychol Sci Pract* 2003;10(2):125-143. [doi: [10.1093/clipsy.bpg015](https://doi.org/10.1093/clipsy.bpg015)]
17. Bishop SR. What do we really know about mindfulness-based stress reduction? *Psychosom Med* 2002;64(1):71-83. [doi: [10.1097/00006842-200201000-00010](https://doi.org/10.1097/00006842-200201000-00010)] [Medline: [11818588](https://pubmed.ncbi.nlm.nih.gov/11818588/)]
18. Kabat-Zinn J, Lipworth L, Burney R. The clinical use of mindfulness meditation for the self-regulation of chronic pain. *J Behav Med* 1985 Jun;8(2):163-190. [doi: [10.1007/bf00845519](https://doi.org/10.1007/bf00845519)]
19. Allen TD, Eby LT, Conley KM, Williamson RL, Mancini VS, Mitchell ME. What do we know about the effects of mindfulness-based training in the workplace? *Ind Organ Psychol* 2015 Dec 17;8(4):652-661. [doi: [10.1017/iop.2015.95](https://doi.org/10.1017/iop.2015.95)]
20. Irving JA, Park-Saltzman J, Fitzpatrick M, Dobkin PL, Chen A, Hutchinson T. Experiences of health care professionals enrolled in mindfulness-based medical practice: a grounded theory model. *Mindfulness* 2012 Sep 19;5(1):60-71. [doi: [10.1007/s12671-012-0147-9](https://doi.org/10.1007/s12671-012-0147-9)]
21. Gotink RA, Chu P, Busschbach JJ, Benson H, Fricchione GL, Hunink MG. Standardised mindfulness-based interventions in healthcare: an overview of systematic reviews and meta-analyses of RCTs. *PLoS One* 2015;10(4):e0124344 [FREE Full text] [doi: [10.1371/journal.pone.0124344](https://doi.org/10.1371/journal.pone.0124344)] [Medline: [25881019](https://pubmed.ncbi.nlm.nih.gov/25881019/)]
22. Flett JA, Hayne H, Riordan BC, Thompson LM, Conner TS. Mobile mindfulness meditation: a randomised controlled trial of the effect of two popular apps on mental health. *Mindfulness* 2018 Oct 31;10(5):863-876. [doi: [10.1007/s12671-018-1050-9](https://doi.org/10.1007/s12671-018-1050-9)]
23. Mani M, Kavanagh DJ, Hides L, Stoyanov SR. Review and evaluation of mindfulness-based iPhone apps. *JMIR Mhealth Uhealth* 2015 Aug 19;3(3):e82 [FREE Full text] [doi: [10.2196/mhealth.4328](https://doi.org/10.2196/mhealth.4328)] [Medline: [26290327](https://pubmed.ncbi.nlm.nih.gov/26290327/)]
24. Cavanagh K, Churchard A, O'Hanlon P, Mundy T, Votolato P, Jones F, et al. A randomised controlled trial of a brief online mindfulness-based intervention in a non-clinical population: replication and extension. *Mindfulness (N Y)* 2018;9(4):1191-1205 [FREE Full text] [doi: [10.1007/s12671-017-0856-1](https://doi.org/10.1007/s12671-017-0856-1)] [Medline: [30100934](https://pubmed.ncbi.nlm.nih.gov/30100934/)]
25. Spijkerman MP, Pots WT, Bohlmeijer ET. Effectiveness of online mindfulness-based interventions in improving mental health: a review and meta-analysis of randomised controlled trials. *Clin Psychol Rev* 2016 Apr;45:102-114 [FREE Full text] [doi: [10.1016/j.cpr.2016.03.009](https://doi.org/10.1016/j.cpr.2016.03.009)] [Medline: [27111302](https://pubmed.ncbi.nlm.nih.gov/27111302/)]
26. Liddy C, Johnston S, Nash K, Ward N, Irving H. Health coaching in primary care: a feasibility model for diabetes care. *BMC Fam Pract* 2014 Apr 03;15:60 [FREE Full text] [doi: [10.1186/1471-2296-15-60](https://doi.org/10.1186/1471-2296-15-60)] [Medline: [24708783](https://pubmed.ncbi.nlm.nih.gov/24708783/)]
27. Hedman A, Karvonen N, Hallberg J, Merilähti J. Designing ICT for health and wellbeing. In: *Ambient Assisted Living and Daily Activities*. Switzerland: Springer; 2014:244-251.
28. Thom DH, Wolf J, Gardner H, DeVore D, Lin M, Ma A, et al. A qualitative study of how health coaches support patients in making health-related decisions and behavioral changes. *Ann Fam Med* 2016 Nov;14(6):509-516 [FREE Full text] [doi: [10.1370/afm.1988](https://doi.org/10.1370/afm.1988)] [Medline: [28376437](https://pubmed.ncbi.nlm.nih.gov/28376437/)]

29. Bennett HD, Coleman EA, Parry C, Bodenheimer T, Chen EH. Health coaching for patients with chronic illness. *Fam Pract Manag* 2010;17(5):24-29 [FREE Full text] [Medline: 21121566]
30. Ghorob A. Health coaching: teaching patients how to fish. *Fam Pract Manag* 2013;20(3):40-42 [FREE Full text] [Medline: 23939739]
31. Carmona-Terés V, Lumillo-Gutiérrez I, Jodar-Fernández L, Rodríguez-Blanco T, Moix-Queralto J, Pujol-Ribera E, et al. Effectiveness and cost-effectiveness of a health coaching intervention to improve the lifestyle of patients with knee osteoarthritis: cluster randomized clinical trial. *BMC Musculoskelet Disord* 2015 Feb 25;16:38 [FREE Full text] [doi: 10.1186/s12891-015-0501-x] [Medline: 25887078]
32. Dwinger S, Dirmaier J, Herbarth L, König HH, Eckardt M, Kriston L, et al. Telephone-based health coaching for chronically ill patients: study protocol for a randomized controlled trial. *Trials* 2013 Oct 17;14:337 [FREE Full text] [doi: 10.1186/1745-6215-14-337] [Medline: 24135027]
33. Huffman MH. Health Coaching: a fresh, new approach to improve quality outcomes and compliance for patients with chronic conditions. *Home Healthc Nurse* 2009 Sep;27(8):490-496. [doi: 10.1097/01.NHH.0000360924.64474.04] [Medline: 19745624]
34. Wong-Rieger D, Rieger FP. Health coaching in diabetes: empowering patients to self-manage. *Can J Diabetes* 2013 Feb;37(1):41-44. [doi: 10.1016/j.jcjd.2013.01.001] [Medline: 24070747]
35. Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. *J Health Soc Behav* 1983 Dec;24(4):385-396. [Medline: 6668417]
36. Cohen S. Perceived stress in a probability sample of the United States. In: *The Claremont Symposium on Applied Social Psychology. The social psychology of health*. Thousand Oaks, California, United States: SAGE Publications; 1988:31-67.
37. Venkatesan A, Rahimi L, Kaur M, Mosunic C. Digital cognitive behavior therapy intervention for depression and anxiety: retrospective study. *JMIR Ment Health* 2020 Aug 26;7(8):e21304 [FREE Full text] [doi: 10.2196/21304] [Medline: 32845246]
38. Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, SciPy 1.0 Contributors. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat Methods* 2020 Mar;17(3):261-272 [FREE Full text] [doi: 10.1038/s41592-019-0686-2] [Medline: 32015543]
39. Silberman JM, Kaur M, Sletteland J, Venkatesan A. Outcomes in a digital weight management intervention with one-on-one health coaching. *PLoS One* 2020;15(4):e0232221 [FREE Full text] [doi: 10.1371/journal.pone.0232221] [Medline: 32353035]
40. Jacobson NS, Truax P. Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. *J Consult Clin Psychol* 1991 Feb;59(1):12-19. [Medline: 2002127]
41. Skipper S, Perktold J. Statsmodels: econometric and statistical modeling with Python. In: *Proceedings of the 9th Python in Science Conference 2010*. 2010 Presented at: 9th Python in Science Conference; June 28 - July 3, 2010; Austin, Texas. [doi: 10.25080/majora-92bf1922-011]
42. Atienza AA, Patrick K. Mobile health: the killer app for cyberinfrastructure and consumer health. *Am J Prev Med* 2011 May;40(5 Suppl 2):151-153. [doi: 10.1016/j.amepre.2011.01.008] [Medline: 21521588]
43. Fleming T, Bavin L, Lucassen M, Stasiak K, Hopkins S, Merry S. Beyond the trial: systematic review of real-world uptake and engagement with digital self-help interventions for depression, low mood, or anxiety. *J Med Internet Res* 2018 Jun 06;20(6):e199 [FREE Full text] [doi: 10.2196/jmir.9275] [Medline: 29875089]
44. Oliver MI, Pearson N, Coe N, Gunnell D. Help-seeking behaviour in men and women with common mental health problems: cross-sectional study. *Br J Psychiatry* 2005 Apr;186:297-301 [FREE Full text] [doi: 10.1192/bjp.186.4.297] [Medline: 15802685]

Abbreviations

- DMHI:** digital mental health intervention
- MBI:** mindfulness-based intervention
- MBSR:** mindfulness-based stress reduction
- PSS:** Perceived Stress Scale

Edited by G Eysenbach; submitted 16.10.20; peer-reviewed by L Tan, A Graham; comments to author 26.11.20; revised version received 26.01.21; accepted 11.04.21; published 25.05.21.

Please cite as:

Venkatesan A, Krymish H, Scharff J, Waber A

Changes in Perceived Stress Following a 10-Week Digital Mindfulness-Based Stress Reduction Program: Retrospective Study

JMIR Form Res 2021;5(5):e25078

URL: <https://formative.jmir.org/2021/5/e25078>

doi: [10.2196/25078](https://doi.org/10.2196/25078)

PMID: [34032571](https://pubmed.ncbi.nlm.nih.gov/34032571/)

©Aarathi Venkatesan, Holly Krymis, Jenny Scharff, Art Waber. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 25.05.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Utilizing Lean Software Methods To Improve Acceptance of Global eHealth Initiatives: Results From the Implementation of the Basic Emergency Care App

Christian Rose¹, MD; Taylor Nichols², BA, MD; Daniel Hackner, BSc; Julia Chang², BA, MD; Steven Straube², BA, MD; Willem Jooste³, MD; Hendry Sawe⁴, MD; Andrea Tenner², BA, MPH, MD

¹Department of Emergency Medicine, Stanford University School of Medicine, Palo Alto, CA, United States

²Department of Emergency Medicine, University of California, San Francisco, San Francisco, CA, United States

³Department of Emergency Medicine, University of Cape Town/Stellenbosch University, Cape Town, South Africa

⁴Department of Emergency Medicine, Muhimbili University of Health and Allied Sciences, Dar es Salaam, United Republic of Tanzania

Corresponding Author:

Christian Rose, MD

Department of Emergency Medicine
Stanford University School of Medicine
900 Welch Road
Suite 350
Palo Alto, CA, 94304
United States
Phone: 1 4159159585
Email: ccrose@stanford.edu

Abstract

Background: Health systems in low- and middle-income countries face considerable challenges in providing high-quality accessible care. eHealth has had mounting interest as a possible solution given the unprecedented growth in mobile phone and internet technologies in these locations; however, few apps or software programs have, as of yet, gone beyond the testing phase, most downloads are never opened, and consistent use is extremely rare. This is believed to be due to a failure to engage and meet local stakeholder needs and the high costs of software development.

Objective: World Health Organization Basic Emergency Care course participants requested a mobile point-of-care adjunct to the primary course material. Our team undertook the task of developing this solution through a community-based participatory model in an effort to meet trainees' reported needs and avoid some of the abovementioned failings. We aimed to use the well-described Lean software development strategy—given our familiarity with its elements and its ubiquitous use in medicine, global health, and software development—to complete this task efficiently and with maximal stakeholder involvement.

Methods: From September 2016 through January 2017, the Basic Emergency Care app was designed and developed at the University of California San Francisco. When a prototype was complete, it was piloted in Cape Town, South Africa and Dar es Salaam, Tanzania—World Health Organization Basic Emergency Care partner sites. Feedback from this pilot shaped continuous amendments to the app before subsequent user testing and study of the effect of use of the app on trainee retention of Basic Emergency Care course material.

Results: Our user-centered mobile app was developed with an iterative participatory approach with its first version available within 6 months and with high acceptance—95% of Basic Emergency Care Course participants felt that it was useful. Our solution had minimal direct costs and resulted in a robust infrastructure for subsequent assessment and maintenance and allows for efficient feedback and expansion.

Conclusions: We believe that utilizing Lean software development strategies may help global health advocates and researchers build eHealth solutions with a process that is familiar and with buy-in across stakeholders that is responsive, rapid to deploy, and sustainable.

(*JMIR Form Res* 2021;5(5):e14851) doi:[10.2196/14851](https://doi.org/10.2196/14851)

KEYWORDS

lean; eHealth; emergency; global health; app development; decision support; primary survey; mHealth; Africa; Tanzania; low- and middle income countries; LMIC

Introduction

eHealth in Low- and Middle-Income Countries

Health systems in low- and middle-income countries continue to face considerable challenges in providing high-quality affordable accessible care [1]. Simultaneously, they have experienced an unprecedented increase in the number of users of mobile phone and internet technologies, as well as a decline in the price of devices and services [2]. Furthermore, most medical professionals report that smartphone and mobile technologies are useful for training and education [3]. As a result, many health program implementers and policy makers are exploring the extent to which eHealth, defined as “the use of information and communication technologies for health” [4], can help address the challenges faced by resource-constrained health markets in terms of the availability, quality, and financing of health care [5-10].

However, few of these initiatives succeed as hoped or progress past the pilot phase. Although eHealth has the potential to strengthen health systems worldwide, consistent use is extremely rare and the evidence base is immature [11]. Consequently, the opportunities to advance knowledge remain limited in scope [12]. Limitations in adoption are likely multifactorial, but they often fail to address the needs of diverse stakeholders [13]. eHealth apps often have unique end-users, and health care scientists and researchers are not often well-versed in how software engineers operate as traditional scientific measures may be too rigid to identify the nuanced differences between user subgroups [14]. Lastly, software and apps are expensive to develop, averaging between a cost of US \$23,000-28,000 (with minimal features) to \$270,000 (for robust apps)—a cost many community-based projects simply cannot afford [15-17].

These limitations in software development are not new, however. Early in the history of software product development and the expansion of the internet, there was concern over excessive amounts of money being spent to develop technologies that did not meet user needs. Years were sometimes spent building apps which would later be found to have been based on false assumptions, were uselessly developed for an outdated technology, or the demand for a solution disappeared by the end of the software production cycle. Development cycles were too long and rarely had early stakeholder buy-in or feedback. In response, developers looked for strategies to remain responsive to an ever evolving market while building solutions as quickly and efficiently as possible.

One of the most commonly employed methods evolved from the Toyota Production System, later coined as Lean, a systematic method for waste reduction [18]. The aim of employing Lean in software development was to use this approach to build solutions that created value for customers, eliminated waste, empowered the developers, and allowed for continuous improvement. Not only has Lean methodology become commonplace in software development, its use has been growing

in global health care and as a validated method to employ evidence-based practices and continuous quality improvement in diverse of practical settings [18-20].

Basic Emergency Care

The 2015 Global Burden of Disease Study [21] found that emergency medical diseases contributed to more than half of all years of life lost globally. While this continues to be a global burden, it is most acutely felt by low- and middle-income countries, which have 4.4 times the disease burden of high-income countries [21]. Moreover, it is estimated that over half of all deaths in these countries are potentially addressable by emergency care [22]. Well-organized emergency care systems play an important role in the delivery of emergency services to and the health outcomes of patients in low-resource settings [23]; however, frontline providers in low- and middle-income countries often lack the basic training to recognize and treat life-threatening conditions [24], and the accessibility of adequate emergency care in low-resource settings is limited [25]. The burden of acute diseases, coupled with the lack of emergency care training in low- and middle-income countries, begets unnecessary mortality.

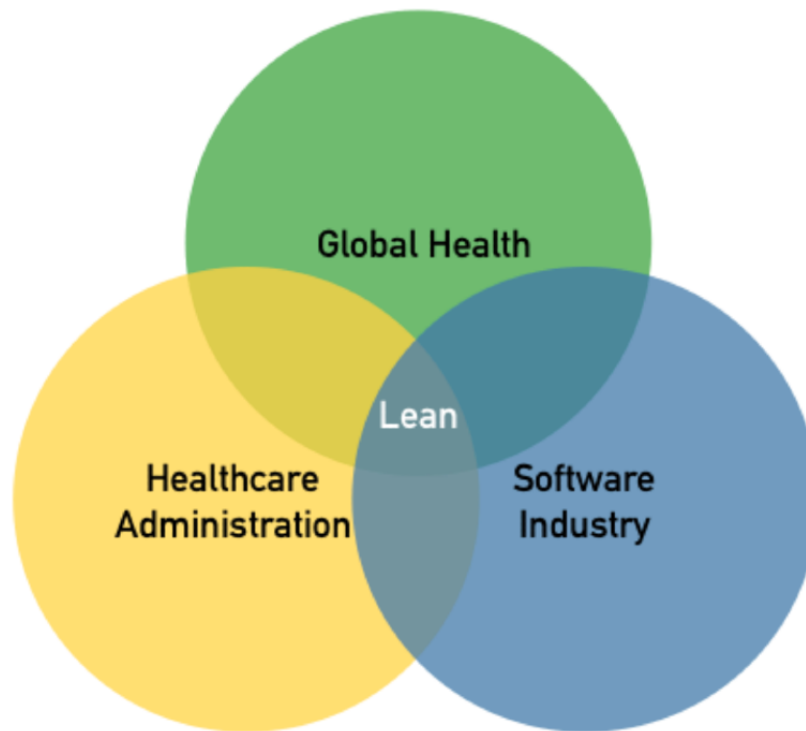
Prioritizing an integrated approach to early recognition and resuscitation substantially reduces morbidity and mortality associated with emergent medical conditions [26,27]. As such, the World Health Organization developed an open-access Basic Emergency Care course to provide standardized training in basic assessment and life-saving techniques using a traditional lecture format. The course covers the Airway, Breathing, Circulation, Disability, and Exposure approach to pediatric and adult emergencies, trauma, respiratory distress, shock, pregnancy, and altered mental status. The Basic Emergency Care course is not, in itself, an eHealth intervention. It is traditional didactic training, consisting of a 5-day series of classes, aimed at teaching participants how to recognize and intervene in a number of emergent conditions (the most significant causes of morbidity and mortality

Participants are trained to conduct a primary survey of emergency situations they may encounter situation, using the mnemonic *Airway, Breathing, Circulation, Disability, and Exposure*, which is, in itself, a concise and standardized algorithm; however, many branch points and deviations as well as condition-specific interventions are possible. But while there are a large number of emergent medical conditions annually, each practitioner may only encounter a few of any particular type in a given year. Diagrammatic representations of the possible emergency situations that one might encounter contained multiple decision points and branches; thus, their complexity made them too unwieldy to be useful. Furthermore, not all emergent conditions take place in a hospital, with many occurring in the field and being managed by emergency medical technicians. Thus, a mobile point-of-care reference was suggested as a means to help individuals recall Basic Emergency Care course training.

While there are many software development models to choose from, such as Waterfall or Scrum, we believe that the breadth of use and familiarity of Lean across health care administration, the software industry, and global health make it the ideal model to be used to address some of the inherent difficulties of global health software development and implementation (Figure 1) [28,29]—each is necessary for developing eHealth interventions and each has robust familiarity with this methodology

[19,30,31]. Global health experts and their stakeholders are necessary to understand the landscape of issues needing to be addressed within a community. Software developers need to be able to communicate with global health experts to understand those needs and the scope of technological solutions. Finally, health care administrators must be able to respond to their community and clinician needs as well as implement the developed solutions to work within their institution.

Figure 1. Stakeholder overlap with Lean methodologies.

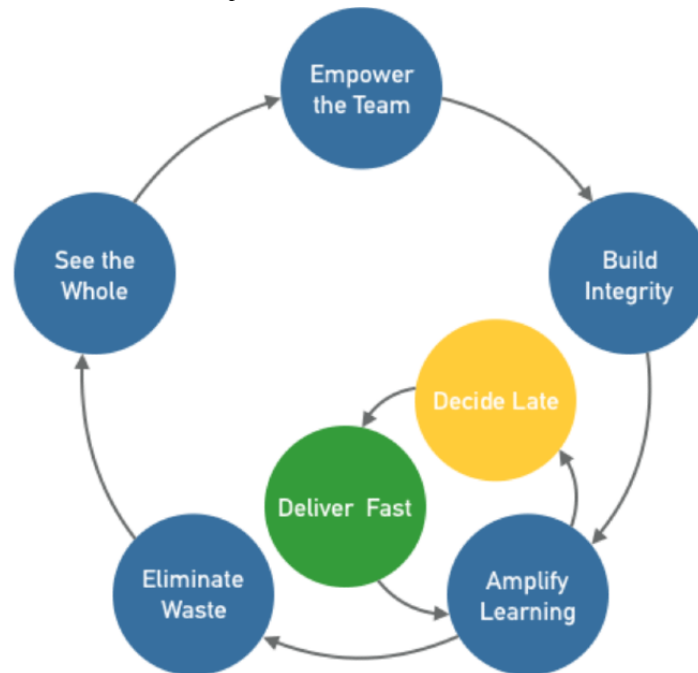


Given that the Lean process makes use of limited resources, empowers local stakeholders, and facilitates remote development, it may prove to be a successful model for wide-scale implementations [30]. We aimed at employing a standardized Lean software development strategy to quickly develop and iterate a community-based global eHealth initiative for diverse stakeholders.

Methods

Lean Strategy

Lean software development has 7 core elements [32]: seeing the whole, empowering the team, building integrity in, amplifying learning, deciding as late as possible, delivering as fast as possible, and eliminating waste (Figure 2).

Figure 2. Lean workflow in global eHealth software development.

Seeing the Whole

Complex systems are composed of networks of interconnected components that influence each other, often in a nonlinear fashion [33]. Health care, specifically global health care, is a complex system. Problems are rarely simple and rarely have clear boundaries [34,35]. In order to develop broad and novel solutions, teams must be able to see the whole and avoid a hammer-looking-for-nails approach. In global health care, this

is often exacerbated by a tendency to focus on squeaky wheels [36] that distract from the true underlying issue—which may be more difficult to address. To address this problem, Lean systems use root-cause analyses. While there are many different types of analyses, one simple approach is called *The 5 Whys* (Figure 3) [37,38]. This is an iterative interrogation technique used to explore the cause-and-effect relationships underlying a particular problem with the goal of determining the root cause by repeating the question “Why?”

Figure 3. The 5 Whys for the Basic Emergency Care app.

- 1: Why are trainees having difficulty with the ABCDE pathway?
A: Because it is complex to remember.
- 2: Why is it complex to remember?
A: Because it has many branching points.
- 3: Why are many branching points difficult to remember?
A: Because they are difficult to follow.
- 4: Why are they difficult to follow?
A: Because they are difficult to visualize.
- 5: Why are they difficult to visualize?
A: Because mapped out it would take up an entire wall.

Initial project planning using this approach began in early 2016 with input from health care providers and software developers. Because groups tend to overestimate the importance of their particular skill set, solutions to a root-cause analysis must be derived from stakeholders without a priori assumptions. For the Basic Emergency Care app, it was participants who referenced their access and regular use of mobile phones for medical purposes which guided the solution.

Empowering the Team

Team empowerment strategies have been shown to improve many elements of health care, from individual management of disease, public health strategies, or improving partnership and sustainability in global health [39,40]. Team empowerment requires: (1) enhancing meaningfulness of work, (2) fostering participation in decision making, (3) facilitating goal accomplishment, and (4) providing autonomy and freedom from bureaucratic restraints that can cause unnecessary delays and bottlenecks [41].

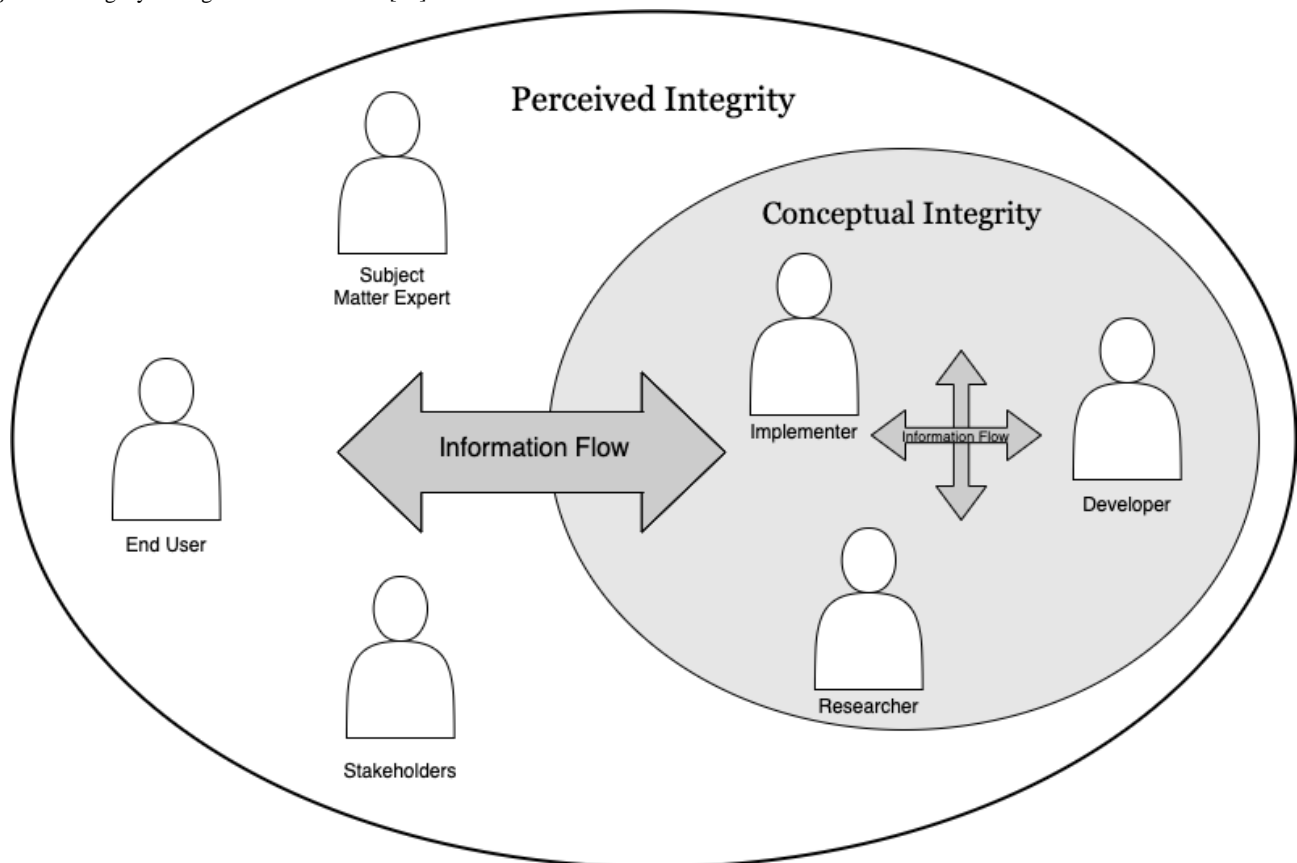
For the Basic Emergency Care app, team empowerment began in mid-2016 with a shared vision of both the problem and solution. Developers, who lacked the experience of emergency department care or global health initiatives, spent time with those who had, in the form of structured meetings and ongoing email communication, throughout the app development and testing process. Efforts were made to make videoconferencing available to developers whenever possible during testing so that helpful features and hurdles to use could be identified and addressed in real time. Participation across the development team was fostered through a variety of free and secure resources, such as Slack (a collaboration hub for projects which allows document sharing and conferencing) and GitHub (a version control system widely used in software engineering which allows for shared to-do lists and assignment of tasks). By fostering shared decision making through open communication and regular feedback, team members were able to develop a plan for addressing their tasks, which was then shared with the team, and suggest changes as necessary based on their role in

the team. This allowed for participation in decision making and autonomy in areas of subject-matter expertise.

Building Integrity In

In the software engineering, *integrity* is described as a freedom from flaw, defect, or decay [42]; a solution should have internal and external consistency. That is, it should make intuitive sense to the end user how to interact with the solution (ie, perceived integrity), and the internal functioning of the solution should be the appropriate use of engineering and organization resources (ie, conceptual integrity). Software is usually expected to evolve gracefully as it adapts to the future. Solutions with integrity have a coherent architecture; score highly on usability and fitness for their purpose; and are maintainable, adaptable, and extensible. Research has shown that integrity comes from wise leadership, relevant expertise, effective communication, and healthy discipline [32]. Thus, integrity is achieved through excellent detailed information flow—from users to the development team and between the development team members (Figure 4) [43].

Figure 4. Integrity through information flow [30].



Achieving integrity for the Basic Emergency Care app first required an empowered team capable of communicating and receiving feedback as described above. During app development from September 2016 through January 2017, all team conversations and amendments to software were recorded using Slack and GitHub, allowing for easy sharing with new members as they joined the team smoothing onboarding and preventing delays when team members transitioned. Furthermore, feedback channels for remote team members and site leads or prototype

testers were maintained to easily provide feedback and suggestions to optimize the solution's fit.

Amplifying Learning

Gathering information is an important step in solving new problems. For complex problems, the preferred approach to a solution is to use the scientific method: observe, create a hypothesis, devise an experiment to test the hypothesis, run the experiment, and see if the results are consistent with the hypothesis [32,44]. Today it is widely accepted that software

design is a problem-solving process that involves discovering solutions through short repeated cycles of investigation, experimentation, and results validation. In Lean development, this is considered the sum of 2 other steps—waiting to as long as possible to finalize a decision or acting quickly and soliciting feedback.

Learning from past interventions has been difficult in prior eHealth initiatives. This is sometimes due to rigid statistical measures that attempt to measure the effect of the software itself rather than how it corresponds to the principles on which it is based [45]. Furthermore, many initiatives do not even make it past the pilot stages, thus limiting the broad awareness of their successes and failures due to publication bias [7,45]. Implementation science measures such as reach, effectiveness, or acceptance are rarely reported, making it difficult to learn from prior endeavors or predict successful strategies. Incorporation of theory-informed frameworks for understanding factors associated with adoption and nonadoption are essential for evaluating eHealth software and generating knowledge that can be applied to other settings [46].

Given that we expected the Basic Emergency Care app to be used in many different settings, we aimed to develop the app with avenues of concomitant feedback that would allow for product improvements and implementation improvements. This included readily available and standardized survey methods and a grounded theory approach to leading focus groups. We conducted these sessions, in Tanzania and South Africa from January 11, 2017 through February 5, 2017. Feedback from each session was used to modify the app; iterative changes to the app were tested in subsequent groups.

Deciding as Late as Possible

Development practices that provide for late decision making are effective in domains that involve uncertainty because they provide an option-based approach. A key strategy for delaying commitments when developing a complex solution is to build a capacity for change into the system. Delayed decisions are valuable because they can be made based on facts rather than speculation. This is also sometimes referred to as concurrent development or just-in-time production [47,48]. Concurrent development allows for a breadth-first approach, to discover big costly problems early on in the development process (before it is too late). Concurrent development, as opposed to sequential development, means a team starts programming the highest value features as soon as high-level conceptual designs have been determined; detailed requirements may still be undergoing investigation. In addition to ensuring against costly mistakes, concurrent development is considered to be the best method for dealing with changing requirements [32]. This is because not only are the big decisions deferred while considering all options, but the little decisions are deferred as well. When change is expected, concurrent development reduces delivery time and overall cost, while improving the performance of the final product. But concurrent software development also has costs. It requires having developers with enough expertise in a particular domain (ie, mobile, machine learning, etc) to anticipate where the emerging design is likely to lead and having collaboration between customers and analysts [32].

During initial user feedback sessions for the Basic Emergency Care app, some participants suggested only allowing the user to progress linearly through the entire pathway thus forcing users to complete the entirety of the primary survey every time they use the app, but a particular subgroup requested the ability to simply search course material as soon as they identified a problem. Given that whether users would be allowed to jump through the app or progress linearly fashion would require significantly different interfaces and functionality (ie, a search bar and the ability to bypass certain checkpoints), it was necessary to test both options and delay decision, until the options had been vetted by domain experts or one had to be deployed. During the first iteration of this process in February 2017, engineers examined the product roadmap to determine whether this feature would even be possible, while end users provided feedback about feature necessity and value.

Delivering as Fast as Possible

While seemingly contradictory, *delivering as fast as possible* complements *deciding as late as possible*. The faster a change or feature can be delivered, the longer decisions can be delayed before testing. Furthermore, without speed, it is not possible to have reliable feedback. In software development, the discovery cycle—design, implement, feedback, improve—is critical for learning. The shorter these cycles, the more can be learned.

For the Basic Emergency Care app navigation, a software solution that made both options possible, was chosen early during app development (in late 2016). Team members tested this solution for bugs. Both options were later deployed to determine the impact of each option on user satisfaction, app use, or recall of the course material.

Eliminating Waste

As the name implies, eliminating waste is the core concept of the Lean development strategy. Waste can come in many forms; from monetary to time. Anything that does not directly add value to the end-user is a waste. There are some particular methods that have been described to see and eliminate waste in software development, called the 7 wastes: partially done work, extra processes, extra features, task switching, waiting, motion, and defects [49]. Eliminating waste is a central tenant of the Lean development strategy. Most authors suggest starting with an approach to eliminate waste. However, in global health, eliminating waste requires a holistic view of the problem, an understanding of what is possible, and stakeholder involvement as well as feedback before assumptions are made. Thus, we have this pivotal step at the end as a reminder to assess the prior work in context. Many of the strategies to eliminate waste are addressed by the other 6 core elements of Lean. Solutions with integrity limit motion and task-switching by allowing for bidirectional feedback from users to the engineering team so that workflows may be taken into account in the solution. Partially done work is limited when decisions are made only when absolutely necessary. Waiting is minimized by delivering as fast as possible. Defects are identified by amplifying learning. Finally, extra features and processes are only identified by seeing the whole and understanding the purpose or utility of each within the greater context of addressing the problem at hand.

For the Basic Emergency Care app, waste could be considered in 2 dimensions: cost and time. Waste in both of these dimensions could result from unnecessary features or delays in feedback or delivery of requested features. Thus, we utilized continuous feedback from stakeholders and iterative testing to ensure that we were not expanding the scope or necessary elements of our solution. This helped us to limit engineering

demands and to give directed instructions to developers, which helped to limit costs and time. We were able to create our pilot app (Figure 5) for the cost of a single Android and iOS app developer (who was hired for a per-project fee of \$3000) and web-hosting (\$96 annually). Development took 6 months (from initial scope description until pilot completion).

Figure 5. Screenshot of the Basic Emergency Care app.

The screenshot shows a mobile application interface for a breathing assessment. At the top, there is a dark blue header with a search bar and a menu icon. Below the header, the section is titled "Section: Breathing". The interface consists of four questions, each in a dark blue rounded rectangle, with corresponding answer buttons. The first question is "Is the patient speaking?" with a "Yes" button. The second question is "Are there abnormal sounds concerning for airway compromise? Examples: Raspy voice, stridor, snoring or gurgling." with a "No" button. The third question is "Is the patient breathing?" with a "Yes" button. The fourth question is "Are there signs of increased work of breathing?" with a "No" button (red) and a "Yes" button (green). At the bottom of the screen, there is a small icon of a person's head and neck.

Results

In a convenience sample of participants from across all levels of training, the majority (56/59, 95%) believed that the Basic Emergency Care app was “useful for their practice.” Responses included that it was an effective means to help “remind (me) to complete the primary survey on all patients” and prevented “missing steps.” A subgroup of physicians trained in emergency medicine (9/30, 30%) reported that they would prefer if the tool allowed “searching conditions” or references for “specific interventions” rather than a rigid algorithm as “I already know what the problem is, but now I want to know how to address it.” This feedback was used to design the subsequent versions of the app to improve acceptance and usability for all target providers.

Discussion

General

Our goal was to develop a mobile app through the use of a model that was well-known to each participant stakeholder in a global health initiative: global health, health care administration, and software engineering. We found that in this setting, the Lean software development model was effective for limiting costs, addressing stakeholder needs and local practices, and bringing solutions to scale quickly.

Online calculators estimated a cost between \$11,700 to \$71,000 to develop an app such as the Basic Emergency Care app with similar data and infrastructure needs [50,51]. However, we were able to build our pilot app for a fraction of that cost at total

monetary cost of \$3096 for software development and annual web-hosting. Of course, we recognize that there are other costs associated with development that must be taken into account. While the monetary cost of our app was low, it took much longer to develop than the estimates given for professional app development. But we believe that this speed was likely balanced by the benefit of continuous feedback in a participatory model which helped to empower the team and build integrity [52-55]. The participatory model emphasized by the Lean software development model is particularly helpful in global health app development where needs span geographic, linguistic, and technological barriers.

There is a poignant need for mobile solutions in low- and middle-income countries. Many health care providers there have mobile devices, and there has been an increase in the development of apps to support care in these locations [2]. However, global health systems are complex and have different practices, system needs, and user demographics [13]. Thus, meaningful resources should not be built entirely external to the primary stakeholders, as many have been to date. However, access to software engineers can be limited for global health interventions that do not often have long-term scalability, need for full-time engineering services, or typical funding. When professional engineering services are employed, time with the engineers is often limited, due to cost or access, and familiarizing them with local needs can take time, limiting their contributions.

However, we do not believe this must necessarily be the case. Given that the Lean model allows for robust stakeholder involvement, rapid iteration, and a holistic approach to problems while sharing a common language for global health, software development, and health care administration, we believe that it should be the model of choice for software development in low- and middle-income countries.

Limitations

This study had several limitations that need to be considered. First, the communities of Tanzania and South Africa may not be representative of other low- and middle-income countries outside of sub-Saharan Africa. The goal is for the Basic Emergency Care app to be available worldwide, though it is currently only available in English-speaking countries. However, part of utilizing a model such as Lean is to promote the practice and skills necessary to generalize solutions to any community.

Acknowledgments

The authors gratefully acknowledge the dedication and input of all participating general practices and trial participants. CR was supported in part by a Veteran's Affairs Office of Academic Affiliations Advanced Fellowship in Informatics.

Conflicts of Interest

None declared.

References

1. Mills A. Health care systems in low- and middle-income countries. *N Engl J Med* 2014 Feb 06;370(6):552-557. [doi: [10.1056/NEJMra1110897](https://doi.org/10.1056/NEJMra1110897)] [Medline: [24499213](https://pubmed.ncbi.nlm.nih.gov/24499213/)]

We believe that the process of *seeing the whole*, in conjunction with the other core elements, make this model particularly well-suited to communities that may be different from those in which we tested our model and software.

Additionally, while our primary university affiliation is in the heart of a major metropolitan software and technology development area (ie, Silicon Valley), other programs may not find it as easy to engage a rounded team of health professionals, engineers, and stakeholders locally in their community. The Lean iterative process requires the timely engagement of team members to effectively suggest, implement, and revise eHealth software. This can be especially challenging when team members operate in different time zones and with different cultural backgrounds. However, as training and education for computer engineers continues to increase, with accessibility in-person and online, there continues to be growth of local engineering capabilities across the globe. The model utilized above is meant to be agnostic to the problem being addressed and the particulars of coding languages.

Similarly, new software development strategies, such as Scrum, continue to gain popularity, and newer models evolve all the time. It's possible that these strategies will prove to be even more efficient than Lean for producing and iterating software products. However, Lean has gained broad acceptability in a multitude of industries across borders and specialties since it was first introduced. This makes it particularly well-suited to maintaining the robust source of knowledge and integrity necessary for consistent global health interventions.

Conclusion

Lean processes have become a standard in the software development industry, but the methodology also has a proven track and has gained broader acceptance in the health care administration and global health communities. Thus, the Lean methodology presents a particularly suited framework which is already available to most of the parties involved in developing global eHealth solutions. Furthermore, its ideals—such as empowering the team, eliminating waste, and having a big-picture view, are implicitly aligned with those of global health interventions. We believe Lean strategies applied to software development for global health initiatives, particularly for low- and middle-income countries, may address some of the concerns regarding the prior limitations of these interventions.

2. Beratarrechea A, Lee AG, Willner JM, Jahangir E, Ciapponi A, Rubinstein A. The impact of mobile health interventions on chronic disease outcomes in developing countries: a systematic review. *Telemed J E Health* 2014 Jan;20(1):75-82 [FREE Full text] [doi: [10.1089/tmj.2012.0328](https://doi.org/10.1089/tmj.2012.0328)] [Medline: [24205809](https://pubmed.ncbi.nlm.nih.gov/24205809/)]
3. Robinson T, Cronin T, Ibrahim H, Jinks M, Molitor T, Newman J, et al. Smartphone use and acceptability among clinical medical students: a questionnaire-based study. *J Med Syst* 2013 Jun;37(3):9936. [doi: [10.1007/s10916-013-9936-5](https://doi.org/10.1007/s10916-013-9936-5)] [Medline: [23526127](https://pubmed.ncbi.nlm.nih.gov/23526127/)]
4. Digital health. World Health Organization. URL: <https://www.who.int/ehealth/en/> [accessed 2019-05-13]
5. Nilsen W, Kumar S, Shar A, Varoquiers C, Wiley T, Riley WT, et al. Advancing the science of mHealth. *J Health Commun* 2012;17 Suppl 1:5-10. [doi: [10.1080/10810730.2012.677394](https://doi.org/10.1080/10810730.2012.677394)] [Medline: [22548593](https://pubmed.ncbi.nlm.nih.gov/22548593/)]
6. Boudreaux ED, Waring ME, Hayes RB, Sadasivam RS, Mullen S, Pagoto S. Evaluating and selecting mobile health apps: strategies for healthcare providers and healthcare organizations. *Transl Behav Med* 2014 Dec;4(4):363-371 [FREE Full text] [doi: [10.1007/s13142-014-0293-9](https://doi.org/10.1007/s13142-014-0293-9)] [Medline: [25584085](https://pubmed.ncbi.nlm.nih.gov/25584085/)]
7. Mechael P, Batavia H, Kaonga N, Searle S, Kwan A, Goldberger A, et al. Barriers and gaps affecting mHealth in low and middle income countries: policy white paper. *Global Health Delivery Online*. 2010. URL: https://www.ghdonline.org/uploads/Barriers_Gaps_to_mHealth_in_LMICs_-_White_Paper_-_May_2010.pdf [accessed 2021-05-04]
8. mHealth for development: the opportunity of mobile technology for health care in the developing world. Relief Web. 2009. URL: https://reliefweb.int/sites/reliefweb.int/files/resources/A695C439809709CF49257689001019B9-UNF_Feb2009.pdf [accessed 2021-05-04]
9. Germanakos P, Mourlas C. A mobile agent approach for ubiquitous and personalized eHealth information systems. 2005 Presented at: UM05 - Personalization for eHealth; July 29; Edinburgh, Scotland URL: <http://cgi.csc.liv.ac.uk/~floriana/UM05-eHealth/UM05-W6.pdf#page=73>
10. Kirwan M, Vandelanotte C, Fenning A, Duncan MJ. Diabetes self-management smartphone application for adults with type 1 diabetes: randomized controlled trial. *J Med Internet Res* 2013 Nov 13;15(11):e235 [FREE Full text] [doi: [10.2196/jmir.2588](https://doi.org/10.2196/jmir.2588)] [Medline: [24225149](https://pubmed.ncbi.nlm.nih.gov/24225149/)]
11. Peiris D, Miranda JJ, Mohr DC. Going beyond killer apps: building a better mHealth evidence base. *BMJ Glob Health* 2018;3(1):e000676 [FREE Full text] [doi: [10.1136/bmjgh-2017-000676](https://doi.org/10.1136/bmjgh-2017-000676)] [Medline: [29527353](https://pubmed.ncbi.nlm.nih.gov/29527353/)]
12. Sondaal SFV, Browne JL, Amoakoh-Coleman M, Borgstein A, Miltenburg AS, Verwijs M, et al. Assessing the effect of mhealth interventions in improving maternal and neonatal care in low- and middle-income countries: a systematic review. *PLoS One* 2016;11(5):e0154664 [FREE Full text] [doi: [10.1371/journal.pone.0154664](https://doi.org/10.1371/journal.pone.0154664)] [Medline: [27144393](https://pubmed.ncbi.nlm.nih.gov/27144393/)]
13. Ben-Zeev D, Schueller SM, Begale M, Duffecy J, Kane JM, Mohr DC. Strategies for mHealth research: lessons from 3 mobile intervention studies. *Adm Policy Ment Health* 2015 Mar;42(2):157-167 [FREE Full text] [doi: [10.1007/s10488-014-0556-2](https://doi.org/10.1007/s10488-014-0556-2)] [Medline: [24824311](https://pubmed.ncbi.nlm.nih.gov/24824311/)]
14. Mohr DC, Weingardt KR, Reddy M, Schueller SM. Three problems with current digital mental health research . . . and three things we can do about them. *Psychiatr Serv* 2017 May 01;68(5):427-429 [FREE Full text] [doi: [10.1176/appi.ps.201600541](https://doi.org/10.1176/appi.ps.201600541)] [Medline: [28412890](https://pubmed.ncbi.nlm.nih.gov/28412890/)]
15. Zhang M, Chew E, Ho CS, Ng BY, Ho R, Cheok CCS. Application of low-cost methodologies for mobile phone app development. *JMIR Mhealth Uhealth* 2014 Dec 09;2(4):e55 [FREE Full text] [doi: [10.2196/mhealth.3549](https://doi.org/10.2196/mhealth.3549)] [Medline: [25491323](https://pubmed.ncbi.nlm.nih.gov/25491323/)]
16. How much does app development cost? *Business of Apps*. URL: <http://www.businessofapps.com/app-developers/research/app-development-cost/> [accessed 2019-03-11]
17. Neagu A. Figuring the costs of mobile app development. *Formotus*. URL: <https://www.formotus.com/blog/figuring-the-costs-of-custom-mobile-business-app-development> [accessed 2019-03-11]
18. Black JR, Miller D. *The Toyota Way to Healthcare Excellence: Increase Efficiency and Improve Quality with Lean*. Chicago: Health Administration Press; Oct 03, 2008.
19. D'Andreamatteo A, Ianni L, Lega F, Sargiacomo M. Lean in healthcare: a comprehensive review. *Health Policy* 2015 Sep;119(9):1197-1209. [doi: [10.1016/j.healthpol.2015.02.002](https://doi.org/10.1016/j.healthpol.2015.02.002)] [Medline: [25737260](https://pubmed.ncbi.nlm.nih.gov/25737260/)]
20. Womack J, Byrne A, Fiume O. *Going lean in health care*. American Academy of Otolaryngology-Head and Neck Surgery.: Institute for Healthcare Improvement URL: <https://www.entnet.org/sites/default/files/GoingLeaninHealthCareWhitePaper-3.pdf> [accessed 2021-05-04]
21. Razzak J, Usmani M, Bhutta Z. Global, regional and national burden of emergency medical diseases using specific emergency disease indicators: analysis of the 2015 Global Burden of Disease Study. *BMJ Glob Health* 2019;4(2):e000733 [FREE Full text] [doi: [10.1136/bmjgh-2018-000733](https://doi.org/10.1136/bmjgh-2018-000733)] [Medline: [30997158](https://pubmed.ncbi.nlm.nih.gov/30997158/)]
22. Debas HT, Donkor P, Gawande A, Jamison DT, Kruk ME, Mock CN. *Essential surgery*. In: *Disease Control Priorities*. Washington, DC: World Bank Publications; 2015.
23. Smith J, Haile-Mariam T. Priorities in global emergency medicine development. *Emerg Med Clin North Am* 2005 Feb;23(1):11-29. [doi: [10.1016/j.emc.2004.09.012](https://doi.org/10.1016/j.emc.2004.09.012)] [Medline: [15663971](https://pubmed.ncbi.nlm.nih.gov/15663971/)]
24. Crisp N, Chen L. Global supply of health professionals. *N Engl J Med* 2014 Mar 06;370(10):950-957. [doi: [10.1056/NEJMra1111610](https://doi.org/10.1056/NEJMra1111610)] [Medline: [24597868](https://pubmed.ncbi.nlm.nih.gov/24597868/)]

25. Hodkinson P, Wallis L. Emergency medicine in the developing world: a Delphi study. *Acad Emerg Med* 2010 Jul;17(7):765-774 [FREE Full text] [doi: [10.1111/j.1553-2712.2010.00791.x](https://doi.org/10.1111/j.1553-2712.2010.00791.x)] [Medline: [20653593](https://pubmed.ncbi.nlm.nih.gov/20653593/)]
26. Reynolds T, Sawe H, Rubiano A, Shin S, Wallis L, Mock C. Strengthening health systems to provide emergency care. In: *Disease Control Priorities: Improving Health and Reducing Poverty*. Washington (DC): The International Bank for Reconstruction and Development/The World Bank; 2018.
27. Hsia R, Razzak J, Tsai AC, Hirshon JM. Placing emergency care on the global agenda. *Ann Emerg Med* 2010 Aug;56(2):142-149. [doi: [10.1016/j.annemergmed.2010.01.013](https://doi.org/10.1016/j.annemergmed.2010.01.013)] [Medline: [20138398](https://pubmed.ncbi.nlm.nih.gov/20138398/)]
28. Balaji S, Murugaiyan M. Waterfall vs v-model vs agile: a comparative study on SDLC. *Int J Inform Tech Bus Manag* 2012 Jun 29:26-30 [FREE Full text]
29. Schwaber K, Beedle M. *Agile Software Development With Scrum*. Saddle River, NJ: Prentice Hall PTR; Oct 2001.
30. Carter P, Desmond J, Akanbobnaab C, Oteng R, Rominski S, Barsan WG, et al. Optimizing clinical operations as part of a global emergency medicine initiative in Kumasi, Ghana: application of Lean manufacturing principals to low-resource health systems. *Acad Emerg Med* 2012 Mar;19(3):338-347 [FREE Full text] [doi: [10.1111/j.1553-2712.2012.01311.x](https://doi.org/10.1111/j.1553-2712.2012.01311.x)] [Medline: [22435868](https://pubmed.ncbi.nlm.nih.gov/22435868/)]
31. Vallon R, Wenzel L, Brüggemann ME, Grechenig T. An Agile and Lean process model for mobile app development: case study into Austrian industry. *J Softw* 2015 Nov;10(11):1245-1264 [FREE Full text] [doi: [10.17706/jsw.10.11.1245-1264](https://doi.org/10.17706/jsw.10.11.1245-1264)]
32. Poppendieck M, Poppendieck T. *Lean Software Development: An Agile Toolkit*. Boston, MA: Addison-Wesley Professional; May 2003.
33. Amaral L, Ottino J. Complex networks. *Eur Phys J B* 2004 Mar 1;38(2):147-162 [FREE Full text] [doi: [10.1140/epjb/e2004-00110-5](https://doi.org/10.1140/epjb/e2004-00110-5)]
34. Hidalgo CA, Blumm N, Barabási AL, Christakis NA. A dynamic network approach for the study of human phenotypes. *PLoS Comput Biol* 2009 Apr 10;5(4):e1000353 [FREE Full text] [doi: [10.1371/journal.pcbi.1000353](https://doi.org/10.1371/journal.pcbi.1000353)] [Medline: [19360091](https://pubmed.ncbi.nlm.nih.gov/19360091/)]
35. Pourbohloul B, Kieny M. Complex systems analysis: towards holistic approaches to health systems planning and policy. *Bull World Health Organ* 2011 Apr 01;89(4):242-242 [FREE Full text] [doi: [10.2471/BLT.11.087544](https://doi.org/10.2471/BLT.11.087544)] [Medline: [21479083](https://pubmed.ncbi.nlm.nih.gov/21479083/)]
36. Senge P. *The fifth discipline, the art and practice of the learning organization*. Nonprofit Manag Leadersh 1991 May;30(5):37-37 [FREE Full text] [doi: [10.1002/pfi.4170300510](https://doi.org/10.1002/pfi.4170300510)]
37. Ohno T. *Toyota Production System: Beyond Large-Scale Production*. Boca Raton, FL: CRC Press; 1988.
38. Serrat O. The five whys technique. In: *Knowledge Solutions: Tools, Methods, and Approaches to Drive Organizational Performance*. Singapore: Springer Singapore; 2017:307.
39. Leffers J, Mitchell E. Conceptual model for partnership and sustainability in global health. *Public Health Nurs* 2011;28(1):91-102. [doi: [10.1111/j.1525-1446.2010.00892.x](https://doi.org/10.1111/j.1525-1446.2010.00892.x)] [Medline: [21198819](https://pubmed.ncbi.nlm.nih.gov/21198819/)]
40. Wallerstein N. What is the evidence on effectiveness of empowerment to improve health? World Health Organization Regional Office for Europe. 2006. URL: https://www.euro.who.int/_data/assets/pdf_file/0010/74656/E88086.pdf [accessed 2021-05-04]
41. MacPhee M, Chang L, Lee D, Spiri W. Global health care leadership development: trends to consider. *J Healthc Leadersh* 2013 Jun;21 [FREE Full text] [doi: [10.2147/jhl.s23010](https://doi.org/10.2147/jhl.s23010)]
42. Sterling C. Building integrity in. *Getting Agile*. URL: <http://www.gettingagile.com/2006/10/31/building-integrity-in/> [accessed 2021-05-04]
43. Fujimoto T, Clark K. *Product Development Performance: Strategy, Organization, and Management in the World Auto Industry*. Cambridge, MA: Harvard Business School; 1991.
44. Reinertsen D. *Managing the Design Factory*. New York, NY: Free Press; 1997.
45. Geng EH, Peiris D, Kruk ME. Implementation science: relevance in the real world without sacrificing rigor. *PLoS Med* 2017 Apr 25;14(4):e1002288 [FREE Full text] [doi: [10.1371/journal.pmed.1002288](https://doi.org/10.1371/journal.pmed.1002288)] [Medline: [28441435](https://pubmed.ncbi.nlm.nih.gov/28441435/)]
46. Greenhalgh T, Wherton J, Papoutsi C, Lynch J, Hughes G, A'Court C, et al. Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. *J Med Internet Res* 2017 Nov 01;19(11):e367 [FREE Full text] [doi: [10.2196/jmir.8775](https://doi.org/10.2196/jmir.8775)] [Medline: [29092808](https://pubmed.ncbi.nlm.nih.gov/29092808/)]
47. Adachi T, Enkawa T, Shih LC. A concurrent engineering methodology using analogies to just-in-time concepts. *Int J Prod Res* 2007 May 03;33(3):587-609 [FREE Full text] [doi: [10.1080/00207549508930167](https://doi.org/10.1080/00207549508930167)]
48. Sohlenius G. Concurrent engineering. *CIRP Ann Manuf Technol* 1992;41(2):645-655 [FREE Full text] [doi: [10.1016/s0007-8506\(07\)63251-x](https://doi.org/10.1016/s0007-8506(07)63251-x)]
49. Poppendieck M. Principles of lean thinking. Software Engineering Group Universidad Nacional de San Luis. 2002. URL: <http://www.sel.unsl.edu.ar/ApuntesMaes/Anteriores/MetodologiasAgiles/LeanThinking.pdf> [accessed 2021-05-04]
50. Estimate my app. Oozou. URL: <https://estimatemyapp.com/> [accessed 2021-05-04]
51. How much does it cost to build an app. Epic AI. URL: <http://www.howmuchtobuildanapp.io/Questions> [accessed 2021-05-04]
52. Cornwall A, Jewkes R. What is participatory research? *Soc Sci Med* 1995 Dec;41(12):1667-1676 [FREE Full text] [doi: [10.1016/0277-9536\(95\)00127-s](https://doi.org/10.1016/0277-9536(95)00127-s)]
53. Wallerstein NB, Duran B. Using community-based participatory research to address health disparities. *Health Promot Pract* 2006 Jul;7(3):312-323. [doi: [10.1177/1524839906289376](https://doi.org/10.1177/1524839906289376)] [Medline: [16760238](https://pubmed.ncbi.nlm.nih.gov/16760238/)]

54. Horowitz CR, Robinson M, Seifer S. Community-based participatory research from the margin to the mainstream: are researchers prepared? *Circulation* 2009 May 19;119(19):2633-2642 [[FREE Full text](#)] [doi: [10.1161/CIRCULATIONAHA.107.729863](https://doi.org/10.1161/CIRCULATIONAHA.107.729863)] [Medline: [19451365](#)]
55. Rowe-Roberts D, Cercos R, Mueller F. Preliminary results from a study of the impact of digital activity trackers on health risk status. *Stud Health Technol Inform* 2014;204:143-148. [Medline: [25087541](#)]

Edited by G Eysenbach; submitted 29.05.19; peer-reviewed by A Joshi, A Kanter; comments to author 14.02.20; revised version received 30.11.20; accepted 13.04.21; published 26.05.21.

Please cite as:

Rose C, Nichols T, Hackner D, Chang J, Straube S, Jooste W, Sawe H, Tenner A

Utilizing Lean Software Methods To Improve Acceptance of Global eHealth Initiatives: Results From the Implementation of the Basic Emergency Care App

JMIR Form Res 2021;5(5):e14851

URL: <https://formative.jmir.org/2021/5/e14851>

doi:[10.2196/14851](https://doi.org/10.2196/14851)

PMID:[33882013](https://pubmed.ncbi.nlm.nih.gov/33882013/)

©Christian Rose, Taylor Nichols, Daniel Hackner, Julia Chang, Steven Straube, Willem Jooste, Hendry Sawe, Andrea Tenner. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 26.05.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Exploring Canadian Children's Social Media Use, Digital Literacy, and Quality of Life: Pilot Cross-sectional Survey Study

Lorie Donelle^{1*}, PhD; Danica Facca^{1*}, BA, MA; Shauna Burke^{2*}, PhD; Bradley Hiebert^{1*}, PhD; Emma Bender^{2*}, BSc, MPT; Stephen Ling^{3*}, BMedSci

¹Arthur Labatt Family School of Nursing, The University of Western Ontario, London, ON, Canada

²School of Health Studies, The University of Western Ontario, London, ON, Canada

³Schulich School of Medicine & Dentistry, The University of Western Ontario, London, ON, Canada

* all authors contributed equally

Corresponding Author:

Danica Facca, BA, MA

Arthur Labatt Family School of Nursing

The University of Western Ontario

1151 Richmond Street

London, ON, N6A3K7

Canada

Phone: 1 5198688675

Email: dfacca2@uwo.ca

Abstract

Background: Understanding social media use and digital literacy among young Canadian children is an increasing area of concern, given the importance of digital inclusion for full and informed participation in evolving educational, civic, corporate, social, and economic spaces.

Objective: The aim of this study was to explore internet and social media knowledge as well as social media use among Canadian children aged between 6 and 10 years.

Methods: We conducted interview surveys with 42 children aged between 6 and 10 years who participated in an after-school health promotion program in an urban community in Southwestern Ontario to understand their digital literacy skills and social media use. The data were analyzed using both quantitative and qualitative methods.

Results: Of the 42 children who participated in this study, 24 (57%) reported that they used social media, specifically YouTube (19/24, 79% reported use), Snapchat (16/24, 67% reported use), and Facebook (8/24, 33% reported use). While using social media, children reported sharing personal information, including videos or pictures of themselves (12/24, 50%), videos or pictures of others (8/24, 33%), and their birthday (12/24, 50%), whereas only one-third (9/24, 38%) of the children believed that only close family and friends had access to the content they shared. When reporting on the quality of life in the context of using social media, most (17/24, 71%) children never felt sad, half (12/24, 50%) never had difficulty making new friends, and nearly one-third (7/24, 30%) indicated that they never had difficulty wanting to play outside.

Conclusions: Owing to the rapidly evolving uptake and use of social media among young Canadians, the implementation of childhood digital health literacy education is vital to best support digital inclusion and well-being in Canada. The findings of our study highlight the need for future research to understand where children receive their digital literacy knowledge from and whether this knowledge is gained through self-directed social media use or observation from other actors, such as parents, siblings, or friends.

(*JMIR Form Res* 2021;5(5):e18771) doi:[10.2196/18771](https://doi.org/10.2196/18771)

KEYWORDS

child; children; internet; social media; digital literacy; digital inclusion; quality of life; mobile phone

Introduction

Background

The infiltration of digital devices across the age continuum supports the need for widespread attention to digital literacy. In general, digital literacy refers to a person's knowledge and skills in using digital devices to perform a variety of tasks, including sharing knowledge and information; communicating with others; participating in web-based communities, such as social media platforms; creating digital content; and analyzing digital content from a critical lens [1,2]. The practical value of digital literacy provides the necessary knowledge to optimize the social benefits and opportunities afforded by digital devices. Fostering digital literacy skills is crucial for ensuring digital inclusion and resolving inequities related to the digital divide [3]. Digital inclusion refers to mending the digital divide in terms of enhancing equitable access to digital devices and the internet so that individuals can meaningfully participate in all sectors of a knowledge- and information-driven society [3,4]. Considering North America's information and data-powered infrastructure, digital literacy skills are, and will continue to be, a crucial component of full and informed participation within evolving educational, civic, corporate, social, and economic spaces [5-7].

Digital literacy and the extent to which someone can effectively use digital devices to navigate and participate in sectors across contemporary society have implications for perceived quality of life, in other words, factors that contribute to life satisfaction among users [8,9]. From a subjective perspective, perceived quality of life is influenced by personality factors and an individual's perception of their position in life in relation to their goals and concerns, including their sense of self-worth, personal development, and feelings of isolation. From an objective perspective, perceived quality of life is influenced by environmental factors, including one's family, job, community, and standard of living. It has been noted in the literature that satisfaction with one's standard of living tends to influence subjective well-being, which is important to keep in mind when considering how advancements in technology and internet connectivity have become material features that contribute to the makeup of one's standard of living [8]. Advancements in technology, internet connectivity, and social media have fundamentally reshaped the way we work, learn, communicate with others, and cultivate our identities; thus, understanding the relationship between perceived quality of life and technology use is a necessary area of research inquiry. For example, researchers have begun to explore the effects of social media use on life satisfaction among preadolescent and adolescent populations in areas related to social (dis)connectedness, relationship conflict, emotional well-being, sleep quality, physical activity, and school performance [10-12].

In North America, for most American adolescents and young adults, being connected to others through digital devices means engagement with social media platforms, such as Facebook and Twitter [12]. Given the ample uptake of digital devices among American youth, it is not surprising that evidence from an American nationwide survey that asked parents about their

children's social media use found that 72% of children aged ≤ 8 years had used digital devices for playing games, watching videos, or using apps, whereas 17% of children used mobile digital devices at least once or more on a daily basis [12]. The investigators also reported that 38% of all children aged ≤ 2 years, 80% of children aged between 2 and 4 years, and 83% of children aged between 5 and 8 years had used a mobile digital device [13].

Recent research also notes that the frequency of social media use among older children, aged 8-11 years, and adolescents, aged 13-17 years, increases with age [14,15]. In particular, patterns of social media use among older children, aged 9-12 years, are seen to be rapidly approaching those of adolescent users [15]. For example, more than one-third of European children aged 9-12 years claimed that they had their own profile for at least one social media site [7]. Research also reports that the types of social media platforms that children and adolescents use evolve with age [15,16]. For example, YouTube was identified as the most popular social media platform for American children, followed by Moshi Monsters and Club Penguin, whereas Facebook, Instagram, and Twitter were identified as the most popular social media platforms for American adolescents [16]. For Canadian adolescents aged 13-17 years, playing web-based games, streaming music or television shows, and connecting with friends were the most common social media activities [17].

Among European children aged 9-12 years, more than one-fourth reported having their social media profile *set to public* [7]. Although European children aged 9-12 years had fewer social media profiles overall than children aged 13-16 years, they were still more likely to have their profiles set to public [7]. Research shows that children who have their social media profiles set to public are twice as likely to also post personal information than those who have their profile *set to private* or *set to partially private* [7]. It may be that children who have their profile set to public are merely unaware of how to change their privacy settings on a particular social media site [7]. As some researchers note, understanding children's and adolescents' awareness and use of the privacy settings on social media is essential for identity formation and maturation, as the internet may offer them a space where they can evade and control adult surveillance [16].

In Canada, between 2009 and 2014, 1.1 million adolescents and young adults aged 15-29 years reported experiences of cyberbullying or cyberstalking [18]. Cyberbullying is a global issue [19], and researchers note that using customized privacy settings and blocking features on social media is a strategic way to protect oneself from such harm [20]. When it comes to examining social relationships, conflict, cyberbullying, and personal safety, most research to date has focused only on adolescents and older children, aged ≥ 12 years, with little insight into younger children [21].

Although evidence shows that older children, preadolescents, and adolescents are engaging with social media alike and suggests that social media is a vital means of communication and reason for their digital device use in North America, there is limited research seeking to understand the rapidly evolving

uptake and use of social media among children aged ≤ 12 years, which creates a significant gap in knowledge concerning their digital literacy [15,19,22,23]. In particular, research concerning social media use and digital literacy among young Canadian children is lacking. Although Canadian researchers continue to provide growing insight into the uptake and nature of digital device use among children aged ≥ 12 years, there is limited research on the nature of social media use and digital literacy among younger children. Research focused on this population segment can explore children's experiences of information sharing, personal safety, cyberbullying, and learning patterns to better understand their concerns, needs, and capabilities associated with digital literacy.

It is important for researchers, educators, and parents to know how young children understand privacy in relation to social media use, given the rise of the privacy paradox: when device users claim to care about their internet privacy but act in ways that contradict this concern. For example, a device user may claim that protecting their personal information is important to them, but then freely give this personal information in exchange for services or convenience. Understanding young children's views on privacy can provide important insights into their behaviors and the factors that influence their rationale for the choices they make when using social media [24]. By seeking a more thorough understanding of digital inclusion as it relates to younger children, researchers, educators, and policy makers can better support Canadian children to enhance their quality of life within the digital age.

Objectives

This pilot study drew on childhood research theories that privilege children's perspectives on their social situations, such as their participation and safety in digital spaces, by recognizing their agency and competence on matters that affect them [25-27]. As recent evidence highlights the increased importance of internet connectivity in the everyday lives of North American children and points toward an increased trajectory of use [13,22], it is vital for researchers to learn from children and work directly with them to understand their experiences. To best recognize and respect children as social actors, this pilot study focused only on young children's self-reported digital literacy and social media use. The research questions that informed this study were as follows:

1. What knowledge of the internet and social media do Canadian children aged 6-10 years demonstrate?
2. What types of digital devices and social media platforms do Canadian children aged 6-10 years use, and how often do they use them?
3. How do Canadian children aged 6-10 years navigate personal safety and information sharing when using social media?

This pilot study aims to identify Canadian children's knowledge of the internet and social media, the types and frequency of social media platforms they use, the types of personal information they share on social media, who has access to the personal information they share on social media, who watches them while they are on social media, who helps them download social media apps, and how they feel about their quality of life

in the context of their social media use. This paper documents the findings from a survey administered through one-on-one interviews conducted among primary school-aged children in Southwestern Ontario.

Methods

Study Design

The pilot study used a nonexperimental cross-sectional survey design that focused on children aged between 6 and 10 years who participated in an after-school health promotion program in an urban community in Southwestern Ontario. The goal of the community-based health promotion program was to foster healthy lifestyle knowledge and skills through play. The after-school health promotion program was offered once a week over the course of 8 weeks and took place in a multipurpose room (eg, gym) within several elementary schools and public housing communities in less affluent areas of the community. In this study, social media was defined as "a set of web applications that enable production, aggregation, sharing, and remixing of content from multiple sources by mass, networked participants" [23]. The nonmedical research ethics board at Western University approved the study (file #107798).

Participants

Eligible participants were recruited from those enrolled in the community-based after-school health promotion program and were included in the study if they were children aged 6-10 years, enrolled in primary school grades 1-3, and English speaking. Convenience sampling was used to recruit participants by sending home a letter of information about the research, a parental consent form, and a child assent form with the children. Inclusion in the study required both a signed parental consent form and a child assent form.

Data Collection

Researchers conducted one-on-one interviews with each child lasting approximately 15 minutes in length during the after-school program's regularly scheduled hours. Interviews were completed in a location away from the after-school program activities to safeguard the children's confidentiality. Demographic data such as the child's age, grade, and gender were also collected at this time. A series of closed- and open-ended questions were constructed to explore children's digital literacy in terms of their knowledge of the internet and social media. Questions about the children's use of social media captured whether others (eg, parents, siblings, and teachers) facilitated their use of social media (eg, who helped to download apps or access information) and monitored their activity. Questions also asked how often the children used social media, shared personal information, and implemented privacy settings.

A modified version of the Pediatric Quality of Life Inventory (PedsQL) version 4.0 Child Report [28] was used to explore children's perceived quality of life in relation to their social media use. The original PedsQL child report questionnaire [28] consisted of 23 items encompassing four dimensions: physical functioning (eight items), emotional functioning (five items), social functioning (five items), and school functioning (five items) and was formatted using a Likert-type scale with

responses ranging from 0 (eg, “it is never a problem”) to 4 (eg, “it is almost always a problem”). This study’s modified version consisted of 18 items: physical functioning (four items), emotional functioning (five items), social functioning (four items), and school functioning (five items). Researchers reverse scored the items using the PedsQL version 4.0 Generic Core Scales [29].

Data Analysis

The data were organized and analyzed using IBM SPSS Statistics [30]. Descriptive analysis of the data was conducted. For categorical variables such as participants’ gender, grade, and types of digital devices and social media participants used, proportions and frequency counts were calculated.

Results

Participants

A total of 42 children were recruited to participate in this study. Overall, 50% (21/42) of the children were identified as male and 50% (21/42) of the children were identified as female. The average age of the children was 7.5 years (SD 0.943). The age range of the children was 6-10 years.

Internet Knowledge

When asked “Do you know what the internet is?” 79% (33/42) of children described the symbols that represent connecting to the internet, including “the symbol with the E” (female participant, age 7 years), “google chrome” (male participant, age 7 years), or “a compass” (male participant, age 7 years). Others described key aspects of digital literacy stating that “you have to be careful” (female participant, age 9 years) while on the internet because it can be “full of strangers” (female participant, age 7 years) inappropriate pictures and a space both

“safe and unsafe” (male participant, age 9 years). Furthermore, several of the children noted the functional role of the internet as a bridge between computer hardware and software, as they described it as what “makes the phone and iPad go” (female participant, age 8 years) while also acknowledging that “without it you can’t do much or download stuff” (male participant, age 8 years). The children also described the internet as an educational space “for reading” (male participant, age 7 years), “for homework” (male participant, age 7 years), to “search stuff” (male participant, age 8 years), and to “explore stuff, discover and talk” (male participant, age 8 years). Most predominantly, however, the children described the internet as a source of entertainment, as they considered it fundamental to access “fun games and cool things” (female participant, age 8 years), “watch YouTube” (male participant, age 7 years), or “make videos” (female participant, age 7 years).

Social Media Knowledge, Type, and Frequency of Use

When asked “Do you know what social media is?” 54% (22/42) of children described social media in reference to the term itself or by a proprietary name such as Facebook or Snapchat. Overall, 57% (24/42) of children reported that they used social media. In terms of frequency, 46% (11/24) of children reported that they always used social media. The predominant reason among children who did not report that they used social media was that their parents did not let them use social media. Regarding the use of specific social media platforms, 79% (19/24) of children reported that they used YouTube; 67% (16/24) of children reported that they used Snapchat; and 33% (8/24) of children reported that they used Facebook, Musically, or other unspecified sites. Table 1 presents the types of social media reported by the children and the number of children who reported using it.

Table 1. Children’s reported social media use (n=24).

Type of social media	Children who reported use, n (%)
YouTube	19 (79)
Snapchat	16 (67)
Facebook	8 (33)
Other	8 (33)
Musically	8 (33)
Skype	7 (29)
Instagram	4 (17)
MySpace	3 (13)
Vine	2 (8)
Twitter	1 (4)

Shared Personal Information on Social Media

When asked “What information have you posted or shared on social media?” about the personal information they posted and shared on social media, 54% (13/24) of children reported that

they shared their names. Similarly, 50% (12/24) of children reported that they shared videos or pictures of themselves. Table 2 presents the type of personal information reported by the children and the number of children who reported sharing it through social media.

Table 2. Personal information shared by children through social media (n=24).

Type of personal information	Children who reported sharing, n (%)
Name	13 (54)
Videos or pictures of self	12 (50)
Birthday	9 (38)
Hobbies	8 (33)
Videos or pictures of family or friends	8 (33)
Hometown	5 (21)

When asked “Do you know who has access to what you post?” 38% (9/24) of children believed that only close family and friends had access to the content they posted on social media. When using social media, 38% (9/24) of children also reported that a family member or friend was sometimes present. When asked “Is your teacher with you when you use social media?” 54% (13/24) of children reported that a teacher was never present when using social media at school, whereas 17% (4/24)

of children stated that a teacher was sometimes present. Even more so, when asked who helped them download social media apps, 42% (10/24) of children stated that no one helped them. [Table 3](#) presents the type of person who helped the children download apps (downloading assistant) and the number of children who reported that type of person helped them within this context.

Table 3. Children’s reported downloading assistance of social media apps (n=24).

Downloading assistant	Children who reported assistance, n (%)
None	10 (42)
Parent	8 (33)
Other	8 (33)
Sibling	4 (17)
Friend	1 (4)

Quality of Life

Children who reported that they used social media were also asked to complete the PedsQL portion of the survey to explore their perceived quality of life. Overall, 57% (24/42) of children answered this section of the survey. When reporting on emotional functioning in the context of social media use within

the past month, 71% (17/24) of children stated that they never felt sad or blue, 63% (15/24) of children indicated that they never felt afraid or scared, and 54% (13/24) of children reported that they never felt angry. [Table 4](#) presents the emotional functioning items of the survey and the number of children who reported that they never had a problem with each item.

Table 4. Children’s reporting on emotional functioning using the Pediatric Quality of Life Inventory (n=24).

Emotional functioning item	Children who reported the item as “never a problem,” n (%)
Feeling sad or blue	17 (70)
Feeling afraid or scared	15 (63)
Feeling angry	13 (54)

When reporting on social functioning in the context of social media use within the past month, 50% (12/24) of children stated that they never had difficulty making new friends, 44% (10/24) of children indicated that they never had difficulty with others wanting to be their friend, 39% (9/24) of children stated that

they never had difficulty talking to other children at recess, and 50% (12/24) of children reported that they never had difficulty talking to their parents or siblings. [Table 5](#) presents the social functioning items of the survey and the number of children who reported that they never had a problem with each item.

Table 5. Children’s reporting on social functioning using the Pediatric Quality of Life Inventory (n=24).

Social functioning item	Children who reported the item as “never a problem,” n (%)
Difficulty making new friends	12 (50)
Difficulty with others wanting to be their friend	10 (44)
Difficulty talking to others at recess	9 (39)
Difficulty talking to parents or siblings	12 (50)

When reporting on school functioning in the context of social media use within the past month, 30% (7/24) of children reported that they never had difficulty paying attention in class, 44% (10/24) of children stated that they never had difficulty participating in class, and 48% (11/24) of children indicated

that they never had difficulty keeping up with schoolwork. [Table 6](#) presents the school functioning items of the survey and the number of children who reported that they never had a problem with each item.

Table 6. Children's reporting on school functioning using the Pediatric Quality of Life Inventory (n=24).

School functioning item	Children who reported the item as "never a problem," n (%)
Difficulty paying attention in class	7 (30)
Difficulty participating in class	10 (44)
Difficulty keeping up with schoolwork	11 (48)

When reporting on physical functioning in the context of social media use within the past month, 48% (11/24) of children indicated that they never had difficulty being active at recess, 57% (13/24) of children reported that they never had difficulty participating in gym class, 30% (7/24) of children stated that they never had difficulty wanting to go outside and play at home,

and 30% (7/24) of children indicated that they never had difficulty wanting to join any sports teams. [Table 7](#) presents the physical functioning items of the survey and the number of children who reported that they never had a problem with each item.

Table 7. Children's reporting on physical functioning using the Pediatric Quality of Life Inventory (n=24).

Physical functioning item	Children who reported the item as "never a problem," n (%)
Difficulty being active at recess	11 (48)
Difficulty participating in gym class	13 (57)
Difficulty wanting to go outside and play at home	7 (30)
Difficulty wanting to join any sports team	7 (30)

Discussion

Principal Findings

Overview

Overall, our principal findings suggest that Canadian children aged between 6 and 10 years are active users of social media. By conducting our survey with one-on-one interviews, we found that Canadian children aged between 6 and 10 years use social media, share their personal information on social media, and download apps without consistent supervision. These findings contribute to the existing knowledge gap in the Canadian literature regarding young children's digital literacy and social media use by incorporating the perspectives of a largely understudied proportion of the Canadian population, which is vital for understanding digital inclusion in Canada.

Social Media Use

Approximately half of the children in the study reported that they used social media. This finding is slightly higher than that reported by Canadian researchers [22], who found that approximately 30% of Canadian students had Facebook accounts and 16% had Twitter accounts. Longitudinal studies in the United Kingdom, which explored child and youth engagement with digital technologies, attest to the rapid growth in social media engagement within the past few years [31]. Most notably, parents from the United Kingdom reported that children aged 3-4 years were using the internet at an average of 8 hours per week, children aged 5-7 years were using the internet for approximately 9 hours per week, and children aged 8-11 years

were using the internet at an average of 13.5 hours per week [31].

Social Media Knowledge

When asked, children in this pilot study described social media as a commercial product by using proprietary names, such as YouTube, Snapchat, and Facebook. The children's identification of YouTube as a preferred social media site remains consistent with the preferences of 48%, 71%, and 81% of British children aged 3-4, 5-7, and 8-11 years, respectively [31]. YouTube's platform may optimize fundamental digital literacy skills and is thus relatively easy for children to navigate, search, and interact with. It also offers children multisensory entertainment, learning opportunities from educational videos, and the ability to create their own content. Multisensory learning activities have been suggested to enhance learners' knowledge and skills with a multisensory experience, including auditory, visual, and kinesthetic stimulation, which may engage and maintain children's attention for long periods [32].

Personal Information Shared on Social Media

What is perhaps most concerning in this study is the nature of personal information and content shared by the children through social media in addition to their lack of awareness regarding who had access to their posted information. Developing evidence indicates that many parents and children are unaware of internet privacy policies and how third parties can collect and use their personal information [33]. Similarly, a large proportion of Canadian youth believe that personal information cannot be collected by commercial corporations if a privacy policy exists [22]. It may also be that children are unaware of the importance

of consent when posting content of others and the permanence of their digital traces when trying to delete previously posted content from sites [22]. Although privacy settings on social media sites are purportedly designed to protect users from potential harm, they are more often than not incomprehensible to children and most adults [7]. Nevertheless, it is important to recognize that sharing personal information and self-produced content on social media is a critical part of identity formation for children and, when handled correctly, can strengthen their relationships with others [19].

Parents and teachers of young children play a significant role in the development of children's digital literacy skills, especially as it relates to personal safety in terms of information sharing, privacy settings, and how to protect oneself from unsolicited or unwanted web-based interactions [22]. The ability of parents and teachers to support the development of digital literacy skills among children is often a limited variable [34,35]. However, the development of effective digital literacy skills among parents is especially relevant, given that young children use digital devices and social media as modeled by their parents [17]. When investigating children's digital literacy, as it relates to personal information sharing and internet safety, researchers found that many parents tend to use a "do what I say not what I do" strategy, whereby they caution their children regarding internet safety but will engage in unsafe internet behaviors and model excessive device use at home [34].

There are important educational opportunities to increase awareness of role model behaviors in the context of digital literacy and internet safety [34]. Parental behavior, along with the family home environment more generally, are influential factors that socialize and shape children's habits and, as such, should be optimized to promote best practices when it comes to digital devices and social media use [17]. Furthermore, education policies in Canada created since the beginning of the COVID-19 pandemic promote elementary students' web-based learning [36,37] and ostensibly increase the influence that teachers have on children's digital literacy skills. Elementary teachers have an opportunity to integrate digital literacy topics into existing learning activities, such as narrative writing in the first grade, to begin solidifying key concepts at an early age.

Canadian researchers claim that Canadian youth "seek both privacy and publicity at the same time by posting information and then seeking to control the various audiences that can access it" [22]. As Canadian youth seem to approach privacy in digital spaces differently than previous generations, the current Canadian legislation "may not only fail to meet their needs but actually increase the kinds of vulnerabilities they face online" [22]. In terms of web-based supervision at home, approximately one-third of the children active on social media reported that a family member or friend was present while they were on social media, whereas when downloading content from the app store, approximately half of the children reported that no one helped them. This could potentially be because of the understanding that children seek privacy from authoritative figures such as parents while using social media and unknowingly choose self-exploration at the expense of safety [38].

Importantly, unsupervised downloading practices raise concern with respect to the terms and conditions children agree to when they download social media apps, such as the monetary and personal information costs levied on app users to access app content. In granting consent to access the content within popular children's social media apps, such as those identified by the children in this study, individuals often consent to the app's collection and sharing of user device information, such as device type identification (eg, smartphone and tablet), device brand identification (eg, Apple and Samsung), device software identification (eg, most up-to-date operating system and previous operating system), device means of connectivity (eg, hardwired internet, wireless modem internet, and wireless cellular internet), and device geolocation to commercialized third parties [33].

The failures of privacy legislation are highlighted in a recent report by the US Federal Trade Commission [33], which documented a significant lack of users', both parent and child, awareness and concern toward personal information exploitation by third-party commercial enterprises among app users. The report noted that the limited transparency and high digital literacy demands of the privacy policies attached to child-focused social media apps [33]. The existing user agreement and consent jargon make it difficult for most parents to make informed decisions regarding the quality and safety of apps for their children and nearly impossible for a child to comprehend when independently downloading their own apps. There is growing awareness of the educational curriculum regarding the development of digital literacy skills and regulations that safeguard information privacy among social media users.

These findings have further implications for areas of study, such as digital health and, by extension, digital health literacy [39,40]. Digital health accounts for the ways in which digital devices permeate every aspect of daily life, in other words, how digital devices affect our overall well-being. [41] Digital health examines the benefits and drawbacks of various technologies aimed at smartphones, smart objects, social media, wearable technologies, websites, forums, and health-related apps [41]. For example, child health researchers have found that social media serves as a support network for health-related matters and a distraction from chronic illness among older children [5].

To effectively engage with and maximize the health benefits afforded by digital devices, users must have sufficient digital health literacy skills [42]. Digital health literacy, also known as eHealth literacy [43], is defined as "the ability to seek, find, understand, and appraise health information from electronic resources and apply the knowledge gained to addressing or solving a health problem" [44]. Simply put, digital health literacy skills integrate the knowledge of multiple literacies, including traditional (writing, reading, and understanding numeracy), computer (managing data privacy and personal information), media (critically assessing mass media sources), scientific (discerning evidence from opinion), information (searching and retrieving relevant information), and health (comprehending health information and systems) literacies [44]. These intersecting literacies further fall into two skill categories, analytical and context specific, given that some demand more sophisticated skills from a user than others [44]. Traditional,

media, and information literacy comprise the analytical skills necessary for a user to critically analyze various sources of information [44]. However, scientific, health, and computer literacy comprise context-specific skills, as they depend on specific issues (eg, discipline terminology), problems (eg, health issues), or contexts (eg, web 2.0) that require nuanced guidance beyond analytical literacies [44].

Considering North America's information and data-driven society, sufficient digital health literacy skills are, and will continue to be, a crucial component of full and informed participation within evolving educational, civic, corporate, social, and economic spaces [5-7]. It is, therefore, essential that digital device users of all ages, including young children, develop proficient digital health literacy skills to best support their overall well-being [45]. This is especially pertinent given that emerging evidence suggests digital natives—that is, youth who grow up with digital technology—lack adequate digital health literacy skills [46].

Quality of Life

Research findings related to social media use and children's emotional well-being, physical activity, social relationships, and school performance are mixed [47-49]. There is ongoing debate in the literature as to whether children's and adolescents' social media use enhances or detracts from their emotional well-being [49,50]. Recent research shows that social media use can be used to support adolescents' self-esteem [50]. The results of this study are consistent with the results of a previous study [50], as more than half of the children who were active on social media reported being happy. As children's happiness may be related to their learning [51], future research that examines the relationship between children's social media use, happiness, and learning outcomes may further elucidate the extent to which social media use and digital literacy skills impact children's well-being.

Regarding physical activity, nearly half of the children in this study reported that they never had difficulty playing outside at recess or participating in gym, which contrasts with previous findings that screen time is negatively associated with physical activity [47]. Half of the children in this study also claimed they did not have difficulty making friends, which is consistent with findings that social media use contributes to friendship development and support among adolescents [52]. However, limited or no access to social media may create an exclusionary social environment in which some children may face inequitable barriers to social or learning opportunities. Thus, further research should examine the role of social media in creating young children's social environments and how participation in such environments influences children's digital literacy skills.

Finally, recent evidence shows that screen time, especially during weekend hours, detracts from good academic performance among adolescents [47]. Additional research is needed to further explore the relationship between screen time and academic performance among very young students, given that almost half of the children in this study reported little to no difficulty with schoolwork completion, which supports previous findings [53]. Furthermore, the transition to web-based learning across Canada in response to the COVID-19 pandemic prompts

a reexamination of screen time's effect on children's academic performance, as elementary students are attending school through the internet from home and thus learning via screen time. Such research may provide insight into the evolving relationship between children's screen time, digital literacy, and academic attainment while simultaneously providing information that may advance digital literacy pedagogies developed for elementary students.

Limitations

This pilot study has several limitations. In this study, data were collected from elementary students involved in a school-based health promotion program within less affluent neighborhoods in Southwestern Ontario. Children from more affluent neighborhoods may report differently on their use of social media. Only children who spoke English as their first language were included in the study; the use of social media apps may be very different for children whose first language is not English. Cultural differences may influence social media use, which was not a focus of this research. Although the children in this study were active and healthy, children with acute or chronic disease or with physical disability may take up social media differently. This pilot study was conducted in an urban after-school program, which was a necessary geographical limitation because of its exploratory nature. The small sample size of this study was not powerful enough to complete statistical analyses beyond descriptive measures. Finally, the cross-sectional design of the study also limited the analysis of children's social media use over time and is not necessarily a representative sample of this population.

Future Research

Future research with this population should consider the possible influence that various social determinants of health, such as citizen status (immigrant or refugee), cultural background, disability, geography, and social class, may have on children's uptake, use, and motivations for engaging with social media. For example, future research that evaluates how children living in a rural community access and use social media is needed to further conceptualize the impact that geographic remoteness and potentially limited access to a reliable internet connection may have on children's engagement with social media. Social forces such as gender norms should be given further attention, as they may impact a child's experience with social media at large as well. Future work with this population should also take into consideration the implementation of longitudinal and/or more substantial cross-sectional design, as they can deepen the understanding of children's relationships with social media and the extent to which this relationship is influenced by social determinants of health, such as age, gender, socioeconomic status, and ethnicity.

This study also has implications for the study of digital health literacy among this population. Although evidence shows youth are engaging with social media, there is limited research seeking to understand the rapidly evolving uptake of, and practices relating to, social media among children aged ≤ 10 years, which creates a significant gap in knowledge concerning children's digital health literacy. For example, the context-specific domains of digital health literacy, namely, health and scientific literacy,

should also be considered as further areas of study, as each contains the skills needed to understand discipline-specific terminology, execute effective search strategies, and assess health information quality.

In light of COVID-19 and the shift to remote e-learning, research that aims to explore young children's digital literacy can contribute to evolving models of preschool and elementary e-learning education by examining inequities and supports in relation to children's engagement and success within a virtual education context. Moreover, examining children's digital health literacy competency within the age of COVID-19 can provide further insight into how children manage their health during a pandemic across a digital landscape of various social media platforms and information sources, in addition to how they understand the consequences of public health policy, such as contact tracing and social distancing, which has affected the structure of their learning environment.

Conclusions

Our findings suggest that Canadian children aged between 6 and 10 years are active social media users. This study highlights how the uptake and adaptation of social media use among particular age groups is constantly evolving and worth tracking over time. Many children demonstrated knowledge of the term

social media, as they understood it in two distinct ways: by the term itself or by a proprietary name such as Facebook. This raises questions as to where children receive their digital literacy knowledge from and whether such knowledge is gained through self-directed exploration or observation from other actors, such as parents, siblings, or teachers.

The digital literacy practices reported by the children in this study have implications for their health, social and learning behaviors, and safety. Children shared their personal information in addition to self-generated content, such as pictures and videos of themselves and others through social media. This has led to the implementation of early childhood digital health literacy education to best guide, support, and empower children when making these choices to fully understand the social consequences and permanency of their actions. Children's reported unsupervised social media use, particularly the downloading of apps, reinforces the notion that young children are self-directed users of digital devices and social media. In turn, user acknowledgment and support should be given to this population, as they are just as engaged as older youth populations and perhaps even more vulnerable to web-based threats due to the oversight of privacy policy development thus far in Canada.

Acknowledgments

The authors would like to thank Matthew Bureau for his assistance in this research and the participants for their contribution.

Conflicts of Interest

None declared.

References

1. Seung-Hyun L. Digital literacy education for the development of digital literacy. *International Journal of Digital Literacy and Digital Competence (IJDLDC)* 2014;5(3):29-43. [doi: [10.4018/ijdlc.2014070103](https://doi.org/10.4018/ijdlc.2014070103)]
2. Jones-Kavalier B, Flannigan S. Connecting the digital dots: literacy of the 21st century. *Teach Librar* 2008;35(3):13-16 [[FREE Full text](#)]
3. Tate T, Warschauer M. The digital divide in language literacy education. In: Thorne S, May S, editors. *Language, Education, and Technology*. New York: Springer; 2017:45-56.
4. Kumpulainen K, Sairanen H, Nordström A. Young children's digital literacy practices in the sociocultural contexts of their homes. *J Early Child Lit* 2020 May 21;20(3):472-499. [doi: [10.1177/1468798420925116](https://doi.org/10.1177/1468798420925116)]
5. Hamm MP, Shulhan J, Williams G, Milne A, Scott SD, Hartling L. A systematic review of the use and effectiveness of social media in child health. *BMC Pediatr* 2014;14:1-15 [[FREE Full text](#)] [doi: [10.1186/1471-2431-14-138](https://doi.org/10.1186/1471-2431-14-138)] [Medline: [24886048](https://pubmed.ncbi.nlm.nih.gov/24886048/)]
6. Perestelo-Perez L, Torres-Castaño A, González-González C, Alvarez-Perez Y, Toledo-Chavarri A, Wagner A, et al. IC-Health Project: development of MOOCs to promote digital health literacy: first results and future challenges. *Sustainability* 2020;12(16):6642 [[FREE Full text](#)] [doi: [10.3390/su12166642](https://doi.org/10.3390/su12166642)]
7. Livingstone S, Olafsson K, Staksrud E. Social networking, age and privacy. *LSE Research Online*. 2011. URL: http://eprints.lse.ac.uk/35849/?from_serp=1 [accessed 2021-01-03]
8. Leung L. Effects of internet connectedness and information literacy on quality of life. *Soc Indic Res* 2009 Oct 1;98(2):273-290. [doi: [10.1007/s11205-009-9539-1](https://doi.org/10.1007/s11205-009-9539-1)]
9. Karimi M, Brazier J. Health, health-related quality of life, and quality of life: what is the difference? *PharmacoEconomics* 2016 Feb 18;34(7):645-649. [doi: [10.1007/s40273-016-0389-9](https://doi.org/10.1007/s40273-016-0389-9)]
10. Buda G, Lukoševičiūtė J, Šalčiūnaitė L, Šmigelskas K. Possible effects of social media use on adolescent health behaviors and perceptions. *Psychol Rep* 2021;124(3):1031-1048. [doi: [10.1177/0033294120922481](https://doi.org/10.1177/0033294120922481)] [Medline: [32366165](https://pubmed.ncbi.nlm.nih.gov/32366165/)]

11. van den Eijnden R, Koning I, Doornwaard S, van Gurp F, ter Bogt T. The impact of heavy and disordered use of games and social media on adolescents' psychological, social, and school functioning. *J Behav Addict* 2018 Sep 10;7(3):697-706. [doi: [10.1556/2006.7.2018.65](https://doi.org/10.1556/2006.7.2018.65)]
12. Smith A, Anderson M. Social media use in 2018. Pew Research Center. 2018. URL: <https://www.pewresearch.org/internet/2018/03/01/social-media-use-in-2018/> [accessed 2020-02-02]
13. Zero to eight: children's media use in America, 2013. Common Sense Media. 2013. URL: <https://www.common SenseMedia.org/research/zero-to-eight-childrens-media-use-in-america-2013> [accessed 2020-02-02]
14. Australian Communications and Media Authority. Like, post, share: young Australians' experience of social media. Analysis & Policy Observatory. 2013. URL: <https://apo.org.au/node/35223> [accessed 2020-02-02]
15. Holloway D. Digital play: the challenge of researching young children's internet use. In: Proceedings of Australian and New Zealand Communication Association Conference. 2014 Presented at: Australian and New Zealand Communication Association Conference; July 9-11, 2014; Melbourne, Australia p. 1-14 URL: <https://ro.ecu.edu.au/ecuworkspost2013/829/>
16. Lenhart A. Pew Research Center. 2015. URL: <https://www.pewinternet.org/2015/04/09/teens-social-media-technology-2015/> [accessed 2020-02-02]
17. Granich J, Rosenberg M, Knuiman M, Timperio A. Individual, social, and physical environment factors associated with electronic media use among childrendentary behavior at home. *J Phys Act Health* 2011;8:613-625. [doi: [10.1123/jpah.8.5.613](https://doi.org/10.1123/jpah.8.5.613)]
18. Study: cyberbullying and cyberstalking among internet users aged 15 to 29 in Canada. Statistics Canada. 2016. URL: <https://www150.statcan.gc.ca/n1/daily-quotidien/161219/dq161219a-eng.htm> [accessed 2020-02-02]
19. Willoughby M. A review of the risks associated with children and young people's social media use and the implications for social work practice. *J Soc Work Pract* 2018 May 08;33(2):127-140. [doi: [10.1080/02650533.2018.1460587](https://doi.org/10.1080/02650533.2018.1460587)]
20. Kwan G, Skoric M. Facebook bullying: an extension of battles in school. *Comput Hum Behav* 2013;29(1):16-25 [FREE Full text] [doi: [10.1016/j.chb.2012.07.014](https://doi.org/10.1016/j.chb.2012.07.014)]
21. Chesley N, Johnson B. Information and communication technology use and social connectedness over the life course. *Sociol Compass* 2014;8:589-602 [FREE Full text] [doi: [10.1111/soc4.12170](https://doi.org/10.1111/soc4.12170)]
22. Steeves V. Young Canadians in a wired world, phase iii: trends and recommendations. MediaSmarts. 2014. URL: <http://mediasmarts.ca/research-policy> [accessed 2020-02-02]
23. Swist T, Collin P, McCormack J, Third A. Social media and the wellbeing of children and young people: a literature review. Western Sydney University. 2015. URL: <https://researchdirect.westernsydney.edu.au/islandora/object/uws:36407> [accessed 2020-02-02]
24. Sunyaev A, Dehling T, Taylor PL, Mandl KD. Availability and quality of mobile health app privacy policies. *J Am Med Inform Assoc* 2015 Apr;22(e1):28-33 [FREE Full text] [doi: [10.1136/amiajnl-2013-002605](https://doi.org/10.1136/amiajnl-2013-002605)] [Medline: [25147247](https://pubmed.ncbi.nlm.nih.gov/25147247/)]
25. Third A, Bellerose D, Dawkins U, Keltie E, Pihl K. Children's rights in the digital age: a download from children around the world. Western Sydney University. 2014. URL: <https://researchdirect.westernsydney.edu.au/islandora/object/uws:28202> [accessed 2020-02-02]
26. Lundy L. Voice is not enough: conceptualising Article 12 of the United Nations Conventional on the Rights of the Child. *Br Educ Res J* 2007;33(6):927-942 [FREE Full text] [doi: [10.1080/01411920701657033](https://doi.org/10.1080/01411920701657033)]
27. James A. Giving voice to children's voices: practices and problems, pitfalls and potentials. *Am Anthropol* 2007;109(2):261-272. [doi: [10.1525/aa.2007.109.2.261](https://doi.org/10.1525/aa.2007.109.2.261)]
28. Varni JW, Seid M, Rode CA. The PedsQL: measurement model for the pediatric quality of life inventory. *Med Care* 1999;37(2):126-139. [Medline: [10024117](https://pubmed.ncbi.nlm.nih.gov/10024117/)]
29. PedsQL review copies - generic core scales. eProvide. URL: <https://eprovide.mapi-trust.org/instruments/pediatric-quality-of-life-inventory/pedsq-ql-review-copies> [accessed 2020-02-02]
30. IBM SPSS Statistics for Windows, version 24.0. IBM. 2016. URL: <https://www.ibm.com/support/pages/how-cite-ibm-spss-statistics-or-earlier-versions-spss> [accessed 2020-02-02]
31. Children and parents: media use and attitudes report. Ofcom. 2016. URL: <https://www.ofcom.org.uk/research-and-data/media-literacy-research/childrens/children-parents-nov16> [accessed 2020-02-02]
32. Hsu HJ. The potential of kinect in education. *Int J Inf Educ Technol* 2011;1(59):365-370 [FREE Full text] [doi: [10.7763/ijiet.2011.v1.59](https://doi.org/10.7763/ijiet.2011.v1.59)]
33. Mobile apps for kids: current privacy disclosures are disappointing. Federal Trade Commission. 2012. URL: <https://www.ftc.gov/reports/mobile-apps-kids-current-privacy-disclosures-are-disappointing> [accessed 2020-02-02]
34. Terras MM, Ramsay J. Family digital literacy practices and children's mobile phone use. *Front Psychol* 2016 Dec 23;7:1-11. [doi: [10.3389/fpsyg.2016.01957](https://doi.org/10.3389/fpsyg.2016.01957)]
35. Hoeschsmann M, DeWaard H. Mapping digital literacy policy and practice in the Canadian education landscape. MediaSmarts. 2015. URL: <https://mediasmarts.ca/sites/default/files/publication-report/full/mapping-digital-literacy.pdf> [accessed 2020-02-02]
36. Learn at home. Government of Ontario. 2020. URL: <https://www.ontario.ca/page/learn-at-home> [accessed 2020-05-01]
37. Online and distributed learning. Government of British Columbia. 2020. URL: <https://www2.gov.bc.ca/gov/content/education-training/k-12/support/classroom-alternatives/online-distributed-learning> [accessed 2020-05-01]
38. Boyd D. It's Complicated – The Social Lives of Networked Teens. New Haven: Yale University Press; 2014:1-296.

39. Hoffman-Goetz L, Donelle L, Ahmed R. Health Literacy in Canada: A Primer for Students. Toronto: Canadian Scholars Press Inc; 2014:1-264.
40. Griebel L, Enwald H, Gilstad H, Pohl A, Moreland J, Sedlmayr M. eHealth literacy research-Quo vadis? *Inform Health Soc Care* 2017 Oct 18;43(4):1-16. [doi: [10.1080/17538157.2017.1364247](https://doi.org/10.1080/17538157.2017.1364247)] [Medline: [29045164](https://pubmed.ncbi.nlm.nih.gov/29045164/)]
41. Lupton D. *Digital Health: Critical and Cross-Disciplinary Perspectives*. London: Routledge; 2017:1-178.
42. Paige SR, Stellefson M, Krieger JL, Anderson-Lewis C, Cheong J, Stopka C. Proposing a transactional model of eHealth literacy: concept analysis. *J Med Internet Res* 2018 Oct 02;20(10):e10175 [FREE Full text] [doi: [10.2196/10175](https://doi.org/10.2196/10175)] [Medline: [30279155](https://pubmed.ncbi.nlm.nih.gov/30279155/)]
43. van der Vaart R, Drossaert C. Development of the digital health literacy instrument: measuring a broad spectrum of Health 1.0 and Health 2.0 skills. *J Med Internet Res* 2017 Jan 24;19(1):e27 [FREE Full text] [doi: [10.2196/jmir.6709](https://doi.org/10.2196/jmir.6709)] [Medline: [28119275](https://pubmed.ncbi.nlm.nih.gov/28119275/)]
44. Norman CD, Skinner HA. eHealth literacy: essential skills for consumer health in a networked world. *J Med Internet Res* 2006 Jun;8(2):e9 [FREE Full text] [doi: [10.2196/jmir.8.2.e9](https://doi.org/10.2196/jmir.8.2.e9)] [Medline: [16867972](https://pubmed.ncbi.nlm.nih.gov/16867972/)]
45. Bautista J. From solving a health problem to achieving quality of life: redefining eHealth literacy. *J Lit Technol* 2015;16(2):33-54 [FREE Full text] [doi: [10.1007/978-88-470-5253-6_4](https://doi.org/10.1007/978-88-470-5253-6_4)]
46. Stellefson M, Hanik B, Chaney B, Chaney D, Tennant B, Chavarria EA. eHealth literacy among college students: a systematic review with implications for eHealth education. *J Med Internet Res* 2011 Dec;13(4):e102 [FREE Full text] [doi: [10.2196/jmir.1703](https://doi.org/10.2196/jmir.1703)] [Medline: [22155629](https://pubmed.ncbi.nlm.nih.gov/22155629/)]
47. Huang X, Zeng N, Ye S. Associations of sedentary behavior with physical fitness and academic performance among Chinese students aged 8–19 years. *Int J Environ* 2019 Nov 14;16(22):4494. [doi: [10.3390/ijerph16224494](https://doi.org/10.3390/ijerph16224494)]
48. Rosen LD, Lim AF, Felt J, Carrier LM, Cheever NA, Lara-Ruiz JM, et al. Media and technology use predicts ill-being among children, preteens and teenagers independent of the negative health impacts of exercise and eating habits. *Comput Human Behav* 2014 Jun;35:364-375 [FREE Full text] [doi: [10.1016/j.chb.2014.01.036](https://doi.org/10.1016/j.chb.2014.01.036)] [Medline: [25717216](https://pubmed.ncbi.nlm.nih.gov/25717216/)]
49. Woods HC, Scott H. #Sleepyteens: social media use in adolescence is associated with poor sleep quality, anxiety, depression and low self-esteem. *J Adolesc* 2016 Dec;51:41-49. [doi: [10.1016/j.adolescence.2016.05.008](https://doi.org/10.1016/j.adolescence.2016.05.008)] [Medline: [27294324](https://pubmed.ncbi.nlm.nih.gov/27294324/)]
50. Gonzales AL, Hancock JT. Mirror, mirror on my Facebook wall: effects of exposure to Facebook on self-esteem. *Cyberpsychol Behav Soc Netw* 2011 Jan 01;14(1-2):79-83. [doi: [10.1089/cyber.2009.0411](https://doi.org/10.1089/cyber.2009.0411)] [Medline: [21329447](https://pubmed.ncbi.nlm.nih.gov/21329447/)]
51. López-Pérez B, Fernández-Castilla B. Children's and adolescents' conceptions of happiness at school and its relation with their own happiness and their academic performance. *J Happiness Stud* 2018 Jun 07;19:1811-1830 [FREE Full text] [doi: [10.1007/s10902-017-9895-5](https://doi.org/10.1007/s10902-017-9895-5)]
52. Vannucci A, Ohannessian CM. Social media use subgroups differentially predict psychosocial well-being during early adolescence. *J Youth Adolescence* 2019 Jun 29;48(8):1469-1493. [doi: [10.1007/s10964-019-01060-9](https://doi.org/10.1007/s10964-019-01060-9)]
53. Kohyama J. Self-reported academic performance and lifestyle habits of school children in Japan. *Int J Child Health Nutr* 2017 Sep 11;6(3):90-97. [doi: [10.6000/1929-4247.2017.06.03.1](https://doi.org/10.6000/1929-4247.2017.06.03.1)]

Abbreviations

PedsQL: Pediatric Quality of Life Inventory

Edited by G Eysenbach; submitted 18.03.20; peer-reviewed by E Neter, Z Ma, JR Bautista; comments to author 25.09.20; revised version received 11.11.20; accepted 12.04.21; published 26.05.21.

Please cite as:

Donelle L, Facca D, Burke S, Hiebert B, Bender E, Ling S

Exploring Canadian Children's Social Media Use, Digital Literacy, and Quality of Life: Pilot Cross-sectional Survey Study

JMIR Form Res 2021;5(5):e18771

URL: <https://formative.jmir.org/2021/5/e18771>

doi: [10.2196/18771](https://doi.org/10.2196/18771)

PMID: [34037525](https://pubmed.ncbi.nlm.nih.gov/34037525/)

©Lorie Donelle, Danica Facca, Shauna Burke, Bradley Hiebert, Emma Bender, Stephen Ling. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 26.05.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Neural Machine Translation–Based Automated Current Procedural Terminology Classification System Using Procedure Text: Development and Validation Study

Hyeon Joo^{1,2}, MHI, MS; Michael Burns², MD, PhD; Sai Saradha Kalidaikurichi Lakshmanan², MS; Yaokun Hu², MS; V G Vinod Vydiswaran^{1,3}, PhD

¹Department of Learning Health Sciences, University of Michigan, Ann Arbor, MI, United States

²Department of Anesthesiology, University of Michigan, Ann Arbor, MI, United States

³School of Information, University of Michigan, Ann Arbor, MI, United States

Corresponding Author:

Hyeon Joo, MHI, MS

Department of Learning Health Sciences

University of Michigan

North Campus Research Center

Ann Arbor, MI, 48109

United States

Phone: 1 7347649826

Email: thejoo@umich.edu

Abstract

Background: Administrative costs for billing and insurance-related activities in the United States are substantial. One critical cause of the high overhead of administrative costs is medical billing errors. With advanced deep learning techniques, developing advanced models to predict hospital and professional billing codes has become feasible. These models can be used for administrative cost reduction and billing process improvements.

Objective: In this study, we aim to develop an automated anesthesiology current procedural terminology (CPT) prediction system that translates manually entered surgical procedure text into standard forms using neural machine translation (NMT) techniques. The standard forms are calculated using similarity scores to predict the most appropriate CPT codes. Although this system aims to enhance medical billing coding accuracy to reduce administrative costs, we compare its performance with that of previously developed machine learning algorithms.

Methods: We collected and analyzed all operative procedures performed at Michigan Medicine between January 2017 and June 2019 (2.5 years). The first 2 years of data were used to train and validate the existing models and compare the results from the NMT-based model. Data from 2019 (6-month follow-up period) were then used to measure the accuracy of the CPT code prediction. Three experimental settings were designed with different data types to evaluate the models. Experiment 1 used the surgical procedure text entered manually in the electronic health record. Experiment 2 used preprocessing of the procedure text. Experiment 3 used preprocessing of the combined procedure text and preoperative diagnoses. The NMT-based model was compared with the support vector machine (SVM) and long short-term memory (LSTM) models.

Results: The NMT model yielded the highest top-1 accuracy in experiments 1 and 2 at 81.64% and 81.71% compared with the SVM model (81.19% and 81.27%, respectively) and the LSTM model (80.96% and 81.07%, respectively). The SVM model yielded the highest top-1 accuracy of 84.30% in experiment 3, followed by the LSTM model (83.70%) and the NMT model (82.80%). In experiment 3, the addition of preoperative diagnoses showed 3.7%, 3.2%, and 1.3% increases in the SVM, LSTM, and NMT models in top-1 accuracy over those in experiment 2, respectively. For top-3 accuracy, the SVM, LSTM, and NMT models achieved 95.64%, 95.72%, and 95.60% for experiment 1, 95.75%, 95.67%, and 95.69% for experiment 2, and 95.88%, 95.93%, and 95.06% for experiment 3, respectively.

Conclusions: This study demonstrates the feasibility of creating an automated anesthesiology CPT classification system based on NMT techniques using surgical procedure text and preoperative diagnosis. Our results show that the performance of the NMT-based CPT prediction system is equivalent to that of the SVM and LSTM prediction models. Importantly, we found that including preoperative diagnoses improved the accuracy of using the procedure text alone.

KEYWORDS

CPT classification; natural language processing; machine learning; neural machine translation

Introduction

Background

In 2017, the administrative costs between insurers and providers in the United States were excessively high, totaling US \$812 billion, US \$2497 per capita, and representing 34.2% of the total health expenditures [1]. Billing and insurance-related expenditures in the United States are twice that of Canada and the Netherlands and quadruple that of Sweden [2-4]. Studies by Himmelstein et al [1,2] identify administrative costs as one of the major drivers of high health care expenditures in the United States and support the necessity of cost reduction. Furthermore, billing and insurance-related activities, a subset of administrative costs, are between 8.4% and 13.9% of the total revenue [5]. There is an opportunity to reduce these administrative costs associated with billing assignments using automation, leveraging machine learning and natural language processing techniques.

To analyze the administrative costs, Tseng et al [6] and O'Malley et al [7] illustrated the process of billing activities (known as *the life of a bill*) from the initial appointment to the time when payment was received. In this process, physicians are first involved in billing activities related to clinical services, even before the patient visit. Next, professional and hospital billing activities occur after the patient visit, where professional coders and billing management teams are involved in coding and managing claims that involve extensive and laborious hospital chart reviews. In particular, Tseng et al [6] found that billing and insurance-related activities carried out by physicians were between US \$6.36 per primary care visit and US \$51.20 per inpatient surgical procedure, which is 11%-31% of the total administrative costs. Similarly, professional and hospital billing teams' administrative costs range from US \$4.22 to US \$45.55 per procedure, accounting for 3%-36% of the total administrative costs.

Automating the coding process could also improve otherwise high rates of medical coding errors. According to the Comprehensive Error Rate Testing report published in the Centers for Medicare and Medicaid Services, Medicare's improper payments were US \$36.2 billion in 2017, 9.5% of the total Medicare payment. The common reasons for improper payments were insufficient documentation errors (64.1%), medical necessity errors (17.5%), incorrect coding errors (13.1%), and no documentation errors (1.7%) [8]. With regard to the effort of evaluating coding errors, numerous studies have associated physicians' limited training and knowledge of billing and insurance-related activities with high errors [9-12]. King et al [10] showed that family physicians' coding accuracy was 52% for established patients and 17% for new patients, as established patients were undercoded, thus failing to report the full services provided, whereas new patients were overcoded, an abuse in reporting medical services not actually performed.

In our study, we developed a two-step neural machine translation (NMT) model to automate current procedural terminology (CPT) coding. This NMT-based model translates manually entered surgical procedure text by a surgeon into a standard CPT description in step 1 and then calculates similarity scores to match the best CPT code in step 2. A single-step automated billing system estimates the likelihood of multiple classes, whereas standardized CPT descriptions from step 1 of the NMT-based model have potential use, in addition to this CPT prediction task. The standardized text from this NMT-based model can improve communication efficiency between physicians and hospital professionals in medical coding processes and ultimately reduce administrative costs by aiding in CPT code classification and reducing medical coding errors. In this study, we demonstrate both the translation performance and CPT prediction accuracy.

Related Work

Advanced machine learning methods have been developed to automate the manual classification of medical codes, including the International Classification of Diseases (ICD) [13-21] and CPT [22-24] codings. In these previous efforts, researchers used narrative clinical notes along with structured data elements to develop machine learning classification algorithms. For traditional machine learning algorithms, Koopman et al [14] proposed a binary support vector machine (SVM) classifier for multiple ICD-10 codes using n-gram features. Perotte et al [13] developed a hierarchy-based SVM by leveraging the hierarchical structure of the ICD-9 codes. Denck et al [16] showed an ensemble of classifier chains to predict billing codes using MRI log data. Virginio and dos Reis [21] and Wu et al [25] used the SVM model to discuss imbalanced data in medical billing data.

With advanced deep neural networks and natural language processing techniques, researchers can apply clinical notes without extensive preprocessing of raw narrative clinical text into a recurrent neural network or convolutional neural network. Xu et al [17] evaluated an ensemble model where unstructured, semistructured, and structural data were trained on text-based convolutional neural networks, bidirectional long short-term memory (LSTM), and decision tree for ICD prediction. Shi et al [26] developed a hierarchical LSTM model with attention techniques to classify clinical notes into ICD codes. Wang et al [27] proposed a label embedding attentive model using label-attentive text representation to improve text classification and applied it to predict ICD codes using clinical notes. Rios and Kavuluru [28] improved the accuracy of ICD prediction by transferring learning with PubMed biomedical abstracts.

NMT has emerged as a state-of-the-art machine learning method for translating text between human languages. The Transformer-based [29] NMT model uses an encoder-decoder framework with self-attention mechanisms to learn the weights of a translation model and understand the complex relationships between the source and target languages [30-32]. Although

NMT models have been adopted in health care to assist health communication such as speech translation or text translation from one language to another [33], these models remain underutilized.

Methods

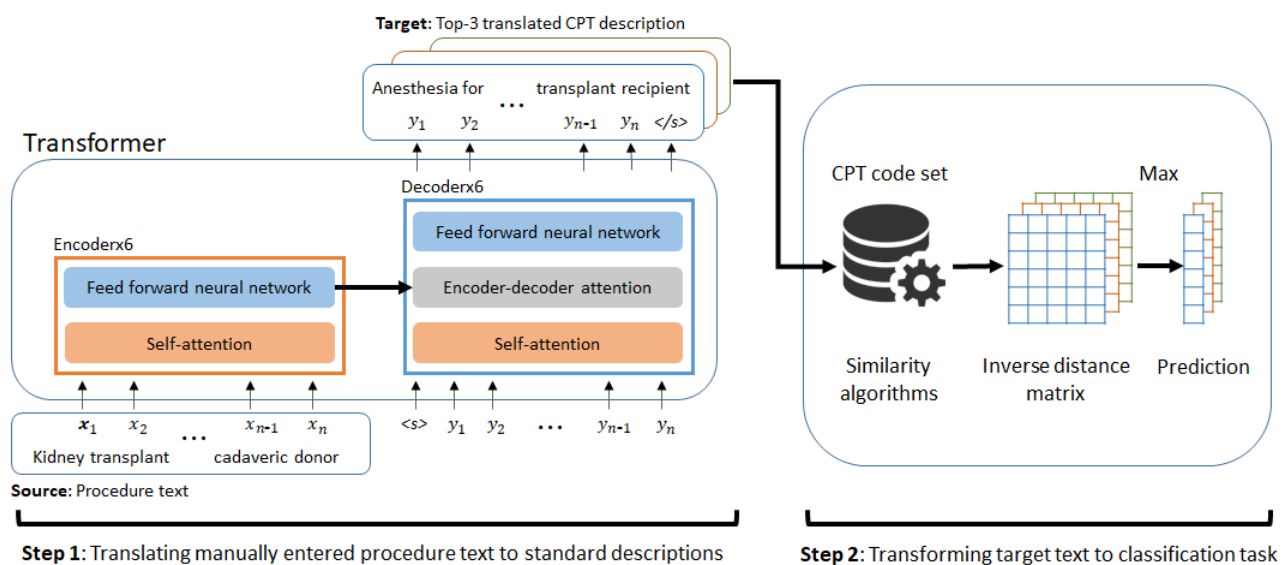
NMT Model Architecture

We developed an NMT-based automated CPT coding system that first translates surgical procedure texts in electronic health records (EHRs) into preferred terms from the Unified Medical Language System (UMLS) [34] and then normalizes the translated preferred term to predict CPT codes. The intuition of this machine translation approach is from text normalization,

the process of transforming noncanonical text into a standard form such as medical concept normalization [35-38] where medical terms are assigned to unique concept codes.

Within medicine, each surgical procedure contains a surgical procedure text and a preoperative diagnosis entered by a surgeon or surgical resident. After completion of the procedure, surgical and anesthesiology CPT codes were assigned by clinical staff and/or professional medical coders. The manually entered texts are the input source, and the preferred terms of the assigned CPT codes are the output target sentences of the NMT model. In our study, surgical procedure texts and preoperative diagnoses were the inputs of the model to predict CPT codes. The architecture of the NMT-based automated CPT prediction system is shown in Figure 1.

Figure 1. The architecture of an automated current procedural terminology coding system based on the Transformer model. CPT: current procedural terminology.



The overall architecture of an NMT-based system is composed of *translation* and *transformation* components. The *translation* component (step 1 in Figure 1) translates surgical procedure text into a standard form using the *Transformer* model [29]. This model uses a set of encoders and decoders using self-attention mechanisms and feed-forwarding neural networks. This Transformer model computes representations between a sequence of source word embeddings, x_i , and a sequence of target word embeddings, y_i , trained from the procedure text and the description of the CPT code.

To train the NMT model, the source and target sentences need to be paired between the manually entered procedure text (source) and the preferred terms of the CPT code (target). This is similar to the development of paired bilingual sentences to train a language translation model. Once trained, the NMT model generates multiple candidate translation outputs ranked by a beam search algorithm. The top three target sentences were retained and processed through step 2: *transformation*.

With the three target sentences, the best CPT code was computed in the transformation step using the Levenshtein and Jaccard distances. The distance was computed between a target sentence and the preferred term of all CPT codes in the UMLS. Each

sample was compared with the CPT descriptions, and its value was stored in a distance matrix, $X \in \mathbb{R}^{M \times N}$, where M is the number of the sample size and N is the number of CPT labels. As closer distances signify better matching to the CPT descriptions, we used an inverse distance matrix, $1/(X + 1)$, to maintain similarity scores for the three target sentences. The final prediction of the CPT code was based on the highest score in the matrix. X is an M by N matrix, where M is the number of the sample size and N is the number of the label size.

$$\text{Distance Matrix} = X \in \mathbb{R}^{M \times N} \quad (1)$$

$$\text{Inverse Distance Matrix}_{i,j} = 1/(x_{i,j} + 1), i = 1, \dots, m; j = 1, \dots, n \quad (2)$$

$$\text{Prediction}_i = \text{argmax}_{j \in J} (1/(x_{i,j} + 1)), i = 1, \dots, m; J = \{1, \dots, n\} \quad (3)$$

The key implication of this two-step NMT-based system is to reframe the translation task to a multiclass classification task using translation and transformation steps. Unlike a single classifier used in other automated coding algorithms [15-17,26,27,39-42], the two-step NMT model translates noncanonical text of human natural language into a normalized

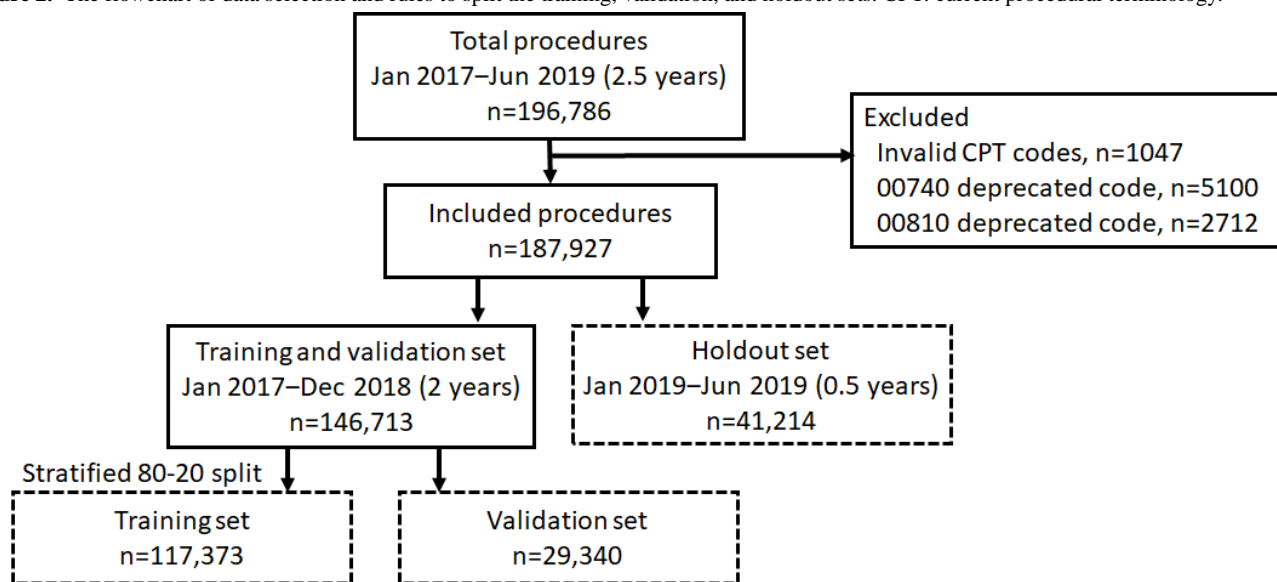
form and then transforms the normalized translation into an appropriate CPT code.

Data

The Multicenter Perioperative Outcomes Group (MPOG) is a nonprofit academic consortium of more than 50 academic and community hospitals across 19 states in the United States and the Netherlands [43,44]. The extent of data in the MPOG is relevant to the field of anesthesiology encompassing perioperative patient care, covering preoperative, intraoperative, and postoperative clinical practice. Through the MPOG infrastructure, a large volume of perioperative data such as patient vital signs, ventilation, medications, laboratory values, and administrative billing data from EHRs at different centers have been systematically aggregated via automated extraction and validated by clinical experts.

We collected all operative procedures performed at Michigan Medicine from January 2017 to June 2019 (2.5 years), resulting in 196,786 operative cases. In these data, we found that 10 unique CPT codes were invalid due to typographic errors and two unique CPT codes (00740: anesthesia for upper gastrointestinal procedures; 00810: anesthesia for lower gastrointestinal procedures) were deprecated in 2018 and replaced with newer CPT codes (00731 and 00732 and 00811, 00812, and 00813, respectively). These invalid and deprecated CPT codes (8859 of operative cases), were excluded from the analysis (Multimedia Appendix 1). The total number of operative procedures used to develop and evaluate the machine learning models used in this study was 187,927. Of the 272 anesthesiology CPT codes in the UMLS, we found 269 unique anesthesiology CPT codes in the final data set. A detailed flowchart of the inclusion and exclusion of the data is shown in Figure 2.

Figure 2. The flowchart of data selection and rules to split the training, validation, and holdout sets. CPT: current procedural terminology.



Training, Validation, and Holdout Data Splits

The final number of operative cases used in this study was 187,927. Procedures performed between January 2019 and June 2019 (6 months) were set aside as the holdout set, and procedures performed in 2017 and 2018 (2 years) were used to train and validate the machine learning models.

The implication of splitting the holdout set based on an explicit date before and after 2019 from the training and validation sets is to measure the behavior of predictive models on unseen future data. The uncertainty of the distribution of surgical procedures should remain separate from training to real predictions. This date-based split technique minimizes data leakage and

information accidentally shared between the training and holdout sets.

The data used for training and validation were stratified based on CPT codes (labels) and randomly split into 80/20 training and validation sets for model development and hyperparameter tuning. The stratified split between the training and validation sets enables maintenance of the same rate of CPT codes in both data sets, allowing an even split of rare surgical procedures.

The prevalence of the surgical procedures per CPT category in the training, validation, and holdout sets is shown in Table 1. The notable difference between the training and the holdout set is that the percentage of upper and lower abdomen procedures was higher in 2019 (holdout set) than in 2017 and 2018 (training set).

Table 1. The prevalence of anesthesiology procedures and services categorized by area of the body in the training, validation, and holdout data sets (N=187,927).

Body part ^a	CPT ^b codes	Data set ^c , n (%)		
		Training set ^d (n=117,373)	Validation set ^d (n=29,340)	Holdout set ^e (n=41,214)
Head	00100-00222	27,863 (23.74)	6964 (23.73)	9110 (22.10)
Neck	00300-00352	9403 (8.01)	2352 (8.01)	2822 (6.84)
Thorax (chest and shoulder)	00400-00474	5554 (4.73)	1388 (4.73)	1710 (4.14)
Intrathoracic	00500-00580	8781 (7.48)	2194 (7.47)	2903 (7.04)
Spine and spinal cord	00600-00670	2606 (2.22)	651 (2.21)	883 (2.14)
Upper abdomen	00700-00797	12,083 (10.29)	3022 (10.29)	5518 (13.38)
Lower abdomen	00800-00882	13,338 (11.36)	3334 (11.36)	6547 (15.88)
Perineum	00902-00952	9853 (8.39)	2464 (8.39)	3001 (7.28)
Pelvis (except hip)	01112-01173	551 (0.46)	137 (0.46)	194 (0.47)
Upper leg (except knee)	01200-01274	2390 (2.03)	595 (2.02)	674 (1.63)
Knee and popliteal area	01320-01444	2629 (2.24)	658 (2.24)	792 (1.92)
Lower leg (below knee)	01462-01522	2117 (1.80)	530 (1.80)	606 (1.47)
Shoulder and axilla	01610-01680	2084 (1.77)	520 (1.77)	668 (1.62)
Upper arm and elbow	01710-01782	785 (0.66)	196 (0.66)	192 (0.46)
Forearm, wrist, and hand	01810-01860	3035 (2.58)	757 (2.58)	870 (2.11)
Radiological procedure	01916-01936	8709 (7.42)	2179 (7.42)	2829 (6.86)
Burn excisions or debridement	01951-01953	505 (0.43)	126 (0.42)	105 (0.25)
Obstetric	01958-01969	4969 (4.23)	1243 (4.23)	1722 (4.17)
Other procedure	01990-01999	118 (0.10)	30 (0.10)	68 (0.16)

^aAnesthesiology current procedural terminology codes are categorized based on the area of the body part.

^bCPT: current procedural terminology.

^cThe percentage may not sum up to 100 because of rounding.

^dThe training and validation sets were stratified and split to maintain the same prevalence of procedures.

^eThe holdout set is new data collected for 6 months to prevent data leakage.

Two types of data, the *operative procedure text* and *preoperative diagnosis*, were extracted from the MPOG. The operative procedure text is a short description of the surgical operation, and the preoperative diagnosis is a disease diagnosis specific to the patient and correlates with the planned surgical procedure. Both text fields are typically entered manually into an EHR system by a surgeon or surgical resident before the surgery. In our training data set, the average length of the procedure text was 5.12 (SD 3.57) words, and the average length of the preoperative diagnosis was 4.12 (SD 2.5) words.

In addition to free-text narratives in EHRs, standard forms of CPT descriptions were extracted from the preferred terms in the UMLS. As the UMLS preserves canonical presentations of

medical concepts, CPT codes and preferred terms maintain a one-to-one association. The average length of the preferred terms in the UMLS was 13.23 (SD 6.44) words.

The basic descriptive statistics of the data sets in [Table 2](#) show the number of tokens (words) recorded in the procedure text, preoperative diagnosis, and preferred terms of CPT from the UMLS in each training, validation, and holdout sets. [Table 2](#) also shows the large variation of the manually entered procedure text with 13,847 unique procedure texts in the training set for the 252 unique CPT codes. This indicates the challenge of CPT coding, even with a single data field, to find the most appropriate CPT code.

Table 2. Descriptive statistics of operative procedure text, preoperative diagnosis, and preferred terms.

Data set	Number of unique CPT ^a codes	Number of unique procedure texts	Number of tokens in procedure text ^b		Number of tokens in preoperative diagnosis ^b		Number of tokens in preferred terms ^b	
			Mean (SD)	Range	Mean (SD)	Range	Mean (SD)	Range
Training	252	13,847	5.12 (3.57)	1-60	4.12 (2.5)	0-15	13.23 (6.44)	3-46
Validation	231	6012	5.15 (3.64)	1-51	4.11 (2.5)	0-13	13.23 (4.45)	3-41
Holdout	224	6731	4.98 (3.52)	1-60	4.01 (2.52)	0-13	13.24 (6.18)	3-41

^aCPT: current procedural terminology.

^bThe unit of descriptive statistics is token (word).

Models: SVM, LSTM, and NMT

The NMT-based automated CPT prediction system is supported by the encoder-decoder methods proposed in the text normalization study by Lusetti et al [45] and the self-attention encoder and decoder in OpenNMT [46], which uses Google's base Transformer model and hyperparameters as shown in [Multimedia Appendix 2](#). This NMT-based system generates three translated descriptions of CPT codes and transforms this translation into CPT prediction using similarity algorithms. Although the system contains translation and transformation components that can be evaluated separately, the primary focus of the systems is to measure the performance of predicting CPT codes based on the NMT model to assist physicians and professional coders. Thus, the outcome measurement is based on the accuracy of the CPT prediction and is compared with other classifiers that were conducted in other studies.

We selected the SVM and LSTM models as the baseline models often adopted in medical billing prediction and as developed in our previous study [24]. For SVM model development, we used the standard Scikit-learn packages in Python (Python Software Foundation) and applied grid search cross-validation for training and tuning hyperparameters. The input features of the SVM model were bigrams extracted from the training data and weighted using the term frequency-inverse document frequency. Owing to the large size of features, we limited the inputs to terms that appeared at least four times in the procedure text and at least 15 times in the combined text of procedure and preoperative diagnosis text.

For the LSTM model development, a sequence of words from the procedure text and preoperative diagnosis text in the training data was fed into the embedding layer. The embedding layer then converted each word in the sequence to a vector representation using a Word2Vec model pretrained on PubMed, PubMed Central, and Wikipedia [47]. The LSTM model was trained on this sequence of vector representations and returned a hidden vector from each state that was passed through a fully connected layer. A final softmax layer was then used to predict the final label. The configurations and hyperparameters of the models are shown in [Multimedia Appendix 2](#).

Experimental Settings

We designed three experimental settings for CPT code prediction with different types of data to evaluate the

performance of NMT, SVM, and LSTM machine learning models.

Experiments 1 and 2 used surgical procedure texts as an input to machine learning models with no demographics or clinical information. As misspellings and use of acronyms in manual data entry are common, raw surgical procedure texts without preprocessing were used for experiment 1 and with preprocessing for experiment 2. The purpose of this experiment was to evaluate the process of text normalization from noncanonical procedure text into a single standard description using a translational model. As a CPT code is primarily determined on surgical procedure text, we focused on the association between procedure texts and standard target sentences in the UMLS by restricting the input of the models. The average length of words of the CPT description in the UMLS is 2.6 times longer than the average length of the procedure text, as shown in [Table 2](#). The performance of top-1 (the best prediction) and top-3 (within the top 3) accuracy will be compared between machine learning models.

Experiment 3 was designed to introduce a preoperative diagnosis in the prediction model. Instead of just the preoperative text, the preoperative diagnosis is also appended to the input of the models. The rationale of this experiment is to evaluate the impact of indirect information on CPT code prediction over that in experiment 2. This incorporation is logical, as surgical procedures can be listed with the same procedure text but coded differently because of the diversity in patient diagnosis. Experiment 3 aimed to determine if this multicoded issue could be mitigated by including the preoperative diagnosis in the models. Especially for the NMT model, where the source and target sentences are paired with the same information, the preoperative diagnosis included in the source would be extraneous. The average length of the combination used in the models was 9.1 (SD 4.55) words.

Preprocessing

Before training the models, the operative procedure and the preoperative diagnosis text were subjected to a series of preprocessing steps, including removing stop words, trimming white spaces, lowering cases, lemmatizing, correcting misspelled medical words, and expanding acronyms ([Multimedia Appendix 3](#)). The spelling correction and acronym expansion were manually reviewed, maintained, and applied to curated preoperative texts and the preoperative diagnosis. The curated version of the procedure text in combination with the

preoperative diagnosis was used as input for model development for experiments 2 and 3.

Evaluation and Performance Metrics

The primary performance metric used to evaluate our system is the accuracy of the documented CPT codes within the MPOG database. We defined two accuracies: top-1 is the accuracy where the true CPT code matches with the first top (most probable) model predicted CPT code, whereas top-3 is the accuracy where the true CPT code matches with any one of the top three most probable CPT codes predicted by the models. The top-1 and top-3 accuracies have different implications. The top-1 accuracy presents the single best-predicted CPT code to measure how models would accurately perform on the medical billing code assignment without additional human effort. In contrast, the top-3 accuracy provides the top three most probable CPT codes, reducing selection by the administrative staff to three probable choices. With this rationale, both top-1 and top-3 translated standard forms can be used to assist and audit billing codes by suggesting appropriate options in real time and improving the efficiency of communication between health care professionals.

The top-1 and top-3 accuracies were evaluated for both validation and holdout sets for the three machine learning models. During evaluation, we repeated the experiment 500 times, bootstrapping 20,000 samples with replacements and selecting the best model after tuning the hyperparameters on the validation set. The SVM and LSTM models returned the final output with the predicted probabilities of the 272 CPT labels. The three highest probabilities were selected for top-3

accuracy. As the output of the NMT model is translation sentences, the top three translations were used for the top-3 accuracy evaluation.

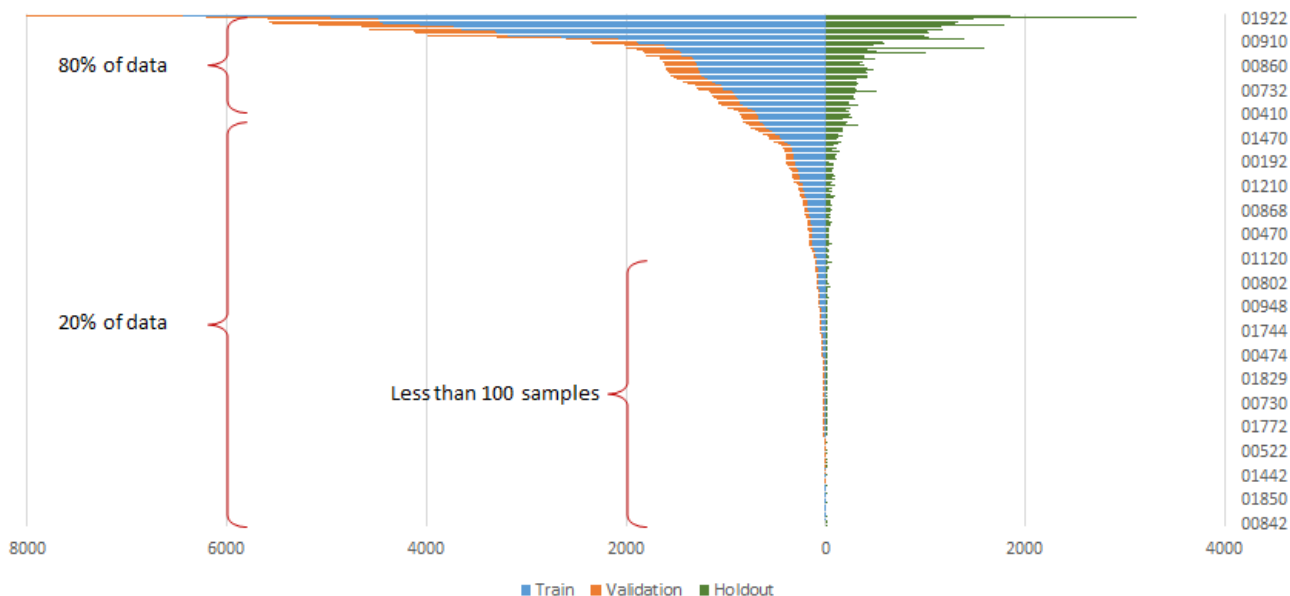
Evaluation of Imbalanced Labels

Class imbalance in medical coding is inevitable because of the nature of hospital services [21,25]. Regular hospital services will show more often in EHRs than in rare cases. For algorithm development, the lack of data in minority classes often creates potential issues and limitations. For example, minority classes that are often underevaluated are critical for patient care. A general approach is to exclude the minority from the data set [17,26,27] and analyze the results on majority classes. It may be valid for evaluating algorithm performance but may raise concerns about implementing the algorithm in clinical settings.

Figure 3 shows the CPT label distribution in our data set between the training, validation, and holdout sets. The top 52 CPT codes accounted for 80.2% (94,216/117,373) of surgical procedures at the University of Michigan, and 220 CPT codes shared the remaining 19.7% (23,157/117,373) of cases. Of these 220 CPT codes, 132 CPT codes, that is, 48.5% (132/272) of the total anesthesiology CPT codes, had less than 100 training samples for 2 years of hospital services.

For the sensitivity analysis of imbalanced data, the anesthesiology CPT codes were split into 10 groups with a mean of 25 (SD 1.3) codes per group to represent the different sizes of samples in the training set (Multimedia Appendix 4). We evaluated and compared the performance of each group between the NMT, SVM, and LSTM models.

Figure 3. The distribution of current procedural terminology codes in the training, testing, and holdout sets, sorted by most to least frequent codes.



Results

Top-1 and Top-3 Accuracy

The top-1 and top-3 accuracies of the NMT, SVM, and LSTM models over three experiments are summarized in [Table 3](#). For top-1 accuracy, the NMT model yielded the highest top-1

accuracy for both experiments 1 and 2 at 81.64% and 81.71% when compared with the SVM model (81.19% and 81.27%) and the LSTM model (80.96% and 81.07%), respectively. The SVM model yielded the highest top-1 accuracy of 84.30% in experiment 3, followed by the LSTM (83.70%) and NMT (82.80%) models.

Table 3. Performance comparison of support vector machine, long short-term memory, and neural machine translation models with raw, curated procedure text and combined, curated procedure text and preoperative diagnosis.

Model	Top-1 accuracy ^a (95% CI)		Top-3 accuracy ^b (95% CI)	
	Validation set	Holdout set	Validation set	Holdout set
Experiment 1: raw procedure text^c				
SVM ^d	83.61 (83.07-84.16)	81.19 (80.63-81.75)	97.56 (97.36-97.76)	95.64 (95.35-95.93)
LSTM ^e	81.86 (81.33-82.40)	80.94 (80.42-81.46)	95.38 (95.06-95.71)	95.72 (95.44-95.99)
NMT ^f	81.68 (81.14-82.21)	81.64 (81.11-82.18)	95.27 (94.96-95.58)	95.60 (95.30-95.89)
Experiment 2: curated procedure text^g				
SVM	83.38 (82.85-83.90)	81.27 (80.72-81.82)	97.45 (97.23-97.67)	95.75 (95.47-96.04)
LSTM	81.81 (81.28-82.34)	81.07 (80.54-81.59)	95.32 (95.00-95.64)	95.67 (95.40-95.95)
NMT	81.72 (81.18-82.26)	81.71 (81.18-82.24)	95.41 (95.11-95.71)	95.69 (95.40-95.98)
Experiment 3: curated procedure text and preoperative diagnosis^h				
SVM	87.62 (87.15-88.09)	84.30 (83.81-84.79)	99.16 (99.03-99.29)	95.88 (95.60-96.15)
LSTM	83.52 (83.00-84.04)	83.70 (83.20-84.20)	95.82 (95.53-96.12)	95.93 (95.65-96.20)
NMT	82.43 (81.90-82.96)	82.80 (82.31-83.29)	94.75 (94.44-95.06)	95.06 (94.77-95.35)

^aTop-1 accuracy is the accuracy of models on the best current procedural terminology code predicted, equivalent to the F1-score micro.

^bTop-3 accuracy is the accuracy of models if the true current procedural terminology is within the top three best codes predicted.

^cRaw procedure text is manually entered by physicians without text preprocessing.

^dSVM: support vector machine.

^eLSTM: long short-term memory.

^fNMT: neural machine translation.

^gCurated procedure text is cleaned text with preprocessing techniques.

^hThe curated procedure text is concatenated with preoperative diagnosis to training models for current procedural terminology prediction.

For top-3 accuracy, the SVM, LSTM, and NMT models achieved 95.64%, 95.72%, and 95.60% accuracy for experiment 1 and 95.75%, 95.67%, and 95.69% accuracy for experiment 2, respectively. As expected, all three models performed significantly higher when comparing top-3 accuracy with top-1 accuracy. The improvement rate of top-3 accuracy over top-1 accuracy in experiment 2 was 17.8%, 18.23%, and 17.1% for the SVM, LSTM, and NMT models, respectively.

By combining the preoperative diagnosis data in experiment 3, the three models improved top-1 accuracy when compared with experiments 1 and 2. The SVM model obtained the most significant enhancement by achieving 84.30% top-1 accuracy, a 3.7% increase from experience 2. The top-1 accuracy of the LSTM and NMT model was 83.70% and 82.80%, which increased by 3.4% and 1.3%, respectively. For top-3 accuracy, the SVM and LSTM model achieved 95.88% and 95.93%, which is 0.1% and 0.3% enhancement from experiment 2, respectively. The top-3 accuracy of the NMT model was 95.06%, which was reduced by 0.6%.

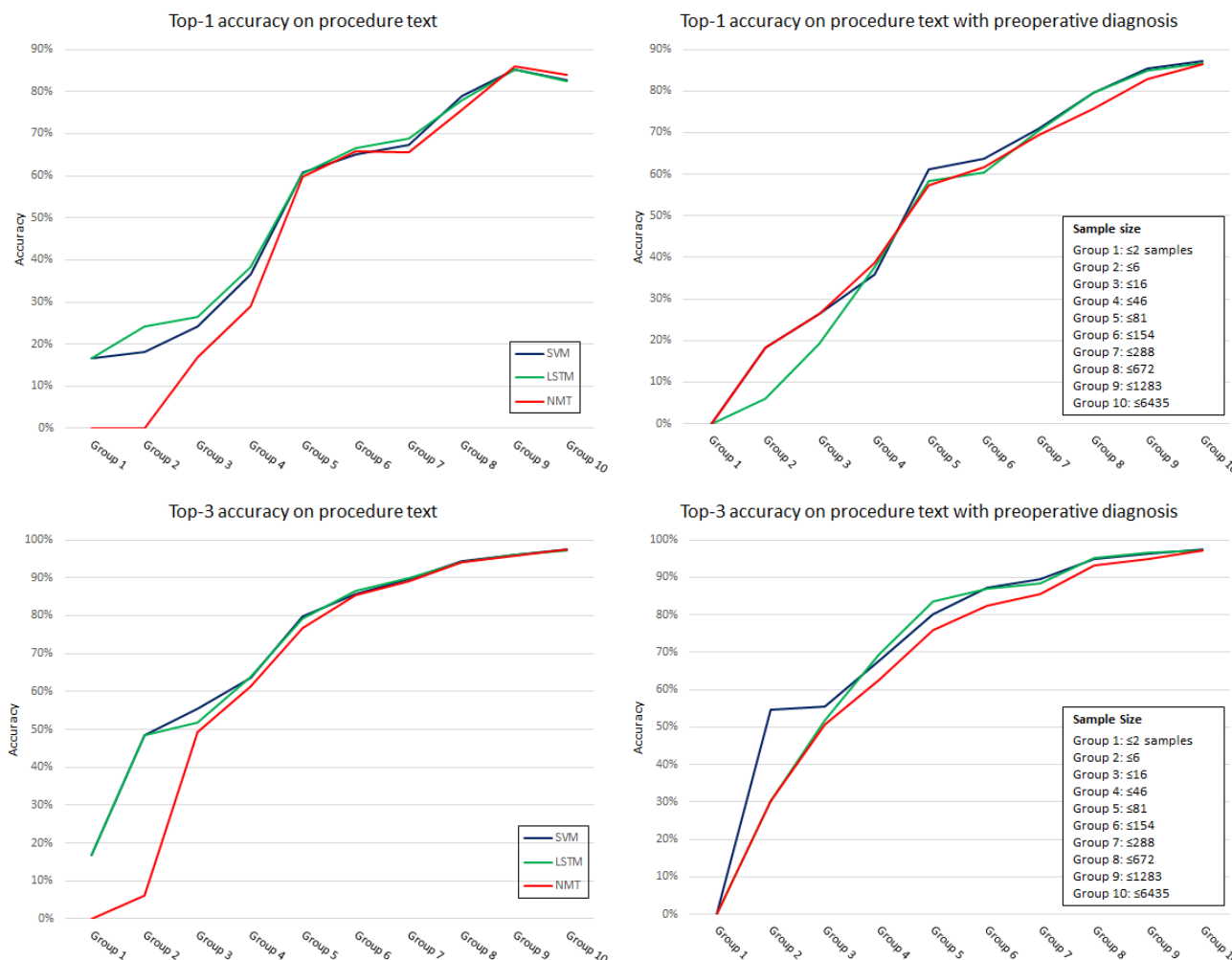
Accuracy by Training Sample Size

We further examined the results in [Table 3](#) based on the training sample size to understand the effect of imbalanced labels. A total of 272 CPT codes were split into 10 groups, and each group contained approximately 27 CPT codes based on the training sample size. For example, group 1 contained CPT codes ≤ 2 samples in the training set, and group 2 included CPT codes that have ≥ 2 samples and ≤ 6 samples. As the group number increased, the size of the samples increased.

The effect of imbalanced labels on performance is illustrated in [Figure 4](#). The top two-line charts show the top-1 accuracy on curated procedure text (top-left) and combined, curated procedure text with a preoperative diagnosis (top-right). The bottom two-line charts show the top-3 accuracy on curated procedure text (bottom-left) and combined text (bottom-right). The top-1 and top-3 accuracies of the NMT model using procedure text were better in group 10 (training sample size >1283) than in other models, but the performance slowly decreased as the sample size decreased. Details on the

performance of each imbalanced label are provided in [Multimedia Appendix 4](#).

Figure 4. The top-1 and top-3 accuracy comparison based on the training sample size. LSTM: long short-term memory; NMT: neural machine translation; SVM: support vector machine.



NMT Result

In addition to the CPT prediction accuracy, the translation performance of imbalanced labels from groups 1 to 10 is summarized in [Figure 5](#). As the sample size increased, the Bilingual Evaluation Understudy (BLEU) scores from all three experiments were significantly improved, and group 10 was close to 0.9. This means that the translation of manually entered procedure texts was close to the preferred terms in the UMLS. The overall BLEU scores of experiments 1, 2, and 3 on the holdout set were 0.872, 0.895, and 0.904, respectively. The detailed BLEU scores are provided in [Multimedia Appendix 5](#).

Examples of translated sentences from our NMT model are presented in [Table 4](#). This table includes the input source text, the output target text, and the gold standard translation used for the NMT model to translate manually entered procedure texts and preoperative diagnoses into standard CPT descriptions. We distinguished the preoperative diagnosis from the procedure text by underlining the source text. The example of 01220 (group 5) demonstrates how additional preoperative diagnosis information ratified the machine translation from manually entered procedure text.

Figure 5. Bilingual Evaluation Understudy scores of imbalanced labels for translating manually entered procedure text into preferred terms in step 1 of the neural machine translation–based model. BLEU: Bilingual Evaluation Understudy.

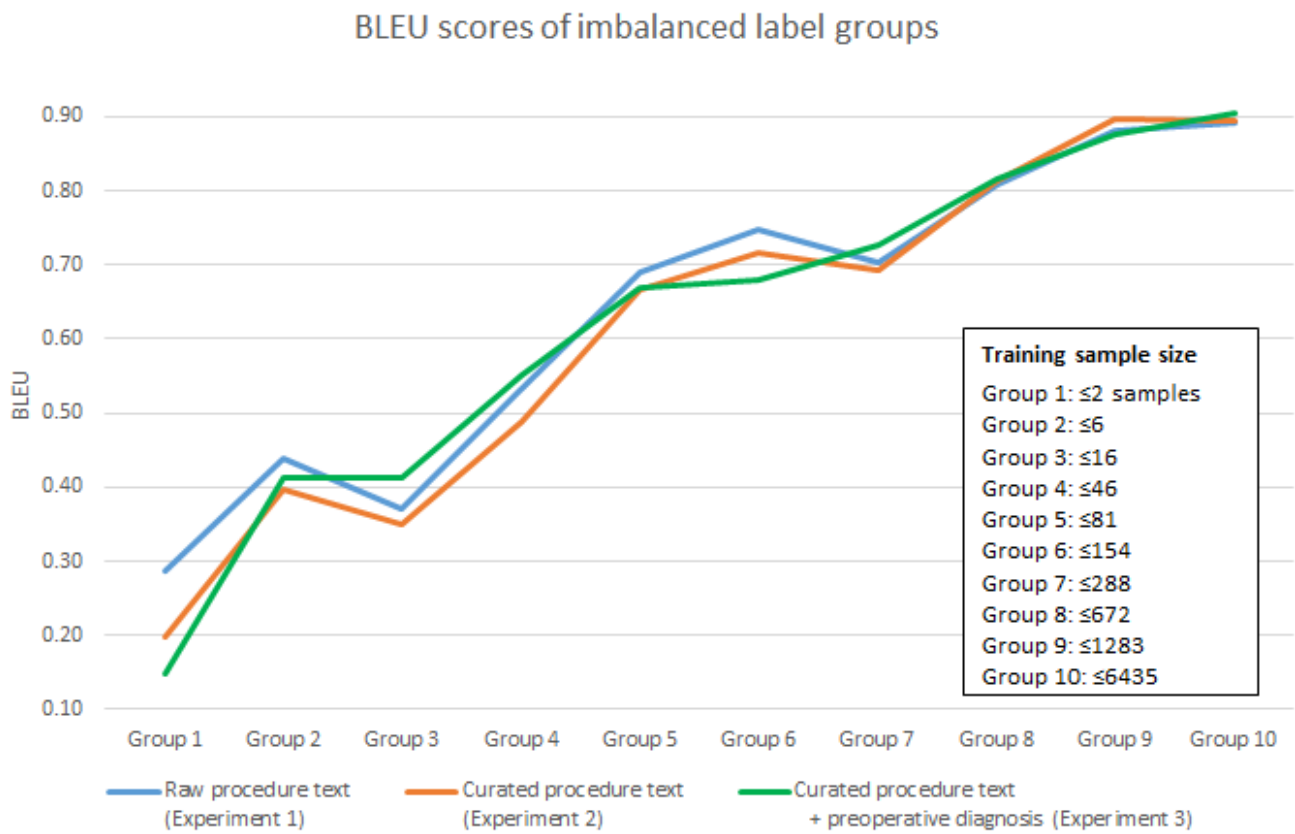


Table 4. The neural machine translation–based model’s source (input) and target (output) translation examples from the holdout data set. The translation task is to convert manually entered procedure text in electronic health records to preferred terms in the Unified Medical Language System.

Current procedural terminology code	Current procedural terminology description ^a	Example source text ^b	Target text ^c
Experiment 1 and 2 with procedure text			
00868 (group 7)	Anesthesia for extraperitoneal procedures in lower abdomen, including urinary tract; renal transplant (recipient)	Right kidney transplant cadaveric donor	Anesthesia for extraperitoneal procedures in lower abdomen including urinary tract renal transplant recipient
01610 (group 9)	Anesthesia for all procedures on nerves, muscles, tendons, fascia, and bursae of shoulder and axilla	Midline wide local excision melanoma iolm sln axillary inguinal	Anesthesia for all procedures on nerves muscles tendons fascia and bursae of shoulder and axilla
00142 (group 10)	Anesthesia for procedures on eye; lens surgery	Left phacoemulsification intraocular lens	Anesthesia for procedures on eye lens surgery
00104 (group 10)	Anesthesia for electroconvulsive therapy	Electroconvulsive therapy	Anesthesia for electroconvulsive therapy
01220 (group 5)	Anesthesia for all closed procedures involving upper two thirds of femur	Left closed reduction cast application	Anesthesia for all closed procedures involving hip joint
Experiment 3 with a combined procedure text and preoperative diagnosis (underlined)^d			
01402 (group 8)	Anesthesia for open or surgical arthroscopic procedures on knee joint; total knee arthroplasty	Right total knee arthroplasty knee arthritis	Anesthesia for open or surgical arthroscopic procedures on knee joint total knee arthroplasty
01922 (group 10)	Anesthesia for noninvasive imaging or radiation therapy	Magnetic resonance imaging cardiac anesthesia hypoplastic left heart syndrome	Anesthesia for noninvasive imaging or radiation therapy
00402 (group 9)	Anesthesia for procedures on the integumentary system on the extremities, anterior trunk, and perineum; reconstructive procedures on breast (eg, reduction or augmentation mammoplasty, muscle flaps)	Left breast tissue expander placement unilateral tissue expander infection	Anesthesia for procedures on the integumentary system on the extremities anterior trunk and perineum reconstructive procedures on breast eg reduction or augmentation mammoplasty muscle flaps
00812 (group 10)	Anesthesia for lower intestinal endoscopic procedures, endoscope introduced distal to duodenum; screening colonoscopy	Colonoscopy anesthesia encounter screening malignant neoplasm colon	Anesthesia for lower intestinal endoscopic procedures endoscope introduced distal to duodenum screening colonoscopy
01200 (group 5)	Anesthesia for all closed procedures involving hip joint	Left closed reduction cast application left prosthetic hip dislocation	Anesthesia for all closed procedures involving hip joint

^aThe current procedural terminology description is the *gold standard* translation.

^bSource text is an example of the preprocessed input of the neural machine translation model.

^cThe target text is the translated sentence from the neural machine translation model.

^dThe preoperative diagnosis in the combined text is underlined to distinguish it from the procedure text.

Discussion

Principal Findings

In this study, we demonstrated the feasibility of predicting anesthesiology CPT codes using a two-step machine translation approach. Our results indicate that the top-1 and top-3 accuracies of the NMT-based model were equivalent to those of the SVM and LSTM models using procedure texts. We also demonstrated that the use of additional information, such as preoperative diagnosis, improves SVM, LSTM, and NMT model performance. Finally, we illustrated that imbalanced labels caused by low relative sample sizes negatively affected model accuracy.

We demonstrated the utility of machine learning models for use in medical billing applications. The automated CPT prediction systems developed in this study can improve the accuracy of medical billing coding by decreasing manual data reviews and

coding errors. Although 81% accuracy for the top-1 assignment remains low, the top-3 accuracies are above 95%. Therefore, recommendations from top-3 accuracy could be used by physicians and professional billing teams to improve coder accuracy and potentially reduce medical coding assignment and processing times. Translational models have the ability to increase communication efficiency between physicians and billing teams by normalizing the inputs used for assignments. Creating CPT prediction systems can also allow increased auditing efforts to find simple errors in cases that were initially undercoded or overcoded.

The addition of preoperative diagnoses to procedure texts in experiment 3 increased the top-1 accuracy of the SVM, LSTM, and NMT models by 3.7%, 3.2%, and 1.3%, respectively, when compared with experiment 2. The increased accuracy across all three models illustrates the importance of including preoperative diagnoses. For example, several cases may contain similar

generic procedure text, such as “exploratory laparotomy,” but entail different operative characteristics requiring differences in billing assignment. By including diagnoses, the models could identify subtle differences beyond the procedure text alone. In addition, there may be significance in the higher improved accuracies for the SVM and LSTM models when compared with the NMT model. The SVM and LSTM models use concatenated procedure text and preoperative diagnosis to predict a CPT code, whereas the NMT-based model uses a paired source and target sentence where the source is concatenated with procedure text and preoperative diagnosis and translated into the target, the preferred term of the CPT code.

This study showed a strong association between the performance of the models and the size of the training samples (Figure 4). When the sample size is above 81 (group 6 or above), all three models' performance is above 60% for top-1 accuracy and above 75% for top-3 accuracy, which is a significant increase compared with smaller sample sizes. Although all three models performed similarly, the NMT model was more sensitive to the sample size. Comparatively, the NMT model showed a strong performance in the larger sample size (groups 9 and 10) but lower performance when the sample size was smaller than 81 (group 5 or below). This may be due to the complex NMT methods based on the Transformer model compared with the SVM and LSTM models.

The NMT model can be applied in medicine beyond billing. Any free-text entries that often include human or systematic errors can be translated into a standard description as a normalization task, mapping clinical terms in medical notes to a standardized vocabulary. Under the recent effort of clinical entity normalization [36,37], our approach using NMT can be utilized for normalization tasks. This approach offers many advantages: (1) the translated sentence is transparent for clinicians and researchers to identify whether NMT works, allowing the model to be adjusted based on their feedback. (2) Human feature engineering effort is minimal as long as the source and target sentences are paired in model training.

This study has demonstrated the use of creating an NMT-based automatic anesthesiology CPT classification system. Its performance is equivalent to that of the SVM and LSTM models and presents itself as another method for machine learning applications in medicine.

Study Limitations

This study has several important limitations that must be considered:

1. The models in this study were developed and evaluated using anesthesiology CPT codes from data collected for about 2.5 years. The holdout data comprised only 6 months, from January to June 2019. Although the proportion of CPT codes in the holdout set is similar to the training data, there is still a risk of hidden seasonal effects or trends in rare cases that may not stand out on initial interpretation.
2. Although limiting to two features (operative procedure text and preoperative diagnosis), it allows us to evaluate the

translation performance from manually entered text to standard form. This may not achieve the best CPT prediction performance. Some complex examples, such as multiple CPT codes assigned to the same procedure text, may not convey enough information with two features. In addition, lower-frequency codes, as shown in Figure 4, may require careful assessment for CPT prediction.

3. The Centers for Medicare and Medicaid Services reported about Medicare's improper payments in the Comprehensive Error Rate Testing report. This implies that our data set may contain inherent medical coding errors [10,48]. To reduce these coding errors, a manual review of CPT codes by coding experts is required to enhance data quality before training models.
4. There are residual CPT codes for procedures when no equivalent or limited documentation is available. The description of the residual codes contains “...not otherwise specified.” These residual codes often lead to undercoding, which fails to capture all clinical procedures and reinforce models to learn undercoding behaviors and negatively affects prediction performance.
5. Anesthesiology CPT codes were chosen for classification because of the limited complexity in assignment. There is one anesthesiology code assigned per operative case, allowing for a multiclass single classification, whereas multiple surgical CPTs are often assigned per operative case. In addition to classification complexity, there is a limited number of CPT codes used for anesthesiology classification compared with surgical CPT codes (<300 vs >5000). The extension of the scope beyond anesthesiology CPT codes requires further evaluation to reconfigure the NMT-based model and to adjust better similarity algorithms.
6. The NMT-based model performed on par with one-step models in terms of accuracy, demonstrating the use of translation models to perform CPT classification. However, without a significant increase in accuracy, the additional processing time may prove significant when applied to larger, more complicated billing classifications, such as ICD and surgical CPT predictions. This may limit the real-world application of translation models for these tasks.
7. The performance of complex models, such as NMT, is more subjective to the small sample size than traditional machine learning models. In Figure 4 and Multimedia Appendix 4, the NMT-based model's performance is lower with fewer training samples and better when more training samples are available.

Conclusions

In this study, we demonstrated an automated anesthesiology CPT classification system based on machine translation techniques using surgical procedure text and preoperative diagnosis. The overall results show that the NMT-based CPT prediction model is equivalent to the SVM and LSTM models. Although the NMT-based model was not significantly outperformed, this new approach enables researchers to normalize manually entered clinical text into a standard form for use in classification tasks.

Acknowledgments

All work and partial funding can be attributed to the Department of Anesthesiology, University of Michigan Medical School (Ann Arbor, Michigan). The research reported in this publication was supported by grant T32GM103730 from the National Institute for General Medical Sciences of the National Institutes of Health (Bethesda, Maryland; to MB). The content of this study is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. Support for underlying EHR data collection was provided in part by the Blue Cross Blue Shield of Michigan (BCBSM; Detroit, Michigan) and Blue Care Network as part of the BCBSM Value Partnerships program for contributing hospitals in the state of Michigan. Although the BCBSM and the MPOG work collaboratively, the opinions, beliefs, and viewpoints expressed by the authors do not necessarily reflect the opinions, beliefs, and viewpoints of BCBSM or any of its employees.

Authors' Contributions

HJ conceived and designed the study. HJ and YH designed and tested the NMT model. HJ, SSKL, and YH acquired and cleaned the data. SSKL and HJ tested the baseline experiments with the SVM and LSTM models. VGVV supervised the study, and MB provided clinical feedback. HJ and MB wrote the initial draft, and HJ, MB, and VGVV made critical revisions to the manuscript. All authors helped with interpreting the results and final review of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Detail counts of data inclusion and exclusion criteria.

[[DOCX File , 18 KB - formative_v5i5e22461_app1.docx](#)]

Multimedia Appendix 2

Best hyperparameters for the support vector machine, long short-term memory, and neural machine translation models.

[[DOCX File , 19 KB - formative_v5i5e22461_app2.docx](#)]

Multimedia Appendix 3

Preprocessing step of the cleaning procedure and preoperative diagnosis text.

[[DOCX File , 18 KB - formative_v5i5e22461_app3.docx](#)]

Multimedia Appendix 4

Performance analysis of imbalanced labels broken down into 10 groups based on sample size. Experiment 1 is exempt from this analysis because the result is similar to that of experiment 2.

[[DOCX File , 26 KB - formative_v5i5e22461_app4.docx](#)]

Multimedia Appendix 5

The detailed corpus-level Bilingual Evaluation Understudy scores of imbalanced labels for translating manually entered procedure text into preferred terms on the validation and holdout sets.

[[DOCX File , 14 KB - formative_v5i5e22461_app5.docx](#)]

References

1. Himmelstein DU, Campbell T, Woolhandler S. Health care administrative costs in the United States and Canada, 2017. *Ann Intern Med* 2020 Jan 21;172(2):134-142 [[FREE Full text](#)] [doi: [10.7326/M19-2818](#)] [Medline: [31905376](#)]
2. Himmelstein DU, Jun M, Busse R, Chevreur K, Geissler A, Jeurissen P, et al. A comparison of hospital administrative costs in eight nations: US costs exceed all others by far. *Health Aff (Millwood)* 2014 Sep;33(9):1586-1594 [[FREE Full text](#)] [doi: [10.1377/hlthaff.2013.1327](#)] [Medline: [25201663](#)]
3. Papanicolaos I, Woskie LR, Jha AK. Health care spending in the United States and other high-income countries. *J Am Med Assoc* 2018 Mar 13;319(10):1024-1039 [[FREE Full text](#)] [doi: [10.1001/jama.2018.1150](#)] [Medline: [29536101](#)]
4. Emanuel EJ. The real cost of the US health care system. *J Am Med Assoc* 2018 Mar 13;319(10):983-985 [[FREE Full text](#)] [doi: [10.1001/jama.2018.1151](#)] [Medline: [29536081](#)]
5. Kahn JG, Kronick R, Kreger M, Gans DN. The cost of health insurance administration in California: estimates for insurers, physicians, and hospitals. *Health Aff (Millwood)* 2005 Nov;24(6):1629-1639 [[FREE Full text](#)] [doi: [10.1377/hlthaff.24.6.1629](#)] [Medline: [16284038](#)]

6. Tseng P, Kaplan RS, Richman BD, Shah MA, Schulman KA. Administrative costs associated with physician billing and insurance-related activities at an academic health care system. *J Am Med Assoc* 2018 Feb 20;319(7):691-697 [FREE Full text] [doi: [10.1001/jama.2017.19148](https://doi.org/10.1001/jama.2017.19148)] [Medline: [29466590](https://pubmed.ncbi.nlm.nih.gov/29466590/)]
7. O'Malley KJ, Cook KF, Price MD, Wildes KR, Hurdle JF, Ashton CM. Measuring diagnoses: ICD code accuracy. *Health Serv Res* 2005 Oct;40(5 Pt 2):1620-1639 [FREE Full text] [doi: [10.1111/j.1475-6773.2005.00444.x](https://doi.org/10.1111/j.1475-6773.2005.00444.x)] [Medline: [16178999](https://pubmed.ncbi.nlm.nih.gov/16178999/)]
8. 2017 Medicare fee-for-service supplemental improper payment data. U.S. Department of Health and Human Services. URL: <https://www.cms.gov/Research-Statistics-Data-and-Systems/Monitoring-Programs/Medicare-FFS-Compliance-Programs/CERT/Downloads/2017-Medicare-FFS-Improper-Payment.pdf> [accessed 2021-05-08]
9. Austin RE, von Schroeder HP. How accurate are we? A comparison of resident and staff physician billing knowledge and exposure to billing education during residency training. *Can J Surg* 2019 Oct 01;62(5):340-346 [FREE Full text] [doi: [10.1503/cjs.008718](https://doi.org/10.1503/cjs.008718)] [Medline: [31550096](https://pubmed.ncbi.nlm.nih.gov/31550096/)]
10. King MS, Sharp L, Lipsky MS. Accuracy of CPT evaluation and management coding by family physicians. *J Am Board Fam Pract* 2001;14(3):184-192 [FREE Full text] [Medline: [11355050](https://pubmed.ncbi.nlm.nih.gov/11355050/)]
11. Achkar MA, Kengeri-Srikantiah S, Yamane BM, Villasmil J, Busha ME, Gebke KB. Billing by residents and attending physicians in family medicine: the effects of the provider, patient, and visit factors. *BMC Med Educ* 2018 Jun 13;18(1):136 [FREE Full text] [doi: [10.1186/s12909-018-1246-7](https://doi.org/10.1186/s12909-018-1246-7)] [Medline: [29895287](https://pubmed.ncbi.nlm.nih.gov/29895287/)]
12. Duszak R, Blackham WC, Kusiak GM, Majchrzak J. CPT coding by interventional radiologists: a multi-institutional evaluation of accuracy and its economic implications. *J Am Coll Radiol* 2004 Oct;1(10):734-740 [FREE Full text] [doi: [10.1016/j.jacr.2004.05.003](https://doi.org/10.1016/j.jacr.2004.05.003)] [Medline: [17411693](https://pubmed.ncbi.nlm.nih.gov/17411693/)]
13. Perotte A, Pivovarov R, Natarajan K, Weiskopf N, Wood F, Elhadad N. Diagnosis code assignment: models and evaluation metrics. *J Am Med Inform Assoc* 2014 Mar 01;21(2):231-237 [FREE Full text] [doi: [10.1136/amiajnl-2013-002159](https://doi.org/10.1136/amiajnl-2013-002159)] [Medline: [24296907](https://pubmed.ncbi.nlm.nih.gov/24296907/)]
14. Koopman B, Zuccon G, Nguyen A, Bergheim A, Grayson N. Automatic ICD-10 classification of cancers from free-text death certificates. *Int J Med Inform* 2015 Nov;84(11):956-965. [doi: [10.1016/j.ijmedinf.2015.08.004](https://doi.org/10.1016/j.ijmedinf.2015.08.004)] [Medline: [26323193](https://pubmed.ncbi.nlm.nih.gov/26323193/)]
15. Baumel T, Nassour-Kassis J, Cohen R, Elhadad M, Elhadad N. Multi-label classification of patient notes : a case study on ICD code assignment. The Workshops of the Thirty-Second AAAI Conference on Artificial Intelligence. 2017. URL: <https://www.aaai.org/ocs/index.php/WS/AAAIW18/paper/view/16881/15610> [accessed 2021-05-08]
16. Denck J, Landschütz W, Nairz K, Heverhagen JT, Maier A, Rothgang E. Automated billing code retrieval from MRI scanner log data. *J Digit Imaging* 2019 Dec 25;32(6):1103-1111 [FREE Full text] [doi: [10.1007/s10278-019-00241-z](https://doi.org/10.1007/s10278-019-00241-z)] [Medline: [31240415](https://pubmed.ncbi.nlm.nih.gov/31240415/)]
17. Xu K, Lam M, Pang J, Gao X, Band C, Mathur P, et al. Multimodal machine learning for automated ICD coding. *Machine Learning for Healthcare*. 2019. URL: <http://proceedings.mlr.press/v106/xu19a/xu19a.pdf> [accessed 2021-05-08]
18. Zeng M, Li M, Fei Z, Yu Y, Pan Y, Wang J. Automatic ICD-9 coding via deep transfer learning. *Neurocomputing* 2019 Jan;324:43-50 [FREE Full text] [doi: [10.1016/j.neucom.2018.04.081](https://doi.org/10.1016/j.neucom.2018.04.081)]
19. Li M, Fei Z, Zeng M, Wu F, Li Y, Pan Y, et al. Automated ICD-9 coding via a deep learning approach. *IEEE/ACM Trans Comput Biol Bioinform* 2019 Jul 1;16(4):1193-1202 [FREE Full text] [doi: [10.1109/TCBB.2018.2817488](https://doi.org/10.1109/TCBB.2018.2817488)] [Medline: [29994157](https://pubmed.ncbi.nlm.nih.gov/29994157/)]
20. Medori J, Fairon C. Machine learning and features selection for semi-automatic ICD-9-CM encoding. In: Proceedings of the NAACL HLT 2010 Second Louhi Workshop on Text and Data Mining of Health Documents. 2010 Presented at: NAACL HLT 2010 Second Louhi Workshop on Text and Data Mining of Health Documents; June, 2010; Los Angeles, California, USA p. 84-89 URL: <https://dl.acm.org/citation.cfm?id=1867735.1867748>
21. Virginio L, dos Reis JC. Automated coding of medical diagnostics from free-text: the role of parameters optimization and imbalanced classes. In: *Data Integration in the Life Sciences*. Switzerland: Springer; 2018:122-134.
22. Lorenzi EC, Brown SL, Sun Z, Heller K. Predictive hierarchical clustering: learning clusters of CPT codes for improving surgical outcomes. *Proceedings of Machine Learning Research*. 2017. URL: <http://proceedings.mlr.press/v68/lorenzi17a.html> [accessed 2021-05-09]
23. Ye J. Construction and utilization of a neural network model to predict current procedural terminology codes from pathology report texts. *J Pathol Inform* 2019;10(1):13 [FREE Full text] [doi: [10.4103/jpi.jpi_3_19](https://doi.org/10.4103/jpi.jpi_3_19)] [Medline: [31057982](https://pubmed.ncbi.nlm.nih.gov/31057982/)]
24. Burns M, Mathis M, Vandervest J, Tan X, Lu B, Colquhoun D, et al. Classification of current procedural terminology codes from electronic health record data using machine learning. *Anesthesiology* 2020 Apr;132(4):738-749 [FREE Full text] [doi: [10.1097/ALN.0000000000003150](https://doi.org/10.1097/ALN.0000000000003150)] [Medline: [32028374](https://pubmed.ncbi.nlm.nih.gov/32028374/)]
25. Wu J, Roy J, Stewart WF. Prediction modeling using EHR data: challenges, strategies, and a comparison of machine learning approaches. *Med Care* 2010 Jun;48(6 Suppl):106-113. [doi: [10.1097/MLR.0b013e3181de9e17](https://doi.org/10.1097/MLR.0b013e3181de9e17)] [Medline: [20473190](https://pubmed.ncbi.nlm.nih.gov/20473190/)]
26. Shi H, Xie P, Hu Z, Zhang M, Xing E. Towards automated ICD coding using deep learning. *arXiv: Computer Science, Computation and Language*. 2017. URL: <http://arxiv.org/abs/1711.04075> [accessed 2021-05-08]
27. Wang G, Li C, Wang W, Zhang Y, Shen D, Zhang X. Joint embedding of words and labels for text classification. In: *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. 2018 Presented at: 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers); July, 2018; Melbourne, Australia p. 2321-2331. [doi: [10.18653/v1/p18-1216](https://doi.org/10.18653/v1/p18-1216)]

28. Rios A, Kavuluru R. Neural transfer learning for assigning diagnosis codes to EMRs. *Artif Intell Med* 2019 May;96:116-122 [FREE Full text] [doi: [10.1016/j.artmed.2019.04.002](https://doi.org/10.1016/j.artmed.2019.04.002)] [Medline: [31164204](https://pubmed.ncbi.nlm.nih.gov/31164204/)]
29. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez A, et al. Attention is all you need. *Advances in Neural Information Processing Systems*. 2017. URL: <http://papers.nips.cc/paper/7181-attention-is-all-you-need.pdf> [accessed 2021-05-09]
30. Sutskever I, Vinyals O, Le QV. Sequence to sequence learning with neural networks. *Advances in Neural Information Processing Systems*. 2014. URL: <http://papers.nips.cc/paper/5346-sequence-to-sequence-learning-with-neural-networks.pdf> [accessed 2021-05-09]
31. Luong M, Pham H, Manning CD. Effective approaches to attention-based neural machine translation. In: *Proceedings of the Conference on Empirical Methods in Natural Language Processing*. 2015 Presented at: Conference on Empirical Methods in Natural Language Processing; September 2015; Lisbon, Portugal. [doi: [10.18653/v1/d15-1166](https://doi.org/10.18653/v1/d15-1166)]
32. Wu Y, Schuster M, Chen Z, Le QV, Norouzi M, Macherey W, et al. Google's neural machine translation system: bridging the gap between human and machine translation. *arXiv: Computer Science, Computation and Language*. 2016. URL: <http://arxiv.org/abs/1609.08144> [accessed 2021-05-09]
33. Dew KN, Turner AM, Choi YK, Bosold A, Kirchhoff K. Development of machine translation technology for assisting health communication: a systematic review. *J Biomed Inform* 2018 Sep;85:56-67. [doi: [10.1016/j.jbi.2018.07.018](https://doi.org/10.1016/j.jbi.2018.07.018)] [Medline: [30031857](https://pubmed.ncbi.nlm.nih.gov/30031857/)]
34. Bodenreider O. The Unified Medical Language System (UMLS): integrating biomedical terminology. *Nucleic Acids Res* 2004 Jan 1;32(Database issue):267-270 [FREE Full text] [doi: [10.1093/nar/gkh061](https://doi.org/10.1093/nar/gkh061)] [Medline: [14681409](https://pubmed.ncbi.nlm.nih.gov/14681409/)]
35. Luo Y, Sun W, Rumshisky A. MCN: A comprehensive corpus for medical concept normalization. *J Biomed Inform* 2019 Apr;92:103132 [FREE Full text] [doi: [10.1016/j.jbi.2019.103132](https://doi.org/10.1016/j.jbi.2019.103132)] [Medline: [30802545](https://pubmed.ncbi.nlm.nih.gov/30802545/)]
36. Luo Y, Sun W, Rumshisky A. A hybrid normalization method for medical concepts in clinical narrative using semantic matching. *AMIA Jt Summits Transl Sci Proc* 2019;2019:732-740 [FREE Full text] [Medline: [31259030](https://pubmed.ncbi.nlm.nih.gov/31259030/)]
37. Ji Z, Wei Q, Xu H. BERT-based ranking for biomedical entity normalization. *arXiv: Computer Science, Information Retrieval*. 2019. URL: <http://arxiv.org/abs/1908.03548> [accessed 2021-05-09]
38. N2C2: National NLP clinical challenges. Harvard Medical School. 2019. URL: <https://n2c2.dbmi.hms.harvard.edu/> [accessed 2020-07-23]
39. Lee JM, Muis AO. Diagnosis code prediction from electronic health records as multilabel text classification: a survey. URL: http://people.cs.pitt.edu/~jlee/papers/cp1_survey_jlee_amuis.pdf [accessed 2021-05-09]
40. Subotin M, Davis AR. A system for predicting ICD-10-PCS codes from electronic health records. URL: <https://www.aclweb.org/anthology/W14-3409.pdf> [accessed 2021-05-09]
41. Lita LV, Yu S, Niculescu S, Bi J. Large scale diagnostic code classification for medical patient records. URL: <https://www.aclweb.org/anthology/I08-2125.pdf> [accessed 2021-05-09]
42. Helwe C, Elbassuoni S, Geha M, Hitti E, Obermeyer CM. CCS coding of discharge diagnoses via deep neural networks. In: *Proceedings of the 2017 International Conference on Digital Health*. 2017 Presented at: DH '17: International Conference on Digital Health; July, 2017; London United Kingdom p. 175-179. [doi: [10.1145/3079452.3079498](https://doi.org/10.1145/3079452.3079498)]
43. Freundlich RE, Kheterpal S. Perioperative effectiveness research using large databases. *Best Pract Res Clin Anaesthesiol* 2011 Dec;25(4):489-498 [FREE Full text] [doi: [10.1016/j.bpa.2011.08.008](https://doi.org/10.1016/j.bpa.2011.08.008)] [Medline: [22099915](https://pubmed.ncbi.nlm.nih.gov/22099915/)]
44. Kheterpal S. Clinical research using an information system: the multicenter perioperative outcomes group. *Anesthesiol Clin* 2011 Sep;29(3):377-388 [FREE Full text] [doi: [10.1016/j.anclin.2011.06.002](https://doi.org/10.1016/j.anclin.2011.06.002)] [Medline: [21871400](https://pubmed.ncbi.nlm.nih.gov/21871400/)]
45. Luseti M, Ruzsics T, Göhring A, Samardžić T, Stark E. Encoder-decoder methods for text normalization. *Association for Computational Linguistics*. 2018. URL: <https://www.aclweb.org/anthology/W18-3902.pdf> [accessed 2021-05-09]
46. Klein G, Kim Y, Deng Y, Senellart J, Rush A. OpenNMT: Open-source toolkit for neural machine translation. In: *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics-System Demonstration*. 2017 Presented at: 55th Annual Meeting of the Association for Computational Linguistics-System Demonstration; July 30 - August 4, 2017; Vancouver, Canada p. 67-72 URL: <https://www.aclweb.org/anthology/P17-4012.pdf> [doi: [10.18653/v1/p17-4012](https://doi.org/10.18653/v1/p17-4012)]
47. Pyysalo S, Ginter F, Moen H, Salakoski T, Ananiadou S. Distributional semantics resources for biomedical text processing. URL: <http://bio.nlplab.org/pdf/pyysalo13literature.pdf> [accessed 2021-05-09]
48. Adams DL, Norman H, Burroughs VJ. Addressing medical coding and billing part II: a strategy for achieving compliance. A risk management approach for reducing coding and billing errors. *J Natl Med Assoc* 2002 Jun;94(6):430-447. [Medline: [12078924](https://pubmed.ncbi.nlm.nih.gov/12078924/)]

Abbreviations

- BCBSM:** Blue Cross Blue Shield of Michigan
- BLEU:** Bilingual Evaluation Understudy
- CPT:** current procedural terminology
- EHR:** electronic health record
- ICD:** International Classification of Diseases

LSTM: long short-term memory
MPOG: Multicenter Perioperative Outcomes Group
NMT: neural machine translation
SVM: support vector machine
UMLS: Unified Medical Language System

Edited by G Eysenbach; submitted 31.07.20; peer-reviewed by J Chen, M Torii; comments to author 14.11.20; revised version received 02.03.21; accepted 19.04.21; published 26.05.21.

Please cite as:

Joo H, Burns M, Kalidaikurichi Lakshmanan SS, Hu Y, Vydishwaran VGV

Neural Machine Translation–Based Automated Current Procedural Terminology Classification System Using Procedure Text: Development and Validation Study

JMIR Form Res 2021;5(5):e22461

URL: <https://formative.jmir.org/2021/5/e22461>

doi: [10.2196/22461](https://doi.org/10.2196/22461)

PMID: [34037526](https://pubmed.ncbi.nlm.nih.gov/34037526/)

©Hyeon Joo, Michael Burns, Sai Saradha Kalidaikurichi Lakshmanan, Yaokun Hu, V G Vinod Vydishwaran. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 26.05.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Bispectral Index Alterations and Associations With Autonomic Changes During Hypnosis in Trauma Center Researchers: Formative Evaluation Study

C Michael Dunham¹, MD; Amanda J Burger¹, PhD; Barbara M Hileman¹, BA; Elisha A Chance¹, BSAS; Amy E Hutchinson¹, MD

St Elizabeth Youngstown Hospital, Youngstown, OH, United States

Corresponding Author:

C Michael Dunham, MD

St Elizabeth Youngstown Hospital

1044 Belmont Avenue

Youngstown, OH, 44501

United States

Phone: 1 330 480 3907

Email: dunham.michael@sbcglobal.net

Abstract

Background: Previous work performed by our group demonstrated that intermittent reductions in bispectral index (BIS) values were found during neurofeedback following mindfulness instructions. Hypnosis was induced to enhance reductions in BIS values.

Objective: This study aims to assess physiologic relaxation and explore its associations with BIS values using autonomic monitoring.

Methods: Each session consisted of reading a 4-minute baseline neutral script and playing an 18-minute hypnosis tape to 3 researchers involved in the BIS neurofeedback study. In addition to BIS monitoring, autonomic monitoring was performed, and this included measures of electromyography (EMG), skin temperature, skin conductance, respiratory rate, expired carbon dioxide, and heart rate variability. The resulting data were analyzed using two-tailed *t* tests, correlation analyses, and multivariate linear regression analyses.

Results: We found that hypnosis was associated with reductions in BIS ($P<.001$), EMG ($P<.001$), respiratory rate ($P<.001$), skin conductance ($P=.006$), and very low frequency power ($P=.04$); it was also associated with increases in expired carbon dioxide ($P<.001$), skin temperature ($P=.04$), high frequency power ($P<.001$), and successive heart interbeat interval difference ($P=.04$) values. Decreased BIS values were associated with reduced EMG measures ($R=0.76$; $P<.001$), respiratory rate ($R=0.35$; $P=.004$), skin conductance ($R=0.57$; $P<.001$), and low frequency power ($R=0.32$; $P=.01$) and with increased high frequency power ($R=-0.53$; $P<.001$), successive heart interbeat interval difference ($R=-0.32$; $P=.009$), and heart interbeat interval SD ($R=-0.26$; $P=.04$) values.

Conclusions: Hypnosis appeared to induce mental and physical relaxation, enhance parasympathetic neural activation, and attenuate sympathetic nervous system activity, changes that were associated with BIS values. Findings from this preliminary formative evaluation suggest that the current hypnosis model may be useful for assessing autonomic physiological associations with changes in BIS values, thus motivating us to proceed with a larger investigation in trauma center nurses and physicians.

(*JMIR Form Res* 2021;5(5):e24044) doi:[10.2196/24044](https://doi.org/10.2196/24044)

KEYWORDS

bispectral index; hypnosis; heart rate variability; electromyography; skin conductance; skin temperature; respiratory rate; expired carbon dioxide; neurofeedback

Introduction

Effects of Hypnosis on Brainwave Physiology

Multiple investigations have shown that a hypnotic state can influence alterations in brainwave activity. Specifically, studies have provided evidence that hypnosis reduces beta brainwave power [1-3]. In addition, multiple studies have shown that hypnosis is associated with increased theta brainwave power [4-6] and enhanced alpha brainwave power [1,6]. These brainwave alterations are indicative of a state of mental relaxation. Of particular relevance are investigations demonstrating that bispectral index (BIS) values substantially decrease during nonpharmacologic hypnosis [7,8]. Other investigators have provided evidence of reductions in BIS values for participants who listened to relaxing music [9], watched relaxing videos [10], or underwent relaxing-guided imagery [11].

Effects of Hypnosis on Heart Rate Variability

When measuring the electrocardiographic R-to-R interval in milliseconds, there is a normal variation from one beat to the next in healthy participants [12,13]. Standard measures of heart rate variability (HRV) include the SD of cardiac interbeat (normal-to-normal; SDNN) time intervals and root mean square of successive cardiac interbeat time intervals (RMSSD). Common HRV measurements also include spectral density measurements of high frequency (HF) power, low frequency (LF) power, very low frequency (VLF) power, LF/HF power ratio, and total power [12,13]. LF power has been shown to be influenced by sympathetic and parasympathetic nervous system activities, whereas HF power better reflects parasympathetic activation [12-15]. Multiple investigators have explored the influence of hypnosis on HRV. With hypnosis, LF power has been shown to decrease [16-18], whereas HF power has been demonstrated to increase [17,19,20]. Accordingly, LF/HF power has been noted to decrease [16,21]. During hypnosis, RMSSD, SDNN, and the interbeat interval mean and SD have been shown to increase [19,20].

Effects of Stress-Relaxation on HRV

Several investigators have demonstrated the effects of stress and relaxation on HRV. Improvements in HRV measurements have been shown in autonomic balance during relaxation therapies and interventions [10,22,23], mindfulness interventions [24], and laparoscopic surgeries with robotic assistance [25]. HRV has also been shown to decrease with psychological stress [15,26], perceived high work stress in nurses [27], and stress in nursing students [28].

Effects of Hypnosis on Other Autonomic Physiology

Several studies have documented the effects of hypnosis on skin conductance, expired carbon dioxide, respiratory rate, electromyography (EMG), and skin temperature values. Two relatively recent studies have demonstrated that hypnosis is associated with reductions in skin conductance values [8,29]. We found that only a single study investigated expired carbon dioxide measurements during hypnosis [30]. The data in the manuscript showed that expired carbon dioxide levels increased and respiratory rates decreased; however, details of the statistical

analysis were not provided. Significant reductions in the respiratory rate have been shown to occur with hypnosis [31,32]. Hypnosis has been demonstrated to reduce EMG tension in the masseter, temporalis, and frontalis muscles [33-35]. Virtually all studies in the literature that assessed the effects of hypnosis on skin temperature involved volitional efforts to alter the temperature or to mitigate temperature responses to heat or cold stressors. However, we found an investigation demonstrating that hypnosis, without any intent to change temperature, was associated with a significant increase in finger temperature [36].

Effects of Stress-Relaxation on Other Autonomic Physiology

Peripheral skin temperature is mediated according to autonomic nervous system regulation; that is, temperature reductions occur with increases in sympathetic neural activity [37]. Reductions in peripheral skin temperature have been associated with psychological stress [38] and mental stress [39,40]. On the other hand, biofeedback relaxation has been found to be associated with increases in peripheral skin temperature [41]. Electrical skin conductance values increase with sympathetic neutrally mediated palmar sweating [37]; multiple investigations have shown that acute mental stress is associated with increases in electrical skin conductance values [42-46]. Evidence in the literature shows that decreased BIS values are associated with decreased skin conductance [47].

Respiratory rate increases with anxiety, whereas mindfulness and relaxation tend to decrease ventilation effort [37]. Investigations have demonstrated that respiratory rates increase with acute mental stress [48-52] and acute fear [53]. In contrast, relaxing music [9], progressive muscle relaxation or galvanic skin resistance biofeedback [54], and yoga training [55] have been associated with significant reductions in respiratory rates. Generally, the expired carbon dioxide concentration has an inverse relationship with the respiratory rate [37]. Expired carbon dioxide has been found to significantly decrease in participants with emotional disorders [56], generalized anxiety disorder [57], acute fear [53], and acute mental stress [50,58]. In contrast, increased expired carbon dioxide has been found to occur with yoga training [55] and progressive muscle relaxation [48].

Surface EMG measures muscle tension beneath the skin at designated anatomical sites [37]. Increased surface EMG activity has been associated with acute mental stress [59], viewing distressing videos [60], and perceptions of negative stress states [61]. Conversely, reduced surface EMG activity has been linked to relaxation interventions [62], relaxation training [63], and meditation [64].

BIS and Monitoring of Coparameters

The National Library of Medicine contains 2607 citations regarding the use of BIS monitoring in humans. Some of these manuscripts describe simultaneous changes in BIS values, HRV, skin conductance, respiratory rate, expired carbon dioxide, and EMG values; however, a smaller number of studies presented correlation analyses between BIS values and these other parameters. The limited relevance of these investigations is that multiple confounding conditions typically exist during each

investigation. For example, study participants usually undergo substantial physical stimulation, such as invasive surgery, airway manipulation, or mechanical ventilation. Often, the research includes participants with brain pathologies, such as traumatic brain injury, cerebrovascular diseases, seizure disorders, or coma. Furthermore, most study participants in these investigations are administered intravenous or inhalation anesthesia, intravenous sedatives, or neuromuscular blocking agents.

Several studies have provided simultaneous measurements of HRV and BIS values; however, these investigations are substantially confounded by multiple factors [65-68]. A study of participants who watched relaxing videos had concomitant decreases in BIS values and increments in HF power measurements, but correlation data between BIS and HF power were not shown [10]. A significant correlation has been found between BIS and EMG values, but this observation was relevant to intraoperative patients [69]. In addition, EMG activity has been demonstrated to contribute considerably to high BIS values in postoperative patients [70].

BIS and respiratory rate values have been shown to simultaneously decrease during music therapy in intensive care unit patients [9]. A study compared simultaneous BIS values, respiratory rate, and expired carbon dioxide measurements in healthy volunteers, but these findings were confounded by the application of a tight-fitting facemask and continuous infusions of intravenous sedatives [71]. Another study monitored serial BIS values, respiratory rate, and expired carbon dioxide measurements; however, these findings were confounded by the need for cataract surgery in moderately sedated participants with varying levels of health status [72]. In a single study, BIS and skin conductance were shown to significantly decrease during hypnosis; however, the authors provided no correlation analysis between the 2 values [8]. In addition, another study demonstrated that BIS and skin conductance values were positively correlated in patients undergoing surgical anesthesia [47].

BIS Neurofeedback Study

Our group demonstrated that trauma center nurses and physician participants could learn to self-regulate brainwave activity using an electroencephalography (EEG)-based BIS monitoring system during neurofeedback immediately after receiving mindfulness instructions [73]. Importantly, most participants also showed improvements in well-being scores after learning brainwave self-regulation. These findings serve as a validation indicator that using the BIS monitor to perform brainwave self-regulation during neurofeedback can be useful. During the 228 neurofeedback learning sessions, participants demonstrated the ability to lower their BIS values following mindfulness instruction (brainwave self-regulation). However, the reductions were often relatively brief; they returned near the baseline value and subsequently decreased again.

First, we sought other evidence to validate the notion that the BIS monitor is a potentially useful tool for performing neurofeedback. In particular, we wanted to use a methodological approach that would likely produce sustained reductions in BIS values. Second, we wanted to use ancillary physiologic

monitoring device data that we could simultaneously compare with BIS values. Our goals were to determine whether hypnosis could produce sustained reductions in BIS values and to organize a process evaluation that would potentially demonstrate that reductions in EEG-based BIS values are associated with a physiological state of relaxation.

Aims

The aims of this feasibility and formative assessment are two-fold. First, we aim to use hypnosis and to describe significant changes that occurred in BIS, HRV, skin conductance, expired carbon dioxide, respiratory rate, EMG, and skin temperature values. Second, we intend to identify any significant relationships between BIS measurements and autonomic physiologic variable values. If the formative assessment appears to be meritorious, we plan to apply a similar approach and process to a larger group of physicians and nurses employed at the same trauma center.

Methods

Assessment Design and Population

This formative assessment was performed in accordance with the recommendations of the Declaration of Helsinki. The protocol was approved by St Elizabeth Youngstown Hospital, Mercy Health Youngstown, LLC's institutional review board (institutional review board organization #0001624). On November 21, 2019, the review board granted expedited approval, and because the evaluation was deemed to have minimal risk, they waived consent (institutional review board approval number: 19-024). Three of the trauma center research investigators (healthy volunteers) who conducted the BIS neurofeedback study participated as participants in this evaluation [73]. There was no intention to compare results between the participants. We intended to compare prehypnotic data with active hypnotic data to determine whether hypnosis could produce sustained reductions in BIS values and whether BIS values are associated with other physiologic measurements of relaxation, forming an approach and process for applying this protocol to a larger group of trauma center physicians and nurses.

Data Variables

Data variables assessed during this evaluation included BIS, EMG (decibels), expired carbon dioxide (mm Hg), respiratory rate (breaths per minute), skin temperature (degrees Fahrenheit), skin conductance (microsiemens), and HRV values. The specific HRV variables were (1) SDNN, (2) RMSSD, (3) LF power, (4) HF power, (5) VLF power, (6) relative LF power (LF power/total power), (7) relative HF power (HF power/total power), and (8) relative VLF power (VLF power/total power). LF power was the absolute power at a band frequency of 0.04-0.15 Hz and measured as ms^2/Hz . HF power was the absolute power at a band frequency of 0.15-0.4 Hz and measured as ms^2/Hz . VLF power was the absolute power at a band frequency of 0.0033-0.04 Hz and measured as ms^2/Hz . Relative LF, HF, and VLF power were quantified as percentages of the total power at the relevant band frequencies. The potential BIS

value range of the system is 0-100; however, awake and alert values typically approach 100 [73].

Signal Sensor Applications

The Bispectral Index Vista Monitoring System (Aspect Medical Systems, Inc) was used to capture BIS and EMG physiological signals. According to the manufacturer's instructions, the BIS sensor was applied to the participant's forehead and temporal fossa. The BIS sensor contained 4 electrodes that corresponded to the international 10-20 EEG system for electrode placement: FPz, FP1, AF7, and FT9. Frontalis muscle EMG activity was measured using the AF7 BIS sensor electrode. Monitoring of expired carbon dioxide respiratory rate was performed using the RespSense capnography monitor (model LS1R-9R, Nonin Medical, Inc). Each participant had a biprong nasal cannula inserted with tubing draped over the ears and was instructed to keep their lips closed.

The ProComp Infiniti encoder hardware system (Thought Technology Ltd) was used to capture skin temperature, skin conductance, and electrocardiographic signals (HRV variables). The skin temperature sensor was secured to the volar surface of the terminal phalanx of the middle finger using a Velcro strap. The 2 skin conductance sensors were secured to the volar surface of the middle phalanx of the ring and index fingers of the opposite hand with Velcro straps. Following the manufacturer's instructions, 3 sensors were applied to the forearms to acquire electrocardiographic signals. The yellow sensor was attached to the volar side of the proximal right forearm. The black sensor was placed on the volar side of the proximal left forearm, and the blue sensor was attached 4 inches distally. Each participant's arms were placed in a relaxed manner on chair armrests. Skin temperature, skin conductance, and electrocardiographic sensors were connected to the ProComp Infiniti encoder.

Signal Processing

The Bispectral Index Vista Monitoring System (hardware and software) transformed BIS and EMG physiological signals into digital outputs. The RespSense capnography monitor, hardware and software, converted carbon dioxide measurements into a digital quantity. As each real-time peak expired carbon dioxide value represents a participant's respiratory exhalation, the respiratory rate was computed by documenting the periodicity of carbon dioxide exhalations relative to time. The ProComp Infiniti encoder hardware system streamed physiological signal information to the BioGraph Infiniti (Thought Technology Ltd) Software, which transformed skin temperature, skin conductance, and electrocardiographic signals (HRV variables) into digital outputs.

Signal Output Storage and Data Harvest

The BIS and EMG digital outputs were routinely captured on the hardware during each session. The cursor review monitor option provided a reading every 10 seconds. The investigators computed the mean BIS and EMG values for each session minute by averaging the 6 values for each minute. The values were entered into Microsoft Excel (Microsoft Corp.). The expired carbon dioxide and respiratory rate values were updated on the monitor hardware every second. The data were exported to Microsoft Excel, and a mean value for each session minute was computed by averaging the 60 expired carbon dioxide and respiratory rate values for each minute. The BioGraph Infiniti software computed and stored a 1-minute mean skin temperature and skin conductance result for each of the HRV variables. Accordingly, a mean 1-minute value was available for all physiological signals. These values were exported to Microsoft Excel, and then, all the data were imported into a statistical software program (SAS System for Windows, release 9.2, SAS Institute Inc).

Signal Quality Assessments and Systems Coordination

Before starting the recording session, broadcasting the neutral script, and playing the hypnosis recording, the investigators examined the quality of the physiological signals. These assessments included an appraisal of the signal quality index and absence of artifacts as displayed on the BIS monitor. Capnography tracing and the stability of the expired carbon dioxide values were also examined. Finally, the researchers assessed the stability of skin temperature, skin conductance, and electrocardiographic tracings using the BioGraph Infiniti software system. The investigators coordinated the activation of the 3 monitoring systems with the initiation of a neutral script sound. As soon as the hypnosis recording ended, recordings within the 3 monitoring systems were immediately discontinued.

Session Audio Scripts

During the 22-minute session, the participants were comfortably seated in an upright chair. The 22-minute session time was chosen to capture 4 minutes of prehypnotic or baseline values and 18 minutes of active hypnosis values. For the first 4 minutes of the session, a neutral script was read that contained neither stressful nor relaxing suggestions. The purpose of this script was to foster a cognitive state of alert, eyes-opened prehypnosis baseline to compare with the 18 minutes when the hypnotic script was played. Immediately following the conclusion of the neutral script, a hypnosis tape in MP3 format found in the public domain was played. While listening to the instructions of the hypnosis tape, the participants' eyes were closed for most of the time. The script timeline milestones for each 22-minute session are presented in [Table 1](#). A summary overview of the procedures performed during the experimental period is presented in [Textbox 1](#).

Table 1. Audio script milestones.

Minute	Audio script	Comments
1	Neutral	Eyes open
2	Neutral	Eyes open
3	Neutral	Eyes open
4	Neutral	Eyes open
5	Start hypnosis	Eyes open
6	Physical relaxation	Eyes closed ^a
7	Physical relaxation	Eyes closed ^a
8	Deepening	Eyes closed ^a
9	Deepening	Eyes closed ^a
10	Mental relaxation	Eyes closed ^a
11	Mental relaxation	Eyes closed ^a
12	Mental relaxation	Eyes closed ^a
13	Deepening	Eyes closed ^a
14	Deepening	Eyes closed ^a
15	Deepening	Eyes closed ^a
16	Deepening	Eyes closed ^a
17	Deepening	Eyes closed ^a
18	Suggestions	Eyes closed ^b
19	Suggestions	Eyes closed ^b
20	Reorientation and awakening	Eyes closed
21	Reorientation and awakening	Eyes closed
22	Reorientation and awakening	Eyes open

^aInfrequent, very brief prompts to open the eyes and then reclose them.

^b“Every day in every way, I am getting better and better.”

Textbox 1. Procedural flow table.

<p>Sensor application (all applied per the manufacturer's instructions)</p> <ul style="list-style-type: none">• Sensors for bispectral index and electromyography applied to the participant's forehead• Carbon dioxide nasal cannula inserted• Skin temperature sensor secured to the participant's finger• Skin conductance sensors secured to the participant's fingers• Electrocardiography sensors applied to the participant's forearms <p>Sensor connections</p> <ul style="list-style-type: none">• Sensors connected to the relevant hardware <p>Signal quality assurance</p> <ul style="list-style-type: none">• Each monitor was assessed for signal quality <p>Session start</p> <ul style="list-style-type: none">• Data recordings began at minute 1 <p>Minutes 1-4, prehypnosis data</p> <ul style="list-style-type: none">• Neutral script <p>Minute 5, hypnosis data</p> <ul style="list-style-type: none">• Hypnosis tape started <p>Minutes 6-7</p> <ul style="list-style-type: none">• Physical relaxation phase <p>Minutes 8-9</p> <ul style="list-style-type: none">• Deepening of the physical relaxation phase <p>Minutes 10-12</p> <ul style="list-style-type: none">• Mental relaxation phase <p>Minutes 13-17</p> <ul style="list-style-type: none">• Deepening of mental and physical relaxation <p>Minutes 18-19</p> <ul style="list-style-type: none">• Suggestive phase <p>Minutes 20-22</p> <ul style="list-style-type: none">• Reorientation and awakening phase <p>Session end</p> <ul style="list-style-type: none">• Hypnosis recording ended• Data recordings ended• Sensors removed from the participant• Data entered into or exported to Microsoft Excel and analyzed with statistical software
--

Statistical Analysis

Results were entered into an Excel 2010 worksheet (Microsoft Corp) and imported into the SAS System for Windows, release 9.2 (SAS Institute Inc). All mean values were accompanied by their SDs. For 2-group interval data comparisons, a two-tailed

t test result and Cohen *d* were computed. Correlation analyses were assessed using the Pearson coefficient procedure. The level of significance was set at $P < .05$.

Results

Overview

Three healthy adult volunteers each completed a 22-minute experimental session. Of the 3 volunteers, there was 1 (33%) male participant and 2 (66%) female participants with ages ranging from 35 to 71 years. All participants were college graduates and single and received monthly compensation from a large hospital corporation. All participants verbalized a feeling of relaxation at the conclusion of their respective sessions.

Autonomic Data Variances

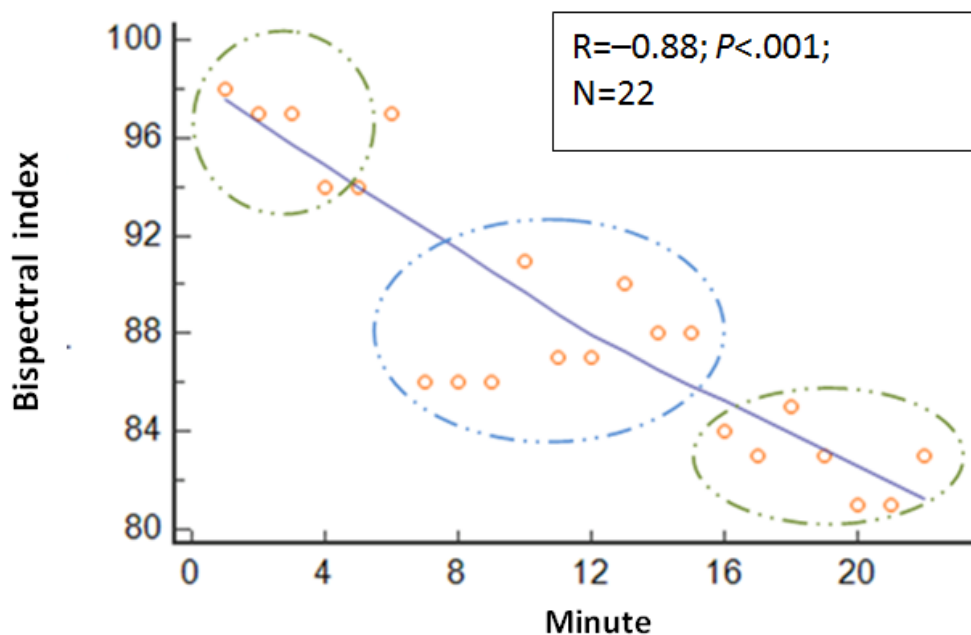
The first 4-minute interparticipant variances were substantial for the following variables: expired carbon dioxide (mean 32.9, SD 5.1 mm Hg), respiratory rate (mean 16.9, SD 2.7 breaths per min), skin temperature (mean 91.4, SD 6.0 °F), skin conductance (mean 3.0, SD 2.1 microsiemens), LF power (mean 576, SD 488 ms²/Hz), HF power (mean 364, SD 301 ms²/Hz), VLF power (mean 433, SD 480 ms²/Hz), total power (mean 1372, SD 913 ms²/Hz), RMSSD (mean 54.2, SD 16.3 ms), and

SDNN (mean 59.0, SD 17.9 ms). Owing to the variance sizes, a mean value was created from the first 4-minute raw values separately for each of these physiologic variables and each participant. Then, all the relevant 22-minute variable data results were divided by the physiologic variable mean value for the first 4 minutes for each participant; that is, all first 4-minute normalized variable values were approximately 1.0, whereas the normalized variable values for minutes 5-22 were relative fractions of 1.0, either ≥ 1.0 or < 1.0 [19].

Changes in Autonomic Variables During Hypnosis

As the session progressed, the BIS values decreased and had a correlation between the 66 BIS values and the session duration in minutes (correlation analysis: raw BIS values=session minutes; $R=-0.70$; $P<.001$). A correlation existed between the 22 mean BIS values and the duration of the session in minutes ($R=-0.88$; $P<.001$; Figure 1). BIS values were lower for hypnosis minutes 7-15 (mean 87.7, SD 6.0) than for prehypnosis minutes 1-4 (mean 96.5, SD 1.7; $P<.001$; Cohen $d=2$). BIS values were also lower for hypnosis minutes 16-22 (mean 82.9, SD 4.1) than for hypnosis minutes 7-15 (mean 87.7, SD 6.0; $P=.002$; Cohen $d=0.9$).

Figure 1. Relationship of the mean bispectral index with time. As the session progressed, the mean bispectral index values steadily decreased.



Several other physiologic variables showed significant changes in the 22 mean values as the session duration in minutes progressed. The R values and P values for these physiologic variables relative to session progression time in minutes (correlation analysis: mean variable values=session minutes) were as follows: EMG, $R=-0.80$ and $P<.001$; expired carbon

dioxide, $R=0.67$ and $P=.001$; respiratory rate, $R=-0.68$ and $P<.001$; skin conductance, $R=-0.82$ and $P<.001$; skin temperature, $R=.66$ and $P=.001$; relative HF power, $R=0.69$ and $P<.001$; relative VLF power, $R=-0.38$ and $P=.08$; and RMSSD, $R=0.64$ and $P=.001$. These relationships are shown in Figures 2-8.

Figure 2. Relationship of the mean EMG with time. As the session progressed, the mean EMG values steadily decreased. EMG: electromyography.

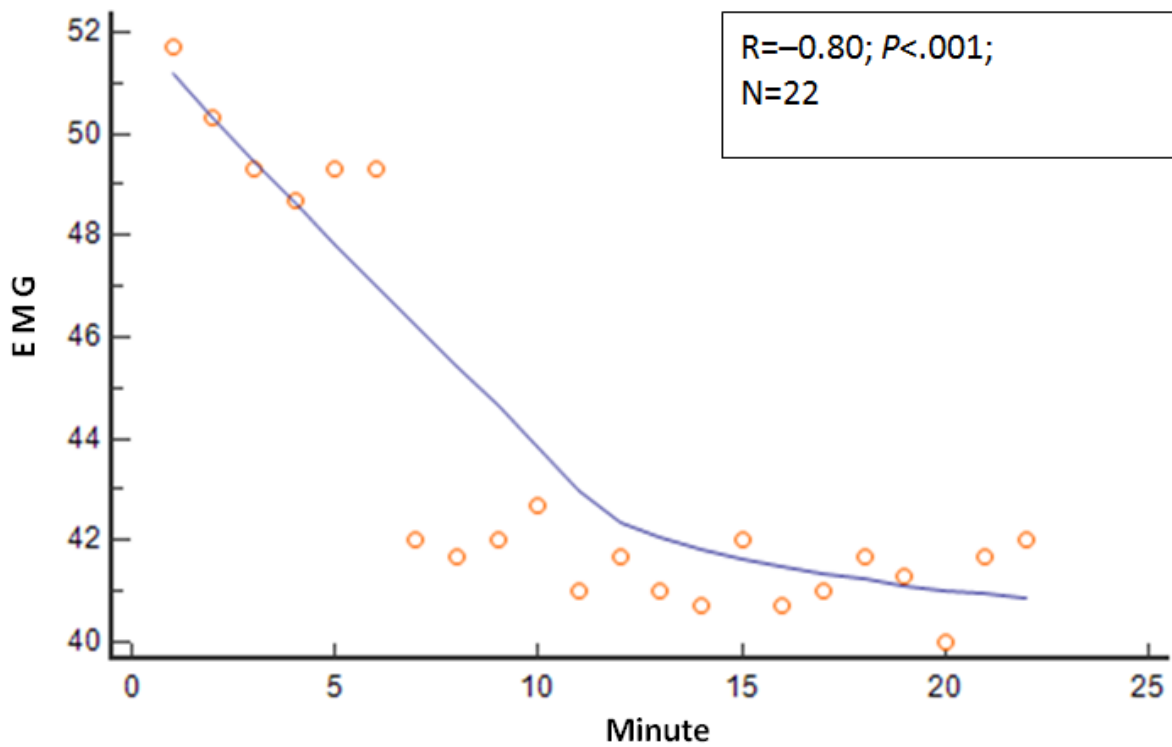


Figure 3. Relationship of the mean respiratory rate with time. As the session progressed, the mean respiratory rate values steadily decreased.

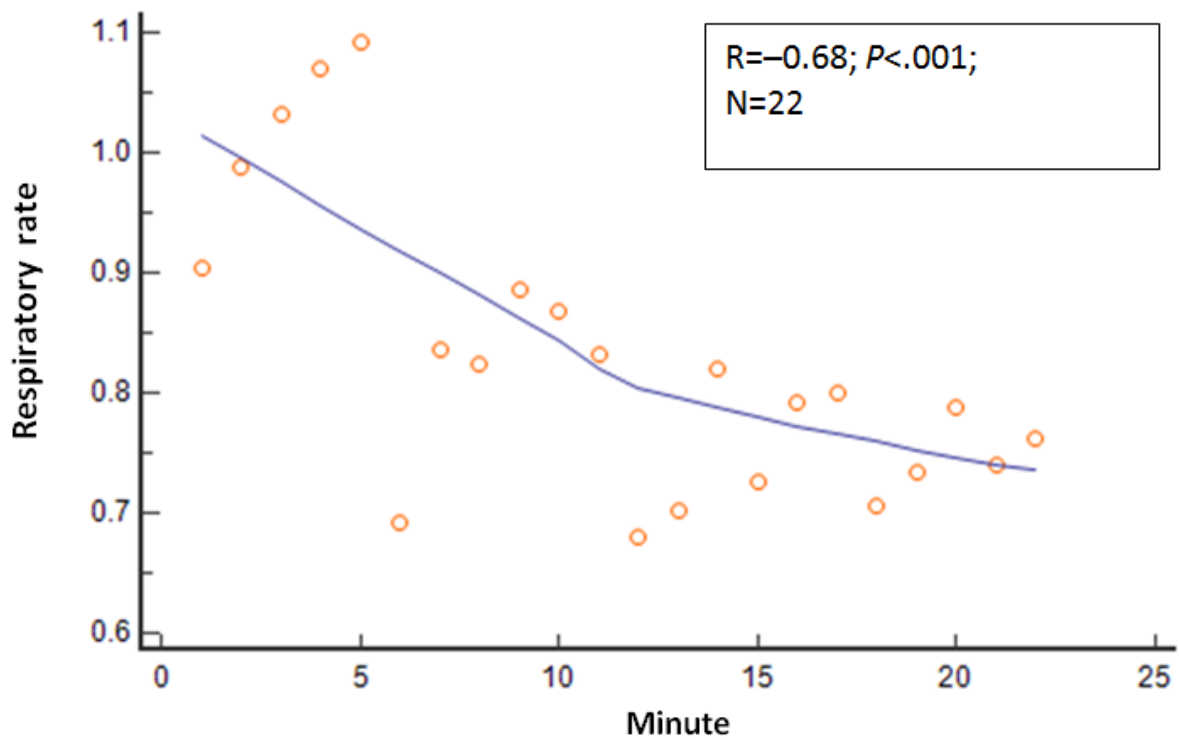


Figure 4. Relationship of the mean expired carbon dioxide with time. As the session progressed, the mean expired carbon dioxide values steadily increased.

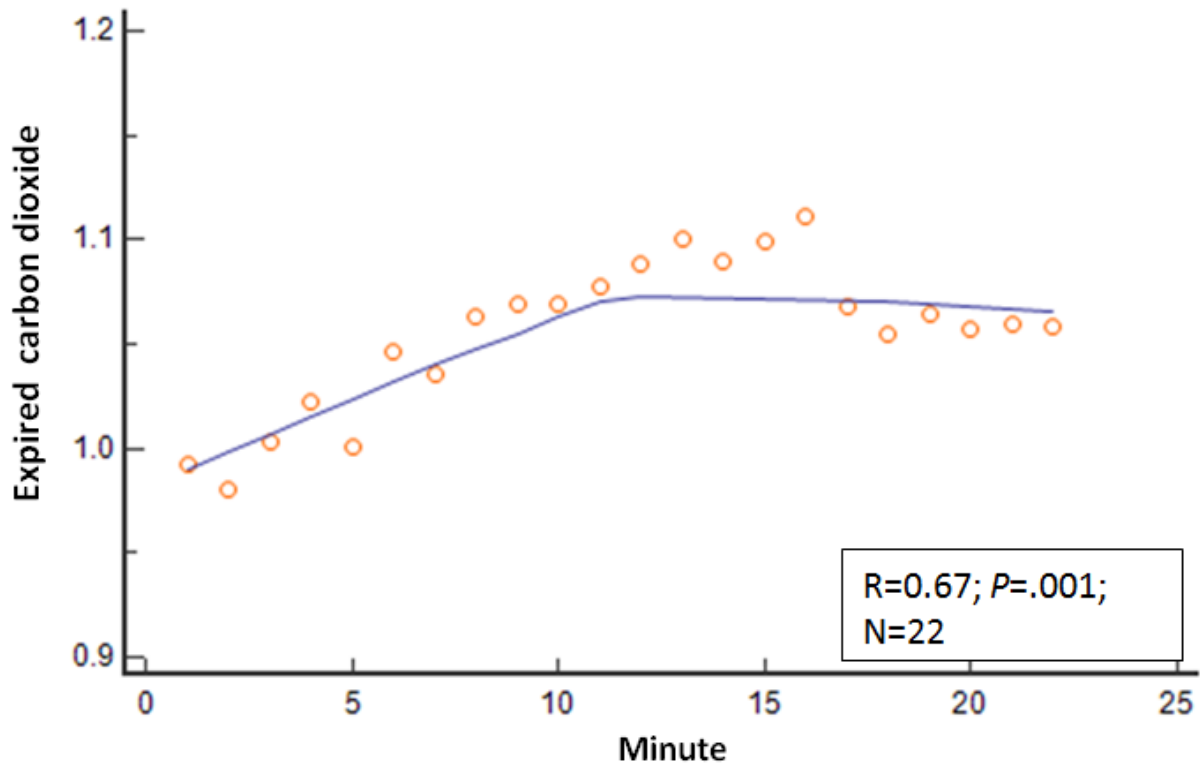


Figure 5. Relationship of the mean skin conductance with time. As the session progressed, the mean skin conductance values steadily decreased.

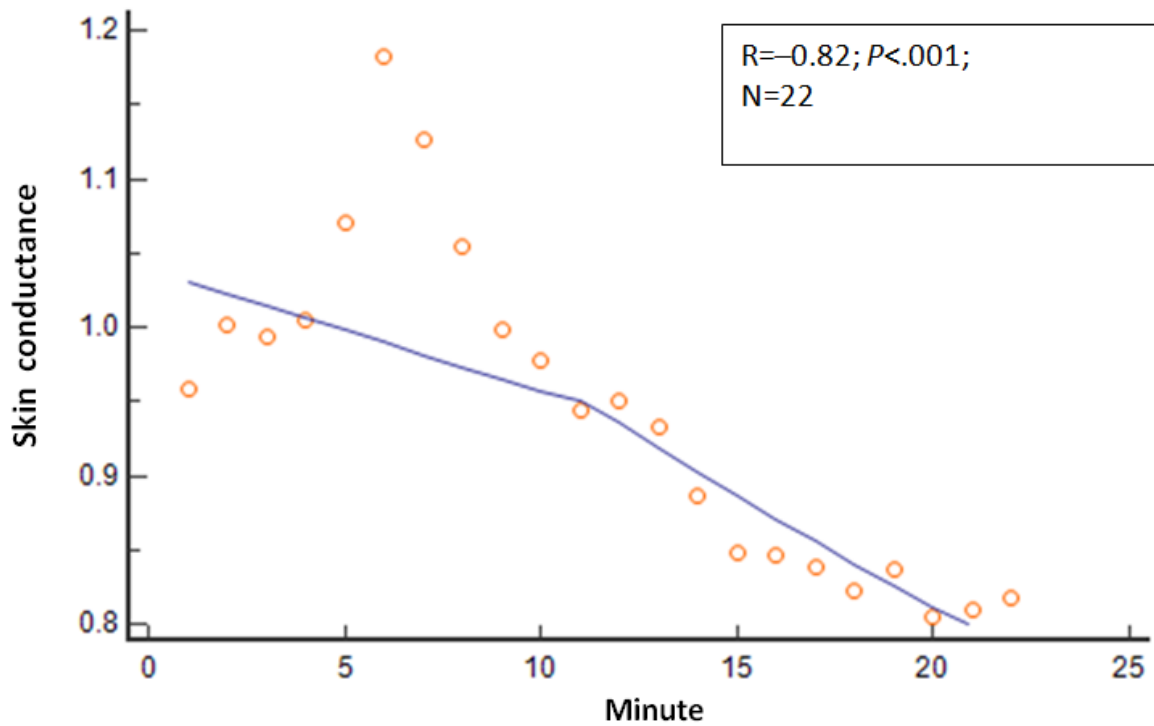


Figure 6. Relationship of the mean skin temperature with time. As the session progressed, the mean skin temperature values steadily increased.

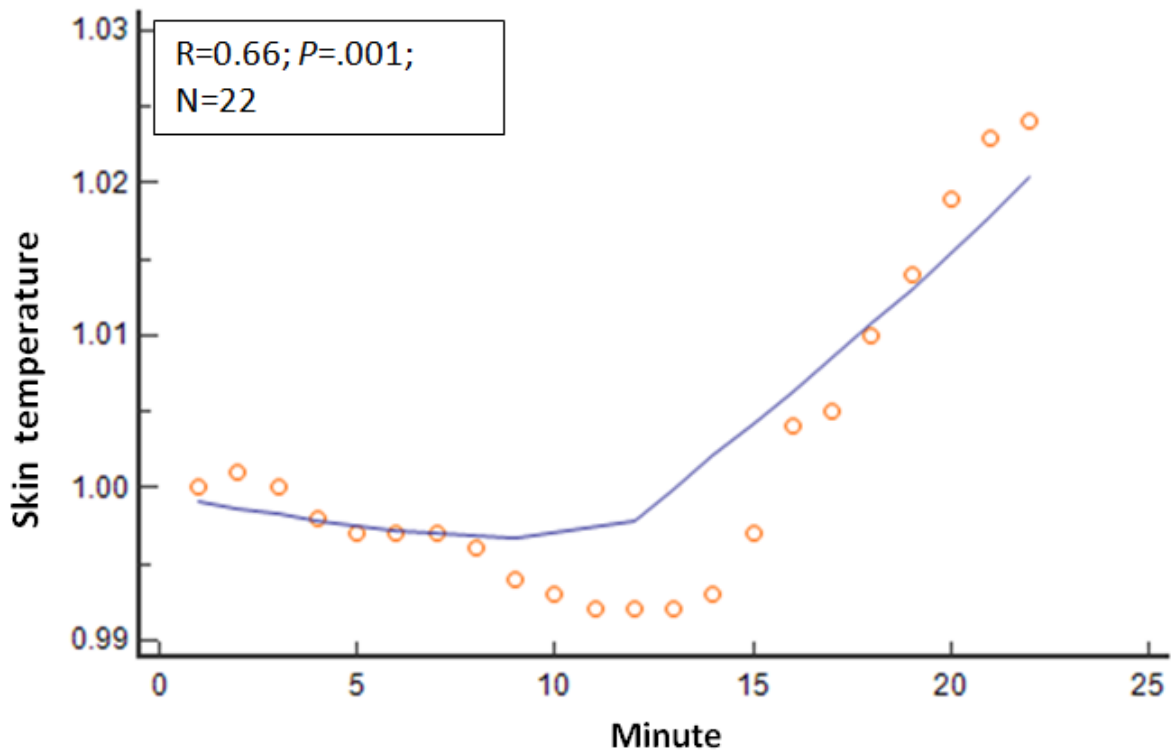


Figure 7. Relationship of the mean relative HF power with time. As the session progressed, the mean relative HF power values steadily increased. HF: high frequency.

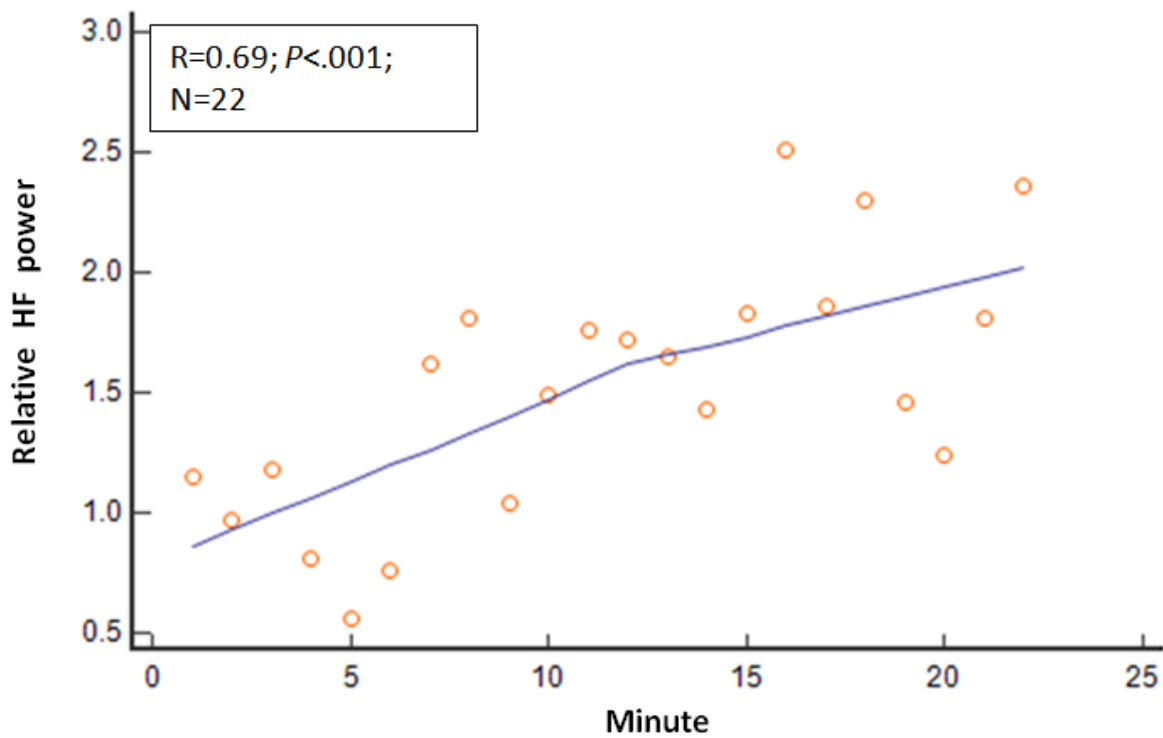
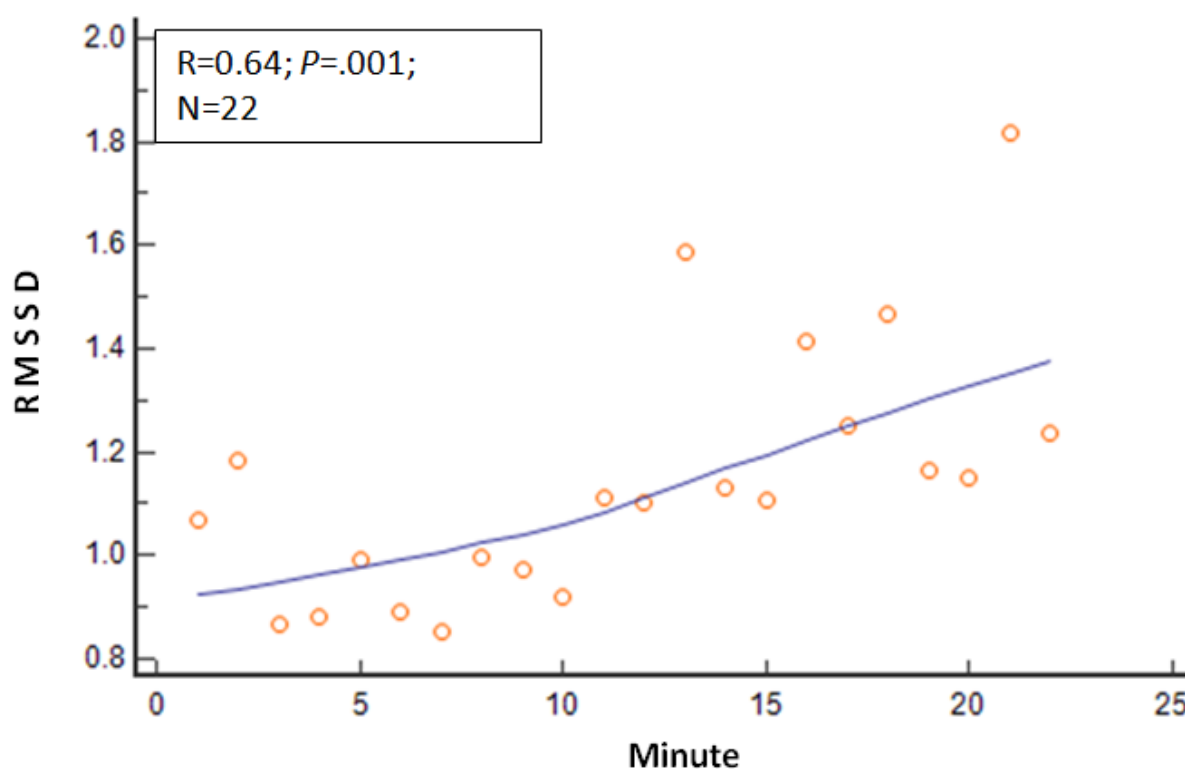


Figure 8. Relationship of the RMSSD with time. As the session progressed, the RMSSD steadily increased. RMSSD: root mean square of successive cardiac interbeat time interval differences.



When compared with prehypnosis (min 1-4), BIS, EMG, respiratory rate, skin conductance, and relative VLF power values decreased during induction and postinduction hypnosis, whereas expired carbon dioxide and relative HF fraction values increased during induction and postinduction hypnosis (min 7-22; Table 2). When compared with prehypnosis (min 1-4),

BIS, EMG, respiratory rate, skin conductance, and relative VLF power values decreased during postinduction hypnosis, whereas expired carbon dioxide, skin temperature, relative HF power, and RMSSD values increased during postinduction hypnosis (min 16-22; Table 3).

Table 2. Significant hypnosis-associated physiologic changes (min 7-22).

Physiological measures	Prehypnosis (min 1-4; n ^a =12), mean (SD)	Hypnosis (min 7-22; n ^a =48), mean (SD)	P value	Cohen d
Bispectral index	96.5 (1.7)	85.6 (5.7)	<.001	2.6
Electromyography (decibel)	50.0 (3.8)	41.4 (2.4)	<.001	2.7
Relative respiratory rate	0.999 (0.11)	0.781 (0.10)	<.001	2.1
Relative skin conductance	0.990 (0.04)	0.906 (0.19)	.006	0.6
Relative very low frequency power	0.892 (0.60)	0.525 (0.38)	.07	0.7
Relative expired carbon dioxide	0.999 (0.02)	1.073 (0.08)	<.001	1.3
Relative high frequency power	1.026 (0.27)	1.743 (1.17)	<.001	0.8

^aData points (number of measurements).

Table 3. Significant postinduction hypnosis-associated physiologic changes (min 16-22).

Physiological measures	Prehypnosis (min 1-4; n ^a =12), mean (SD)	Postinduction hypnosis (min 16-22; n ^a =21), mean (SD)	P value	Cohen <i>d</i>
Bispectral index	96.5 (1.7)	82.9 (4.1)	<.001	4.3
Electromyography (decibel)	50.0 (3.8)	41.2 (1.9)	<.001	2.9
Relative respiratory rate	0.999 (0.11)	0.760 (0.09)	<.001	2.4
Relative skin conductance	0.990 (0.04)	0.825 (0.21)	.002	1.1
Relative very low frequency power	0.892 (0.60)	0.467 (0.35)	.04	0.9
Expired carbon dioxide	0.999 (0.02)	1.068 (0.08)	<.001	1.2
Relative skin temperature	0.999 (0.00)	1.014 (0.03)	.04	0.7
Relative high frequency power	1.026 (0.27)	1.933 (1.28)	.005	1.0
Relative root mean square of successive cardiac interbeat time intervals	0.999 (0.17)	1.357 (0.70)	.04	0.7

^aData points (number of measurements).

Significant BIS Correlations With Other Autonomic Variables

Correlation analyses of the 66 BIS values during minutes 1-22 showed multiple significant associations with other physiological variables. When BIS values decreased, EMG, respiratory rate, skin conductance, and relative LF power values decreased (Table 4). When BIS values decreased, HF power,

relative HF power, RMSSD, and SDNN values increased (Table 4). Multivariate linear regression analysis showed that the 66 BIS values during minutes 1-22 were simultaneously and independently associated with EMG ($P<.001$), skin conductance ($P<.001$), and relative LF power ($P=.03$), with a total $R^2=0.71$. HF power had positive univariate associations with SDNN ($R=0.78$; $P<.001$) and RMSSD ($R=0.85$; $P<.001$).

Table 4. Significant bispectral index correlations with other autonomic variables (min 1-22).

Physiological measures	R value	P value
Electromyography	0.76	<.001
Respiratory rate	0.35	.004
Skin conductance	0.57	<.001
Relative low frequency power	0.32	.01
High frequency power	-0.41	<.001
Relative high frequency power	-0.53	<.001
Root mean square of successive cardiac interbeat time intervals	-0.32	.009
SD of cardiac interbeat (normal-to-normal) time intervals	-0.26	.04

Discussion

Principal Findings

The goal of this formative and feasibility assessment was to determine whether hypnosis could produce sustained reductions in BIS values. The results demonstrate that hypnosis produced sustained reductions in BIS values in healthy volunteers and that the BIS values steadily decreased linearly as the hypnosis session progressed. Another goal of this formative assessment was to organize and evaluate a process that could identify significant changes and correlations between BIS values and other measurements of physiologic relaxation. In implementing the complex and coordinated data collection process described in the *Methods* section, we were able to identify significant changes and associations between BIS values and EMG, respiratory rate, skin conductance and temperature, expired carbon dioxide, and HRV values. These findings indicate that

when mental relaxation occurs, physical relaxation occurs simultaneously. These associations also demonstrate that hypnosis fostered a state of mental and physical relaxation in this small group of healthy volunteers and indicates that the process could be feasibly applied to a larger homogeneous group of trauma center physicians and nurses.

Changes in BIS Values During Hypnosis

As the hypnosis session progressed, the BIS values progressively decreased. Compared with the prehypnosis state, BIS values were lower during hypnosis induction and even lower during postinduction hypnosis. The mean BIS values during hypnosis induction and postinduction were similar to the BIS values reported during stage I sleep [74]. The BIS reductions that we found during hypnosis are corroborated by another similar investigation by Almeida-Marques [8]. In their study, the mean BIS values before hypnosis were also 97, but they averaged 77 in the hypnotic state, which is lower than that in this

investigation. However, their sample size was larger and consisted of patients with fibromyalgia, and they employed a hypnotherapist for their sessions rather than using a prerecorded hypnosis script [8]. Supplementary evidence for BIS reductions is the evidence provided during investigations of participants watching relaxing videos [10] or trying relaxing-guided imagery [11]. The brainwave physiology of BIS reductions, and similarity to stage I sleep, implies that such participants were in a relative state of cognitive or mental relaxation [73-75].

Changes in HRV During Hypnosis

Hypnosis was associated with changes in HRV in this assessment. In particular, relative HF power and RMSSD values increased, whereas relative VLF power values decreased. The influence of hypnosis on HRV has been explored by several investigators [16-21]. With hypnosis, LF power has been shown to decrease [16-18], whereas HF power has been demonstrated to increase [17,19,20]. Accordingly, LF/HF power has been noted to decrease [16,21]. During hypnosis, RMSSD, SDNN, and the interbeat interval mean and SD have been shown to increase [19,20].

Contrary to other investigations [16-21], we did not observe any significant decreases in LF and LF/HF power or increases in SDNN and the interbeat interval in this assessment. This discrepancy could be related to the very small number of participants in this study. HRV changes in a larger sample would be more revealing. Similar to other investigations, the increases seen in relative HF power and RMSSD values suggest that there was an increase in parasympathetic neural activation during hypnosis [13,17,19,20]. The interpretation of the decrease in relative VLF power during hypnosis is somewhat uncertain. Two recent publications specifically stated that the association between VLF power and autonomic neural function is unclear [76,77]. However, both investigations demonstrated that nonrandom eye movement sleep was associated with reductions in VLF power, when compared with the awake state. Furthermore, the reductions in VLF power in these studies were associated with reductions in LF and LF/HF power, suggesting that reductions in VLF likely imply that a reduction in sympathetic neural activity existed.

Changes in Other Autonomic Variables During Hypnosis

This assessment provides evidence that hypnosis is associated with increases in expired carbon dioxide and skin temperature and decreases in EMG, respiratory rate, and skin conductance values. Nearly all studies investigating the effects of hypnosis on skin temperature involved focused attempts to warm or cool. Only one other investigation demonstrated that neutral hypnotic induction was associated with passive significant increases in skin temperature in pediatric patients [36]. In this assessment, we were able to demonstrate passive significant increases in skin temperature with neutral hypnotic induction in healthy adult volunteers. As in this assessment, 2 other investigations, one in patients with fibromyalgia and the other in healthy adults, demonstrated a decrease in skin conductance during hypnosis [8,29]. The increases in skin temperature and reductions in skin conductance are consistent with a decrease in sympathetic neural activity during hypnosis.

The finding that hypnosis was associated with an increase in expired carbon dioxide is relatively novel. Only 1 study, published more than 30 years ago, has investigated expired carbon dioxide measurements during hypnosis [30]. In that study, the authors compared the results of 27 patients with hyperventilation syndrome with those of 10 healthy controls, who were hospital workers [30]. The results demonstrated that both the patient and control groups had increased expired carbon dioxide and decreased respiratory rates during the deep relaxation phase of hypnosis, similar to the findings of this assessment. Other investigations of healthy adult volunteers also demonstrated that the respiratory rate significantly decreased during hypnosis, similar to the findings of this assessment [31,32]. Although 1 study was performed in France and the other in the United States, the authors reported identical decreases in respiratory rates from 18 to 14 breaths per minute [31,32]. Other hypnosis investigations demonstrated significant reductions in facial EMG measurements in healthy volunteers during hypnosis, as we did in this assessment [33-35]. The increases in expired carbon dioxide values and decreases in the respiratory rate and EMG values are consistent with the onset of relaxation during hypnosis.

Significant Correlations of BIS Values With Other Autonomic Variables

A unique aspect of this assessment is exploring the associations of BIS values with other autonomic measurements. All other investigations that measured BIS during hypnosis reported other measures of physiologic relaxation as coparameters, not as correlations, or were confounded. We identified several significant associations between BIS values and other physiological variables over the 22 session minutes. When BIS values decreased, EMG, respiratory rate, skin conductance, and relative LF power values decreased. However, as BIS values decreased, HF power, relative HF power, RMSSD, and SDNN values increased. The simultaneous reductions in BIS, respiratory rate, and EMG values suggest that lower BIS values connote the presence of a state of physiologic relaxation. The concomitant decreases in BIS values and increases in HF power, relative HF power, RMSSD, and SDNN values imply that BIS reductions are associated with increased parasympathetic neural activation [13]. The large positive associations that we found between HF power with SDNN and RMSSD further support the notion that increased parasympathetic neural activity parallels reductions in BIS values. The concurrent reductions in BIS, skin conductance, and relative LF power values suggest that sympathetic neural activation is decreased [13]. Multivariate linear regression analysis showed that BIS values during the 22-minute session were simultaneously and independently associated with EMG, skin conductance, and relative LF power. These findings are consistent with the notion that reductions in BIS values are principally related to a combined state of relaxation and blunted sympathetic neural responses; however, multiple analyses also suggest that parasympathetic neural activation is also present.

Duration of the Experimental Procedure

More than 20 publications were cited at the beginning of this manuscript regarding the impact of hypnosis on relevant

physiological processes. Of those manuscripts, 8 provided the duration (in minutes) of the control period plus the hypnosis phase. The total duration was 10-30 minutes in 4 manuscripts [1,3,19,36] and 40-45 minutes in the other 4 manuscripts [17,29,30,33]. Hypnosis audiotapes were used in 4 of the investigations. Of relevance, 2 meta-analyses of randomized controlled trials provided substantial documentation regarding the duration of hypnosis sessions in the included studies [78,79]. The duration of the hypnosis period was documented to be 3-20 minutes in 14 studies and 23-30 minutes in 8 trials. Hypnosis audiotapes were used in 8 of these 22 studies. The aforementioned evidence lends credibility to the duration of this experimental design and the use of hypnosis audiotapes.

Limitations

The principal limitation of this formative assessment is the small number of participants; however, the statistical significance of the findings is compelling. Another limitation is that the authors did not compare results between the participants because 3 participants are an insufficient number to compare. Comparisons between age, sex, and other epidemiological information is considered thought provoking in a larger sample. A posthypnosis observation period was not evaluated to determine if the hypnosis-associated physiologic changes were sustained for any period afterward, which would be informative. Finally, the authors did not obtain data on medications, supplements, or

caffeine intake, which could affect participants' physiological variables in a resting state.

Conclusions

The preliminary formative assessment results indicated that the 2 assessment objectives were met. These findings suggest that hypnosis may manifest as a state of mental and physical relaxation, enhance parasympathetic nervous system activation, and attenuate sympathetic nervous system activation, observations that were associated with reductions in BIS values. They also imply that reduced BIS values were associated with a state of physiologic relaxation, increased parasympathetic nervous system activity, and decreased sympathetic nervous system stimulation. In addition to the findings of improvements in participant well-being in our BIS neurofeedback study [73], this evaluation lends further support to the notion that the BIS monitor may be a valid device for use in neurofeedback investigations. Finally, the BIS monitor may subsequently be demonstrated to be considered as a *hypnometer* for continuously measuring hypnotic depth, as proposed by the former Harvard University investigator Solomon Gilbert Diamond [19]. The findings of this preliminary formative evaluation indicate that this hypnosis model may be useful for assessing autonomic physiological associations with changes in BIS values. These observations from this exploratory process motivated us to proceed with a larger and similarly designed investigation that will include trauma center nurses and physician participants.

Acknowledgments

The authors would like to thank Marina C Hanes, BA, ELS, for copyediting the manuscript.

Authors' Contributions

CMD, AJB, BMH, EAC, and AEH contributed to the conception and design of the assessment. CMD, BMH, and EAC conducted the monitoring sessions and documented the data results. CMD performed statistical analyses. CMD, AJB, BMH, EAC, and AEH reviewed the initial data results. CMD wrote the first draft of the manuscript. All authors contributed to manuscript revision and read and approved the final submitted version.

Conflicts of Interest

None declared.

References

1. Ulett GA, Akpınar S, Itil TM. Quantitative EEG analysis during hypnosis. *Electroencephalogr Clin Neurophysiol* 1972 Oct;33(4):361-368. [doi: [10.1016/0013-4694\(72\)90116-2](https://doi.org/10.1016/0013-4694(72)90116-2)] [Medline: [4115694](https://pubmed.ncbi.nlm.nih.gov/4115694/)]
2. Batty MJ, Bonnington S, Tang B, Hawken MB, Gruzelier JH. Relaxation strategies and enhancement of hypnotic susceptibility: EEG neurofeedback, progressive muscle relaxation and self-hypnosis. *Brain Res Bull* 2006 Dec 11;71(1-3):83-90. [doi: [10.1016/j.brainresbull.2006.08.005](https://doi.org/10.1016/j.brainresbull.2006.08.005)] [Medline: [17113932](https://pubmed.ncbi.nlm.nih.gov/17113932/)]
3. Fingelkurts AA, Fingelkurts AA, Kallio S, Revonsuo A. Cortex functional connectivity as a neurophysiological correlate of hypnosis: an EEG case study. *Neuropsychologia* 2007 Apr 08;45(7):1452-1462. [doi: [10.1016/j.neuropsychologia.2006.11.018](https://doi.org/10.1016/j.neuropsychologia.2006.11.018)] [Medline: [17208259](https://pubmed.ncbi.nlm.nih.gov/17208259/)]
4. Sabourin ME, Cutcomb SD, Crawford HJ, Pribram K. EEG correlates of hypnotic susceptibility and hypnotic trance: spectral analysis and coherence. *Int J Psychophysiol* 1990 Dec;10(2):125-142. [doi: [10.1016/0167-8760\(90\)90027-b](https://doi.org/10.1016/0167-8760(90)90027-b)]
5. Graffin NF, Ray WJ, Lundy R. EEG concomitants of hypnosis and hypnotic susceptibility. *J Abnorm Psychol* 1995 Feb;104(1):123-131. [doi: [10.1037/0021-843x.104.1.123](https://doi.org/10.1037/0021-843x.104.1.123)] [Medline: [2272860](https://pubmed.ncbi.nlm.nih.gov/2272860/)]
6. Williams JD, Gruzelier JH. Differentiation of hypnosis and relaxation by analysis of narrow band theta and alpha frequencies. *Int J Clin Exp Hypn* 2001 Jul;49(3):185-206. [doi: [10.1080/00207140108410070](https://doi.org/10.1080/00207140108410070)] [Medline: [11430154](https://pubmed.ncbi.nlm.nih.gov/11430154/)]
7. Burkle CM, Jankowski CJ, Torsher LC, Rho EH, Degnim AC. BIS monitor findings during self-hypnosis. *J Clin Monit Comput* 2005 Dec 25;19(6):391-393. [doi: [10.1007/s10877-005-6539-9](https://doi.org/10.1007/s10877-005-6539-9)] [Medline: [16437289](https://pubmed.ncbi.nlm.nih.gov/16437289/)]

8. Almeida-Marques FX, Sánchez-Blanco J, Cano-García FJ. Hypnosis is more effective than clinical interviews. *Int J Clin Exp Hypn* 2018 Jan 10;66(1):3-18. [doi: [10.1080/00207144.2018.1396104](https://doi.org/10.1080/00207144.2018.1396104)] [Medline: [29319461](https://pubmed.ncbi.nlm.nih.gov/29319461/)]
9. Jaber S, Bahloul H, Guétin S, Chanques G, Sebbane M, Eledjam J. [Effects of music therapy in intensive care unit without sedation in weaning patients versus non-ventilated patients]. *Ann Fr Anesth Reanim* 2007 Jan;26(1):30-38. [doi: [10.1016/j.annfar.2006.09.002](https://doi.org/10.1016/j.annfar.2006.09.002)] [Medline: [17085009](https://pubmed.ncbi.nlm.nih.gov/17085009/)]
10. Tsutsumi M, Nogaki H, Shimizu Y, Stone TE, Kobayashi T. Individual reactions to viewing preferred video representations of the natural environment: a comparison of mental and physical reactions. *Jpn J Nurs Sci* 2017 Jan 10;14(1):3-12. [doi: [10.1111/jjns.12131](https://doi.org/10.1111/jjns.12131)] [Medline: [27160351](https://pubmed.ncbi.nlm.nih.gov/27160351/)]
11. Hudetz JA, Hudetz AG, Reddy DM. Effect of relaxation on working memory and the Bispectral Index of the EEG. *Psychol Rep* 2004 Aug 31;95(1):53-70. [doi: [10.2466/pr0.95.1.53-70](https://doi.org/10.2466/pr0.95.1.53-70)] [Medline: [15460358](https://pubmed.ncbi.nlm.nih.gov/15460358/)]
12. Ernst G. Heart-rate variability-more than heart beats? *Front Public Health* 2017 Sep 11;5:240 [FREE Full text] [doi: [10.3389/fpubh.2017.00240](https://doi.org/10.3389/fpubh.2017.00240)] [Medline: [28955705](https://pubmed.ncbi.nlm.nih.gov/28955705/)]
13. Shaffer F, Ginsberg JP. An overview of heart rate variability metrics and norms. *Front Public Health* 2017 Sep;5:258 [FREE Full text] [doi: [10.3389/fpubh.2017.00258](https://doi.org/10.3389/fpubh.2017.00258)] [Medline: [29034226](https://pubmed.ncbi.nlm.nih.gov/29034226/)]
14. Karim N, Hasan JA, Ali SS. Heart rate variability – a review. *J Basic Appl Sci.* 2011. URL: <https://vlibrary.emro.who.int/imemr/heart-rate-variability-a-review-2/> [accessed 2021-05-03]
15. Fauquet-Alekhine P, Rouillac L, Berton J, Granry JC. Heart rate vs stress indicator for short term mental stress. *Br J Med Med Res* 2016 Jan 10;17(7):1-11. [doi: [10.9734/BJMMR/2016/27593](https://doi.org/10.9734/BJMMR/2016/27593)]
16. Hippel CV, Hole G, Kaschka WP. Autonomic profile under hypnosis as assessed by heart rate variability and spectral analysis. *Pharmacopsychiatry* 2001 May;34(3):111-113. [doi: [10.1055/s-2001-14279](https://doi.org/10.1055/s-2001-14279)] [Medline: [11434402](https://pubmed.ncbi.nlm.nih.gov/11434402/)]
17. Aubert AE, Verheyden B, Beckers F, Tack J, Vandenberghe J. Cardiac autonomic regulation under hypnosis assessed by heart rate variability: spectral analysis and fractal complexity. *Neuropsychobiology* 2009;60(2):104-112. [doi: [10.1159/000239686](https://doi.org/10.1159/000239686)] [Medline: [19776654](https://pubmed.ncbi.nlm.nih.gov/19776654/)]
18. Yüksel R, Ozcan O, Dane S. The effects of hypnosis on heart rate variability. *Int J Clin Exp Hypn* 2013 Apr;61(2):162-171. [doi: [10.1080/00207144.2013.753826](https://doi.org/10.1080/00207144.2013.753826)] [Medline: [23427840](https://pubmed.ncbi.nlm.nih.gov/23427840/)]
19. Diamond SG, Davis OC, Howe RD. Heart-rate variability as a quantitative measure of hypnotic depth. *Int J Clin Exp Hypn* 2008 Jan 27;56(1):1-18. [doi: [10.1080/00207140701672961](https://doi.org/10.1080/00207140701672961)] [Medline: [18058483](https://pubmed.ncbi.nlm.nih.gov/18058483/)]
20. Chen X, Yang R, Ge L, Luo J, Lv R. Hypnosis in the treatment of major depression: an analysis of heart rate variability. *Int J Clin Exp Hypn* 2017 Dec 09;65(1):52-63. [doi: [10.1080/00207144.2017.1246873](https://doi.org/10.1080/00207144.2017.1246873)] [Medline: [27935459](https://pubmed.ncbi.nlm.nih.gov/27935459/)]
21. DeBenedittis G, Cigada M, Bianchi A, Signorini MG, Cerutti S. Autonomic changes during hypnosis: a heart rate variability power spectrum analysis as a marker of sympatho-vagal balance. *Int J Clin Exp Hypn* 1994 Apr;42(2):140-152. [doi: [10.1080/00207149408409347](https://doi.org/10.1080/00207149408409347)] [Medline: [8200716](https://pubmed.ncbi.nlm.nih.gov/8200716/)]
22. Markil N, Whitehurst M, Jacobs PL, Zoeller RF. Yoga Nidra relaxation increases heart rate variability and is unaffected by a prior bout of Hatha yoga. *J Altern Complement Med* 2012 Oct;18(10):953-958. [doi: [10.1089/acm.2011.0331](https://doi.org/10.1089/acm.2011.0331)] [Medline: [22866996](https://pubmed.ncbi.nlm.nih.gov/22866996/)]
23. Pal GK, Ganesh V, Karthik S, Nanda N, Pal P. The effects of short-term relaxation therapy on indices of heart rate variability and blood pressure in young adults. *Am J Health Promot* 2014 Sep;29(1):23-28. [doi: [10.4278/ajhp.130131-quan-52](https://doi.org/10.4278/ajhp.130131-quan-52)] [Medline: [24200249](https://pubmed.ncbi.nlm.nih.gov/24200249/)]
24. Shearer A, Hunt M, Chowdhury M, Nicol L. Effects of a brief mindfulness meditation intervention on student stress and heart rate variability. *Int J Stress Manag* 2016 May;23(2):232-254. [doi: [10.1037/a0039814](https://doi.org/10.1037/a0039814)]
25. Heemskerk J, Zandbergen HR, Keet SW, Martijnse I, van Montfort G, Peters RJ, et al. Relax, it's just laparoscopy! A prospective randomized trial on heart rate variability of the surgeon in robot-assisted versus conventional laparoscopic cholecystectomy. *Dig Surg* 2014 Sep 25;31(3):225-232. [doi: [10.1159/000365580](https://doi.org/10.1159/000365580)] [Medline: [25277215](https://pubmed.ncbi.nlm.nih.gov/25277215/)]
26. Kim H, Cheon E, Bai D, Lee YH, Koo B. Stress and heart rate variability: a meta-analysis and review of the literature. *Psychiatry Investig* 2018 Mar;15(3):235-245 [FREE Full text] [doi: [10.30773/pi.2017.08.17](https://doi.org/10.30773/pi.2017.08.17)] [Medline: [29486547](https://pubmed.ncbi.nlm.nih.gov/29486547/)]
27. Borchini R, Veronesi G, Bonzini M, Gianfagna F, Dashi O, Ferrario M. Heart rate variability frequency domain alterations among healthy nurses exposed to prolonged work stress. *Int J Environ Res Public Health* 2018 Jan 11;15(1):113 [FREE Full text] [doi: [10.3390/ijerph15010113](https://doi.org/10.3390/ijerph15010113)] [Medline: [29324640](https://pubmed.ncbi.nlm.nih.gov/29324640/)]
28. Nakayama N, Arakawa N, Ejiri H, Matsuda R, Makino T. Heart rate variability can clarify students' level of stress during nursing simulation. *PLoS One* 2018 Apr 5;13(4):e0195280 [FREE Full text] [doi: [10.1371/journal.pone.0195280](https://doi.org/10.1371/journal.pone.0195280)] [Medline: [29621278](https://pubmed.ncbi.nlm.nih.gov/29621278/)]
29. Kekecs Z, Szekely A, Varga K. Alterations in electrodermal activity and cardiac parasympathetic tone during hypnosis. *Psychophysiology* 2016 Feb 21;53(2):268-277. [doi: [10.1111/psyp.12570](https://doi.org/10.1111/psyp.12570)] [Medline: [26488759](https://pubmed.ncbi.nlm.nih.gov/26488759/)]
30. Freeman LJ, Conway A, Nixon PG. Physiological responses to psychological challenge under hypnosis in patients considered to have the hyperventilation syndrome: implications for diagnosis and therapy. *J R Soc Med* 1986 Feb 29;79(2):76-83 [FREE Full text] [doi: [10.1177/014107688607900206](https://doi.org/10.1177/014107688607900206)] [Medline: [3081708](https://pubmed.ncbi.nlm.nih.gov/3081708/)]
31. VandeVusse L, Hanson L, Berner MA, White Winters JM. Impact of self-hypnosis in women on select physiologic and psychological parameters. *J Obstet Gynecol Neonatal Nurs* 2010 Mar;39(2):159-168. [doi: [10.1111/j.1552-6909.2010.01103.x](https://doi.org/10.1111/j.1552-6909.2010.01103.x)] [Medline: [20409116](https://pubmed.ncbi.nlm.nih.gov/20409116/)]

32. Boselli E, Musellec H, Martin L, Bernard F, Fusco N, Guillou N, et al. Effects of hypnosis on the relative parasympathetic tone assessed by ANI (Analgesia/Nociception Index) in healthy volunteers: a prospective observational study. *J Clin Monit Comput* 2018 Jun 20;32(3):487-492. [doi: [10.1007/s10877-017-0056-5](https://doi.org/10.1007/s10877-017-0056-5)] [Medline: [28825157](https://pubmed.ncbi.nlm.nih.gov/28825157/)]
33. Miller LS, Cross HJ. Hypnotic susceptibility, hypnosis, and EMG biofeedback in the reduction of frontalis muscle tension. *Int J Clin Exp Hypn* 1985 Jul 31;33(3):258-272. [doi: [10.1080/00207148508406654](https://doi.org/10.1080/00207148508406654)] [Medline: [4030155](https://pubmed.ncbi.nlm.nih.gov/4030155/)]
34. Manns A, Zuazola RV, Sirhan RM, Quiroz M, Rocabado M. Relationship between the tonic elevator mandibular activity and the vertical dimension during the states of vigilance and hypnosis. *Cranio* 1990 Apr;8(2):163-170. [doi: [10.1080/08869634.1990.11678310](https://doi.org/10.1080/08869634.1990.11678310)] [Medline: [2073696](https://pubmed.ncbi.nlm.nih.gov/2073696/)]
35. Al-Enaizan N, Davey KJ, Lyons MF, Cadden SW. Effect of hypnosis on masseter EMG recorded during the 'resting' and a slightly open jaw posture. *J Oral Rehabil* 2015 Nov 07;42(11):840-846. [doi: [10.1111/joor.12316](https://doi.org/10.1111/joor.12316)] [Medline: [26059538](https://pubmed.ncbi.nlm.nih.gov/26059538/)]
36. Dikel W, Olness K. Self-hypnosis, biofeedback, and voluntary peripheral temperature control in children. *Pediatrics* 1980 Sep;66(3):335-340. [Medline: [7422423](https://pubmed.ncbi.nlm.nih.gov/7422423/)]
37. Khazan I. *Biofeedback and Mindfulness in Everyday Life – Practical Solutions for Improving Your Health and Performance*. New York: W. W. Norton & Company; 2019:1-288.
38. Herborn KA, Graves JL, Jerem P, Evans NP, Nager R, McCafferty DJ, et al. Skin temperature reveals the intensity of acute stress. *Physiol Behav* 2015 Dec 01;152(Pt A):225-230 [FREE Full text] [doi: [10.1016/j.physbeh.2015.09.032](https://doi.org/10.1016/j.physbeh.2015.09.032)] [Medline: [26434785](https://pubmed.ncbi.nlm.nih.gov/26434785/)]
39. Shusterman V, Barnea O. Spectral characteristics of skin temperature indicate peripheral stress-response. *Biofeedback Self Regul* 1995 Dec;20(4):357-367. [doi: [10.1007/BF01543790](https://doi.org/10.1007/BF01543790)] [Medline: [8695701](https://pubmed.ncbi.nlm.nih.gov/8695701/)]
40. Yamakoshi T, Yamakoshi K, Tanaka S, Nogawa M, Park SB, Shibata M, et al. Feasibility study on driver's stress detection from differential skin temperature measurement. In: *Proceedings of the 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. 2008 Presented at: 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society; August 20-25, 2008; Vancouver, BC, Canada p. 1076-1079. [doi: [10.1109/iembs.2008.4649346](https://doi.org/10.1109/iembs.2008.4649346)] [Medline: [19162849](https://pubmed.ncbi.nlm.nih.gov/19162849/)]
41. Moser DK, Dracup K, Woo MA, Stevenson LW. Voluntary control of vascular tone by using skin-temperature biofeedback-relaxation in patients with advanced heart failure. *Altern Ther Health Med* 1997 Jan;3(1):51-59. [Medline: [8997805](https://pubmed.ncbi.nlm.nih.gov/8997805/)]
42. Jacobs SC, Friedman R, Parker JD, Tofler GH, Jimenez AH, Muller JE, et al. Use of skin conductance changes during mental stress testing as an index of autonomic arousal in cardiovascular research. *Am Heart J* 1994 Dec;128(6):1170-1177. [doi: [10.1016/0002-8703\(94\)90748-x](https://doi.org/10.1016/0002-8703(94)90748-x)] [Medline: [7985598](https://pubmed.ncbi.nlm.nih.gov/7985598/)]
43. Lin H, Lin H, Lin W, Huang AC. Effects of stress, depression, and their interaction on heart rate, skin conductance, finger temperature, and respiratory rate: sympathetic-parasympathetic hypothesis of stress and depression. *J Clin Psychol* 2011 Oct 26;67(10):1080-1091. [doi: [10.1002/jclp.20833](https://doi.org/10.1002/jclp.20833)] [Medline: [21905026](https://pubmed.ncbi.nlm.nih.gov/21905026/)]
44. Liu J, Verhulst S, Massar S, Chee M. Sleep deprived and sweating it out: the effects of total sleep deprivation on skin conductance reactivity to psychosocial stress. *Sleep* 2015 Jan 01;38(1):155-159 [FREE Full text] [doi: [10.5665/sleep.4346](https://doi.org/10.5665/sleep.4346)] [Medline: [25325448](https://pubmed.ncbi.nlm.nih.gov/25325448/)]
45. Hinrichs R, Michopoulos V, Winters S, Rothbaum AO, Rothbaum BO, Ressler KJ, et al. Mobile assessment of heightened skin conductance in posttraumatic stress disorder. *Depress Anxiety* 2017 Jun 21;34(6):502-507 [FREE Full text] [doi: [10.1002/da.22610](https://doi.org/10.1002/da.22610)] [Medline: [28221710](https://pubmed.ncbi.nlm.nih.gov/28221710/)]
46. Kim AY, Jang EH, Choi KW, Jeon HJ, Byun S, Sim JY, et al. Skin conductance responses in Major Depressive Disorder (MDD) under mental arithmetic stress. *PLoS One* 2019 Apr 3;14(4):e0213140 [FREE Full text] [doi: [10.1371/journal.pone.0213140](https://doi.org/10.1371/journal.pone.0213140)] [Medline: [30943195](https://pubmed.ncbi.nlm.nih.gov/30943195/)]
47. Storm H, Shafiei M, Myre K, Raeder J. Palmar skin conductance compared to a developed stress score and to noxious and awakening stimuli on patients in anaesthesia. *Acta Anaesthesiol Scand* 2005 Jul;49(6):798-803. [doi: [10.1111/j.1399-6576.2005.00665.x](https://doi.org/10.1111/j.1399-6576.2005.00665.x)] [Medline: [15954962](https://pubmed.ncbi.nlm.nih.gov/15954962/)]
48. Schleifer LM, Ley R. End-tidal PCO₂ as an index of psychophysiological activity during VDT data-entry work and relaxation. *Ergonomics* 1994 Feb;37(2):245-254. [doi: [10.1080/00140139408963642](https://doi.org/10.1080/00140139408963642)] [Medline: [8119258](https://pubmed.ncbi.nlm.nih.gov/8119258/)]
49. Masaoka Y, Homma I. Anxiety and respiratory patterns: their relationship during mental stress and physical load. *Int J Psychophysiol* 1997 Sep;27(2):153-159. [doi: [10.1016/s0167-8760\(97\)00052-4](https://doi.org/10.1016/s0167-8760(97)00052-4)] [Medline: [9342646](https://pubmed.ncbi.nlm.nih.gov/9342646/)]
50. Ley R, Yelich G. Fractional end-tidal CO₂ as an index of the effects of stress on math performance and verbal memory of test-anxious adolescents. *Biol Psychol* 1998 Sep;49(1-2):83-94. [doi: [10.1016/s0301-0511\(98\)00028-3](https://doi.org/10.1016/s0301-0511(98)00028-3)] [Medline: [9792486](https://pubmed.ncbi.nlm.nih.gov/9792486/)]
51. Widjaja D, Orini M, Vlemincx E, Van Huffel S. Cardiorespiratory dynamic response to mental stress: a multivariate time-frequency analysis. *Comput Math Methods Med* 2013;2013:451857 [FREE Full text] [doi: [10.1155/2013/451857](https://doi.org/10.1155/2013/451857)] [Medline: [24386006](https://pubmed.ncbi.nlm.nih.gov/24386006/)]
52. Murayama Y, Hu L, Sakatani K. Relation between prefrontal cortex activity and respiratory rate during mental stress tasks: a near-infrared spectroscopic study. *Adv Exp Med Biol* 2016;923:209-214. [doi: [10.1007/978-3-319-38810-6_28](https://doi.org/10.1007/978-3-319-38810-6_28)] [Medline: [27526145](https://pubmed.ncbi.nlm.nih.gov/27526145/)]

53. Kreibig SD, Wilhelm FH, Roth WT, Gross JJ. Cardiovascular, electrodermal, and respiratory response patterns to fear- and sadness-inducing films. *Psychophysiology* 2007 Sep;44(5):787-806. [doi: [10.1111/j.1469-8986.2007.00550.x](https://doi.org/10.1111/j.1469-8986.2007.00550.x)] [Medline: [17598878](https://pubmed.ncbi.nlm.nih.gov/17598878/)]
54. Khanna A, Paul M, Sandhu JS. A study to compare the effectiveness of GSR biofeedback training and progressive muscle relaxation training in reducing blood pressure and respiratory rate among highly stressed individuals. *Indian J Physiol Pharmacol* 2007;51(3):296-300. [Medline: [18341229](https://pubmed.ncbi.nlm.nih.gov/18341229/)]
55. Himashree G, Mohan L, Singh Y. Yoga practice improves physiological and biochemical status at high altitudes: a prospective case-control study. *Altern Ther Health Med* 2016 Sep;22(5):53-59. [Medline: [27622961](https://pubmed.ncbi.nlm.nih.gov/27622961/)]
56. Blom EH, Serlachius E, Chesney MA, Olsson EM. Adolescent girls with emotional disorders have a lower end-tidal CO₂ and increased respiratory rate compared with healthy controls. *Psychophysiology* 2014 May 27;51(5):412-418 [FREE Full text] [doi: [10.1111/psyp.12188](https://doi.org/10.1111/psyp.12188)] [Medline: [24571123](https://pubmed.ncbi.nlm.nih.gov/24571123/)]
57. Conrad A, Isaac L, Roth WT. The psychophysiology of generalized anxiety disorder: 1. Pretreatment characteristics. *Psychophysiology* 2008 May;45(3):366-376. [doi: [10.1111/j.1469-8986.2007.00601.x](https://doi.org/10.1111/j.1469-8986.2007.00601.x)] [Medline: [18221449](https://pubmed.ncbi.nlm.nih.gov/18221449/)]
58. Schleifer LM, Spalding TW, Kerick SE, Cram JR, Ley R, Hatfield BD. Mental stress and trapezius muscle activation under psychomotor challenge: a focus on EMG gaps during computer work. *Psychophysiology* 2008 May;45(3):356-365. [doi: [10.1111/j.1469-8986.2008.00645.x](https://doi.org/10.1111/j.1469-8986.2008.00645.x)] [Medline: [18282206](https://pubmed.ncbi.nlm.nih.gov/18282206/)]
59. Zheng R, Yamabe S, Nakano K, Suda Y. Biosignal analysis to assess mental stress in automatic driving of trucks: palmar perspiration and masseter electromyography. *Sensors (Basel)* 2015 Mar 02;15(3):5136-5150 [FREE Full text] [doi: [10.3390/s150305136](https://doi.org/10.3390/s150305136)] [Medline: [25738768](https://pubmed.ncbi.nlm.nih.gov/25738768/)]
60. McGlynn FD, Bichajian C, Tira DE, Lundeen HC, Mahan PE, Nicholas BV. The effect of experimental stress and experimental occlusal interference on masseteric EMG activity. *J Craniomandib Disord* 1989;3(2):87-92. [Medline: [2607002](https://pubmed.ncbi.nlm.nih.gov/2607002/)]
61. Rissén D, Melin B, Sandsjö L, Dohns I, Lundberg U. Surface EMG and psychophysiological stress reactions in women during repetitive work. *Eur J Appl Physiol* 2000 Oct 13;83(2-3):215-222. [doi: [10.1007/s004210000281](https://doi.org/10.1007/s004210000281)] [Medline: [11104063](https://pubmed.ncbi.nlm.nih.gov/11104063/)]
62. Jones DM, Lewis R. Effects of structured relaxation on retrieval from long term storage: an EMG study. *Biol Psychol* 1980 Aug;11(1):37-44. [doi: [10.1016/0301-0511\(80\)90024-1](https://doi.org/10.1016/0301-0511(80)90024-1)] [Medline: [7018600](https://pubmed.ncbi.nlm.nih.gov/7018600/)]
63. Kappes BM. Sequence effects of relaxation training, EMG, and temperature biofeedback on anxiety, symptom report, and self-concept. *J Clin Psychol* 1983 Mar;39(2):203-208. [doi: [10.1002/1097-4679\(198303\)39:2<203::aid-jclp2270390209>3.0.co;2-h](https://doi.org/10.1002/1097-4679(198303)39:2<203::aid-jclp2270390209>3.0.co;2-h)] [Medline: [6341411](https://pubmed.ncbi.nlm.nih.gov/6341411/)]
64. Narayan R, Kamat A, Khanolkar M, Kamat S, Desai SR, Dhume RA. Quantitative evaluation of muscle relaxation induced by Kundalini yoga with the help of EMG integrator. *Indian J Physiol Pharmacol* 1990 Oct;34(4):279-281. [Medline: [2100290](https://pubmed.ncbi.nlm.nih.gov/2100290/)]
65. Kanaya N, Hirata N, Kurosawa S, Nakayama M, Namiki A. Differential effects of propofol and sevoflurane on heart rate variability. *Anesthesiology* 2003 Jan;98(1):34-40 [FREE Full text] [doi: [10.1097/0000542-200301000-00009](https://doi.org/10.1097/0000542-200301000-00009)] [Medline: [12502976](https://pubmed.ncbi.nlm.nih.gov/12502976/)]
66. Lin YT, Wu HT, Tsao J, Yien HW, Hseu SS. Time-varying spectral analysis revealing differential effects of sevoflurane anaesthesia: non-rhythmic-to-rhythmic ratio. *Acta Anaesthesiol Scand* 2014 Feb 10;58(2):157-167. [doi: [10.1111/aas.12251](https://doi.org/10.1111/aas.12251)] [Medline: [24410106](https://pubmed.ncbi.nlm.nih.gov/24410106/)]
67. Liu Q, Ma L, Chiu R, Fan S, Abbod M, Shieh J. HRV-derived data similarity and distribution index based on ensemble neural network for measuring depth of anaesthesia. *PeerJ* 2017;5:e4067 [FREE Full text] [doi: [10.7717/peerj.4067](https://doi.org/10.7717/peerj.4067)] [Medline: [29158992](https://pubmed.ncbi.nlm.nih.gov/29158992/)]
68. Tsuchiya S, Kanaya N, Hirata N, Kurosawa S, Kamada N, Edanaga M, et al. Effects of thiopental on bispectral index and heart rate variability. *Eur J Anaesthesiol* 2006;23(6):454-459. [doi: [10.1017/s0265021506000159](https://doi.org/10.1017/s0265021506000159)] [Medline: [16507201](https://pubmed.ncbi.nlm.nih.gov/16507201/)]
69. Panousis P, Heller AR, Burghardt M, Bleyl JU, Koch T. The effects of electromyographic activity on the accuracy of the Narcotrend monitor compared with the Bispectral Index during combined anaesthesia. *Anaesthesia* 2007 Sep;62(9):868-874 [FREE Full text] [doi: [10.1111/j.1365-2044.2007.05145.x](https://doi.org/10.1111/j.1365-2044.2007.05145.x)] [Medline: [17697211](https://pubmed.ncbi.nlm.nih.gov/17697211/)]
70. Hayashi K, Sawa T. The fundamental contribution of the electromyogram to a high bispectral index: a postoperative observational study. *J Clin Monit Comput* 2019 Dec 3;33(6):1097-1103. [doi: [10.1007/s10877-018-00244-1](https://doi.org/10.1007/s10877-018-00244-1)] [Medline: [30607805](https://pubmed.ncbi.nlm.nih.gov/30607805/)]
71. Hagiwara A, Matsuura N, Ichinohe T. Comparison of changes in respiratory dynamics immediately after the start of propofol sedation with or without midazolam. *J Oral Maxillofac Surg* 2018 Jan;76(1):52-59. [doi: [10.1016/j.joms.2017.05.038](https://doi.org/10.1016/j.joms.2017.05.038)] [Medline: [28672136](https://pubmed.ncbi.nlm.nih.gov/28672136/)]
72. Yıldırım AA, Abitağaoğlu S, Turan G, Karip C, Akgün N, Boybeyi DM, et al. Integrated Pulmonary Index (IPI) monitorization under sedation in cataract surgery with phacoemulsification technique. *Int Ophthalmol* 2019 Sep 4;39(9):1949-1954. [doi: [10.1007/s10792-018-1024-x](https://doi.org/10.1007/s10792-018-1024-x)] [Medline: [30284695](https://pubmed.ncbi.nlm.nih.gov/30284695/)]
73. Dunham CM, Burger AL, Hileman BM, Chance EA, Hutchinson AE, Kohli CM, et al. Brainwave self-regulation during bispectral index neurofeedback in trauma center nurses and physicians after receiving mindfulness instructions. *Front Psychol* 2019 Sep 26;10:2153 [FREE Full text] [doi: [10.3389/fpsyg.2019.02153](https://doi.org/10.3389/fpsyg.2019.02153)] [Medline: [31616348](https://pubmed.ncbi.nlm.nih.gov/31616348/)]

74. Dunham CM, McClain JV, Burger A. Comparison of Bispectral Index™ values during the flotation restricted environmental stimulation technique and results for stage I sleep: a prospective pilot investigation. *BMC Res Notes* 2017 Nov 29;10(1):640 [FREE Full text] [doi: [10.1186/s13104-017-2947-4](https://doi.org/10.1186/s13104-017-2947-4)] [Medline: [29187246](https://pubmed.ncbi.nlm.nih.gov/29187246/)]
75. Dunham CM, Burger AL, Hileman BM, Chance EA. Learning receptive awareness via neurofeedback in stressed healthcare providers: a prospective pilot investigation. *BMC Res Notes* 2018 Sep 04;11(1):645 [FREE Full text] [doi: [10.1186/s13104-018-3756-0](https://doi.org/10.1186/s13104-018-3756-0)] [Medline: [30180909](https://pubmed.ncbi.nlm.nih.gov/30180909/)]
76. Huang R, Lai C, Lee S, Wang W, Tseng L, Chen Y, et al. Scaling exponent values as an ordinary function of the ratio of very low frequency to high frequency powers in heart rate variability over various sleep stages. *Sleep Breath* 2016 Sep 2;20(3):975-985. [doi: [10.1007/s11325-016-1320-5](https://doi.org/10.1007/s11325-016-1320-5)] [Medline: [27039097](https://pubmed.ncbi.nlm.nih.gov/27039097/)]
77. Utriainen KT, Airaksinen JK, Polo OJ, Scheinin H, Laitio RM, Leino KA, et al. Alterations in heart rate variability in patients with peripheral arterial disease requiring surgical revascularization have limited association with postoperative major adverse cardiovascular and cerebrovascular events. *PLoS One* 2018 Sep 13;13(9):e0203519 [FREE Full text] [doi: [10.1371/journal.pone.0203519](https://doi.org/10.1371/journal.pone.0203519)] [Medline: [30212552](https://pubmed.ncbi.nlm.nih.gov/30212552/)]
78. Tefikow S, Barth J, Maichrowitz S, Beelmann A, Strauss B, Rosendahl J. Efficacy of hypnosis in adults undergoing surgery or medical procedures: a meta-analysis of randomized controlled trials. *Clin Psychol Rev* 2013 Jul;33(5):623-636. [doi: [10.1016/j.cpr.2013.03.005](https://doi.org/10.1016/j.cpr.2013.03.005)] [Medline: [23628907](https://pubmed.ncbi.nlm.nih.gov/23628907/)]
79. Fisch S, Brinkhaus B, Teut M. Hypnosis in patients with perceived stress - a systematic review. *BMC Complement Altern Med* 2017 Jun 19;17(1):323 [FREE Full text] [doi: [10.1186/s12906-017-1806-0](https://doi.org/10.1186/s12906-017-1806-0)] [Medline: [28629342](https://pubmed.ncbi.nlm.nih.gov/28629342/)]

Abbreviations

- BIS:** bispectral index
- EEG:** electroencephalography
- EMG:** electromyography
- HF:** high frequency
- HRV:** heart rate variability
- LF:** low frequency
- RMSSD:** root mean square of successive cardiac interbeat time intervals
- SDNN:** SD of cardiac interbeat (normal-to-normal)
- VLF:** very low frequency

Edited by G Eysenbach; submitted 01.09.20; peer-reviewed by J Alqahtani, WF Tan; comments to author 20.01.21; revised version received 10.02.21; accepted 13.04.21; published 26.05.21.

Please cite as:

Dunham CM, Burger AJ, Hileman BM, Chance EA, Hutchinson AE

Bispectral Index Alterations and Associations With Autonomic Changes During Hypnosis in Trauma Center Researchers: Formative Evaluation Study

JMIR Form Res 2021;5(5):e24044

URL: <https://formative.jmir.org/2021/5/e24044>

doi: [10.2196/24044](https://doi.org/10.2196/24044)

PMID: [34037529](https://pubmed.ncbi.nlm.nih.gov/34037529/)

©C Michael Dunham, Amanda J Burger, Barbara M Hileman, Elisha A Chance, Amy E Hutchinson. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 26.05.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Digitalizing a Brief Intervention to Reduce Intrusive Memories of Psychological Trauma for Health Care Staff Working During COVID-19: Exploratory Pilot Study With Nurses

Laura Singh^{1,2*}, PhD; Marie Kanstrup^{3,4*}, PhD; Katherine Depa³, MA; Ann-Charlotte Falk⁵, PhD; Veronica Lindström^{5,6,7}, PhD; Oili Dahl⁶, PhD; Katarina E Göransson^{8,9}, PhD; Ann Rudman^{3,10}, PhD; Emily A Holmes^{1,3}, PhD

¹Department of Psychology, Uppsala University, Uppsala, Sweden

²Swedish Collegium for Advanced Study, Uppsala, Sweden

³Division of Psychology, Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden

⁴Functional Area Medical Psychology, Karolinska University Hospital, Stockholm, Sweden

⁵Department for Health Promoting Science, Sophiahemmet University, Stockholm, Sweden

⁶Division of Nursing, Department of Neurobiology, Care Sciences and Society, Karolinska Institutet, Huddinge, Sweden

⁷Samariten Ambulance, Stockholm, Sweden

⁸Emergency and Reparatative Medicine Theme, Karolinska University Hospital, Stockholm, Sweden

⁹Department of Medicine, Solna, Karolinska Institutet, Stockholm, Sweden

¹⁰School of Education, Health and Social Studies, Dalarna University, Falun, Sweden

*these authors contributed equally

Corresponding Author:

Laura Singh, PhD

Department of Psychology

Uppsala University

Box 1225

Uppsala, 75142

Sweden

Phone: 46 018 471 2118

Email: laura.singh@psyk.uu.se

Abstract

Background: The COVID-19 pandemic has accelerated the worldwide need for simple remotely delivered (digital) scalable interventions that can also be used preventatively to protect the mental health of health care staff exposed to psychologically traumatic events during their COVID-19-related work. We have developed a brief behavioral intervention that aims to reduce the number of intrusive memories of traumatic events but has only been delivered face-to-face so far. After digitalizing the intervention materials, the intervention was delivered digitally to target users (health care staff) for the first time. The adaptation for staff's working context in a hospital setting used a co-design approach.

Objective: The aims of this mixed method exploratory pilot study with health care staff who experienced working in the pandemic were to pilot the intervention that we have digitalized (for remote delivery and with remote support) and adapted for this target population (health care staff working clinically during a pandemic) to explore its ability to reduce the number of intrusive memories of traumatic events and improve related symptoms (eg, posttraumatic stress) and participant's perception of their functioning, and to explore the feasibility and acceptability of both the digitalized intervention and digitalized data collection.

Methods: We worked closely with target users with lived experience of working clinically during the COVID-19 pandemic in a hospital context (registered nurses who experienced intrusive memories from traumatic events at work; N=3). We used a mixed method design and exploratory quantitative and qualitative analysis.

Results: After completing the digitalized intervention once with remote researcher support (approximately 25 minutes) and a brief follow-up check-in, participants learned to use the intervention independently. All 3 participants reported zero intrusive memories during week 5 (primary outcome: 100% digital data capture). Prior to study inclusion, two or more intrusions in the week were reported preintervention (assessed retrospectively). There was a general pattern of symptom reduction and improvement in perceived functioning (eg, concentration) at follow-up. The digitalized intervention and data collection were perceived as

feasible and rated as acceptable (eg, all 3 participants would recommend it to a colleague). Participants were positive toward the digital intervention as a useful tool that could readily be incorporated into work life and repeated in the face of ongoing or repeated trauma exposure.

Conclusions: The intervention when delivered remotely and adapted for this population during the pandemic was well received by participants. Since it could be tailored around work and daily life and used preventatively, the intervention may hold promise for health care staff pending future evaluations of efficacy. Limitations include the small sample size, lack of daily intrusion frequency data in the week before the intervention, and lack of a control condition. Following this co-design process in adapting and improving intervention delivery and evaluation, the next step is to investigate the efficacy of the digitalized intervention in a randomized controlled trial.

(*JMIR Form Res* 2021;5(5):e27473) doi:[10.2196/27473](https://doi.org/10.2196/27473)

KEYWORDS

intrusive memories; psychological trauma; prevention; pilot trial; COVID-19; digital intervention; remote delivery; cognitive science; person-based approach; mixed methods; co-design; health care staff

Introduction

The mental health of health care staff exposed to stressful and traumatic events during their work in the COVID-19 pandemic is a number one research priority internationally [1]. This work-related exposure to psychologically traumatic events may have serious effects for staff, such as symptoms related to posttraumatic stress [2]. Brief interventions that prevent the buildup or recurrence of such symptoms and can be tailored around working life and delivered remotely are urgently needed.

Here, we focus on a brief intervention to target one focal symptom that can arise after exposure to traumatic events—intrusive memories [3]. Intrusive memories are defined as recurrent distressing sensory-perceptual impressions of the traumatic event that intrude into the mind involuntarily [3], typically in the form of visual images [4]. They are a core clinical feature [5] of posttraumatic stress disorder (PTSD) [3] and constitute a promising target for novel interventions [6]. Intrusive memories can be distressing in their own right and can impair work functioning, for example, by disturbing concentration [7]. One concern for health care staff is how such mental health symptoms might affect their ability to deliver high quality patient care [2]. A barrier for staff is how to fit time for treatment into an already overly burdened schedule.

Exposure to psychologically traumatic events presents a problem for health care staff working in the pandemic and will continue to be a problem once the pandemic is over. Before the pandemic, experiencing different traumatic events (either as a direct threat to themselves or a witnessed threat to patients) could lead to mental health difficulties such as PTSD in health care staff [8] as shown in studies including rescue workers [9], obstetricians [10], critical care nurses [11], and emergency nurses [12,13]. During the COVID-19 pandemic, health care staff have experienced much higher levels of exposure to potentially traumatic events and already reported increased posttraumatic stress symptoms [2,14,15]. For example, 35% of health care workers exposed to COVID-19 in China reported moderate to severe PTSD symptoms 1 month after onset of the COVID-19 pandemic [16]; 26% of health care workers in Italy scored above the cutoff for PTSD [17], 40% of intensive care unit (ICU) staff in the United Kingdom reported clinically significant levels of PTSD symptoms [2], and 64% of nurses in Jordan experienced

acute stress disorder during the pandemic [18]. PTSD symptoms can impair work performance: 27% of medical workers who reported PTSD symptoms said it interfered with their work functioning [19] and 20% considered changing their job [20]. In addition to staff turnover, PTSD symptoms have also been related to burnout in health care staff [21].

Interventions to support health care staff need to be suitable for delivery in their work context. This includes being suitable for ongoing exposure to stressful events in the line of work (ie, repeatable); preventative use to keep staff well and working (ie, prevent the buildup of symptoms); scalability, simplicity, and brevity; and remote delivery to reduce the risk of virus transmission during a pandemic. The development of new interventions requires adaptation and detailed feedback from people with lived experience [22]. In that light, we will report a pilot study with health care staff (N=3) working with patients with COVID-19.

Good evidence-based treatments for PTSD exist (eg, trauma-focused cognitive behavioral therapy and eye movement desensitization and reprocessing [23,24]). However, there is a lack of available therapists with the prerequisite training to deliver such treatments. Further, some health care staff perceive stigma related to mental health problems [20] and can be reluctant to undertake weekly psychotherapy given increased time demands in a pandemic. Little evidence exists for treatment effectiveness for people with ongoing trauma exposure, such as health care staff working in the pandemic [23,24]. Evidence-based treatment guidelines for PTSD [23,24] suggest that, when current treatments are lacking or ineffective, there may be utility in targeting single symptoms. We have proposed intrusive memories of trauma as a targeted single symptom [4,6]. For example, intrusive memories were reported by 65% of emergency nurses [12]. Drawing on cognitive science (eg, on mental imagery [25] and memory consolidation and reconsolidation), we have developed a behavioral intervention aimed to limit the occurrence or recurrence of intrusive memories.

The procedure takes approximately half an hour and is delivered according to a clear protocol and administered typically in one or two guided sessions (thereafter self-administered, if needed). Delivery currently requires guided support by someone trained

in the intervention but does not always require a fully qualified mental health professional. The intervention consists of several components including a brief memory reminder to moments within the trauma (hotspots), training in mental rotation, and engaging in a visuospatial cognitive task (the computer game Tetris) for a specific time while actively using mental rotation (ie, planning ahead and visualizing in the mind's eye how to rotate and move upcoming Tetris blocks to fit them into a horizontal line). The cognitive task aspect can be delivered on the participant's own smartphone. This intervention does not require a detailed discussion of the traumatic event.

The rationale underlying the intervention includes the following. Engaging in the visuospatial task is hypothesized to compete for limited working memory resources [25] with mental imagery (sensory) aspects of the trauma memory. This in turn is hypothesized to limit the storage or restorage [26] of the sensory representations of trauma [27] and reduces the subsequent number of times that memory intrudes involuntarily.

The intervention can be used on the same day the traumatic event occurred (day 1 protocol) [28-30]. For intrusive memories from older traumatic events [31,32], there are seemingly minor but important procedural differences including the type of memory reminder instructions and the time between the memory reminder and the task (see Visser et al [26]). In protocols for older memories, participants are instructed that they will briefly bring to mind the visual image from a specific intrusive memory and then play the computer game Tetris using mental rotation for at least 20 minutes. For a more detailed discussion regarding the intervention, see Iyadurai et al [4] and Singh et al [6].

Early studies have shown that the intervention may prevent intrusive memory occurrence in patients soon after trauma [29,30] (eg, by 62% compared to attention placebo control in motor vehicle accident survivors [29]). In a study with patients with more diverse trauma types in a Swedish emergency department, participants in the intervention condition reported 48% fewer intrusive memories compared to attention placebo control at week 1 following the intervention and 90% fewer at week 5 [28]. Promising results in terms of established intrusion reductions have also been shown in small-scale case series research with refugees [32], patients with complex PTSD [31], and in a person with bipolar disorder and PTSD [33].

To date the intervention has been delivered in person. So that it can be delivered remotely, we have first taken steps to digitalize the instructions for this brief intervention [34]. Here, participants (clinicians, researchers, students) were generally positive toward the materials created, noting that they were clear, concise, and helpful. In addition, participants shared potential concerns about remote delivery, mainly regarding the need for real-time communication with target users [34].

As there is currently no gold standard for a systematic stepwise approach to use for developing a successful intervention, it has been recommended to collect and apply relevant steps suggested from different methodologies, to use them with a "flexible" approach [35], and to iteratively make adjustments based on stakeholders' [36] and target users' input [37]. Thus, the critical next steps in our intervention development following our initial work with digitalizing the intervention [34] involves getting

feedback from target users. Therefore, we now piloted the intervention with digital study procedures and when adapted for the health care work context during a pandemic, and obtained target user feedback.

The aims of this mixed methods exploratory pilot study with health care staff with experience of working in the pandemic were to:

- Pilot the intervention procedures that we have digitalized (for remote delivery and with remote support) and adapted for this target population (health care staff working during a pandemic) to explore its ability to reduce the number of intrusive memories of trauma (primary outcome: week 5 diary post intervention) and improve related symptoms (eg, posttraumatic stress symptoms) and participant's perception of their functioning (eg, concentration)
- Explore the feasibility and acceptability of both the digitalized intervention procedures and the digitalized data collection (eg, primary outcome measure)

Methods

Participants

Participants (N=3, all female) were Swedish registered nurses, all with a university education and in full-time employment, who had worked clinically in the ICU or ambulance service during the COVID-19 pandemic in the spring of 2020 and were still currently working. They were all around 50 years of age and had >30 years of experience in their work. They had specialized in anesthesia and intensive care or ambulance care. To protect participant anonymity, most demographic characteristics have been omitted.

The inclusion criteria were being 18 years or older; doing clinical work during the COVID-19 pandemic in hospital care facilities (eg, ICU, intermediate care, or ward), experiencing at least one traumatic event in relation to their clinical work as health care staff during the COVID-19 pandemic that met Diagnostic and Statistical Manual of Mental Disorders (Fifth Edition; DSM-5) [3] Criterion A for PTSD within the last 3 months, reporting memory of the trauma, experiencing at least two intrusive memories of work-related traumatic events during the COVID-19 pandemic during the week before inclusion, willing and able to briefly write these down, being fluent in Swedish, being alert and oriented, having access to a smartphone, having sufficient physical mobility to use their smartphone, willing and able to provide informed consent, and completing study procedures and willing to be contacted during the study. Exclusion criteria (consistent with our previous study with patients) were current intoxication or loss of consciousness >5 minutes.

Measures and Materials

Participant and Traumatic Event Characteristics

Details of the traumatic events causing intrusive memories were assessed with a bespoke item ("Select which of the following category(s) best fits for the traumatic event(s) you have experienced during COVID-19 within the last three months and experience intrusive memories from") and then 11 categories

(including a category for *other* events) were presented. Event categories were based on existing literature on traumatic events in health care staff (eg, a traumatic or tragic death of a patient

[13]). Experience of prior psychological trauma was assessed with the Life Events Checklist for DSM-5 [38] (see Table 1).

Table 1. Characteristics of the worked-related traumatic events that participants had intrusive memories from and prior experiences of traumatic events reported per participant.

Variable	Participant ^a	Participant	Participant
Time(s) of trauma	Between 1 and 3 months ago	Between 1 and 3 months; within the last 24 hours	Between 1 and 3 months ago; ongoing exposure
Traumatic event causing intrusive memories	Traumatic or tragic patient death; situation in which patient care did not work as planned	Traumatic or tragic patient death	Situation in which patient care did not work as planned; confronted with distressed family members of patients; <i>other</i> category: relative who did not dare to say farewell to a patient who was critically ill
Retrospectively rated number of intrusions in the week before study participation ^b	2	3	3
Prior experiences of traumatic event types (LEC-5 ^{c,d})	6	12	10

^aParticipant numbers are omitted to preserve anonymity.

^bAssessed via a single item retrospective rating.

^cLEC-5: Life Events Checklist for Diagnostic and Statistical Manual of Mental Disorders (5th Edition).

^dA list of 17 traumatic event types. Here, we report the number of event types endorsed as “happened to me,” “witnessed it,” or “part of my job” [39].

Primary Outcome Measure

Number of Intrusive Memories of Trauma Post Intervention (Week 5)

The number of intrusive memories of the traumatic events was assessed with a digital adaptation of the pen-and-paper diary used in our previous work [28]. Instead of ticking a box for each intrusive memory during four time periods (morning, afternoon, evening, and night) on a paper diary, participants received four digital links per day (after each period of the day had passed) via SMS text message and email from the electronic platform SmartTrial [40] version 2020.1. Participants recorded their intrusive memories with this digital diary for 7 days starting 1 month after the intervention (ie, day 29 post intervention, week 5). In each link, they were asked how many intrusive memories they had during that time period (eg, in the morning) on a 9-point scale from 0 to *more than 7* (a follow-up question to specify the number appeared if they selected *more than 7*). The link also included a brief description of what intrusive memories are: “Intrusive memories are IMAGES from a traumatic event that pop suddenly into your mind, when you DO NOT WANT them to. (They are NOT the same as deliberately choosing to think about the event or thinking about it in words.) Please record EVERY intrusive memory you have had - even if it is the same one popping up several times. If you did not have any, please CHOOSE 0.” In addition to this brief description, participants had received more detailed instructions prior to commencing the diary (eg, an information video about the symptom *intrusive memories* and researcher support).

Secondary Outcome Measures (Including Intrusive Memory Measures)

Number of Intrusive Memories of Trauma Immediately Post Intervention (Week 1)

The number of intrusive memories of the traumatic events was also assessed in week 1 (starting after having received the intervention on day 1) in another identical 7-day digital diary.

Intrusion Questionnaire

The intrusion questionnaire (eg, [41]) was used to assess the frequency of intrusive or unwanted memories in the previous week (7-point scale from *never* to *many times a day*, with a follow-up question to specify the number if necessary) and the characteristics of the intrusive or unwanted memories (ie, distress,nowness, reliving, and disconnectedness) and whether different triggers are associated with the intrusive or unwanted memories of the traumatic events (101-point scale from 0 *not at all* to 100 *very strongly*). The retest reliability of the four scales assessing characteristics of intrusions ranges between 0.61 and 0.72 [42].

Impact of Event Scale-Revised: Posttrauma Intrusion and Avoidance Symptoms

The intrusion and avoidance subscales of the Impact of Event Scale-Revised (IES-R) [43] were used to assess the degree of subjective distress of posttrauma intrusion and avoidance symptoms. The IES-R shows high internal consistency ($\alpha=.96$) and agreement with other measures of posttraumatic stress (eg, PTSD checklist: $r=0.84$) [44].

Posttraumatic Stress Disorder Checklist for DSM-5

Participants' current symptoms of PTSD were assessed via the PTSD Checklist for DSM-5 (PCL-5) short version [45]. The PCL-5 short version accounts for 94.1% ($r=0.97$) of the variance in the original 20-item validated PCL-5 version [45] and has

been specifically recommended for remote digital assessment after trauma [46].

Distress and Vividness of Intrusive Trauma Memories During Diary Weeks

Two self-rated items assessed participants' level of distress and vividness associated with the intrusions (11-point scales from 0 *not at all* to 10 *extremely*). Ratings were collected within the diary at the end of week 1 and week 5.

Self-Rated Initial Intrusions (Baseline)

One item was used to assess how many intrusive memories the participant had experienced in the week prior to entering the study, from 2 to more than 7. If more than 7, a free-text response field to specify the number of intrusive memories was presented. This was followed by three self-rated items measuring the level of distress, vividness, or concentration disruption associated with the intrusions (11-point scales from 0 *not at all* to 10 *extremely*).

Other Prespecified Outcome Measures (Including Functioning Measures)

Self-Rated Concentration Disruption

Participants rated their perceived level of concentration disruption associated with intrusions with a bespoke item adapted from Holmes et al [7] (11-point scale from 0 *not at all* to 10 *extremely*).

Self-Rated Impact of Intrusive Memories on Functioning

A bespoke item was used to assess their perceived impact on daily functioning associated with the intrusions ("During the previous month how much did your intrusive memories of the traumatic event affect your functioning (social, occupational, or other important areas e.g. relationships with other people, work, parenting, schoolwork, housework, volunteer work etc.)?") [4] (11-point scale from 0 *not at all* to 10 *extremely*).

Credibility/Expectancy Questionnaire of Doing the Intervention

Before the intervention, participants completed the Credibility/Expectancy Questionnaire [47], which included 5 ratings of treatment expectancy measuring to what degree the participant finds the intervention credible (wording adapted for this study).

Subjective Units of Distress

Subjective units of distress (SUD) were collected three times during the intervention process to measure participants' level of distress (11-point scales from 0 *no distress* to 10 *worst possible distress*).

Coping

Participants perceived coping during the COVID-19 pandemic was assessed via two free-text response field questions (eg, are there any specific factors you think have made it more difficult

or easier for you to handle the COVID-19 situation and its consequences?).

Adverse Events

Adverse events [48] were assessed via a free-text response field asking about the occurrence of any health problems since the last contact.

Feedback Questionnaire About Participation

A feedback questionnaire consisting of nine bespoke items assessed participant's experience of study participation. Items included (eg, "How acceptable was it to do the task?"; 11-point scale from 0 *not at all* to 10 *extremely*) and questions about what has happened since the study with a yes or no response (eg, "Have you had any psychological or medical treatment since you did the task?") and items with a free-text response field (eg, "Do you have any other comments?").

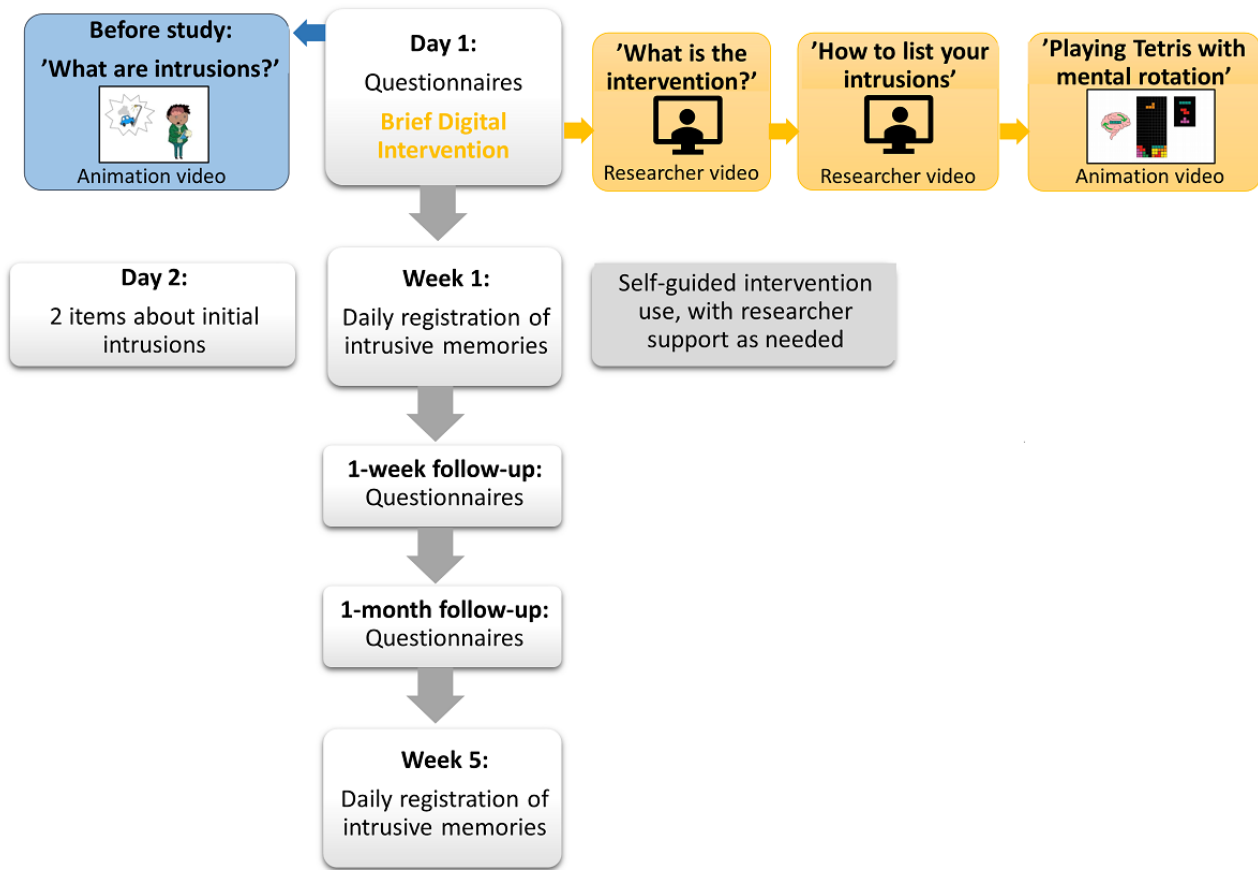
Procedure

Recruitment and Instructions for Study Procedure

Participants were all recruited from the professional network of the research team (author AR). Data collection occurred from July 8 to August 28, 2020. Participants received study information materials via email and postally, and provided their written and informed consent prior to study procedures. They were contacted by the researcher (author MK) to set up a time for a remote digital meeting using Zoom [49] (Zoom Communications Inc; premium university account). In the session, which was scheduled around their work, participants received instructions on how to complete measures and the intervention via an online platform, and help to complete the intervention as needed, and they were encouraged to give feedback on all aspects of study procedures.

The intervention was administered fully remotely and digitally. Remote researcher support during the procedure and delivery of the intervention was provided by a clinical psychologist with extensive experience in delivering the intervention to clinical participants (MK; see description of training in the Training to Deliver the Intervention section). During the remote meeting, participants received links to each step of the study (baseline measures; intervention package), and researchers (authors MK and KD) observed and took notes on feedback and questions that arose. The baseline and intervention session took approximately 1 hour in total (see Figure 1 for an overview of study procedures), depending on how much discussion arose. Of this, the intervention procedure itself took approximately 25 minutes. The overall approach here, with the intervention adapted for ongoing trauma exposure and self-use in a work context, was to deliver one intervention session with researcher support to teach participants how to do the intervention by targeting one selected intrusive memory and to promote continued use of the intervention to target any remaining or different intrusive memories and new intrusive memories should new traumatic events occur, with the option for researcher support in the initial week as needed.

Figure 1. Flow of study procedures. Note that, in contrast to the planned follow-up randomized controlled trial, for these pilot participants, participation ended after the week 5 diary was complete.



Intervention Procedure

The intervention procedure commenced after the completion of baseline measures (see Figure 1). Participants watched a video (researcher video 1: 1:51 minutes) of a trained researcher giving an overview of the three parts of this intervention, that is, (1) a *brief memory reminder* so the chosen intrusive image is held in working memory before (2) engaging in a visuospatial interference task (playing the computer game *Tetris* via tetris.com) for *at least 20 min*, and (3) actively using *mental rotation* (ie, planning ahead and visualizing in their mind's eye how to rotate and move upcoming Tetris blocks to fit them into a horizontal line). Next, they were instructed to briefly list their intrusive memories. They watched a video (researcher video 2: 2:03 minutes) on how to write this list and received written instructions in the platform. Instructions included "Please list your intrusive memories, using only a few words to describe what you 'see' when the intrusive memory pops up, e.g. 'patient on ventilator'." They were instructed to save this list (eg, screenshot) and then choose which memory to target in this session (eg, the most frequent or most distressing one). They were then instructed to "gently and briefly bring the chosen memory to mind" so they could *see it* in their mind's eye (ie, for it to become active in working memory prior to the visuospatial interference task).

Next, a video and written instructions followed on how to play Tetris with mental rotation and how to access the game (via Tetris.com) and adjust the necessary settings (*ghost piece* off;

animation video 3: 2:43 minutes). In this online version of Tetris, the game runs in *marathon mode*. Participants could choose to play on their smartphone or computer. Time between memory activation and game play was approximately 10 minutes (ie, the hypothesized time gap for memory to become malleable [26]). Prior to starting, they were reminded of their chosen memory and instructed to play for at least 20 minutes using mental rotation instructions and then to return to the platform.

Follow-up Procedures

After the intervention, participants were encouraged to keep using the intervention as needed and in a way that would fit with their work demands. They received instructions for completing an online daily diary registration on the number of intrusive memories over week 1 (see the Secondary Outcome Measures section). Incoming data was monitored approximately daily for the number of intrusive memories, and whenever intrusive memories were reported, participants had the option for *booster* doses, that is, they were reminded by the researcher (by SMS text message, email, or phone) to use the intervention themselves for the remaining intrusive memories and were given the option for researcher support (see the Results section).

At the 1-week and 1-month follow-ups, participants completed self-report questionnaires via the electronic platform. After the 1-month follow-up, they commenced with the primary outcome—daily registrations on the number of intrusive memories during week 5. Participants received an end-of-study letter that included a graph depicting their change in the number

of intrusions over time. They were invited to a phone call with a researcher (MK) for additional feedback on study procedures to help refine these prior to starting the main randomized controlled trial (RCT).

Training to Deliver the Intervention

The researcher delivering the intervention (MK) had received prior detailed training in delivering the intervention and how to obtain the primary outcome. This included theoretical and procedural knowledge; observing and being observed by a supervisor (author EAH) in role plays; receiving in-vivo and group supervision; and in this study, receiving regular ongoing supervision (from EAH) as necessary.

Data Analysis

Quantitative Analysis

We conducted an exploratory analysis of quantitative results using a descriptive approach. Analyses and graphs were performed in Excel Professional Plus Version 16.0.5065.1000 (Microsoft Corporation).

Qualitative Analysis

A flexible approach building on content analysis as outlined by Bengtsson [50] was used. We conducted a qualitative analysis based on the notes that were taken by the researchers (authors MK and KD) during observing and discussing with participants in the digital intervention session with the researchers, from researcher-participant contact during the follow-up period, and open-ended questions in the electronic platform. The feedback was then organized based on emerging themes as an iterative process among the researchers.

Participant feedback was systematically evaluated to determine whether changes should be implemented, following the suggested person-based approach created by Yardley et al [51] and Bradbury et al [52]. This is a methodological approach for developing digital interventions (see also Gamble et al [34] for a detailed description of how this was carried out, including a description of how we applied the must have, should have, could have, won't have [MoSCoW] method for prioritization [52,53]).

Ethical Approval

This study was approved by the Swedish Ethical Review Board before the start of the study: 2020-03085. All participants provided their written and informed consent in accordance with the guidelines of the Declaration of Helsinki; signed consent

forms were returned digitally and on paper. These pilot cases were done as a precursor to the planned RCT registered under ClinicalTrials.gov NCT04460014 (July 7, 2020).

Results

Characteristics of Traumatic Event Exposure and of the Intrusive Memories

Two participants experienced one recurring visual intrusive memory of a work-related traumatic event, and 1 experienced three separate recurring intrusive memories (ie, different visual scenes). Examples of these intrusive memories' content included an image of patient's faces, image of patient's relatives, or image of a dead person (see Table 1 for additional information about the index traumatic events associated with these intrusions).

Aim 1: Pilot the Digitalized and Adapted Intervention

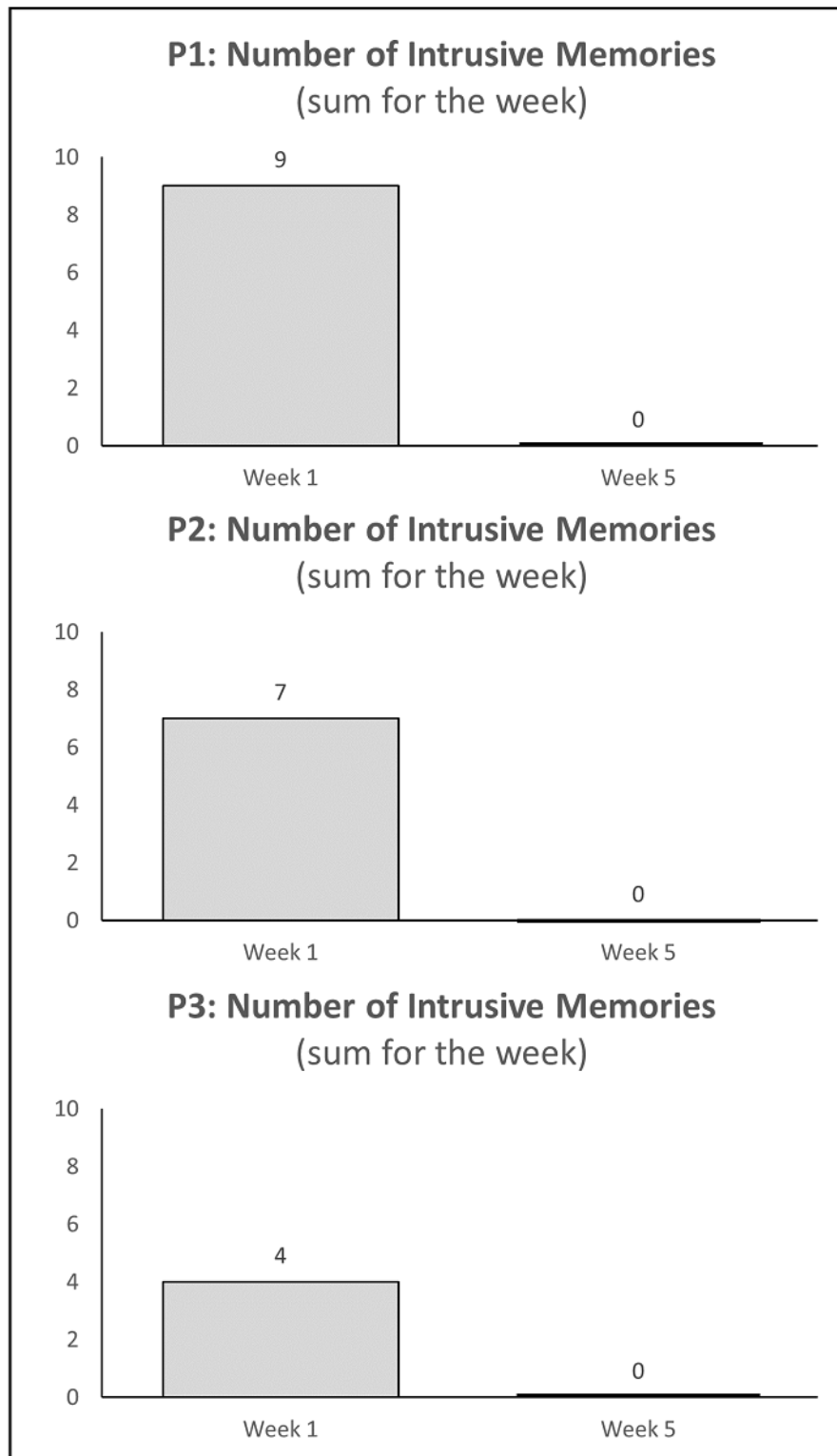
All 3 participants completed all study procedures. In the past, we had used paper diaries, and here, the daily digital data capture of the number of intrusive memories via an electronic link (SMS text message or email) worked well (ie, the digital diary was completed by all 3 participants, and the feedback was positive). Our 3 participants answered 100% of the intrusive memory diary links, which were sent out four times a day for 7 days in a row, twice (primary outcome: week 5; secondary outcome: week 1 post intervention). The digital data capture for other questionnaires (secondary outcomes) was also successful (ie, questionnaires were completed by participants), with only two ratings missing for 1 participant (intrusion questionnaire: unwanted memory frequency and concentration disruption on day 2).

Primary Outcome

Number of Intrusive Memories of Trauma Post Intervention (Week 5)

In contrast to baseline and week 1 postintervention levels (see the Secondary Outcomes: Number of Intrusive Memories [Week 1] section), all 3 participants reported 0 intrusive memories throughout the week 5 diary (Figure 2). Self-reported diary accuracy ratings were high (10, 6, and 10 out of 10). During discussion with the research team, the participant who rated a 6 explained that she was unsure if she should have also noted having a *normal* (ie, not intrusive) memory of the event when she no longer experienced intrusions.

Figure 2. Graphs for visual inspection of the number of intrusive memories of trauma (sum for the week) during week 1 (secondary outcome) and week 5 (primary outcome) following the initial intervention session for each participant. After learning how to do the intervention (guided by the researcher) and practicing it independently during week 1, all 3 participants reported zero intrusions by week 5.

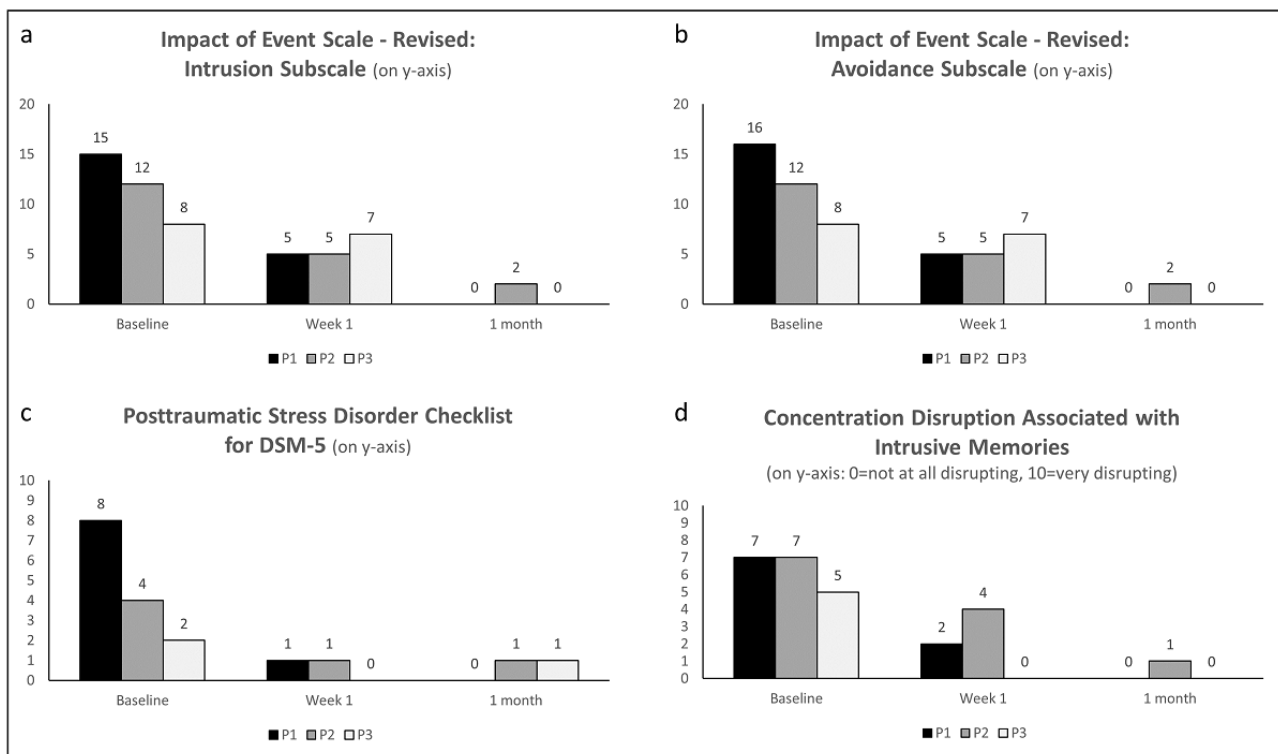


Secondary Outcomes

Multimedia Appendix 1 Table S1 shows secondary outcome data for each participant at all assessed time points, and Figure

3 displays bar graphs showing examples of secondary and other outcomes.

Figure 3. Bar graphs showing examples of intrusion-related symptoms (posttraumatic stress) and participant's perception of their functioning in terms of concentration disruption (all secondary and other outcomes). All 3 participants reported reductions in intrusion (a) and avoidance (b) symptoms on the Impact of Event Scale-Revised from baseline to week 1 and the 1-month follow-up. Participant's total scores on the Posttraumatic Stress Disorder Checklist for DSM-5 (c) decreased from baseline to week 1 and the 1-month follow-up. All 3 participants reported that their perceived concentration disruption associated with having an intrusive memory (d) decreased from baseline to week 1 and the 1-month follow-up. DSM-5: Diagnostic and Statistical Manual of Mental Disorders (5th Edition).



Number of Intrusive Memories of Trauma Immediately Post Intervention (Week 1)

In week 1, participants reported 9 (P1), 7 (P2), and 4 (P3) intrusive memories in total. Self-reported diary accuracy ratings were high for P1 and P3 (10/10), and medium for P2 (4/10).

We discuss lessons learned during the first week on a case-by-case basis. For all 3 participants, the intervention was delivered once with researcher support (approximately 25 minutes), and they subsequently used the intervention on their smartphone on their own. Participants received encouragement about intervention use from the researchers via 1 to 4 SMS text messages or emails (all participants) and one phone call (1 participant).

P1 reported 1 intrusive memory in the first half of the week and then a spike in intrusive memories on day 5 (7 intrusive memories). In communication with the researchers, she described that on this day she experienced an event at work similar to the one represented in her intrusive memory, which triggered these intrusions. She felt too tired to use the intervention straight after that work shift but successfully used it for that intrusion the day after.

P2 completed the game play part of the intervention on her computer on day 1. She reported 4 intrusive memories on the same day and 3 more on day 2. On day 2, she was instructed over the phone to repeat the intervention and encouraged by researchers to use her smartphone instead and to bring the image

to mind before the game play. From day 3 to day 7 she reported no more intrusive memories.

P3 reported 2 intrusive memories on day 1 and 1 each on day 2 and day 3. From day 4 to day 7 she was intrusion-free, thus showing a general decrease of intrusion frequency over the week. She kept using the intervention during week 1 and wrote that she was “incredibly impressed by how such a simple thing could make such a huge difference.”

Intrusion Questionnaire

From baseline to week 1, the frequency of intrusive or unwanted memories of the traumatic event during the previous week slightly decreased from *twice a week* for 2 participants and remained the same (*twice a week*) for 1 participant. It then decreased to *never* after 1 month for all 3 participants (in line with the electronic diary data). The characteristics of the intrusive or unwanted memories—distress,nowness, and reliving—decreased from baseline to week 1 for all 3 participants. P1 and P2 reported a decrease in disconnectedness and in different triggers associated with the intrusive or unwanted memories from baseline to week 1, whereas P3 reported an increase from baseline to week 1. At 1 month, none of our 3 participants reported intrusions anymore, so characteristics were not rated.

Impact of Event Scale-Revised: Posttrauma Intrusion and Avoidance Symptoms

From baseline to 1 month, participants' scores on the intrusion subscale decreased (from a range of 8-15 to a range of 0-2);

similarly on the avoidance subscale, all scores decreased to 0 after 1 month (Figure 3a and b).

Posttraumatic Stress Disorder Checklist for DSM-5

From baseline to week 1, total scores on the PCL-5 decreased for all 3 participants and remained low at 1 month (range 0-1; Figure 3c). A similar pattern was shown in the intrusion subscale. P1 showed a decrease from 5 at baseline to 0 at 1 month. P2 and P3 showed a slight decrease from 2 at baseline to 1 at 1 month. Participants' scores on the remaining subscales were already very low at baseline (0 or 1, with the exception of P1's score of 3 on the avoidance subscale) and remained at or decreased to a very low level at 1 month.

Distress and Vividness of Intrusive Trauma Memories During Baseline and Diary Weeks

From baseline to week 1, vividness and distress associated with the intrusive memories decreased. P1 showed a strong decrease of vividness (7 to 3) and distress (7 to 2) from baseline to week 1, P2 showed a slight decrease from baseline to week 1 (vividness 10 to 9, distress 5 to 4), and P3 showed a strong decrease of vividness (10 to 2) and a slight decrease in distress (1 to 0) from baseline to week 1. At 1 month, none of our 3 participants reported intrusions anymore, thus related vividness or distress was not rated.

Other Prespecified Outcomes

Here we present outcomes regarding functioning and intrusive memories as used in our previous work [4,7,28]. We also present ratings of credibility and expectancy for the intervention, SUD experienced during the intervention procedure, coping, and adverse events (see Multimedia Appendix 1 Table S2 for details and remaining prespecified outcome data for each participant at all assessed time points).

Self-Rated Concentration Disruption

At baseline, concentration disruption associated with having an intrusive memory was medium to high (5 and 7 out of 10) and reported as lasting for 1 to 5 minutes per intrusive memory. At 1 month, concentration disruption had decreased to 0 or 1 out of 10 (Figure 3d). At week 5, none of our 3 participants reported intrusions anymore and thus did not rate associated concentration disruption.

Self-Rated Impact of Intrusive Memories on Functioning

From baseline to 1 month post intervention, 2 participants reported that the impact of intrusive memories on functioning decreased (from 5 [P1] or 2 [P3] out of 10 at baseline to 0 at 1 month), and it remained low for P2 (1 out of 10 at baseline and 1 month).

Credibility/Expectancy Questionnaire of Doing the Intervention

Credibility ratings taken after a brief description of the intervention prior to engaging in it were low to midrange for all 3 participants (38 out of 50 by P1, 35 by P2, and 21 by P3), and this was also reflected in qualitative feedback (eg, P3 wrote at the 1-month follow-up, "...[I] didn't believe at all that this would work!").

Subjective Units of Distress

During the intervention session with the researcher, participants completed distress ratings (SUD) before and after describing their intrusive memories, and after playing Tetris. For 2 participants, distress increased (P1 from 2 to 6; P3 from 0 to 1) but did not reach ceiling after describing and choosing an intrusion to target with the intervention, indicating successful emotional memory activation (note, P2's ratings decreased from 2 to 0). Critically, after game play, distress decreased or remained at zero for all 3 participants (P1: 1; P2: 0; P3: 0).

Coping

In the end of the intervention session, we asked participants if any specific situations or factors made it difficult to cope with the COVID-19 situation at work and which made it easier to cope. They noted that, for example, support from the employer and help and support from colleagues made it easier to cope, and a lack of knowledge about COVID-19, lack of competence, difficulties in accessing and using personal protective equipment, and not being able to give person-centered care made it more difficult.

Adverse Events

There were no reported adverse events within the platform at week 1 and the 1-month follow-up.

Aim 2: Feasibility and Acceptability of the Digitalized Intervention

Intervention

The digitalized intervention (approximately 25 minutes) was delivered with the option for researcher support (via video call). There was no dropout in our 3 participants. The session with the researcher and subsequent use of the intervention was flexibly scheduled into participant's daily work and life. After this initial contact, limited researcher support was needed and none of the 3 participants requested additional support on how to use the intervention. Only 1 participant received a phone call on day 2; for the others, only brief encouragement to use the intervention via SMS text message or email was used.

Outcome Data Capture

All primary and secondary outcome data was successfully collected digitally or remotely. Digital procedures (eg, SMS text message) for reminding participants to fill in the daily registration on intrusions were used if needed.

Feedback Questionnaire About Participation

In a feedback questionnaire completed at 1 month, the intervention task was rated by participants 1 and 3 to be easy, and low in terms of how upsetting it was (see Table 2). One participant (P2) rated the task itself low in terms of easiness, noting later that the Tetris game play component was not easy. All 3 participants rated the intervention task to be acceptable. All 3 participants had used the intervention several times on their own after the initial session with the researcher. For instance, P1 reported that the work context triggered her intrusive memories and found it helpful to be able to use the intervention on her own in this setting. Participants had all mentioned the intervention to others, for example, to their

friends or colleagues. None of the 3 participants reported receiving any other treatment since they took part in the study.

Critically, all 3 indicated that they would highly recommend the intervention task to friends or colleagues.

Table 2. Responses on feedback questionnaire about participation from each participant at 1 month.

Feedback questionnaire about participation	P1	P2	P3
"If a friend or colleague has gone through a similar event, how likely is it that you would recommend this task?" ^a	10	10	10
"How easy did you find it to do the task?" ^{a,b}	9	1	10
"How upsetting did you find it to do the task?" ^b	1	0	0
"How much did you appreciate having something to do?" ^b	9	1	2
"How acceptable did you find the task?" ^b	9	10	10
"Received other treatment due to the traumatic event" ^c	No	No	No
"Did task on their own (i.e., after the researcher-led session)" ^c	Yes	Yes	Yes
"How often?" ^c	3-4	8	Daily for the first 2 weeks, after that every now and then
"Have you talked to others (e.g. friends, colleagues) about the task?" ^d	Yes	Yes	Yes

^aRated on an 11-point scale from 0 (not at all) to 10 (extremely/very much).

^bIn this study, *task* refers to the intervention as a whole. The wording in this questionnaire is adapted for use in the randomized controlled trial, where *task* can refer to either intervention or control.

^cFree-text response field.

^dYes/no response.

Open-ended Feedback From Participants

In the open-ended feedback collected within the platform, the 3 participants expressed overall positive feedback about their outcomes after the intervention, that their experience with using the platform was overall good, and that they found the instructions and videos to be clear and helpful. P3 wrote as part of the open-ended feedback question given at the 1-month follow-up that the intervention was "excellent" in her case. P1 wrote in the platform both at the 1-week and 1-month follow-ups that she had not experienced any more intrusive memories since she had used the intervention on her own in the beginning of week 1. She noted that even when choosing to talk about the event of the memory with her colleagues she no longer experienced intrusive memories. Additionally, when contacted for a follow-up via phone call, participants reported that seeing their results gave them a sense of self-achievement (P1) and that they have learned a tool that they can use in future situations (P3).

In terms of impact on work functioning, P1 reported in the follow-up phone call that when she had intrusive memories it affected her empathy and her interactions with patients and relatives: "I did not dare to let them in, I was afraid something similar would happen and I would get more intrusive memories."

Analysis of Themes in the Open-ended Feedback From Participants

We systematically analyzed participant feedback for changes or improvement, and three overarching themes emerged:

1. Getting the data right: This refers to minor changes in data collection (eg, clarifying the wording to some questions,

2. removing items that were poorly understood, and adding a baseline diary; ensure quality of data).
2. Doing the digital intervention right: This refers to any changes in instructions, adding more information to further emphasize the important aspects, and for tech or intervention procedure (ensure intervention fidelity).
3. Feeling that study participation is alright: This refers to tailoring study procedures (eg, adding examples of staff categories in questions or examples) and adapting data collection as necessary (ensure participants feel included, can relate, and that participant burden is minimized).

Changes Made to the Study Materials Based on Feedback

Based on participant feedback, 18 changes were made (see [Multimedia Appendix 1](#), Table S3 for a complete list). Eight were categorized under the theme "Getting the data right," 7 under the theme "Doing the intervention right," and 2 under "Feeling that participation is alright." Based on the MoSCoW priority [52,53], 9 of the implemented changes were labeled as being a "Must Have," such as adding "colleagues" as an example for the question on sources of social support, since health care professionals must abide by patient confidentiality rules, thus not being able to seek social support for work-related incidents in friends and family. Four points of feedback were labeled as being a "Should Have" and 6 as a "Could Have."

Some changes that were suggested were not implemented, such as altering the original or translated wording of an official or pre-established questionnaire. [Table 3](#) shows three selected examples of the implemented changes. For example, row 3 shows that P1 stated that it was difficult to recollect how many intrusions she experienced in the previous week. Based on this feedback, we proposed to our pilot participants as part of our

follow-up phone call that future participants could complete a baseline diary that would begin at the time of enrollment. Pilot participants agreed that this would be a good addition to the study, and all reacted positively (eg, that it would provide a better baseline value of how many intrusions one experiences before compared to after the intervention), and they did not feel that adding an additional diary would be burdensome for participants. Thus, participants will only have to recall how many intrusions they have had within the last few hours and provide more accurate intrusion counts. This feedback was categorized as belonging to the theme “Getting the data right,” as this relates to participants’ being able to more accurately report the number of intrusions they have experienced.

Further, for the daily registration of intrusive memories, 1 participant said that it was easy to press the incorrect option indicating how many intrusions she had (eg, P1 accidentally pressed “1” when she meant to press “0” for the week 5 diary). Overall, even though the digital intrusive memory registration was said to be a feasible and acceptable data collection method, there is a need for researchers to closely monitor and check in to verify the accuracy, as a participant can easily select the wrong option without being able to edit their response. Furthermore, we decided to take out the question about “diary accuracy,” as low responses given by some of the participants were due to confusion on what this really meant (ie, how precisely did they count their work-related intrusions or how well or how swiftly they entered their data into the platform after having received the link).

Table 3. Examples of changes made to study materials and procedure following participant feedback [54].

Theme	Participant feedback as expressed by one or more participants	Changes made to the study based on feedback
Doing the digital intervention right	It is unclear what the next step is after filling out the baseline questionnaires, and after completing the assigned task. “ <i>What happens now?</i> ” (P1)	Added flowcharts that show the participant journey at the beginning and end of each module link that is sent out. It includes a green arrow with the text “You are here” to demonstrate where they are in the study and the next step.
Getting the data right	It is difficult to remember exactly how many intrusions one had in the last week (for the retrospective rating at baseline). (P1)	We kept this question as part of inclusion criteria and ask for a general estimate (ie, “have you had at least 2 intrusive memories in the last week”), but also added a baseline (week -1) daily electronic registration/diary of intrusive memories that participants are asked to complete during the week prior to filling in the baseline questionnaires and completing the intervention/control session.
Feeling that participation is alright	It looks like there are going to be a lot of questions in the SmartTrial list of questionnaires. [You] could lump some of the single-item, or shorter questionnaires together. (P3)	We did this with four questionnaires, where we combined two of them together twice. We also removed one of the work-related questionnaires.

Discussion

Summary of Findings

In this study with health care staff (3 registered nurses) who experienced intrusive memories in the context of the COVID-19 pandemic, we piloted a simple and brief intervention (here, approximately 25 minutes) with digitalized study procedures (for remote delivery and with remote support) that was adapted for this population (health care staff working during a pandemic). We explored whether the intervention reduced the number of intrusive memories of trauma (primary outcome) and improved related symptoms and participant’s perception of their functioning. All 3 participants reported a reduction in their intrusive memories (to zero) at week 5 post intervention, a greater reduction than expected; though, this has to be interpreted with caution given the small sample size. They also reported a reduction in other related symptoms—posttraumatic stress symptoms on the IES-R intrusion and avoidance subscales and the PCL-5. Of interest, all 3 participants perceived that their functioning improved (eg, they reported less disruption of their concentration). For example, P1 described that having intrusive memories affected her empathy and her interactions with patients and their relatives. This highlights that a reduction of

intrusive memories might be beneficial for health care staff and for their patients and care.

We also explored the feasibility and acceptability of both the digitalized intervention and digitalized data collection. Participants perceived the intervention as feasible and rated it as acceptable (ie, they rated acceptability as 9, 10, and 10 out of 10). There was no dropouts (ie, all 3 participants completed the intervention), indicating favorable acceptability according to National Institute of Health and Care Excellence guidelines [23]; though, the small sample size is again noted. The primary outcome measure—a daily digital diary to capture intrusions—was successfully completed with zero missing data. Further, the secondary outcome measures (including another daily digital diary and digital questionnaires) were successfully collected remotely. Participants were overall positive about taking part in the study, the reduction in intrusion and other symptoms, and their perceived improvement in functioning (see the Open-ended Feedback From Participants section). They were also positive about the intervention as a useful tool that could be used in work and daily life. Our analysis of participant feedback led to changes in study materials and procedures in preparation for a subsequent RCT (see the Adjustments Prior to the RCT section).

Of particular interest is the notion that participants readily engaged in repeated self-use of the intervention to reduce remaining intrusive memories, those not targeted in the session with the researcher, or when their work triggered intrusive memories. This indicates high levels of user engagement with the intervention [55]. They described the intervention as a tool they could use in future situations and that they felt empowered by it. This is particularly important given the ongoing pandemic, which exposes health care staff working with patients with COVID-19 to potentially repeated traumatic events over sustained periods of time. The fact that this intervention is brief (only about 25 minutes) and can be used flexibly (eg, around shift work), on one's own device, and does not require attending scheduled sessions with a mental health specialist is key for adoption given the high workload in this population. It also requires little input from the researchers (eg, one initial guided session, then little encouragement via SMS text message or email and only one phone call was needed).

To support health care staff during the pandemic and beyond, and to address broader critical challenges of reaching people affected by trauma at scale [56], interventions need to be not only simple and swift to deliver remotely but also repeatable. Clinical trials to test the effectiveness of this intervention are needed and should be adapted as necessary to different settings and trauma groups. If the intervention approach were proven effective, it may hold some useful features for implementation. For example, given that the intervention is brief and once learned can be self-administered, it could be used again (in effect as a booster session) following new trauma exposure. Thus, it may be particularly suitable for health care staff facing repeated or ongoing exposure to traumatic events in their lines of work and could be one of several possible measures to promote sustainability at work and well-being [57]. Psychological interventions to reduce trauma symptoms during ongoing exposure to trauma (eg, war or so called "frontline" work during a pandemic [1]) are currently lacking [24].

We note the number of intrusions reported at baseline was 2 to 3 per week, which is indicative of 8 to 12 per month. This may sound like a low number, but to the contrary, such levels can potentially cause significant distress and functional impairment; the gold standard assessment tool for PTSD (Clinical Administered PTSD Scale for DSM-5) states that an intrusion frequency of *only two intrusive memories per month* indicates a severity rating at the "moderate/threshold" level [58]. The intrusion reduction reported by our participants was to zero and perhaps reduced more than we might have expected given previous studies (eg, intrusion reduction in intervention condition vs control of 62% [29] and 48% [28,30] during week 1 and 90% during week 5 post intervention [28]). Therefore, we interpreted these results with caution given the small sample size and design used here. More generally, a symptom reduction of 50% may represent a clinically significant change, and such a drop in the number of intrusions might be a goal for this intervention approach rather than expecting total elimination.

In terms of perception of functioning, we note that participant's self-rated concentration improved as their intrusions reduced. This is important as participants reported at baseline that their concentration was disrupted for approximately 1 to 5 minutes

each time they experienced an intrusion. Why might having an intrusion, which disrupts concentration, at times lead to a problem of functioning at work? First, having even a brief but sudden and unplanned lapse in concentration after an intrusion can have the potential to interfere with the type of work-related duties that require focused attention, such as monitoring patients who are critically ill on a respirator. In our previous work, we have seen indication that intrusive memories can have a significant impact on people's perception of their ability to concentrate [7,32]. Second, when the content of intrusive memories is of work-related traumatic events (eg, a difficult resuscitation, a typical situation reported by nurses as cause for intrusions [12]), the triggers for these intrusions are typically in work-related environments. This means that health care staff might be more likely to find work settings in which psychological trauma has occurred difficult, as those settings can trigger distressing memories, which could potentially lead to increased absentee rates due to avoidance of such work and reminders. The idea that having intrusions can lead to avoiding triggers of those intrusions is illustrated by one of our participants that related experiencing intrusions to their negative anticipations about going to work. Third, the content of reported intrusions can have a direct link to the ability to function on tasks at work that are similar to the intrusion content. That is, say the intrusive memory content is an image of a patient's face and a tube, then this intrusive memory can be specifically triggered by other patient's faces and tubes, rendering maximum disruption just at those specific situations at work where there is patient contact or fitting of a tube. By alleviating intrusive memories with specific content, the participant potentially regains or improves their ability to perform work tasks related to that content. For example, as described by our participants, the content of an intrusive memory (eg, including an image of a specific patient) from a traumatic situation involving a patient had a negative effect on their interactions with new patients (eg, perceived reduced capacity for providing empathic person-centered care). Thus, not having intrusive memories might also aid functioning at work. Were this to be the case, intrusion reduction techniques could be one among many strategies that will be needed to support staff and ultimately help reduce stress, burnout [2], and staff turnover [20].

To take steps that might help prevent burnout among health care staff is of high importance due to the fact that even before COVID-19 there was a global shortage of nurses [59]. The World Health Organization states that to address the nursing shortage by 2030, both educational efforts to increase the total number of nurse graduates together with an improved capacity to both employ and retain nurses in the health care system is necessary. Units such as high dependency departments are stressful environments and can lead to an even greater incidence of burnout [60], which might be further exacerbated during the COVID-19 pandemic [61,62]. Therefore, interventions that mitigate the impact of traumatic events are vital.

Limitations

We noted several limitations of the current mixed methods exploratory pilot study. First and most important, since our study included a small sample (N=3), any changes in

quantitative data reported by our 3 participants have to be interpreted with great caution.

Second, as a precursor to a planned RCT, we did not preregister this pilot study separately but did preregister the RCT before commencing this pilot (ClinicalTrials.gov NCT04460014, July 7, 2020), which reflects the outcomes reported here.

Third, we piloted only our intervention procedures in this study and can therefore only interpret reported intrusive memory data (primary outcome) compared to baseline assessment, not compared to a placebo control group. We will use an attention placebo control task, which we have piloted previously [28], in our planned RCT.

Fourth, we assessed the number of intrusive memories in the week prior to the intervention with a single time point retrospective rating. However, participants reported that it was difficult to remember how many intrusive memories they had during the previous week, posing a risk of the numbers we obtained through retrospective ratings being an under- or overestimation of the actual number of intrusions at baseline and making it difficult to compare numbers obtained through a retrospective rating at baseline with numbers obtained through a daily diary during week 1 and week 5. Following this limitation in this pilot study and because participants were positive toward completing an additional diary to monitor their intrusive memories at baseline, we will add a baseline diary in the planned RCT.

Strengths

Strengths of this study include our use of different methodologies as part of stepwise intervention development [35] and including relevant organization stakeholders (end users with lived experience of health care staff working with patients with COVID-19 during the pandemic) as part of the development and implementation process [36]. Exploring the use of the digitalized intervention using a co-design approach together with the target user's feedback has paved the way for optimizing intervention delivery for the next steps. Following Bird et al's [63] checklist for feasibility of mental health interventions, we note that this first digitalized form of the intervention—once proven effective in a clinical trial—holds promise as a feasible intervention as it is cost saving, not time consuming, simple, applicable to the population of interest, matches prioritized goals, and no adverse events have hitherto been reported.

Adjustments Prior to the RCT

Based on the target user feedback and lessons learned from this pilot study, we will make the following adjustments to the main RCT. In addition to adding a diary assessing the number of intrusive memories during the week before the intervention, we will also make changes to the diary itself. We will remove the accuracy rating at the end of the diary we had used in this pilot based on our previous pen-and-paper diary because it was unclear for our participants how to rate the item in relation to the digital diary. By collecting the diary data electronically (four times a day), we automatically obtain data on how compliant participants completed it (ie, how often they recorded their intrusive memories) in a more objective way compared to a

subjective rating scale. Thus, we will use this data to report diary accuracy in the main RCT. Furthermore, being at work may trigger intrusive memories in this population of health care staff whose work also represents the context in which the traumatic events occurred. Therefore, we will add two ratings assessing the number of days at work and the number of night shifts in the previous week at the end of each diary. These ratings may help the researcher understand patterns of intrusions reported by participants (eg, high frequency of intrusions at night but none during the day could be because a person works night shifts).

We will also adapt the item assessing perceived functional impairment at baseline and follow-ups. Because participants found it difficult to answer our question on perceived functional impairment (since it included many different areas of functioning) and because our questionnaire measures of work-related stress did not seem to capture impairment at work due to intrusive memories very well, we will split this item into perceived work-related functional impairment (rating and an option for a free-text response on how work-related functioning was impaired because of having intrusive memories) and perceived functional impairment in other areas (eg, relationships with other people, social life, schoolwork, housework, or volunteer work). When asked in a phone call after week 5, participants were positive to this change. They also provided examples of perceived work-related functional impairment (see the Clinical Implications section).

We will also remove questionnaires on work functioning that did not reveal any effects in this pilot study to reduce participant burden. For example, a questionnaire assessing intention to leave the profession or workplace will be removed completely, and a questionnaire assessing burnout symptoms will only be assessed at baseline and the 6-month follow-up in the main RCT, since the effects of having intrusive memories on burnout might only unfold after longer time periods. Additionally, we will slightly change the wording of the intrusion questionnaire [41] from *unwanted memories* to *intrusive memories* to keep wording consistent across different measures in our study and because pilot participants did not understand the questionnaire as referring to intrusive memories (ie, the main focus of the study). Further, we will take out one irrelevant item (ie, a feedback question about appreciating having something to do, which we have used with emergency department patients but which does not make sense for this population), and we will use a modified version of the Credibility/Expectancy Questionnaire in the main RCT.

There was also some feedback from our participants we could not implement because of the limitations of the electronic data collection platform we are using in this study. For instance, we could not embed instruction videos directly in the electronic platform. Instead, participants had to click on a link directing them to a separate site on which we host the videos, which impaired study flow for participants. We also had to use a freely available version of *Tetris* (via tetris.com with the appropriate settings; see the Methods and Materials section), which included a 5 second commercial before being able to play, again impairing study flow. However, none of these issues were essential for intervention delivery or acceptability to our pilot

participants. Finally, we learned from P2 that playing the game on the computer made it difficult to achieve flow. In the RCT, we will strongly encourage smartphone use for the game play task.

Clinical Implications

As illustrated with participants P1 and P3, the recurrence of one specific intrusive memory (eg, recurrent images of a specific patient's face) can directly impact related emotion and behavior (eg, having intrusions was reported as having an impact on the capacity to provide empathic patient-centered care or experiencing intrusions was related to negative anticipations about going to work). The perceived work-related impact from having intrusions described by P1 (see the Open-ended Feedback From Participants section) reflects a typical avoidance circle commonly seen in PTSD where similar stimuli as the trauma stimuli (here patients) triggered subtle avoidance behavior for P1. For P3, intrusive memories made her feel reluctant about going to work and worried that something similar would happen. P3 noted that prior to the intervention it felt as if her mental "backpack was totally full" and that experiencing intrusive memories warned her that she might experience more events at work. She was worried that encountering similar events would cause additional intrusions, that is, intrusive imagery in only *one* intrusive memory (from an event not considered particularly traumatic in itself, according to P3) brought on a sense of current threat and a strong negative emotional response [64,65]. P2 described that, after the intervention, the intrusive memories were just "normal memories."

In this study, the intervention was delivered by a researcher who had been trained and had extensive experience on delivering the intervention with clinical participants (MK) and who received additional supervision as needed (from EAH). As training and supervision are currently deemed crucial to deliver the intervention adequately, we note that next steps in addition to the RCT include creating more standardized training material that can be delivered remotely (eg, in the form of an online training course) to researchers and procedures for role play, observation and feedback, and *in vivo* supervision.

Further, in this study, participants were recruited from within the professional network of the researchers. A potential barrier

for recruiting a large sample for an RCT could be the heavy burden placed on health care staff during the pandemic, which might prevent them from engaging in a research study. Thus, recruitment strategies for an RCT need careful consideration and optimization.

Conclusion

In this mixed method exploratory pilot study, we piloted a brief intervention that we have digitalized (for remote delivery and with remote support) and adapted for this target population (health care staff working during a pandemic). We explored whether the intervention reduced the number of intrusive memories of trauma (primary outcome) a month after the intervention session and improved related symptoms and participant's perception of their functioning. Our participants (3 registered nurses who had experienced intrusive memories in the context of their work during the COVID-19 pandemic) reported zero intrusive memories at week 5 post intervention, a greater reduction than expected, which has to be interpreted with caution given the small sample size. Participants also reported a reduction in other related clinical symptoms (eg, posttraumatic stress symptoms) and perceived that their functioning improved (eg, they reported less concentration disruption).

We also explored the feasibility and acceptability of both the digitalized intervention and the digitalized data collection. After an initial session with the researcher, all 3 participants continued to use the intervention independently and perceived it as feasible and acceptable. The primary outcome measure (a daily digital intrusion diary) and secondary outcome measures (eg, digital questionnaires) were successfully completed. Should the intervention prove effective in future clinical trials, it could hold promise as a helpful tool, which because of its brevity and digital format, could be incorporated into health care staff's work life and be repeated as needed in the face of numerous intrusive memories or ongoing trauma exposure. One next step will be to investigate the efficacy of the intervention in reducing the number of intrusive memories in an RCT with health care staff with exposure to traumatic events during the COVID-19 pandemic.

Acknowledgments

We thank Evelina Kontio for support with the formatting of tables; Katy Metcalf for her animation work; Beau Gamble for a discussion around guiding principles; Lalitha Iyadurai for helpful input in the preparation stages of, for example, adapting study materials for health care staff; and Michelle Moulds for input in moving from this pilot study to the RCT.

This project was mainly supported by grants to EAH from the Swedish Research Council (2017-00957 and 2020-00873), the Oak Foundation (OKAY-18-442), AFA Insurance (200342), and the Lupina Foundation. LS was supported by a Swiss National Science Foundation grant (P2BEP1_184378) and a Thunberg Fellowship by the Swedish Collegium for Advanced Study. MK was supported by a FO Medical Psychology grant from Karolinska Institutet and Karolinska University Hospital. KEG was supported by a grant from FO Akut, Karolinska University Hospital. AR was supported by a grant from AFA Insurance (200311). Funders were not involved in the study design, collection, analysis and interpretation of data; writing of the paper; or decision to submit for publication.

Authors' Contributions

LS contributed to study preparation; was responsible for data monitoring and analysis, tables, and graphs; and wrote the first draft of the manuscript with MK.

MK contributed to study preparation; was responsible for intervention delivery, data collection, and qualitative data analysis; and wrote the first draft of the manuscript with LS.

KD provided assistance during data collection and contributed to qualitative data analysis and write up under supervision.

ACF, VL, OD, KEG, and AR all provided nursing expertise during study preparation and throughout recruitment, data collection, and write up.

EAH conceived the study, including study design, interpretation of data, and write up, and supervised the study.

Conflicts of Interest

EAH reports serving on the board of trustees of the charity MQ: Transforming Mental Health but receives no remuneration for this role. EAH receives royalties from books and occasional fees for workshops and invited addresses; she receives occasional consultancy fees from the Swedish agency for health technology assessment and assessment of social services. The other authors have no conflicts of interest to declare.

Multimedia Appendix 1

Supplementary materials.

[[DOCX File, 55 KB - formative_v5i5e27473_app1.docx](#)]

References

1. Holmes EA, O'Connor RC, Perry VH, Tracey I, Wessely S, Arseneault L, et al. Multidisciplinary research priorities for the COVID-19 pandemic: a call for action for mental health science. *Lancet Psychiatry* 2020 Jun;7(6):547-560 [FREE Full text] [doi: [10.1016/S2215-0366\(20\)30168-1](https://doi.org/10.1016/S2215-0366(20)30168-1)] [Medline: [32304649](https://pubmed.ncbi.nlm.nih.gov/32304649/)]
2. Greenberg N, Weston D, Hall C, Caulfield T, Williamson V, Fong K. Mental health of staff working in intensive care during Covid-19. *Occup Med (Lond)* 2021 Apr 09;71(2):62-67 [FREE Full text] [doi: [10.1093/occmed/kqaa220](https://doi.org/10.1093/occmed/kqaa220)] [Medline: [33434920](https://pubmed.ncbi.nlm.nih.gov/33434920/)]
3. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders : DSM-5 Fifth ed.* Arlington, VA: American Psychiatric Publishing; 2013:A.
4. Iyadurai L, Visser RM, Lau-Zhu A, Porcheret K, Horsch A, Holmes EA, et al. Intrusive memories of trauma: a target for research bridging cognitive science and its clinical application. *Clin Psychol Rev* 2019 Apr;69:67-82 [FREE Full text] [doi: [10.1016/j.cpr.2018.08.005](https://doi.org/10.1016/j.cpr.2018.08.005)] [Medline: [30293686](https://pubmed.ncbi.nlm.nih.gov/30293686/)]
5. Kupfer DJ, Regier DA. Neuroscience, clinical evidence, and the future of psychiatric classification in DSM-5. *Am J Psychiatry* 2011 Jul;168(7):672-674. [doi: [10.1176/appi.ajp.2011.11020219](https://doi.org/10.1176/appi.ajp.2011.11020219)] [Medline: [21724672](https://pubmed.ncbi.nlm.nih.gov/21724672/)]
6. Singh L, Espinosa L, Ji JL, Moulds ML, Holmes EA. Developing thinking around mental health science: the example of intrusive, emotional mental imagery after psychological trauma. *Cogn Neuropsychiatry* 2020 Sep;25(5):348-363. [doi: [10.1080/13546805.2020.1804845](https://doi.org/10.1080/13546805.2020.1804845)] [Medline: [32847486](https://pubmed.ncbi.nlm.nih.gov/32847486/)]
7. Holmes EA, Ghaderi A, Eriksson E, Lauri KO, Kukacka OM, Mamish M, et al. 'I Can't Concentrate': a feasibility study with young refugees in Sweden on developing science-driven interventions for intrusive memories related to trauma. *Behav Cogn Psychother* 2017 Mar;45(2):97-109 [FREE Full text] [doi: [10.1017/S135246581600062X](https://doi.org/10.1017/S135246581600062X)] [Medline: [28229806](https://pubmed.ncbi.nlm.nih.gov/28229806/)]
8. Alden LE, Regambal MJ, Laposa JM. The effects of direct versus witnessed threat on emergency department healthcare workers: implications for PTSD criterion A. *J Anxiety Disord* 2008 Dec;22(8):1337-1346. [doi: [10.1016/j.janxdis.2008.01.013](https://doi.org/10.1016/j.janxdis.2008.01.013)] [Medline: [18325730](https://pubmed.ncbi.nlm.nih.gov/18325730/)]
9. Berger W, Coutinho ESF, Figueira I, Marques-Portella C, Luz MP, Neylan TC, et al. Rescuers at risk: a systematic review and meta-regression analysis of the worldwide current prevalence and correlates of PTSD in rescue workers. *Soc Psychiatry Psychiatr Epidemiol* 2012 Jun;47(6):1001-1011 [FREE Full text] [doi: [10.1007/s00127-011-0408-2](https://doi.org/10.1007/s00127-011-0408-2)] [Medline: [21681455](https://pubmed.ncbi.nlm.nih.gov/21681455/)]
10. Wahlberg Å, Andreen Sachs M, Johannesson K, Hallberg G, Jonsson M, Skoog Svanberg A, et al. Post-traumatic stress symptoms in Swedish obstetricians and midwives after severe obstetric events: a cross-sectional retrospective survey. *BJOG* 2017 Jul;124(8):1264-1271. [doi: [10.1111/1471-0528.14259](https://doi.org/10.1111/1471-0528.14259)] [Medline: [27562912](https://pubmed.ncbi.nlm.nih.gov/27562912/)]
11. Salmon G, Morehead A. Posttraumatic stress syndrome and implications for practice in critical care nurses. *Crit Care Nurs Clin North Am* 2019 Dec;31(4):517-526. [doi: [10.1016/j.cnc.2019.07.007](https://doi.org/10.1016/j.cnc.2019.07.007)] [Medline: [31685118](https://pubmed.ncbi.nlm.nih.gov/31685118/)]
12. Kleim B, Bingisser M, Westphal M, Bingisser R. Frozen moments: flashback memories of critical incidents in emergency personnel. *Brain Behav* 2015 Jul;5(7):e00325. [doi: [10.1002/brb3.325](https://doi.org/10.1002/brb3.325)] [Medline: [26221567](https://pubmed.ncbi.nlm.nih.gov/26221567/)]
13. Adriaenssens J, de Gucht V, Maes S. The impact of traumatic events on emergency room nurses: findings from a questionnaire survey. *Int J Nurs Stud* 2012 Nov;49(11):1411-1422. [doi: [10.1016/j.ijnurstu.2012.07.003](https://doi.org/10.1016/j.ijnurstu.2012.07.003)] [Medline: [22871313](https://pubmed.ncbi.nlm.nih.gov/22871313/)]
14. Kang L, Ma S, Chen M, Yang J, Wang Y, Li R, et al. Impact on mental health and perceptions of psychological care among medical and nursing staff in Wuhan during the 2019 novel coronavirus disease outbreak: a cross-sectional study. *Brain Behav Immun* 2020 Jul;87:11-17 [FREE Full text] [doi: [10.1016/j.bbi.2020.03.028](https://doi.org/10.1016/j.bbi.2020.03.028)] [Medline: [32240764](https://pubmed.ncbi.nlm.nih.gov/32240764/)]

15. Zhang W, Wang K, Yin L, Zhao W, Xue Q, Peng M, et al. Mental health and psychosocial problems of medical health workers during the COVID-19 epidemic in China. *Psychother Psychosom* 2020;89(4):242-250 [FREE Full text] [doi: [10.1159/000507639](https://doi.org/10.1159/000507639)] [Medline: [32272480](https://pubmed.ncbi.nlm.nih.gov/32272480/)]
16. Lai J, Ma S, Wang Y, Cai Z, Hu J, Wei N, et al. Factors associated with mental health outcomes among health care workers exposed to coronavirus disease 2019. *JAMA Netw Open* 2020 Mar 02;3(3):e203976 [FREE Full text] [doi: [10.1001/jamanetworkopen.2020.3976](https://doi.org/10.1001/jamanetworkopen.2020.3976)] [Medline: [32202646](https://pubmed.ncbi.nlm.nih.gov/32202646/)]
17. Di Tella M, Romeo A, Benfante A, Castelli L. Mental health of healthcare workers during the COVID-19 pandemic in Italy. *J Eval Clin Pract* 2020 Dec;26(6):1583-1587. [doi: [10.1111/jep.13444](https://doi.org/10.1111/jep.13444)] [Medline: [32710481](https://pubmed.ncbi.nlm.nih.gov/32710481/)]
18. Shahrouf G, Dardas LA. Acute stress disorder, coping self-efficacy and subsequent psychological distress among nurses amid COVID-19. *J Nurs Manag* 2020 Oct;28(7):1686-1695 [FREE Full text] [doi: [10.1111/jonm.13124](https://doi.org/10.1111/jonm.13124)] [Medline: [32767827](https://pubmed.ncbi.nlm.nih.gov/32767827/)]
19. Laposa J, Alden L. Posttraumatic stress disorder in the emergency room: exploration of a cognitive model. *Behav Res Ther* 2003 Jan;41(1):49-65. [doi: [10.1016/s0005-7967\(01\)00123-1](https://doi.org/10.1016/s0005-7967(01)00123-1)] [Medline: [12488119](https://pubmed.ncbi.nlm.nih.gov/12488119/)]
20. Laposa JM, Alden LE, Fullerton LM. Work stress and posttraumatic stress disorder in ED nurses/personnel. *J Emerg Nurs* 2003 Feb;29(1):23-28. [doi: [10.1067/men.2003.7](https://doi.org/10.1067/men.2003.7)] [Medline: [12556825](https://pubmed.ncbi.nlm.nih.gov/12556825/)]
21. Mealer M, Burnham EL, Goode CJ, Rothbaum B, Moss M. The prevalence and impact of post traumatic stress disorder and burnout syndrome in nurses. *Depress Anxiety* 2009;26(12):1118-1126 [FREE Full text] [doi: [10.1002/da.20631](https://doi.org/10.1002/da.20631)] [Medline: [19918928](https://pubmed.ncbi.nlm.nih.gov/19918928/)]
22. Muller I, Santer M, Morrison L, Morton K, Roberts A, Rice C, et al. Combining qualitative research with PPI: reflections on using the person-based approach for developing behavioural interventions. *Res Involv Engagem* 2019;5:34 [FREE Full text] [doi: [10.1186/s40900-019-0169-8](https://doi.org/10.1186/s40900-019-0169-8)] [Medline: [31807316](https://pubmed.ncbi.nlm.nih.gov/31807316/)]
23. Post-traumatic stress disorder. National Institute for Health and Care Excellence. 2018. URL: <https://www.nice.org.uk/guidance/ng116> [accessed 2021-05-12]
24. Psychological, psychosocial and other non-pharmacological treatments of posttraumatic stress disorder (PTSD) in adults. Statens Beredning för Medicinsk och Social Utvärdering. 2019. URL: <https://www.sbu.se/sv/publikationer/sbu-kommentar/psykologisk-psykosocial-och-annan-icke-farmakologisk-behandling-av-posttraumatiskt-stressyndrom-ptsd-hos-vuxna/?pub=41678> [accessed 2021-05-12]
25. Baddeley AD, Andrade J. Working memory and the vividness of imagery. *J Exp Psychol Gen* 2000 Mar;129(1):126-145. [doi: [10.1037//0096-3445.129.1.126](https://doi.org/10.1037//0096-3445.129.1.126)] [Medline: [10756490](https://pubmed.ncbi.nlm.nih.gov/10756490/)]
26. Visser RM, Lau-Zhu A, Henson RN, Holmes EA. Multiple memory systems, multiple time points: how science can inform treatment to control the expression of unwanted emotional memories. *Philos Trans R Soc Lond B Biol Sci* 2018 Mar 19;373(1742):20170209 [FREE Full text] [doi: [10.1098/rstb.2017.0209](https://doi.org/10.1098/rstb.2017.0209)] [Medline: [29352036](https://pubmed.ncbi.nlm.nih.gov/29352036/)]
27. Grey N, Holmes EA. "Hotspots" in trauma memories in the treatment of post-traumatic stress disorder: a replication. *Memory* 2008 Oct;16(7):788-796. [doi: [10.1080/09658210802266446](https://doi.org/10.1080/09658210802266446)] [Medline: [18720224](https://pubmed.ncbi.nlm.nih.gov/18720224/)]
28. Kanstrup M, Singh L, Göransson KE, Widoff J, Taylor RS, Gamble B, et al. Reducing intrusive memories after trauma via a brief cognitive task intervention in the hospital emergency department: an exploratory pilot randomised controlled trial. *Transl Psychiatry* 2021 Jan 11;11(1):30. [doi: [10.1038/s41398-020-01124-6](https://doi.org/10.1038/s41398-020-01124-6)] [Medline: [33431807](https://pubmed.ncbi.nlm.nih.gov/33431807/)]
29. Iyadurai L, Blackwell SE, Meiser-Stedman R, Watson PC, Bonsall MB, Geddes JR, et al. Preventing intrusive memories after trauma via a brief intervention involving Tetris computer game play in the emergency department: a proof-of-concept randomized controlled trial. *Mol Psychiatry* 2018 Mar;23(3):674-682 [FREE Full text] [doi: [10.1038/mp.2017.23](https://doi.org/10.1038/mp.2017.23)] [Medline: [28348380](https://pubmed.ncbi.nlm.nih.gov/28348380/)]
30. Horsch A, Vial Y, Favrod C, Harari MM, Blackwell SE, Watson P, et al. Reducing intrusive traumatic memories after emergency caesarean section: a proof-of-principle randomized controlled study. *Behav Res Ther* 2017 Jul;94:36-47 [FREE Full text] [doi: [10.1016/j.brat.2017.03.018](https://doi.org/10.1016/j.brat.2017.03.018)] [Medline: [28453969](https://pubmed.ncbi.nlm.nih.gov/28453969/)]
31. Kessler H, Holmes EA, Blackwell SE, Schmidt A, Schweer JM, Bücker A, et al. Reducing intrusive memories of trauma using a visuospatial interference intervention with inpatients with posttraumatic stress disorder (PTSD). *J Consult Clin Psychol* 2018 Dec;86(12):1076-1090. [doi: [10.1037/ccp0000340](https://doi.org/10.1037/ccp0000340)] [Medline: [30507232](https://pubmed.ncbi.nlm.nih.gov/30507232/)]
32. Kanstrup M, Kontio E, Geranmayeh A, Olofsdotter Lauri K, Moulds ML, Holmes EA. A single case series using visuospatial task interference to reduce the number of visual intrusive memories of trauma with refugees. *Clin Psychol Psychother* 2021 Jan;28(1):109-123. [doi: [10.1002/cpp.2489](https://doi.org/10.1002/cpp.2489)] [Medline: [32525244](https://pubmed.ncbi.nlm.nih.gov/32525244/)]
33. Iyadurai L, Hales SA, Blackwell SE, Young K, Holmes EA. Targeting intrusive imagery using a competing task technique: a case study. *Behav Cogn Psychother* 2020 Nov;48(6):739-744. [doi: [10.1017/S1352465820000296](https://doi.org/10.1017/S1352465820000296)] [Medline: [32594968](https://pubmed.ncbi.nlm.nih.gov/32594968/)]
34. Gamble B, Depa K, Holmes E, Kanstrup M. Digitalizing a brief intervention to reduce intrusive memories of psychological trauma: qualitative interview study. *JMIR Ment Health* 2021 Feb 22;8(2):e23712 [FREE Full text] [doi: [10.2196/23712](https://doi.org/10.2196/23712)] [Medline: [33616540](https://pubmed.ncbi.nlm.nih.gov/33616540/)]
35. O' Cathain A, Croot L, Duncan E, Rousseau N, Sworn K, Turner KM, et al. Guidance on how to develop complex interventions to improve health and healthcare. *BMJ Open* 2019 Aug 15;9(8):e029954 [FREE Full text] [doi: [10.1136/bmjopen-2019-029954](https://doi.org/10.1136/bmjopen-2019-029954)] [Medline: [31420394](https://pubmed.ncbi.nlm.nih.gov/31420394/)]

36. Movsisyan A, Arnold L, Evans R, Hallingberg B, Moore G, O'Cathain A, et al. Adapting evidence-informed complex population health interventions for new contexts: a systematic review of guidance. *Implement Sci* 2019 Dec 17;14(1):105 [FREE Full text] [doi: [10.1186/s13012-019-0956-5](https://doi.org/10.1186/s13012-019-0956-5)] [Medline: [31847920](https://pubmed.ncbi.nlm.nih.gov/31847920/)]
37. Hawkins J, Madden K, Fletcher A, Midgley L, Grant A, Cox G, et al. Development of a framework for the co-production and prototyping of public health interventions. *BMC Public Health* 2017 Sep 04;17(1):689 [FREE Full text] [doi: [10.1186/s12889-017-4695-8](https://doi.org/10.1186/s12889-017-4695-8)] [Medline: [28870192](https://pubmed.ncbi.nlm.nih.gov/28870192/)]
38. Weathers FW, Blake DD, Schnurr PP, Kaloupek DG, Marx BP, Keane TM. Life Events Checklist for DSM-5 (LEC-5). US Department of Veterans Affairs. 2013. URL: https://www.ptsd.va.gov/professional/assessment/te-measures/life_events_checklist.asp [accessed 2021-05-12]
39. Weathers FW, Litz BT, Keane TM, Palmieri PA, Marx BP, Schnurr PP. PTSD Checklist for DSM-5 (PCL-5). US Department of Veterans Affairs. 2013. URL: <https://www.ptsd.va.gov/professional/assessment/adult-sr/ptsd-checklist.asp> [accessed 2021-05-12]
40. SMART-TRIAL. URL: <https://www.smart-trial.com> [accessed 2021-05-12]
41. Hackmann A, Ehlers A, Speckens A, Clark DM. Characteristics and content of intrusive memories in PTSD and their changes with treatment. *J Trauma Stress* 2004 Jun;17(3):231-240. [doi: [10.1023/B:JOTS.0000029266.88369.fd](https://doi.org/10.1023/B:JOTS.0000029266.88369.fd)] [Medline: [15253095](https://pubmed.ncbi.nlm.nih.gov/15253095/)]
42. Speckens A, Ehlers A, Hackmann A, Clark D. Changes in intrusive memories associated with imaginal reliving in posttraumatic stress disorder. *J Anxiety Disord* 2006;20(3):328-341. [doi: [10.1016/j.janxdis.2005.02.004](https://doi.org/10.1016/j.janxdis.2005.02.004)] [Medline: [16564436](https://pubmed.ncbi.nlm.nih.gov/16564436/)]
43. Weiss DS, Marmar CR. Impact of Event Scale--Revised. In: *Assessing Psychological Trauma and PTSD: A Handbook for Practitioners*. New York, NY: Guilford Press; 1997:399-411.
44. Creamer M, Bell R, Failla S. Psychometric properties of the Impact of Event Scale - Revised. *Behav Res Ther* 2003 Dec;41(12):1489-1496. [doi: [10.1016/j.brat.2003.07.010](https://doi.org/10.1016/j.brat.2003.07.010)] [Medline: [14705607](https://pubmed.ncbi.nlm.nih.gov/14705607/)]
45. Price M, Szafranski DD, van Stolk-Cooke K, Gros DF. Investigation of abbreviated 4 and 8 item versions of the PTSD Checklist 5. *Psychiatry Res* 2016 May 30;239:124-130. [doi: [10.1016/j.psychres.2016.03.014](https://doi.org/10.1016/j.psychres.2016.03.014)] [Medline: [27137973](https://pubmed.ncbi.nlm.nih.gov/27137973/)]
46. Price M, van Stolk-Cooke K, Brier ZMF, Legrand AC. mHealth solutions for early interventions after trauma: improvements and considerations for assessment and intervention throughout the acute post-trauma period. *Mhealth* 2018;4:22. [doi: [10.21037/mhealth.2018.06.03](https://doi.org/10.21037/mhealth.2018.06.03)] [Medline: [30148137](https://pubmed.ncbi.nlm.nih.gov/30148137/)]
47. Devilly GJ, Borkovec TD. Psychometric properties of the credibility/expectancy questionnaire. *J Behav Ther Exp Psychiatry* 2000 Jun;31(2):73-86. [doi: [10.1016/s0005-7916\(00\)00012-4](https://doi.org/10.1016/s0005-7916(00)00012-4)] [Medline: [11132119](https://pubmed.ncbi.nlm.nih.gov/11132119/)]
48. 1. Glossary. Good Clinical Practice Network. URL: <https://ichgcp.net/1-glossary> [accessed 2021-05-12]
49. Zoom Video Communications. URL: <https://ki-se.zoom.us/> [accessed 2021-05-12]
50. Bengtsson M. How to plan and perform a qualitative study using content analysis. *NursingPlus Open* 2016;2:8-14. [doi: [10.1016/j.npls.2016.01.001](https://doi.org/10.1016/j.npls.2016.01.001)]
51. Yardley L, Morrison L, Bradbury K, Muller I. The person-based approach to intervention development: application to digital health-related behavior change interventions. *J Med Internet Res* 2015 Jan 30;17(1):e30 [FREE Full text] [doi: [10.2196/jmir.4055](https://doi.org/10.2196/jmir.4055)] [Medline: [25639757](https://pubmed.ncbi.nlm.nih.gov/25639757/)]
52. Bradbury K, Watts S, Arden-Close E, Yardley L, Lewith G. Developing digital interventions: a methodological guide. *Evid Based Complement Alternat Med* 2014;2014:561320. [doi: [10.1155/2014/561320](https://doi.org/10.1155/2014/561320)] [Medline: [24648848](https://pubmed.ncbi.nlm.nih.gov/24648848/)]
53. Kuhn J. Decrypting the MoSCoW analysis. *itSM Solutions* 2009;5(44):1-2.
54. Bradbury K, Morton K, Band R, van Woezik A, Grist R, McManus RJ, et al. Using the Person-Based Approach to optimise a digital intervention for the management of hypertension. *PLoS One* 2018;13(5):e0196868 [FREE Full text] [doi: [10.1371/journal.pone.0196868](https://doi.org/10.1371/journal.pone.0196868)] [Medline: [29723262](https://pubmed.ncbi.nlm.nih.gov/29723262/)]
55. Yeager CM, Benight CC. If we build it, will they come? Issues of engagement with digital health interventions for trauma recovery. *Mhealth* 2018;4:37. [doi: [10.21037/mhealth.2018.08.04](https://doi.org/10.21037/mhealth.2018.08.04)] [Medline: [30363749](https://pubmed.ncbi.nlm.nih.gov/30363749/)]
56. Ruzek JI, Yeager CM. Mobile mental health interventions for trauma survivors. *Mhealth* 2019;5:9. [doi: [10.21037/mhealth.2019.02.03](https://doi.org/10.21037/mhealth.2019.02.03)] [Medline: [31019962](https://pubmed.ncbi.nlm.nih.gov/31019962/)]
57. Sustainable development goals. United Nations: Department of Economic and Social Affairs. URL: <https://sdgs.un.org/goals> [accessed 2021-05-12]
58. Weathers FW, Bovin MJ, Lee DJ, Sloan DM, Schnurr PP, Kaloupek DG, et al. The Clinician-Administered PTSD Scale for DSM-5 (CAPS-5): development and initial psychometric evaluation in military veterans. *Psychol Assess* 2018 Mar;30(3):383-395 [FREE Full text] [doi: [10.1037/pas0000486](https://doi.org/10.1037/pas0000486)] [Medline: [28493729](https://pubmed.ncbi.nlm.nih.gov/28493729/)]
59. State of the World's Nursing 2020: Investing in Education, Jobs and Leadership. Geneva, Switzerland: World Health Organization; 2020.
60. Friganović A, Selić P, Ilić B, Sedić B. Stress and burnout syndrome and their associations with coping and job satisfaction in critical care nurses: a literature review. *Psychiatr Danub* 2019 Mar;31(Suppl 1):21-31 [FREE Full text] [Medline: [30946714](https://pubmed.ncbi.nlm.nih.gov/30946714/)]
61. Ross J. The exacerbation of burnout during COVID-19: a major concern for nurse safety. *J Perianesth Nurs* 2020 Aug;35(4):439-440 [FREE Full text] [doi: [10.1016/j.jopan.2020.04.001](https://doi.org/10.1016/j.jopan.2020.04.001)] [Medline: [32741522](https://pubmed.ncbi.nlm.nih.gov/32741522/)]

62. Kok N, van Gurp J, Teerenstra S, van der Hoeven H, Fuchs M, Hoedemaekers C, et al. Coronavirus disease 2019 immediately increases burnout symptoms in ICU professionals: a longitudinal cohort study. *Crit Care Med* 2021 Mar 01;49(3):419-427. [doi: [10.1097/CCM.0000000000004865](https://doi.org/10.1097/CCM.0000000000004865)] [Medline: [33555778](https://pubmed.ncbi.nlm.nih.gov/33555778/)]
63. Bird VJ, Le Boutillier C, Leamy M, Williams J, Bradstreet S, Slade M. Evaluating the feasibility of complex interventions in mental health services: standardised measure and reporting guidelines. *Br J Psychiatry* 2014;204:316-321. [doi: [10.1192/bjp.bp.113.128314](https://doi.org/10.1192/bjp.bp.113.128314)] [Medline: [24311549](https://pubmed.ncbi.nlm.nih.gov/24311549/)]
64. Ehlers A, Hackmann A, Steil R, Clohessy S, Wenninger K, Winter H. The nature of intrusive memories after trauma: the warning signal hypothesis. *Behav Res Ther* 2002 Sep;40(9):995-1002. [doi: [10.1016/s0005-7967\(01\)00077-8](https://doi.org/10.1016/s0005-7967(01)00077-8)] [Medline: [12296496](https://pubmed.ncbi.nlm.nih.gov/12296496/)]
65. Ji JL, Heyes SB, MacLeod C, Holmes EA. Emotional mental imagery as simulation of reality: fear and beyond—a tribute to Peter Lang. *Behav Ther* 2016 Sep;47(5):702-719 [FREE Full text] [doi: [10.1016/j.beth.2015.11.004](https://doi.org/10.1016/j.beth.2015.11.004)] [Medline: [27816082](https://pubmed.ncbi.nlm.nih.gov/27816082/)]

Abbreviations

DSM-5: Diagnostic and Statistical Manual of Mental Disorders (Fifth Edition)

ICU: intensive care unit

IES-R: Impact of Event Scale-Revised

MoSCoW: must have, should have, could have, won't have

PCL-5: Posttraumatic Stress Disorder Checklist for Diagnostic and Statistical Manual of Mental Disorders (Fifth Edition)

PTSD: posttraumatic stress disorder

RCT: randomized controlled trial

SUD: subjective units of distress

Edited by G Eysenbach; submitted 27.01.21; peer-reviewed by A Woolard; comments to author 19.02.21; revised version received 01.03.21; accepted 17.03.21; published 26.05.21.

Please cite as:

Singh L, Kanstrup M, Depa K, Falk AC, Lindström V, Dahl O, Göransson KE, Rudman A, Holmes EA

Digitalizing a Brief Intervention to Reduce Intrusive Memories of Psychological Trauma for Health Care Staff Working During COVID-19: Exploratory Pilot Study With Nurses

JMIR Form Res 2021;5(5):e27473

URL: <https://formative.jmir.org/2021/5/e27473>

doi: [10.2196/27473](https://doi.org/10.2196/27473)

PMID: [33886490](https://pubmed.ncbi.nlm.nih.gov/33886490/)

©Laura Singh, Marie Kanstrup, Katherine Depa, Ann-Charlotte Falk, Veronica Lindström, Oili Dahl, Katarina E Göransson, Ann Rudman, Emily A Holmes. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 26.05.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

The Effect of Training on Participant Adherence With a Reporting Time Frame for Momentary Subjective Experiences in Ecological Momentary Assessment: Cognitive Interview Study

Cheng K Fred Wen^{1*}, MPH, PhD; Doerte U Junghaenel^{1*}, PhD; David B Newman^{2*}, PhD; Stefan Schneider^{1*}, PhD; Marilyn Mendez^{1*}, BSc; Sarah E Goldstein^{1*}, BA; Sarah Velasco^{1*}, MA; Joshua M Smyth^{3*}, PhD; Arthur A Stone^{1,4*}, PhD

¹Dornsife Center for Self-Report Science, University of Southern California, Los Angeles, CA, United States

²Department of Psychiatry, University of California, San Francisco, San Francisco, CA, United States

³Department of Biobehavioral Health, Pennsylvania State University, University Park, PA, United States

⁴Department of Psychology, University of Southern California, Los Angeles, CA, United States

* all authors contributed equally

Corresponding Author:

Cheng K Fred Wen, MPH, PhD
Dornsife Center for Self-Report Science
University of Southern California
635 Downey Way
Verna and Peter Dauterive Hall 405H
Los Angeles, CA, 90089
United States
Phone: 1 2138212894
Email: chengkuw@usc.edu

Abstract

Background: Ecological momentary assessment (EMA) has the potential to minimize recall bias by having people report on their experiences in the moment (momentary model) or over short periods (coverage model). This potential hinges on the assumption that participants provide their ratings based on the reporting time frame instructions prescribed in the EMA items. However, it is unclear what time frames participants actually use when answering the EMA questions and whether participant training improves participants' adherence to the reporting instructions.

Objective: This study aims to investigate the reporting time frames participants used when answering EMA questions and whether participant training improves participants' adherence to the EMA reporting timeframe instructions.

Methods: Telephone-based cognitive interviews were used to investigate the research questions. In a 2×2 factorial design, participants (n=100) were assigned to receive either basic or enhanced EMA training and randomized to rate their experiences using a momentary (at the moment you were called) or a coverage (since the last phone call) model. Participants received five calls over the course of a day to provide ratings; after each rating, participants were immediately interviewed about the time frame they used to answer the EMA questions. A total of 2 raters independently coded the momentary interview responses into time frame categories (Cohen κ=0.64, 95% CI 0.55-0.73).

Results: The results from the momentary conditions showed that most of the calls referred to the period during the call (57/199, 28.6%) or just before the call (98/199, 49.2%) to provide ratings; the remainder were from longer reporting periods. Multinomial logistic regression results indicated a significant training effect ($\chi^2_1=16.6$; $P<.001$) in which the enhanced training condition yielded more reports within the intended reporting time frames for momentary EMA reports. Cognitive interview data from the coverage model did not lend themselves to reliable coding and were not analyzed.

Conclusions: The results of this study provide the first evidence about adherence to EMA instructions to reporting periods and that enhanced participant training improves adherence to the time frame specified in momentary EMA studies.

(JMIR Form Res 2021;5(5):e28007) doi:[10.2196/28007](https://doi.org/10.2196/28007)

KEYWORDS

ecological momentary assessment; EMA; cognitive interview; participant training; reporting period

Introduction

Background

In recent decades, behavioral and psychological researchers have increasingly used ecological momentary assessment (EMA) [1,2] and similar methods (eg, the experience sampling method [3,4]) to measure subjective experiences in daily life. In contrast to traditional self-report instruments, in which participants are asked to provide global evaluations or to retrospectively summarize their experience over a period (eg, 30 days), EMA studies typically ask respondents to answer questions about their experiences just before they were prompted or over short periods. Repeated assessments of participant experiences within a day and over multiple days afford the opportunity to examine various within-person processes [5-7] and changes that occur over short periods [8]. Moreover, by assessing participants in their everyday environments, EMA data are thought to be ecologically valid and to have reduced recall bias [9,10].

When inquiring about participant experiences, there are two commonly implemented approaches: the momentary model and the coverage model [1]. Studies that use the momentary model typically aim to capture participant experiences at the current moment or at the moment just before they were prompted. For example, studies assessing happiness at the moment may use questions such as, “How happy were you feeling right before the prompt?” In contrast, studies implementing a coverage model do not mean to capture current experience but instead capture experiences over a relatively short period. For example, a coverage approach might instruct respondents to report their experiences over the last 30 minutes or since the last time they were prompted (eg, “How happy were you, overall, since the last time you were beeped?”). Both momentary and coverage models have been used in EMA studies to assess a variety of experiences, including affective states [11,12], hunger [13,14], and pain [15,16].

An important assumption of the EMA data is that participants adhere to the stated time frame for their self-report ratings. However, there is no current evidence that participants follow the time frame instructions as intended by the researchers. When asked to report on their experience *right before the prompt*, some participants may provide an overall rating of their experiences over longer periods extending beyond the current moment [17]. Similarly, when asked to report about their experience *since the last prompt*, some participants may only provide information about a portion of the prescribed time frame. A recent study on the time frames that participants used to complete end-of-day diaries illustrates the possibility that people may not follow the designated reporting instructions [18]. Participants were randomized to one of four versions of an end-of-day question meant to assess their affective states over the course of the day. They subsequently indicated the periods they considered while making their ratings. Although all instructions directed participants to rate their affective states over the course of the day (defined in several ways), participants

reported using a wide range of time frames, many of which were not consistent with the intended period. For example, only 34.7% (49/141) of individuals reported using time frames consistent with the instructions “In the last day, I felt...” versus 96.5% (141/146) who were consistent when the instructions were “Today, I felt...” Although the instructional wording was similar, the time frames generated by the two instructions varied significantly and suggested that attention should be paid to participants’ interpretations of the diary instructions.

Objectives

Pertinent to this study, the findings highlight the possibility that participants may report what *makes sense* to them based on the contextual information available in the survey [10] or on conversational norms [19]. This could be problematic as participants’ self-report ratings might not reflect their experiences from a time frame that was intended by the specific research protocol. Interpreting data from outside of the intended reporting time frame could further introduce errors when examining the relationship between EMA data and data from other sources (eg, accelerometers). To our knowledge, there has been no research on how well participants adhere to the time frame instructions used in EMA studies. Therefore, the primary aim of this study is to describe the time frame participants use when answering both the momentary and coverage EMA questions. Owing to the methodological challenges that we experienced when analyzing the reported time frame data for the coverage model in this study (which are described in detail later in this report), we only report results pertinent to the momentary EMA model.

One factor that could influence participants’ adherence to the designated time frame is training. Providing training is a common practice in EMA studies to minimize participant confusion and improve data quality in EMA study protocols. Existing evidence from daily diary studies suggests that training can reduce missing data and improve the internal consistency of diary ratings [20] and that iterative training (eg, providing feedback during the training) is helpful in guiding participants to report on their momentary experience more precisely [21]. A recent systematic review of EMA studies also suggests the importance of providing comprehensive training, including providing examples and opportunities to practice before data collection [22]. However, there is no evidence pertaining to the effectiveness of training on whether participants adhere to the time frame instructions used in the EMA. Therefore, the secondary aim of this study is to evaluate (compared with participants training with *basic instructions* that included just the study procedure) whether training with *enhanced instructions*, which included the opportunity for practice and feedback from the research team in addition to basic instructions, is more effective in generating participant self-reports that are consistent with the intended reporting time frame.

Methods

Participant Recruitment

Study participants were recruited through Amazon Mechanical Turk (MTurk) between October 2 and October 18, 2019. The study invitation was only available to registered MTurk workers who had already completed a minimum of 500 approved human intelligence tasks, had a minimum human intelligence task approval rate of 99%, and lived in the Pacific Time (PT) time zones. MTurk workers who met these qualifications were provided with a link to a web-based screener survey in which they provided demographic information and answered five questions that determined their eligibility to participate in the study. Eligibility criteria included being at least 21 years of age, living in the PT time zone, being fluent in English, having access to a phone, and being willing and available to receive and answer five phone calls on the day of the study. Participants were limited to those who resided in the PT time zone and had access to a phone because the interviews were conducted over the phone by members of the research team located in this time zone. Eligible participants were then provided with a detailed study description. Interested participants who agreed to participate were asked to provide their contact information (ie, email address and phone number) to the study team. All MTurk workers who participated in the screening survey, regardless of whether they met the eligibility criteria or provided contact information, were compensated with US \$0.50 through MTurk. Interested participants who completed the full study protocol, which included the five phone calls on the subsequent day, were compensated with an additional US \$10 through MTurk.

Study Procedure

Participants were randomized to one of four experimental conditions in a 2×2 factorial design, with a reporting time frame factor (momentary or coverage model) and a training condition factor (basic or enhanced training). The *EMA prompts* in this study were administered using phone calls. We decided to administer the EMA prompts using phone calls because it allowed study staff to conduct semistructured cognitive interviews immediately after participants answered the EMA questions to probe for the time frames that participants used, which would not be feasible to use device-based EMA reports. The EMA items asked participants to report their anxiety, happiness, hunger, and pain. These experience domains were chosen because they are emotional and sensory experiences that are commonly assessed in both coverage and momentary model EMA studies but we acknowledge that this is only a small set of the states and behaviors that could have been chosen.

In all experimental conditions, participants received five phone calls from the research team on one weekday (Monday through Friday) between 9 AM and 5 PM PT. A day before the start of the study, participants were reminded via text messages and through email that a study staff member would contact them shortly after 9 AM PT on the next day. The first call was the introductory call, which took place between 9 AM and 10 AM on the day of the study. During this call, participants were introduced to the study procedures and were informed that all conversations were audiorecorded. If a participant initially did

not answer the phone for the introductory phone call, the researcher called two additional times that were at least 15 minutes apart before removing the participant from the study. During the introductory call, participants randomized to the *basic training* condition were read a script that described the overall study procedure with some information about the content of the subsequent calls ([Multimedia Appendix 1](#) provides the full script). Participants randomized to the *enhanced training* condition received the same information as those in the *basic training* condition but were also read a more detailed introductory script. The script for the enhanced training included presentation and practice of the questions that would be asked (ie, how happy, anxious, hungry, or how much pain they felt on a scale from 0 [ie, not at all] to 100 [ie, extremely]) and real-time feedback, to ensure that participants understood the rating scale and time frame as intended by the research team ([Multimedia Appendix 2](#) provides full scripts for participants in the momentary model condition with enhanced training and [Multimedia Appendix 3](#) provides full scripts for participants in the coverage model with enhanced training). After the training, participants were informed that they would subsequently receive four cognitive interview calls at random times throughout the day. The participants were instructed not to call back if they missed a call from the researcher.

After the introductory call, the research team divided the rest of the participant's day (9 AM to 5 PM) into four 2-hour time windows. The subsequent four cognitive interview calls were scheduled to occur at a randomly selected time within each of these time windows, with a restriction such that no two calls occurred within 45 minutes of each other. If a participant did not answer the phone, a member of the research team called two additional times that were at least 15 minutes apart. All calls were conducted in a semistructured manner. At the beginning of each phone call, the participants were asked for permission to start recording the conversation. After they agreed, participants were asked to provide a rating for one of the four experience domains (ie, level of pain, hunger, anxiousness, and happiness) on a 0-100 scale. In the momentary condition, the wording of the questions as read aloud to the study participants by a research team member was, "On a scale from zero to one hundred, how XX did you feel right before the phone call (momentary model)/since the last call (coverage model), where zero means not at all XX and a hundred means extremely XX?" Participants rated only one of the experience domains (ie, pain, hunger, anxiousness, and happiness) per call, with the order being randomized across calls to minimize the effect of the time of day and question ordering. After the participants provided the rating, they were asked three structured open-ended questions: "What came to mind as you answered this question?" "How did you come up with the rating you provided?" and "There was likely a particular time that you were thinking about when you answered the question. When was that?" After the participants responded to the three open-ended questions, the research staff thanked them and concluded the phone conversation. During the interview, the research staff did not probe participants for further details on their answers, except for clarification purposes, because this interview protocol was intended to solicit participants' own perspectives and thought processes. All conversations between the research staff and

participants were audiorecorded. The introductory call and cognitive interview calls were conducted by different members of the research team to ensure the blinding of participants' training conditions. All study procedures were approved by the institutional review board of the University of Southern California.

Coding of Cognitive Interview Data

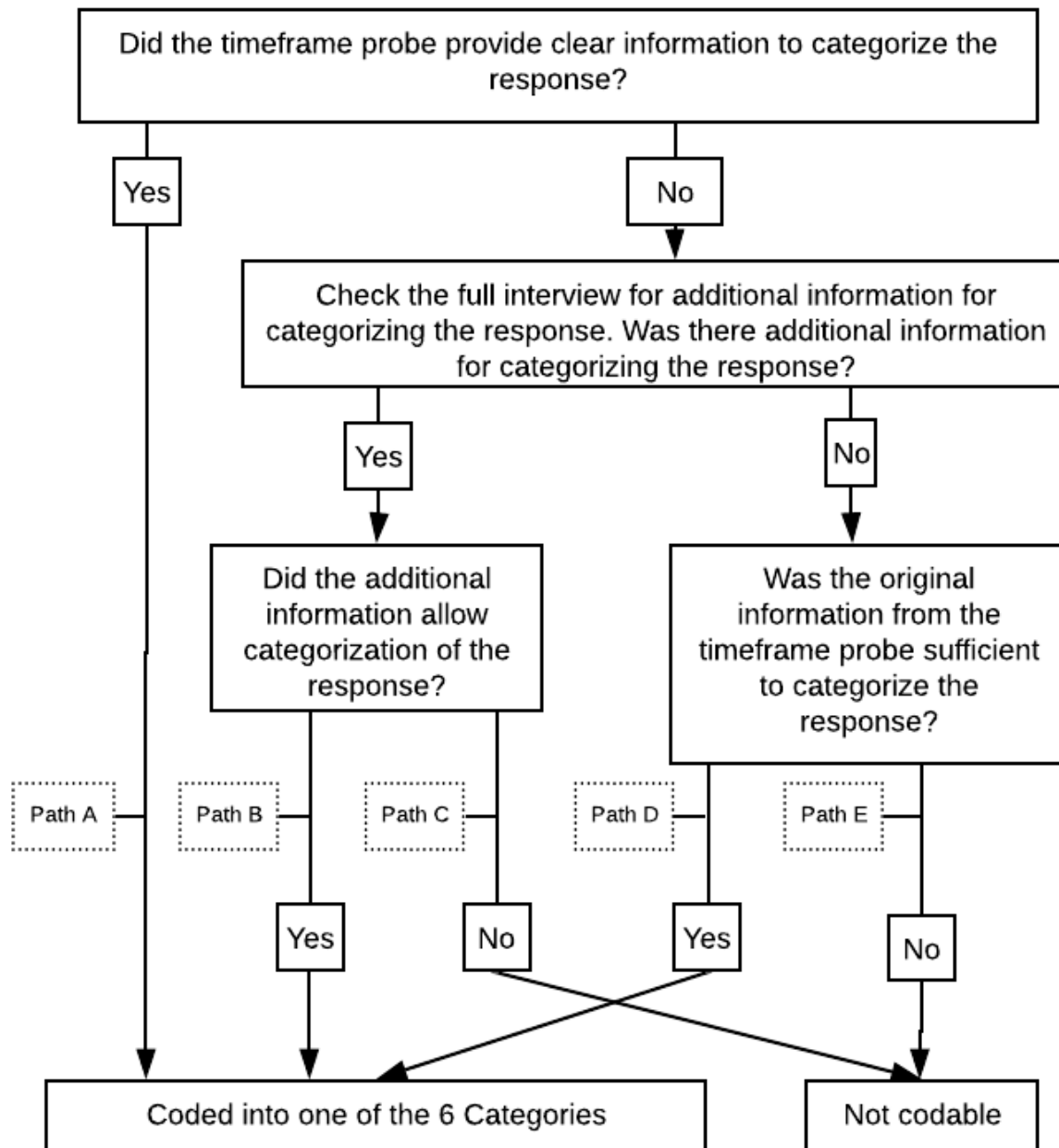
The primary aim of this study was to examine the reporting time frames that study participants used for their ratings. A total of two raters (CKFW and DUJ) first reviewed participants' responses to the time frame probe (ie, the question of "There was likely a particular time that you were thinking about when you answered the question. When was that?") because this prompt specifically asked respondents what time frame they had in mind when completing their rating. We first describe the coding of the momentary model, and the flowchart in [Figure 1](#) summarizes the coding decision process. Briefly, the two raters independently coded the time frame probe responses in the momentary model condition into one of seven categories: *during the call*, *right before the call*, *within 5 minutes of the call*, *within 5-15 minutes of the call*, *within 15-60 minutes of the call*, *more than 60 minutes prior to the call*, or *not codable* (path A, [Figure 1](#)). The intervals for the first six categories were determined based on an initial review of participant responses and were designed to capture the variety of reporting time frames used by the participants. If participants' responses to the time frame probe did not provide clear information to categorize the response, the coders reviewed the full interview for additional information that could facilitate the coding of the response. For example, when a participant responded to the time frame probe asking about their hunger level with statements such as "I ate something an hour ago," the raters reviewed the full interview for additional information to determine whether the participant reported their hunger level 60 minutes ago when they ate something or whether the participant reported their hunger level right before the call and simply mentioned eating something an hour ago. If the full interview contained additional evidence that clarified the participant's response to the time frame probe (eg, participant mentioned "thinking about his lunch when coming up with how hungry he was right before of the interview" when asked how he came up with the rating), then

this additional information was used by the raters to code the reporting time frame (path B).

If raters still could not determine the reporting time frame after reviewing all the available information, the participant's response was categorized as *not codable* (path C). Continuing with the previous example, if the participant mentioned other time frames (eg, *recently* or *in a couple of hours*) in addition to the *an hour ago* response to the time frame probe, then it was not clear which of these time frames was the one used for the rating, and the reporting time frame was therefore designated as *not codable*. If the full interview did not provide additional information, then the time frame categories were coded based on the participant's response to the time frame probe, provided that it contained sufficient information for coding. Continuing with "I ate something an hour ago," if the raters did not find additional information from the full interview, then the reporting time frame for this interview was coded as 60 minutes ago (path D); however, if the participant provided a response such as "I ate something not too long ago," instead of a concrete time point like "an hour ago," then this interview was rated as *not codable* (path E). Each rater coded all interviews independently before comparing the results. The two raters discussed discrepancies in their coding and attempted to reach a consensus.

We now turn to coding for the coverage model condition. Unlike the momentary model, the two raters found that the responses to the time frame probe in the coverage model were generally vague and challenging to code. For example, respondents provided answers such as "9 AM" when they were instructed to rate their experience since the last phone call. It was not clear whether the participant meant to report their rating at 9 AM (ie, rating at a single time point) or from 9 AM (ie, an aggregated rating based on the whole time between the reported time point and the current time) based on the available information, that is, it was difficult to determine whether the entire period was used for the response. In fact, when designing the interview protocol, we did not anticipate that the responses would be as vague as they were, and additional probes were not included as part of the protocol. After consulting with the full research team, we determined that the reporting time frame for the coverage model could not be reliably coded. Therefore, we eliminated the coverage model data from further considerations in this study.

Figure 1. Decision flowchart for categorizing participants' reporting time frame.



Statistical Analyses

A subset of 10% (20/199) of momentary model interviews was randomly selected as training samples [23] for two practice rounds before the remaining interviews were coded. Interrater reliability was assessed with both the raw percentage of agreement and with Cohen κ using the 180 interviews that were not included in the training subset. All disagreements were resolved by consensus after Cohen κ was calculated. The resulting time frame categories were used for the primary aim of the study, which describes the reporting time frame participants used when answering the EMA questions. As an exploratory aim, we tested whether the intensity of self-report ratings (ie, self-reported levels of pain, hunger, anxiousness, and happiness) differed by time frame category using multilevel

regression models. Multilevel regression models also included demographic variables, training conditions, sequential call numbers (ie, call #1, #2, #3, and #4), and the experience domain as covariates to control for potential confounding effects.

The time frame categories were also used as the primary dependent variable in multinomial logistic regression models for examining the second aim of the study: to determine whether the probability of certain time frame categories differed by training condition (ie, basic training or enhanced training). Interviews coded as *not codable* were excluded. The analytic model also included demographic variables (ie, age, gender, race, education level, marital status, and annual household income level) as well as sequential call numbers (ie, call #1, #2, #3, and #4) and the experience domain (ie, pain, hunger,

anxiousness, and happiness) as covariates to control for potential confounding effects. Interactions among the covariates were not explored, as this went beyond the main goals of the study. Models with categorical outcomes (time frame categories) were conducted using Proc GENMOD, and models with continuous outcomes (self-report ratings) were conducted using Proc MIXED in SAS (version 9.4; SAS Institute).

Results

Participant Characteristics

Of the 672 MTurk workers who accessed the screening survey, 94 exited the survey before the study description was presented, 388 did not meet the eligibility criteria (one was aged <21 years, 67 did not reside in the PT time zone, 125 reported not having access to a phone, and 195 did not wish to receive five phone calls), and 47 were not interested in participating. All 143 eligible and interested participants provided their contact information to the research team, but 43 did not respond to

contact attempts made by the research team for the introductory call. This group of people was, on average, younger (mean age 35.1, SD 8.9 years) than the 100 who responded to contact attempts (mean age 40.1, SD 12.8 years; $t_{141}=-2.34$, two-tailed; $P=.02$) but did not differ in other demographic characteristics. The participant characteristics of the remaining 100 participants who participated in the full study protocol are presented in [Table 1](#). The participant characteristics did not significantly differ between the four experimental conditions. Participants who were randomized to the basic training conditions spent an average of 145.9 (SD 27.4) seconds, whereas those randomized to the enhanced training conditions spent an average of 430.6 (SD 62.2) seconds on the introductory call. Of 400 scheduled calls for cognitive interviews, 359 (89.7%) were answered at the first contact, 29 (7.2%) were answered at the second attempt, and 11 (2.8%) were answered at the third and final attempt. The research team was not able to reach one participant for 1 (0.3%) occasion after three attempts of contact.

Table 1. Participant demographic characteristics.

Variable	Population (N=100)	Momentary model		Coverage model	
		Basic training (n=25)	Enhanced training (n=25)	Basic training (n=25)	Enhanced training (n=25)
Age (years), mean (SD)	40.1 (12.8)	36.4 (9.2)	42.6 (13.3)	38.0 (11.9)	43.3 (15.4)
Gender, n (%)					
Female	50 (50)	14 (56)	12 (48)	14 (56)	10 (40)
Male	50 (50)	11 (44)	13 (52)	11 (44)	15 (60)
Ethnicity, n (%)					
Non-Hispanic	83 (83)	19 (76)	21 (84)	21 (84)	22 (88)
Hispanic	17 (17)	6 (24)	4 (16)	4 (16)	3 (12)
Race, n (%)					
African American	2 (2)	0 (0)	1 (4)	1 (4)	0 (0)
Asian	9 (9)	1 (4)	2 (8)	2 (8)	4 (16)
Native American	2 (2)	0 (0)	1 (4)	0 (0)	1 (4)
Mixed	12 (12)	3 (12)	1 (4)	4 (16)	4 (16)
White	75 (75)	21 (84)	20 (80)	18 (72)	16 (64)
Marital status, n (%)					
Never married	38 (38)	10 (40)	7 (28)	9 (36)	12 (48)
Living with partner	11 (11)	4 (16)	4 (16)	1 (4)	2 (8)
Married	34 (34)	7 (28)	9 (36)	12 (48)	6 (24)
Divorced	12 (12)	3 (12)	3 (12)	3 (12)	3 (12)
Separated	1 (1)	0 (0)	1(4)	0 (0)	0 (0)
Widowed	4 (4)	1 (4)	1(4)	0 (0)	2 (8)
Education, n (%)					
High school graduate	5 (5)	1 (4)	0 (0)	3 (12)	1 (4)
Some college	26 (26)	9 (36)	8 (32)	6 (24)	3 (12)
College graduate	60 (60)	10 (40)	16 (64)	14 (56)	20 (80)
Master's degree	6 (6)	5 (20)	1 (4)	0 (0)	0 (0)
Doctoral degree	3 (3)	0 (0)	0 (0)	2 (8)	1 (4)
Employment status, n (%)					
Employed full time	56 (56)	16 (64)	13(52)	15 (60)	12 (48)
Employed part time	20 (20)	4 (16)	7 (28)	7 (28)	2 (8)
Self-employed	7 (7)	2 (8)	3 (12)	0 (0)	2 (8)
Homemaker	5 (5)	1 (4)	0 (0)	1 (4)	3 (12)
Retired	6 (6)	0 (0)	2 (8)	1 (4)	3 (12)
Unemployed	4 (4)	1 (4)	0 (0)	1 (4)	2 (8)
Unable to work	2 (2)	1 (4)	0 (0)	0 (0)	1 (4)
Income (US \$), n (%)					
<20,000	13(13)	8 (32)	2 (8)	2 (8)	1 (4)
20,000-34,999	16 (16)	3 (12)	5 (20)	3 (12)	5 (20)
35,000-49,999	20 (20)	4 (16)	5 (20)	7 (28)	4 (16)
50,000-74,999	18 (18)	4 (16)	2 (8)	4 (12)	8 (32)
>75,000	33 (33)	6 (24)	11 (44)	9 (36)	7 (28)

Reliability of Coding Participants Reporting Time Frame

Of the 200 interviews conducted for the momentary condition, one call was excluded because of administrative errors (the call was accidentally not recorded). The two raters independently coded all the remaining 199 interviews. They agreed on 74.9% (134/179) of the interviews, resulting in a *moderate* interrater reliability [24] (Cohen $\kappa=0.64$, 95% CI 0.55-0.73). All disagreements were resolved by consensus after Cohen κ was calculated.

Reporting Time Frame Categories in Momentary Model Conditions

Of the 199 cognitive interviews collected in the momentary model conditions, 57 (28.6%) were coded as *during the call*,

98 (49.3%) were coded as *right before the call*, 10 (5.1%) were coded as *within 5 minutes of the call*, 8 (4.0%) were coded as *within 5-15 minutes of the call*, 7 (3.5%) were coded as *within 15-60 minutes of the call*, 13 (6.5%) were coded as *more than 60 minutes prior to the call*, and 6 (3.0%) were coded as *not codable* (Table 2). The raters reviewed the full interview for additional information in 36.2% (72/199) of calls. After the full interview review, 8.3% (6/72) of interviews remained uncodable. Of those coded as *not codable*, the raters were unable to identify a reporting time frame for four interviews because the participant did not provide sufficient information over the entire interview. The raters were unable to identify a reporting time frame for the remaining two interviews because participants reported multiple and conflicting time frames.

Table 2. Reporting time frame categories by training condition (momentary condition).

Reporting time frame categories	Overall (N=199), n (%)	Training condition, n (%)	
		Basic training (n=100)	Enhanced training (n=99)
During call or right now	57 (28.6)	40 (40.0)	17 (17.2)
Just before the call	98 (49.2)	23 (23.0)	75 (75.8)
≤5 min before call	10 (5.0)	5 (5.0)	5 (5.1)
>5 min, but ≤15 min before call	8 (4.0)	7 (7.0)	1 (1.0)
>15 min, but ≤60 min before call	7 (3.5)	7 (7.0)	0 (0.0)
>60 min before call	13 (6.5)	12 (12.0)	1 (1.0)
Not codable	6 (3.0)	6 (6.0)	0 (0.0)

Distribution of Momentary Reporting Time Frame Category by Training Condition

Multinomial logistic regression models showed a significant effect of training on the momentary reporting time frame categories ($\chi^2_1=16.6$; $P<.001$). None of the covariates included in the analytic model were significant. The proportions of reporting time frame categories by training assignments are presented in Table 2.

Discussion

Principal Findings

Data collected using EMA methods have the potential to provide a fine-grained understanding of a wide range of psychological, behavioral, and medical phenomena as they occur in daily life. Despite the growing interest in EMA for assessing participant experiences, little attention has been paid to whether participants adhere to the reporting time frame instructions in EMA studies. In this study, we examined this question for commonly used EMA items with momentary reporting instructions. Unfortunately, the data were not suitable for examining the question for coverage model EMA items, which is discussed further in the following section. For the momentary condition, the results showed that although nearly half of the EMA data were reported to be from the intended time frame (ie, just before the call), participants also drew on other reporting time frames, including during the call or other time frames before the moment

before the call. The study also revealed that compared with a basic training procedure, an enhanced training protocol with detailed explanations and opportunities for practice was effective in improving participants' adherence to momentary time frame instructions.

First, we consider our inability to reliably code coverage model interviews. Compared with the momentary model, there were many more ways in which a participant could engage in responding to the coverage model questions. When rating their experience under the coverage model, participants could have considered the proximity and duration of a relevant event related to the inquired experiences and could have reported their rating based on a single moment, shorter or longer periods that covered only parts or all the time between two phone calls. Our interview protocol was not fully prepared to handle these complexities, as we used an open-ended approach to elicit responses, and participants' statements were often too ambiguous to confidently categorize their responses. Future investigations of how participants answer coverage model questions will need to incorporate additional probes in the interview protocol that are designed to solicit more specific information about the proximity and duration of experiences.

In terms of the results for the momentary model, this study provided detailed information about the time frames used in the basic and enhanced training conditions so that readers may consider what *they* view as acceptable adherence to momentary EMA instructions. We grouped together the categories of *during the call/right now*, *just before the call*, and *within 5 minutes of*

the call in our discussion in the following paragraph because, in our opinion, these seem to be reasonable time frames for momentary EMA questions. However, other researchers may find this categorization too liberal for their purposes (eg, associating affective states with ambulatory monitoring of heart rate the minute before the prompt) and instead adopt a more stringent rule where only those time frames that included how participants felt *just before the call* are considered valid. Researchers with other goals for their EMA studies may choose to adopt more liberal rules.

For participant training, the results for the momentary data showed that when participants were provided with basic training, only 68.0% (68/100) of the interviews were coded into the three categories of less than 5 minutes before the call, just before the call, and during the call; only 23.0% (23/100) of the interviews were coded as just before the call. In our view, these percentages are less than ideal and suggest that simply stating the intended reporting time frame as part of the EMA questions without more detailed participant training is not sufficient to achieve high adherence rates to the intended time frame. A possible explanation for this could be that instead of considering the literal meaning of *right before the call*, participants may have responded based on conversational norms [10,19]. For example, when a participant was asked to rate their hunger level *right before the call*, perhaps they may have thought that we meant to capture their hunger level around the time they last ate something before the phone call because this information may be more in line with what is worth communicating based on conversational norms in daily life (or what the participant thought the researchers really wanted to know).

When participants were provided with enhanced training before data collection in the momentary condition, significantly more calls (75/99, 76%) had a reporting time frame coded as just before the call and 98% (97/99) had a reporting time frame coded as less than 5 minutes before the call, just before the call, and during the call. This is an excellent level of adherence to the instructions, assuming that a reporting time frame from within 5 minutes of the prompt is appropriate for one's research goals and argues that researchers should ensure that participants are thoroughly trained in momentary data collection. This is especially important as self-reports for momentary EMA because longer-than-momentary reporting time frames (≥ 2 hours) have previously been shown to yield systematically higher rating levels compared with immediate ratings [25]. These level differences could bias the EMA reports if a sizable portion of a sample did not adhere to the intended reporting time frames.

Although this study provides new information regarding participants' self-reported time frames in EMA studies, it has limitations. First, although the cognitive interview phone calls were conducted according to a typical EMA schedule, we acknowledge the possibility of the mode of administration

effects. For example, participants' responses to a telephone interviewer could differ from responses provided without an interviewer. It is possible that participants may tend to provide more socially desirable responses to the interviewer [26]. Another example of the administration effect could be that the study participant picks up subtle and unintended verbal cues from the research team members as they read the items aloud (eg, emphasis on specific parts of the sentence). Both examples may have introduced some bias in the resulting data. Second, it is challenging to verify how and whether the participants made their ratings based on the intended time frame. For example, it is possible that some participants responded with *right before the prompt* as their answer to the time frame probe by simply repeating the time frame instruction from the question regardless of the time frame they actually used. Future studies extending this line of investigation could incorporate additional interview procedures (eg, using the *think-aloud* method) that solicit the details of the thought processes leading up to their ratings for both the momentary and coverage models. A third limitation is that we only assessed the time frame associated with each domain once per participant. An essential characteristic of EMA is that it is comprised repeated assessments of the same constructs, and this feature of EMA studies may alter the time frames used as a study progresses. For this reason, the results presented here may only be generalizable to the first few EMA prompts. Future studies that inquire about participants' reporting time frames on many occasions would provide additional information as to whether participants improve their adherence to the desired time frame or drift into broader time frames throughout study participation. Conducting cognitive interviews in a random and intermittent manner within an existing EMA protocol may potentially be an opportunity to expand on this question. The fourth limitation is the generalizability of the results to participants recruited from the general public. The training effect may be more robust in MTurk workers than in the general population, possibly because MTurk workers tend to be better educated [27] and more attentive to instructions [28]. Future studies that involve participants from a more diverse pool of participants would be able to further expand on the training effect documented in this study.

Conclusions

In summary, this study provides evidence that participants do not reliably use the momentary time frame intended for EMA protocols when brief instructions are provided; rather, they provide evidence that respondents often appear to use longer periods. The results also indicate that training participants with detailed time frame definitions and providing opportunities to practice EMA reports during training substantially improved participants' adherence to the time frame instructions. Adherence levels in coverage EMA were not codable in this study; therefore, this remains to be a question for future research.

Conflicts of Interest

AAS is a senior scientist with the Gallup Organization and a consultant with Adelphi Values, Inc.

Multimedia Appendix 1

Actual training script for groups assigned to receive basic training for both momentary and coverage model conditions.

[[DOCX File , 14 KB - formative_v5i5e28007_app1.docx](#)]

Multimedia Appendix 2

Actual training script for groups assigned to enhanced training in the momentary model.

[[DOCX File , 15 KB - formative_v5i5e28007_app2.docx](#)]

Multimedia Appendix 3

Actual training script for groups assigned to enhanced training in the coverage model.

[[DOCX File , 16 KB - formative_v5i5e28007_app3.docx](#)]

References

1. Shiffman S, Stone AA, Hufford MR. Ecological momentary assessment. *Annu Rev Clin Psychol* 2008 Apr;4(1):1-32. [doi: [10.1146/annurev.clinpsy.3.022806.091415](#)] [Medline: [18509902](#)]
2. Stone AA, Shiffman S. Ecological momentary assessment (EMA) in behavioral medicine. *Ann Behav Med* 1994;16(3):199-202. [doi: [10.1093/abm/16.3.199](#)]
3. Csikszentmihalyi M, Larson R. Validity and reliability of the experience-sampling method. In: *Flow and the Foundations of Positive Psychology*. Dordrecht: Springer; 2014:35-54.
4. Csikszentmihalyi M, Larson R. Validity and reliability of the Experience-Sampling Method. *J Nerv Ment Dis* 1987 Sep;175(9):526-536. [doi: [10.1097/00005053-198709000-00004](#)] [Medline: [3655778](#)]
5. Affleck G, Zautra A, Tennen H, Armeli S. Multilevel daily process designs for consulting and clinical psychology: a preface for the perplexed. *J Consult Clin Psychol* 1999 Oct;67(5):746-754. [doi: [10.1037/0022-006x.67.5.746](#)]
6. Bolger N, Davis A, Rafaeli E. Diary methods: capturing life as it is lived. *Annu Rev Psychol* 2003 Feb;54(1):579-616. [doi: [10.1146/annurev.psych.54.101601.145030](#)] [Medline: [12499517](#)]
7. Smyth JM, Juth V, Ma J, Sliwinski M. A slice of life: ecologically valid methods for research on social relationships and health across the life span. *Soc Personal Psychol Compass* 2017 Sep 26;11(10):e12356. [doi: [10.1111/spc3.12356](#)]
8. Stone AA, Shiffman S, Atienza A, Nebeling L. *The Science of Real-Time Data Capture: Self-reports in health research*. New York: Oxford University Press; 2007:1-416.
9. Bradburn N, Rips L, Shevell S. Answering autobiographical questions: the impact of memory and inference on surveys. *Science* 1987 Apr 10;236(4798):157-161. [doi: [10.1126/science.3563494](#)] [Medline: [3563494](#)]
10. Schwarz N. Why researchers should think “real time”: a cognitive rationale. In: *Handbook of Research Methods for Studying Daily Life*. New York: Guilford Press; 2013:1-676.
11. Dunton GF, Liao Y, Dzibur E, Leventhal AM, Huh J, Gruenewald T, et al. Investigating within-day and longitudinal effects of maternal stress on children's physical activity, dietary intake, and body composition: protocol for the MATCH study. *Contemp Clin Trials* 2015 Jul;43:142-154 [FREE Full text] [doi: [10.1016/j.cct.2015.05.007](#)] [Medline: [25987483](#)]
12. Trull TJ, Solhan MB, Tragesser SL, Jahng S, Wood PK, Piasecki TM, et al. Affective instability: measuring a core feature of borderline personality disorder with ecological momentary assessment. *J Abnorm Psychol* 2008 Aug;117(3):647-661. [doi: [10.1037/a0012532](#)] [Medline: [18729616](#)]
13. Grange DL, Gorin A, Catley D, Stone AA. Does momentary assessment detect binge eating in overweight women that is denied at interview? *Eur Eat Disorders Rev* 2001 Sep;9(5):309-324. [doi: [10.1002/erv.409](#)]
14. Stein RI, Kenardy J, Wiseman CV, Douchis JZ, Arnow BA, Wilfley DE. What's driving the binge in binge eating disorder?: a prospective examination of precursors and consequences. *Int J Eat Disord* 2007 Apr;40(3):195-203. [doi: [10.1002/eat.20352](#)] [Medline: [17103418](#)]
15. Fischer S, Doerr JM, Strahler J, Mewes R, Thieme K, Nater UM. Stress exacerbates pain in the everyday lives of women with fibromyalgia syndrome--the role of cortisol and alpha-amylase. *Psychoneuroendocrinology* 2016 Jan;63:68-77. [doi: [10.1016/j.psyneuen.2015.09.018](#)] [Medline: [26431802](#)]
16. Burns JW, Gerhart JI, Bruehl S, Peterson KM, Smith DA, Porter LS, et al. Anger arousal and behavioral anger regulation in everyday life among patients with chronic low back pain: relationships to patient pain and function. *Health Psychol* 2015 May;34(5):547-555 [FREE Full text] [doi: [10.1037/hea0000091](#)] [Medline: [25110843](#)]
17. Hurlburt RT, Heavey CL. Investigating pristine inner experience: implications for experience sampling and questionnaires. *Conscious Cogn* 2015 Jan;31:148-159. [doi: [10.1016/j.concog.2014.11.002](#)] [Medline: [25486341](#)]
18. Stone AA, Wen CK, Schneider S, Junghaenel DU. Evaluating the effect of daily diary instructional phrases on respondents' recall time frames: survey experiment. *J Med Internet Res* 2020 Feb 21;22(2):e16105 [FREE Full text] [doi: [10.2196/16105](#)] [Medline: [32130129](#)]
19. Grice HP. Logic and conversation. *Speech Acts* 1975:41-58. [doi: [10.1163/9789004368811_003](#)]

20. Smith S, Amtmann D, Askew R, Gewandter JS, Hunsinger M, Jensen MP, et al. Pain intensity rating training: results from an exploratory study of the ACTION PROTECCT system. *Pain* 2016 May;157(5):1056-1064. [doi: [10.1097/j.pain.0000000000000502](https://doi.org/10.1097/j.pain.0000000000000502)] [Medline: [27058680](https://pubmed.ncbi.nlm.nih.gov/27058680/)]
21. Hurlburt RT. Iteratively apprehending pristine experience. *J Conscious Stud.* 2009. URL: <http://hurlburt.faculty.unlv.edu/hurlburt-2009c.pdf> [accessed 2021-05-13]
22. Heron K, Everhart R, McHale S, Smyth J. Using mobile-technology-based Ecological Momentary Assessment (EMA) methods with youth: a systematic review and recommendations. *J Pediatr Psychol* 2017 Nov 01;42(10):1087-1107. [doi: [10.1093/jpepsy/jsx078](https://doi.org/10.1093/jpepsy/jsx078)] [Medline: [28475765](https://pubmed.ncbi.nlm.nih.gov/28475765/)]
23. Campbell JL, Quincy C, Osserman J, Pedersen OK. Coding in-depth semistructured interviews. *Sociol Methods Res* 2013 Aug 21;42(3):294-320. [doi: [10.1177/0049124113500475](https://doi.org/10.1177/0049124113500475)]
24. McHugh ML. Interrater reliability: the kappa statistic. *Biochem Med* 2012;276-282. [doi: [10.11613/bm.2012.031](https://doi.org/10.11613/bm.2012.031)]
25. Walentynowicz M, Schneider S, Stone AA. The effects of time frames on self-report. *PLoS One* 2018 Aug 9;13(8):e0201655 [FREE Full text] [doi: [10.1371/journal.pone.0201655](https://doi.org/10.1371/journal.pone.0201655)] [Medline: [30092034](https://pubmed.ncbi.nlm.nih.gov/30092034/)]
26. Richman WL, Kiesler S, Weisband S, Drasgow F. A meta-analytic study of social desirability distortion in computer-administered questionnaires, traditional questionnaires, and interviews. *J Appl Psychol* 1999 Oct;84(5):754-775. [doi: [10.1037/0021-9010.84.5.754](https://doi.org/10.1037/0021-9010.84.5.754)]
27. Keith MG, Tay L, Harms PD. Systems perspective of Amazon Mechanical Turk for organizational research: review and recommendations. *Front Psychol* 2017 Aug 08;8:1359 [FREE Full text] [doi: [10.3389/fpsyg.2017.01359](https://doi.org/10.3389/fpsyg.2017.01359)] [Medline: [28848474](https://pubmed.ncbi.nlm.nih.gov/28848474/)]
28. Hauser DJ, Schwarz N. Attentive Turkers: MTurk participants perform better on online attention checks than do subject pool participants. *Behav Res Methods* 2016 Mar 12;48(1):400-407. [doi: [10.3758/s13428-015-0578-z](https://doi.org/10.3758/s13428-015-0578-z)] [Medline: [25761395](https://pubmed.ncbi.nlm.nih.gov/25761395/)]

Abbreviations

- EMA:** ecological momentary assessment
MTurk: Mechanical Turk
PT: Pacific Time

Edited by G Eysenbach; submitted 17.02.21; peer-reviewed by L König, V Boesen; comments to author 07.04.21; revised version received 14.04.21; accepted 17.04.21; published 26.05.21.

Please cite as:

Wen CKF, Junghaenel DU, Newman DB, Schneider S, Mendez M, Goldstein SE, Velasco S, Smyth JM, Stone AA
The Effect of Training on Participant Adherence With a Reporting Time Frame for Momentary Subjective Experiences in Ecological Momentary Assessment: Cognitive Interview Study
JMIR Form Res 2021;5(5):e28007
URL: <https://formative.jmir.org/2021/5/e28007>
doi: [10.2196/28007](https://doi.org/10.2196/28007)
PMID: [34037524](https://pubmed.ncbi.nlm.nih.gov/34037524/)

©Cheng K Fred Wen, Doerte U Junghaenel, David B Newman, Stefan Schneider, Marilyn Mendez, Sarah E Goldstein, Sarah Velasco, Joshua M Smyth, Arthur A Stone. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 26.05.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Patient-Provider Text Messaging and Video Calling Among Case-Managed Patients Living With HIV: Formative Acceptability and Feasibility Study

Virginia A Fonner^{1,2}, MPH, PhD; Samuel Kennedy², MPH, MDiv; Rohan Desai³, BSc; Christie Eichberg², BA; Lisa Martin², RN; Eric G Meissner², MD, PhD

¹Division of Global and Community Health, Department of Psychiatry and Behavioral Sciences, Medical University of South Carolina, Charleston, SC, United States

²Division of Infectious Diseases, Medical University of South Carolina, Charleston, SC, United States

³College of Medicine, Medical University of South Carolina, Charleston, SC, United States

Corresponding Author:

Virginia A Fonner, MPH, PhD

Division of Global and Community Health

Department of Psychiatry and Behavioral Sciences

Medical University of South Carolina

176 Croghan Spur

Suite 104

Charleston, SC, 29407

United States

Phone: 1 8438761463

Email: fonner@musc.edu

Abstract

Background: Patient-provider communication is critical for engaging and retaining people living with HIV in care, especially among medically case-managed patients in need of service coordination and adherence support. Expanding patient-provider communication channels to include mobile health modalities, such as text messaging and video calling, has the potential to facilitate communication and ultimately improve clinical outcomes. However, the implementation of these communication modalities in clinical settings has not been well characterized.

Objective: The purpose of this study is to understand patient and provider perspectives on the acceptability of and preferences for using text messaging and video calling as a means of communication; perceived factors relevant to adoption, appropriateness, and feasibility; and organizational perspectives on implementation within an HIV clinic in South Carolina.

Methods: We conducted 26 semistructured in-depth interviews among patients receiving case management services (n=12) and clinic providers (n=14) using interview guides and content analysis informed by the Proctor taxonomy of implementation outcomes and the Consolidated Framework for Implementation Research. Participants were purposefully sampled to obtain maximum variation in terms of age and gender for patients and clinic roles for providers. The data were analyzed using quantitative and qualitative content analyses.

Results: Most patients (11/12, 92%) and providers (12/14, 86%) agreed that they should have the capacity to text message and/or video call each other. Although consensus was not reached, most preferred using a secure messaging app rather than standard text messaging because of the enhanced security features. Perceived benefits to adoption included the added convenience of text messaging, and potential barriers included the cost and access of smartphone-based technology for patients. From an organizational perspective, some providers were concerned that offering text messaging could lead to unreasonable expectations of instant access and increased workload.

Conclusions: Patients and providers perceived text messaging and video calling as acceptable, appropriate, and feasible and felt that these expanded modes of communication could help meet patients' needs while being safe and not excessively burdensome. Although patients and providers mostly agreed on implementation barriers and facilitators, several differences emerged. Taking both perspectives into account when using implementation frameworks is critical for expanding mobile health-based communication, especially as implementation requires active participation from providers and patients.

KEYWORDS

HIV; mHealth; text messaging; video calling; implementation science; mobile phone

Introduction

Background

Provider communication with patients outside of office visits can be critical to retaining people living with HIV in care, which is vital to sustaining health and preventing onward HIV transmission. In the United States, particularly in the southeastern United States where the HIV epidemic is now concentrated [1,2], rates of retention in care remain suboptimal. In South Carolina, approximately 20,000 individuals are living with HIV, but only 53% are retained in care [3].

Traditionally, providers, including case managers who have frequent interactions with patients regarding all aspects of clinical care, have relied on landline telephones to reach patients. These efforts can consume a considerable amount of provider time [4], especially for care involving chronic diseases [5] such as HIV. However, as technology has evolved, traditional means of communication have faced significant challenges, including people's disinclination to answer a blocked or unfamiliar number and lack of a functioning voicemail box. Mobile technology has the potential to overcome these challenges and improve communication among patients experiencing barriers to care [6]. Enhancing patient-provider communication can work to build trust, foster more patient engagement, and improve health outcomes along the HIV care cascade [7-10].

Communication using mobile technology has become almost universal. Currently, 96% of American adults report owning a cell phone and 81% own a smartphone [11]. Mobile health (mHealth), defined as "medical and public health practice supported by mobile devices" [12], offers the potential to connect health care providers and patients, enhance in-clinic interactions, and improve care access. Studies have demonstrated that mHealth interventions implemented in HIV clinics are acceptable and can improve clinical outcomes [13-17]. However, many of these interventions either used automated one-way forms of communication (eg, appointment or medication reminders) or scripted two-way communication (eg, check-ins requiring confirmation of message receipt), and several studies have found no significant effect on clinical outcomes [18-20]. A recent systematic review found that more effective text messaging interventions allowed for interactivity, two-way communication, and links to support [21]. Several studies have also examined the feasibility and acceptability of video calling in the context of HIV, mostly to support telehealth visits [22], partner notification services [23], and counseling related to HIV testing [24]. There is a gap in understanding whether using mHealth modalities, including text messaging and video calling, simply as additional channels of communication between patients and providers in an unscripted, bidirectional manner could help improve communication without unduly increasing staff burden.

Objectives

Implementing mHealth-based communication in a clinical setting presents unique challenges, as these modalities require adoption by both providers and patients. Although some HIV-related mHealth interventions have described participant and provider perspectives on acceptability and feasibility [25-27], few have used an implementation science approach [28]. Implementation science focuses on understanding how evidence-based interventions can be integrated into health-related policies and practices [29,30]. In this study, we sought to understand patient and provider acceptability of and preferences for expanding communication channels to include mHealth modalities within an HIV clinic in South Carolina using an implementation science perspective. The specific research questions were as follows:

1. Are text messaging and video calling acceptable forms of communication for patients and providers, and what are the preferred modalities?
2. What are the perceived benefits and barriers to adoption, appropriateness, and feasibility of expanding patient-provider communication to include text messaging and video calling?
3. From an organizational perspective, what factors will influence the implementation process?

Methods

Context

The research was conducted within an academic medical institution in Charleston, South Carolina, housing a Ryan White HIV clinic that delivers comprehensive outpatient medical care for approximately 1200 people living with HIV. A multidisciplinary team works within the clinic to provide clinical care, social support, and case management services. Approximately 20% of the clinic population receives case management consisting of the coordination of medical and social services tailored to meet individual clients' needs. Patients are eligible for case management if they meet certain eligibility criteria established by the Ryan White Program (Part B), including having been diagnosed with HIV; income at or below 550% of the federal poverty level; and a demonstrated need for receiving assistance with social, community, legal, or financial barriers to care [31]. Case management has proven effective in improving clinical outcomes, including medication adherence and viral suppression [32].

This study was conducted as formative research to inform a subsequent intervention involving the provision of bidirectional text and video calling capabilities between case-managed patients and providers, specifically case managers and pharmacists, to enhance communication and ultimately improve retention in care and medication adherence.

Participants, Sampling, and Recruitment

Patients were eligible to participate if they were aged 18 years or above, were receiving medical case management services through the HIV clinic, and reported having a cell phone capable of text messaging and/or video calling. Patients were recruited for the study immediately following their routine clinic visit and were purposefully sampled to obtain maximum variation in terms of age, gender, and race. Compensation was provided to the patients in the form of a US \$50 gift card. Providers at the HIV clinic, including those providing medical services (eg, physicians, physician assistants, and pharmacists), social services (eg, case managers, social workers, and outreach coordinators), and administrative support (eg, nurse administrators, coordinators, and program support staff), were recruited via email. The providers were not compensated for their participation. The sample size was determined based on data saturation. The study was approved by the institutional review board of the Medical University of South Carolina. All participants provided verbal consent and gave permission to audiorecord the interviews.

Potential Intervention Modalities Under Study

As part of the formative research development process, we sought to understand whether patients and providers would prefer using the standard SMS text messaging feature that is available on all cell phones or using an encrypted, secure app available only on smartphones and computers. The research team chose to use the secure messaging app, QliqSOFT (QliqSOFT, Inc), hereafter referred to as Qliq, as it provides secure text messaging for use within health care settings and offers a Health Insurance Portability and Accountability Act (HIPAA)-compliant platform that is free for patients [33]. The app is passcode protected, includes the ability to delete messages if a device is lost or stolen, and keeps contacts separate from contacts stored on the mobile device. For video calling, participants were also presented with a choice of a *standard* option, such as using *FaceTime*, or use of an encrypted app, although no specific encrypted video calling app was specified.

Theoretical Orientation

We drew upon 2 implementation science frameworks to inform our study: (1) the Proctor taxonomy of implementation outcomes [34] and (2) the Consolidated Framework for Implementation Research (CFIR) [35]. We chose to focus on 4 of the 8 implementation outcomes outlined in the Proctor taxonomy, including acceptability (degree to which the intervention is agreeable, satisfactory, or has relative advantage), adoption (the initial decision, intent, or action to try an intervention), appropriateness (perceived intervention fit, relevance, or compatibility), and feasibility (extent to which an intervention can be conducted or successfully used) [34], as these were most salient to addressing our first 2 research questions and were most applicable to understanding factors relevant to implementation before intervention rollout. In addition, the Proctor taxonomy provides a flexible, practical framework for categorizing and understanding implementation aspects relevant to both providers and patients. We also used CFIR, a comprehensive, explanatory model of implementation addressing 5 major domains, including intervention

characteristics, outer setting, inner setting, characteristics of individuals, and process [35]. CFIR is particularly relevant to understanding intervention implementation from an organizational perspective, which comprises our third research question.

In-depth Interviews

Data collection consisted of semistructured in-depth interviews with both patients and providers. The interviews took place in a private room at the clinic and lasted for 30-60 minutes. Interview guides containing open-ended questions were developed with a focus on understanding preferences for, and acceptability of, a text-based intervention within the clinic (Multimedia Appendix 1). During the interviews, participants watched a brief demonstration video of Qliq and were asked about their initial thoughts and preferences of using Qliq versus SMS text messaging. Interview guides for providers were also developed using domains from the CFIR as a guide. In addition, at the end of the qualitative interviews, patients and providers were asked to verbally provide their responses to 4 statements using a 5-point Likert scale ranging from *strongly disagree* to *strongly agree*. The 4 statements were as follows: (1) "I would like to use this form of communication frequently," (2) "Providers and patients should have the capacity to text and/or video call with each other," (3) "Trying to implement a system to allow texting and/or video calling between patients and providers is too complex and not worth the time and risk," and (4) "Most patients would be interested in communicating by text and/or video with their case manager or pharmacist." All interviews were audiorecorded and transcribed verbatim by a third party.

Analysis

Data were analyzed using both qualitative and quantitative content analyses [36,37]. Data on preferences for intervention modality (standard SMS text messaging vs encrypted secure messaging) were extracted from interview transcripts and quantitatively tabulated. Similarly, the Likert-scale statements relating to acceptability asked at the end of the in-depth interviews were analyzed quantitatively, with scores between patients and providers averaged by question and compared using two-tailed *t* tests. Quantitative analysis was performed using Stata version 16 [38].

For the qualitative analysis, emerging themes and technology preferences were discussed by the study team during data collection. Following the completion of data collection, we followed the steps outlined in the qualitative content analysis, including data preparation, categorization, and reporting [36]. All interview transcripts were read and discussed by 2 members of the study team. For data categorization, a coding scheme was developed using both inductive and deductive methods. Open, inductive coding was applied to the data regarding acceptability, barriers and facilitators to adoption, appropriateness, and feasibility. These codes were then further categorized and mapped onto the Proctor implementation framework. A second layer of deductive codes was applied to the data generated from provider interviews using predefined categories and constructs from CFIR. Data were coded using Atlas.ti (version 8). Memos

were drafted to keep track of the categorization decisions, and the findings were iteratively discussed with the study team.

The research team consisted of a behavioral scientist trained in qualitative research methods (VAF), 2 research assistants (SK and CE) and a medical student (RD) with no affiliation to the HIV clinic and who received training in qualitative methods specific to the project, a research nurse who served as the study coordinator (LM), and a physician scientist who provided clinical care at the Medical University of South Carolina HIV clinic (EGM). Data collection and analysis were conducted by the study team members with no direct affiliation to the HIV clinic (VAF, SK, CE, and RD). Coding was conducted by a member of the research team (VAF), and a second member of the research team (RD) conducted an audit trail, which included

reviewing transcripts, coding schemes, and memos to ensure conformability and authenticity. To enhance the trustworthiness and transparency of the data, authentic citations drawn from interview transcripts were used to create tables to report the results ([Multimedia Appendices 2 and 3](#)).

Results

Sample Characteristics

We conducted 26 interviews with 12 patients and 14 providers ([Table 1](#)). The patients were predominantly African American (10/12, 83%), aged between 23 and 57 years, and all reported having access to a smartphone. Providers included medical providers (6/14, 43%), social support providers (5/14, 36%), and support staff (3/14, 21%).

Table 1. Sample characteristics.

Characteristic	Value
Patients (n=12)	
Gender, n (%)	
Male	5 (42)
Female	6 (50)
Gender fluid	1 (8)
Race, n (%)	
African American	10 (83)
Native American or White	1 (8)
White	1 (8)
Age (years), mean (SD)	39.5 (9.8)
Smartphone access, n (%)	
Yes	12 (100)
No	0 (0)
Years enrolled in clinic, mean (SD)	6.4 (4.3)
Providers (n=14), n (%)	
Role	
Medical provider	6 (43)
Social support provider	5 (36)
Support staff	3 (21)
Gender	
Male	3 (21)
Female	11 (79)

Acceptability

Most patients (11/12, 92%) and providers (12/14, 86%) either strongly agreed or agreed with the statement that they “should have the capacity to text and/or video call” each other. Similarly, participants mostly agreed that patients would be interested in

texting and/or video calling providers, specifically case managers and/or pharmacists, and felt that implementing texting messaging and/or video calling was “worth the time and risk” ([Table 2](#)). No statistical differences were found between patients’ and providers’ perceptions of acceptability.

Table 2. Acceptability of text messaging and video calling between patients and providers.

Statement	Participants, mean (SD) ^a	Providers, mean (SD) ^a	P value
"I would like to use this form of communication frequently"	4.4 (0.9)	4.2 (1.0)	.48
"Providers and patients should have the capacity to text and/or video call with each other"	4.4 (0.8)	4.0 (0.9)	.22
"Trying to implement a system to allow texting and/or video calling between patients and providers is too complex and not worth the time and risk"	1.7 (0.8)	1.9 (0.5)	.30
"Most patients would be interested in communicating by text and/or video with their case manager or pharmacist"	4.3 (0.6)	4.5 (0.6)	.52

^aLikert scale ranged from 1 (strongly disagree) to 5 (strongly agree).

In general, patients felt that communication with their health care team was already quite good and that being able to text or video call would further enhance their ability to connect. When describing the potential benefits of being able to text with a case manager, one participant said:

Texting is fast, go...it gets straight good, easy for you to reply back. Any questions, just [text]...and then it probably might open up the dialogue a little bit more.

Providers, especially social support providers, expressed frustration with their ability to reach patients using landline telephones. One social service provider said:

There are so many patients that I have...where I am not able to get ahold of because they don't have voicemail set up. Their voicemail is full. They might have called me, but when I try to call them back, I can't get them for whatever reason. So that's very frustrating. It's very time consuming. I want to be an efficient [position redacted], and when doing a lot of phone tag takes up a bulk of your day when you have other patients waiting on you to get to them, it's just a nuisance.

Providers mostly felt that having the ability to text patients would help alleviate this challenge.

Preferences

With regard to preferring SMS text messaging to Qlik, the secure messaging app, 83% (10/12) of patients endorsed a preference for Qlik (Table 3). For providers, preferences were more varied.

Table 3. Patients' and providers' preferences for SMS text messaging versus Qlik.

Provider preference for text messaging modality	Patients' preferences (n=12), n (%)	Providers' preferences (n=14), n (%)
Prefers SMS text messaging	2 (17)	3 (21)
Prefers secure app (Qlik)	10 (83)	5 (37)
Prefers use of the web-based patient portal (MyChart)	0 (0)	2 (14)
Prefers that both methods be available (SMS text messaging and Qlik)	0 (0)	2 (14)
No preference	0 (0)	2 (14)

Perceived Benefits and Barriers to Adoption

Patients and providers agreed on the perceived benefits of text messaging, which included efficiency, convenience, and ease of use, irrespective of modality (Qlik vs SMS text messaging; Multimedia Appendix 2). The ability to engage in asynchronous

Qlik was favored over SMS text messaging (5/14, 37% vs 3/14, 21%, respectively), although a sizable proportion either had no preference (2/14, 14%) or a preference for having both SMS text messaging and Qlik available (2/14, 14%) to reach as many patients as possible. In other words, providers, specifically most social service providers, were deferential toward patients' preferences. One social support provider voiced:

I just think that I'm kind of torn between both [using Qlik or SMS]...I think it's more about privacy for the client, but it's also convenient for us.

In addition, 14% (2/14) of providers felt that adding a text messaging platform as a form of communication, including either SMS text messaging or Qlik, was duplicative of the secure messaging function built into the existing patient portal to which patients already had access, called MyChart. Those favoring the patient portal were primarily medical providers who frequently interacted with MyChart as part of their clinical duties. Of those favoring Qlik, both patients and providers highlighted the security features and enhanced privacy, and patients liked other features of Qlik, such as the ability to see when a message had been read. Among those favoring SMS text messaging, convenience was cited as the primary reason. Patients' and providers' preferences for SMS text messaging versus Qlik mirrored their preference for video calls, that is, those favoring SMS text messaging typically favored using FaceTime, and those favoring Qlik typically preferred a similar secure app for video calls. Some patients and most medical providers did not see a need for video calling and therefore did not have a preference for either method.

conversations through text messaging was seen as a clear advantage by both groups. Patients specifically cited the inability to answer calls while at work and hesitation to answer calls because of privacy concerns related to HIV. One participant stated:

As a patient, we go through our day with people and doing things. And people don't really know about our situations. So a lot of people feel, when they do get a pop-up phone call, they can't really answer the phone and speak freely because of who they're around.

Although providers also felt that having the ability to text would be efficient and convenient, some worried that these benefits could become drawbacks by creating unreasonable expectations of instant access and overutilization by patients, thus leading to increased workload. A few other barriers were mentioned. Some patients and providers cited the impersonality of texting as a barrier. As a medical provider voiced, "people don't see my concern over text." For this reason, a few participants preferred video calling over text messaging. One participant said:

Video chat I think would be better, like I said, that way you can visually see them and try to get body language and stuff like that, communicating, that's means a lot.

Some providers and patients also mentioned cost and cell phone access, specifically smartphone access, as potential barriers.

Appropriateness

From both patients' and providers' perspectives, allowing text messaging and video calling was perceived as appropriate because of its potential ability to improve communication while being relatively safe and not overly burdensome ([Multimedia Appendix 2](#)). Most patients articulated a desire for texting, and many providers were enthusiastic about reaching patients using this modality, as it was perceived to be mutually beneficial. When describing her preferences for text messaging (Qliq vs SMS), one social support provider said:

Whatever will make my work life easier. Again, when it comes to the client, it's about making sure that they get what they need.

The appropriateness of video calling was less clear for both the patients and providers. However, for providers who field questions about medication and adherence, such as pharmacists, video calling was perceived as offering the distinct benefit of being able to see which pill a patient had questions about and to directly observe a patient taking their medication in cases where daily adherence was in question. Some social service providers also felt that video calling would be useful for providing social support and counseling, as it would enable visual contact.

Feasibility

Two issues regarding feasibility emerged: (1) privacy or security issues and (2) concern over message content ([Multimedia Appendix 2](#)). With regard to privacy and security, patients and providers were well aware of the stigmatizing nature of HIV and were in agreement that nothing directly mentioning HIV should be sent over text. Providers were concerned about patients sending protected health information over text, whereas patients did not share this concern. One social support provider explained:

...You can control what you send out to patients, but you can't control what patients send back to you.

Patients mostly trusted their health care providers to be discreet, although there was some concern about how their contact details would be stored in the providers' phones and whether the providers' phones would be used outside of the office setting. Some patients also had privacy concerns over video calling because of the potential of being overheard when discussing sensitive health-related matters. For text messaging, patients favored Qliq, the secure app, in large part because of the enhanced security features. One patient stated:

You have kids, you have family members who doodle in your phone. They can read the text, everybody else's text. But when it comes to your health and your privacy, there should be an app that is created just for that.

With regard to message content, providers agreed that texting should be used for short, simple messages, such as checking-in with patients or confirming logistics for transportation to a clinic visit. Patients mostly agreed to use text messaging for such short messages, although some thought it would be helpful if clinical information, such as laboratory results, could also be sent over text.

CFIR-Specific Implementation Constructs

Provider-specific themes relevant to CFIR mainly fell into 3 domains: (1) intervention characteristics; (2) the inner setting, specifically the implementation climate; and (3) the outer setting, specifically patient needs and resources ([Multimedia Appendix 3](#)). With regard to intervention characteristics, providers, specifically social support providers, mostly felt that texting offered a significant relative advantage to the status quo in that it would both reduce the *phone-tag* burden and would increase efficiency. However, some worried that having the ability to text and video call would be complex to implement, as it could lead to significant disruptions and task them with yet another mode of communication to manage throughout the day. Providers felt that the source of the intervention was internal and had been developed to respond to patient needs and provider frustrations while trying to communicate with patients using traditional modalities.

With regard to the inner setting, there was a discernable tension for change raised by a number of providers, specifically those associated with providing case management and outreach services. These providers already had access to basic cell phones (*flip phones*) for patient communication but found texting on these devices cumbersome; hence, providers felt that smartphone access was needed to facilitate text-based communication. Importantly, one provider felt that using a secure app for text messaging with patients would be *overkill* and unnecessarily complicate the process of text messaging. However, most providers felt that the enhanced security measures of Qliq were of value, both to themselves and patients, and offered benefits over standard SMS text messaging.

With regard to the outer setting, providers, especially social support providers, were keenly aware of patient needs and perceived text-based communication as a means to better meet

patient needs. Providers' perceptions of patients' needs were mostly accurate, although a few differences emerged. Several providers felt that patients would prioritize convenience over security and were worried about those without smartphone access. One social support provider said:

I think also knowing the majority of our patients and their tech savvy skills, I think texting [SMS] would be much easier for them...not all of them [patients] are gonna have smartphones, unfortunately.

In addition, several providers felt that patients already had the ability to send secure text messages to providers through the web-based patient portal and thus viewed expanding text messaging capabilities as redundant. However, patients consistently cited logistic difficulties using the patient portal (eg, trouble remembering passwords). One patient voiced:

I see [text messaging] as a better improvement 'cause when I go through the patient care app like every six months you gotta get a new password and it's like, "I don't want to get a new password. Just show me when the next doctor's appointment is coming up."

Some providers were aware of these difficulties and were concerned that using a secure app for texting would result in similar logistical challenges.

Discussion

Principal Findings

In this formative study of 26 patients living with HIV and HIV care providers within a South Carolina-based clinic, we found high acceptability and feasibility for expanding patient-provider communication channels to include text messaging and video calling. Below, we discuss the implications of our findings for each specific research question.

Acceptability and Preferences for mHealth Communication Modalities

The high acceptability of text messaging found in our study resonates with the findings from prior mHealth interventions implemented in HIV clinics that used either SMS or a form of secure messaging [20,26,27,39]. To our knowledge, our study is one of the first to directly compare preferences for standard versus secure text messaging among patients living with HIV and providers. Although we did not find unanimous consensus, patients' and providers' overall preference for using secure messaging because of privacy concerns mirrors prior insights regarding technology use, stigma, and privacy among people living with HIV [40]. However, enhanced security comes at the cost of additional requirements (eg, smartphone use) and logistical hurdles (eg, passwords), which have been identified as barriers to technology use [41,42]. Therefore, there is a need to consider issues surrounding security, access, and convenience when developing best practices related to the use of text messaging within HIV clinics. Recently, efforts have been made to clarify best practices regarding text messaging use in health care settings [43-45]. For the intervention informed by this formative study, the clinic ultimately chose to use the secure messaging app but also allowed for standard SMS text

messaging for patients without smartphones to maximize both security and access.

Benefits and Barriers to Adoption, Appropriateness, and Feasibility

We found that patients and providers agreed that text messaging could be a faster, more efficient mode of communication and seemed appropriate given its potential to increase patient engagement. There was less consensus on the benefits of video calling, although both patients and providers saw its utility in specific circumstances, such as allowing for directly observed therapy and fostering social support in times of need. Our findings are supported by other studies that have examined patients' and providers' perspectives on mHealth among people living with HIV, specifically with regard to the perceived benefits of convenience and ease of use and the perceived barriers of cost or access and confidentiality or security concerns [25,28].

What is novel from our findings is that providers' acceptance of and enthusiasm for using the technology was often contingent upon patients' ability and willingness to use it. With regard to technology preference, many providers were deferential to patients' preferred modality (SMS text messaging vs Qlik). In addition, patients factored in providers' use of the technology (eg, storage of contact details in provider phone and location of phone use) into their assessment of feasibility.

Given our study's findings of patient-provider interdependence related to intervention acceptability, more theoretical understanding is needed concerning the implementation of mHealth interventions dependent upon patient-provider interaction. Although many implementation science frameworks include a construct related to patient needs, few, if any, elevate the importance of this construct to the forefront. One recent study to understand the implementation of patient-centered care transformation interventions made adjustments to CFIR throughout their analysis, including promoting *patient needs as resources* to its own domain as opposed to a subdomain of the *outer setting*, which investigators felt reflected the prioritization of patients and patient satisfaction currently being emphasized within health-related interventions [46]. Although this change is notable, gaps remain in understanding health interventions whose effectiveness depends on uptake and use among both patients and providers and on the interplay between them.

Organizational Factors Relevant for Implementation

Providers perceived that adding text messaging and video calling as communication channels was an internally generated idea to overcome frustrations with the existing channels (eg, using landline phones). Some staff were concerned that the very features that made text messaging appealing, such as its convenience and ease of use, could create false expectations of instant access and overutilization, thus creating implementation challenges by disrupting existing workflows and increasing staff burden. However, most felt that the benefits of improved patient engagement and communication outweighed the potential drawbacks.

Results from our formative study show that patients' and providers' intentions to use text messaging and video calling

exist on a continuum, highlighting that one size may not fit all. For example, although many providers were enthusiastic, several providers expressed reticence to use text messaging and/or video calling with patients, mostly as a result of failing to see the advantage over existing platforms or because of concerns that the intervention would increase workload and reduce efficiency. However, if interventions such as offering text messaging and/or video calling were to be rolled out clinic wide, uneven use of the intervention might result in technology preference mismatches, for example, patients who prefer texting assigned to providers who prefer to call and vice versa. This finding suggests that implementation strategies are needed to ensure that patients and providers are aware of and can access all possible communication methods, paying specific attention to patient preferences.

Limitations

Our study has several limitations. The generalizability of our findings might be limited, given that the data came from one specific HIV clinic in South Carolina. However, provider and patient participants were purposefully chosen to maximize variation in terms of professional position and sociodemographic characteristics, respectively, to capture a diverse set of experiences and perceptions. Although not formally tracked, almost all patients and providers who were invited to join the study agreed to participate. Our sample predominantly consisted of African American patients, which reflects the racial makeup of the clinic population more generally and the disproportionate burden of HIV found among this group. Contributing to this limitation, our sample size was relatively small, and we were unable to tease out the differences in patient perspectives by specific characteristics, such as age, race, and place of residence (urban vs rural), despite known disparities in technology use across such factors [47-49]. This limitation also highlights the need to recognize that preferences for communication differ among individuals, and allowing choice is important.

In addition, patient participants were selected from case-managed patients who were physically present at the clinic

to attend their regular medical appointments. Although all case-managed patients demonstrated a need for supportive services to help them stay engaged in care, we did not specifically sample patients according to their medication adherence status or past clinic attendance. We are also missing perspectives from patients who are disengaged from medical care—a critical group to reach. However, we did receive provider perspectives from outreach staff who specifically work to reengage patients in care and who spoke on the potential benefits and challenges of using text messaging and video calling to communicate among this population. Owing to logistical constraints, we were unable to return transcripts for member checking, although the findings were informally discussed with the providers during intervention development. In addition, all patients in our sample had access to smartphones, whereas smartphone access is not universal among the clinical population. Finally, our study took place before the implementation of the text messaging and video calling intervention; therefore, the findings are based on perceptions of hypothetical use and barriers to implementation. Implementation findings gathered after the completion of the intervention will help convey a more complete understanding.

Conclusions

We found broad enthusiasm, acceptability, and feasibility for the implementation of bidirectional text messaging between patients and providers, with a preference for doing so using secure means. There was less consensus about the appropriateness of video calling, although both groups acknowledged its utility in certain circumstances. Engagement in text messaging and video calling within an HIV clinic setting requires buy-in from patients and providers. The expansion of communication channels will only be effective in enhancing communication if the services are used by both parties. Implementing a bidirectional intervention raises important questions for implementation research, as it requires a thorough understanding of the organizational, provider, and patient needs and finding implementation strategies and options that work for all.

Acknowledgments

This study was sponsored by Viiv Healthcare. Viiv Healthcare reviewed and approved this manuscript before submission but was not involved in data collection or analysis. The authors sincerely thank all participants and clinic staff who gave their time and participated in this study. This publication was also supported in part by the Health Resources and Services Administration of the US Department of Health and Human Services as part of the National Telehealth Center of Excellence Award (U66 RH31458) and the South Carolina Clinical & Translational Research Institute (National Institutes of Health/National Center for Advancing Translational Sciences grant UL1TR001450). The contents are those of the authors and do not necessarily represent the official views of, nor an endorsement, by Health Resources and Services Administration, Health and Human Services, or the US Government.

Authors' Contributions

EGM and VAF designed the study. VAF conducted the primary analysis and prepared the initial draft of the results. RD conducted a review of the raw data and an audit trail. SK and CE conducted the interviews, and LM served as the study coordinator and was responsible for participant recruitment, informed consent, and remuneration. All the authors reviewed and provided input on the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Semistructured in-depth interview guides.

[[DOCX File , 19 KB - formative_v5i5e22513_app1.docx](#)]

Multimedia Appendix 2

Exemplary quotes related to implementation outcomes using Proctor's taxonomy.

[[DOCX File , 15 KB - formative_v5i5e22513_app2.docx](#)]

Multimedia Appendix 3

Exemplary quotes related to Consolidated Framework for Implementation Research-specific domains and constructs.

[[DOCX File , 15 KB - formative_v5i5e22513_app3.docx](#)]

References

1. Reif SS, Whetten K, Wilson ER, McAllaster C, Pence BW, Legrand S, et al. HIV/AIDS in the Southern USA: a disproportionate epidemic. *AIDS Care* 2014 Aug 14;26(3):351-359. [doi: [10.1080/09540121.2013.824535](https://doi.org/10.1080/09540121.2013.824535)] [Medline: [23944833](https://pubmed.ncbi.nlm.nih.gov/23944833/)]
2. HIV in the United States by region 2019. Centers for Disease Control and Prevention. 2019. URL: <https://www.cdc.gov/hiv/statistics/overview/geographicdistribution.html> [accessed 2021-04-29]
3. An epidemiologic profile of HIV and AIDS in South Carolina 2018. Division of Surveillance and Technical Support Bureau of Disease Control and South Carolina Department of Health and Environmental Control. 2018. URL: <https://scdhec.gov/sites/default/files/media/document/2018%20Epi%20Profile.pdf> [accessed 2021-04-29]
4. Chen MA, Hollenberg JP, Michelen W, Peterson JC, Casalino LP. Patient care outside of office visits: a primary care physician time study. *J Gen Intern Med* 2011 Jan 2;26(1):58-63 [FREE Full text] [doi: [10.1007/s11606-010-1494-7](https://doi.org/10.1007/s11606-010-1494-7)] [Medline: [20811956](https://pubmed.ncbi.nlm.nih.gov/20811956/)]
5. Farber J, Siu A, Bloom P. How much time do physicians spend providing care outside of office visits? *Ann Intern Med* 2007 Nov 20;147(10):693-698. [doi: [10.7326/0003-4819-147-10-200711200-00005](https://doi.org/10.7326/0003-4819-147-10-200711200-00005)] [Medline: [18025445](https://pubmed.ncbi.nlm.nih.gov/18025445/)]
6. Henny KD, Wilkes AL, McDonald CM, Denson DJ, Neumann MS. A rapid review of eHealth interventions addressing the continuum of HIV Care (2007-2017). *AIDS Behav* 2018 Jan 5;22(1):43-63 [FREE Full text] [doi: [10.1007/s10461-017-1923-2](https://doi.org/10.1007/s10461-017-1923-2)] [Medline: [28983684](https://pubmed.ncbi.nlm.nih.gov/28983684/)]
7. Flickinger T, Saha S, Moore R, Beach M. Higher quality communication and relationships are associated with improved patient engagement in HIV care. *J Acquir Immune Defic Syndr* 2013 Jul 01;63(3):362-366 [FREE Full text] [doi: [10.1097/QAI.0b013e318295b86a](https://doi.org/10.1097/QAI.0b013e318295b86a)] [Medline: [23591637](https://pubmed.ncbi.nlm.nih.gov/23591637/)]
8. Johnson BT, Redding CA, DiClemente RJ, Mustanski BS, Dodge B, Sheeran P, et al. A network-individual-resource model for HIV prevention. *AIDS Behav* 2010 Dec 23;14(Suppl 2):204-221 [FREE Full text] [doi: [10.1007/s10461-010-9803-z](https://doi.org/10.1007/s10461-010-9803-z)] [Medline: [20862606](https://pubmed.ncbi.nlm.nih.gov/20862606/)]
9. Brion J. The patient-provider relationship as experienced by a diverse sample of highly adherent HIV-infected people. *J Assoc Nurses AIDS Care* 2014;25(2):123-134. [doi: [10.1016/j.jana.2013.01.006](https://doi.org/10.1016/j.jana.2013.01.006)] [Medline: [23809659](https://pubmed.ncbi.nlm.nih.gov/23809659/)]
10. Finitis DJ, Pellowski JA, Johnson BT. Text message intervention designs to promote adherence to antiretroviral therapy (ART): a meta-analysis of randomized controlled trials. *PLoS One* 2014 Feb 5;9(2):e88166 [FREE Full text] [doi: [10.1371/journal.pone.0088166](https://doi.org/10.1371/journal.pone.0088166)] [Medline: [24505411](https://pubmed.ncbi.nlm.nih.gov/24505411/)]
11. Mobile fact sheet. Pew Research Center. 2021. URL: <https://www.pewresearch.org/internet/fact-sheet/mobile/> [accessed 2019-04-29]
12. World Health Organization. mHealth: New Horizons for Health Through Mobile Technologies: Second Global Survey on eHealth. Geneva: World Health Organization; 2011.
13. Escobar-Viera C, Zhou Z, Morano JP, Lucero R, Lieb S, McIntosh S, et al. The Florida Mobile Health Adherence Project for People Living With HIV (FL-mAPP): longitudinal assessment of feasibility, acceptability, and clinical outcomes. *JMIR Mhealth Uhealth* 2020 Jan 08;8(1):e14557 [FREE Full text] [doi: [10.2196/14557](https://doi.org/10.2196/14557)] [Medline: [31913127](https://pubmed.ncbi.nlm.nih.gov/31913127/)]
14. McNairy ML, Lamb MR, Gachuhi AB, Nuwagaba-Biribonwoha H, Burke S, Mazibuko S, et al. Effectiveness of a combination strategy for linkage and retention in adult HIV care in Swaziland: the Link4Health cluster randomized trial. *PLoS Med* 2017 Nov 7;14(11):e1002420 [FREE Full text] [doi: [10.1371/journal.pmed.1002420](https://doi.org/10.1371/journal.pmed.1002420)] [Medline: [29112963](https://pubmed.ncbi.nlm.nih.gov/29112963/)]
15. Lester RT, Ritvo P, Mills EJ, Kariri A, Karanja S, Chung MH, et al. Effects of a mobile phone short message service on antiretroviral treatment adherence in Kenya (WelTel Kenya1): a randomised trial. *Lancet* 2010 Nov 27;376(9755):1838-1845. [doi: [10.1016/S0140-6736\(10\)61997-6](https://doi.org/10.1016/S0140-6736(10)61997-6)] [Medline: [21071074](https://pubmed.ncbi.nlm.nih.gov/21071074/)]

16. Hirsch-Moverman Y, Daftary A, Yuengling K, Saito S, Ntoane M, Frederix K, et al. Using mHealth for HIV/TB treatment support in Lesotho: enhancing patient-provider communication in the START study. *J Acquir Immune Defic Syndr* 2017 Jan 01;74 Suppl 1:37-43 [FREE Full text] [doi: [10.1097/QAI.0000000000001202](https://doi.org/10.1097/QAI.0000000000001202)] [Medline: [27930610](https://pubmed.ncbi.nlm.nih.gov/27930610/)]
17. Rana AI, van den Berg JJ, Lamy E, Beckwith CG. Using a mobile health intervention to support HIV treatment adherence and retention among patients at risk for disengaging with care. *AIDS Patient Care STDS* 2016 Apr;30(4):178-184 [FREE Full text] [doi: [10.1089/apc.2016.0025](https://doi.org/10.1089/apc.2016.0025)] [Medline: [27028183](https://pubmed.ncbi.nlm.nih.gov/27028183/)]
18. Christopoulos K, Riley E, Carrico A, Tulsy J, Moskowitz J, Dilworth S. A randomized controlled trial of a text messaging intervention to promote virologic suppression and retention in care in an urban safety-net human immunodeficiency virus clinic: the Connect4Care trial. *Clin Infect Dis* 2018;67(5):4001. [doi: [10.1093/cid/ciy156](https://doi.org/10.1093/cid/ciy156)]
19. Mbuagbaw L, Thabane L, Ongolo-Zogo P, Lester RT, Mills EJ, Smieja M, et al. The Cameroon Mobile Phone SMS (CAMPS) trial: a randomized trial of text messaging versus usual care for adherence to antiretroviral therapy. *PLoS One* 2012 Dec 6;7(12):e46909 [FREE Full text] [doi: [10.1371/journal.pone.0046909](https://doi.org/10.1371/journal.pone.0046909)] [Medline: [23236345](https://pubmed.ncbi.nlm.nih.gov/23236345/)]
20. Mao L, Buchanan A, Wong HT, Persson A. Beyond mere pill taking: SMS reminders for HIV treatment adherence delivered to mobile phones of clients in a community support network in Australia. *Health Soc Care Community* 2018 Jul 15;26(4):486-494. [doi: [10.1111/hsc.12544](https://doi.org/10.1111/hsc.12544)] [Medline: [29336111](https://pubmed.ncbi.nlm.nih.gov/29336111/)]
21. Shah R, Watson J, Free C. A systematic review and meta-analysis in the effectiveness of mobile phone interventions used to improve adherence to antiretroviral therapy in HIV infection. *BMC Public Health* 2019 Jul 09;19(1):915 [FREE Full text] [doi: [10.1186/s12889-019-6899-6](https://doi.org/10.1186/s12889-019-6899-6)] [Medline: [31288772](https://pubmed.ncbi.nlm.nih.gov/31288772/)]
22. Dandachi D, Dang BN, Lucari B, Teti M, Giordano TP. Exploring the attitude of patients with HIV about using telehealth for HIV care. *AIDS Patient Care STDS* 2020 Apr 01;34(4):166-172. [doi: [10.1089/apc.2019.0261](https://doi.org/10.1089/apc.2019.0261)] [Medline: [32324481](https://pubmed.ncbi.nlm.nih.gov/32324481/)]
23. Udeagu CN, Shah S, Toussaint MM, Pickett L. Sociodemographic differences in clients preferring video-call over in-person interview: a pilot study of HIV tele-partner services. *AIDS Behav* 2017 Nov 17;21(11):3078-3086. [doi: [10.1007/s10461-016-1586-4](https://doi.org/10.1007/s10461-016-1586-4)] [Medline: [27752874](https://pubmed.ncbi.nlm.nih.gov/27752874/)]
24. Maksut JL, Eaton LA, Siembida EJ, Driffin DD, Baldwin R. A test of concept study of at-home, self-administered HIV testing with web-based peer counseling via video chat for men who have sex with men. *JMIR Public Health Surveill* 2016 Dec 14;2(2):e170 [FREE Full text] [doi: [10.2196/publichealth.6377](https://doi.org/10.2196/publichealth.6377)] [Medline: [27974287](https://pubmed.ncbi.nlm.nih.gov/27974287/)]
25. Baranoski AS, Meuser E, Hardy H, Closson EF, Mimiaga MJ, Safren SA, et al. Patient and provider perspectives on cellular phone-based technology to improve HIV treatment adherence. *AIDS Care* 2014 Jan 06;26(1):26-32 [FREE Full text] [doi: [10.1080/09540121.2013.802282](https://doi.org/10.1080/09540121.2013.802282)] [Medline: [23742640](https://pubmed.ncbi.nlm.nih.gov/23742640/)]
26. Swendeman D, Farmer S, Mindry D, Lee SJ, Medich M. HIV care providers' attitudes regarding mobile phone applications and web-based dashboards to support patient self-management and care coordination: results from a qualitative feasibility study. *J HIV AIDS* 2016 Oct;2(4):A [FREE Full text] [doi: [10.16966/2380-5536.127](https://doi.org/10.16966/2380-5536.127)] [Medline: [28066820](https://pubmed.ncbi.nlm.nih.gov/28066820/)]
27. Saberi P, Siedle-Khan R, Sheon N, Lightfoot M. The use of mobile health applications among youth and young adults living with HIV : focus group findings. *AIDS Patient Care STDS* 2016 Jun;30(6):254-260 [FREE Full text] [doi: [10.1089/apc.2016.0044](https://doi.org/10.1089/apc.2016.0044)] [Medline: [27214751](https://pubmed.ncbi.nlm.nih.gov/27214751/)]
28. Kemp CG, Velloza J. Implementation of eHealth interventions across the HIV care cascade: a review of recent research. *Curr HIV/AIDS Rep* 2018 Dec 1;15(6):403-413 [FREE Full text] [doi: [10.1007/s11904-018-0415-y](https://doi.org/10.1007/s11904-018-0415-y)] [Medline: [30171519](https://pubmed.ncbi.nlm.nih.gov/30171519/)]
29. Schackman B. Implementation science for the prevention and treatment of HIV/AIDS. *J Acquir Immune Defic Syndr* 2010 Dec;55 Suppl 1(Suppl 1):27-31 [FREE Full text] [doi: [10.1097/QAI.0b013e3181f9c1da](https://doi.org/10.1097/QAI.0b013e3181f9c1da)] [Medline: [21045596](https://pubmed.ncbi.nlm.nih.gov/21045596/)]
30. Glasgow R, Eckstein E, Elzarrad M. Implementation science perspectives and opportunities for HIV/AIDS research: integrating science, practice, and policy. *J Acquir Immune Defic Syndr* 2013 Jun 01;63 Suppl 1:26-31. [doi: [10.1097/QAI.0b013e3182920286](https://doi.org/10.1097/QAI.0b013e3182920286)] [Medline: [23673882](https://pubmed.ncbi.nlm.nih.gov/23673882/)]
31. SC Ryan White Part B-HIV/AIDS Program: medical case management standards. SC Department of Health and Environmental Control. 2018. URL: https://scdhc.gov/sites/default/files/docs/Health/docs/stdhiv/RyanWhite/PartB/Ryan%20White%20Part%20B%20Medical%20Case%20Management%20Standards_Final_6-27-18.pdf [accessed 2021-04-29]
32. López JD, Shacham E, Brown T. The impact of the Ryan White HIV/AIDS Medical Case Management Program on HIV clinical outcomes: a longitudinal study. *AIDS Behav* 2018 Sep 24;22(9):3091-3099. [doi: [10.1007/s10461-018-2124-3](https://doi.org/10.1007/s10461-018-2124-3)] [Medline: [29691681](https://pubmed.ncbi.nlm.nih.gov/29691681/)]
33. QliqSOFT. URL: <https://www.qliqsoft.com/> [accessed 2021-04-29]
34. Proctor E, Silmere H, Raghavan R, Hovmand P, Aarons G, Bunger A, et al. Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. *Adm Policy Ment Health* 2011 Mar 19;38(2):65-76 [FREE Full text] [doi: [10.1007/s10488-010-0319-7](https://doi.org/10.1007/s10488-010-0319-7)] [Medline: [20957426](https://pubmed.ncbi.nlm.nih.gov/20957426/)]
35. 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](https://doi.org/10.1186/1748-5908-4-50)] [Medline: [19664226](https://pubmed.ncbi.nlm.nih.gov/19664226/)]
36. Elo S, Kyngäs H. The qualitative content analysis process. *J Adv Nurs* 2008 Apr;62(1):107-115. [doi: [10.1111/j.1365-2648.2007.04569.x](https://doi.org/10.1111/j.1365-2648.2007.04569.x)] [Medline: [18352969](https://pubmed.ncbi.nlm.nih.gov/18352969/)]
37. White MD, Marsh EE. Content analysis: a flexible methodology. *Libr Trends* 2006;55(1):22-45. [doi: [10.1353/lib.2006.0053](https://doi.org/10.1353/lib.2006.0053)]

38. Stata Statistical Software: Release 16. StataCorp LLC. 2019. URL: <https://www.stata.com/> [accessed 2021-04-29]
39. Murray MC, O'Shaughnessy S, Smillie K, Van Borek N, Graham R, Maan EJ, WelTel BC1 Study Team. Health care providers' perspectives on a weekly text-messaging intervention to engage HIV-positive persons in care (WelTel BC1). *AIDS Behav* 2015 Oct 22;19(10):1875-1887. [doi: [10.1007/s10461-015-1151-6](https://doi.org/10.1007/s10461-015-1151-6)] [Medline: [26297567](https://pubmed.ncbi.nlm.nih.gov/26297567/)]
40. Sherman E, Clauson KA, Elrod S, Eckardt P, Alkhateeb F, Jacobs R. Patient perceptions of text messaging to improve antiretroviral therapy adherence: a qualitative study of patients in a Ryan White Clinic. *J Int Assoc Provid AIDS Care* 2018 Feb 23;17:A [FREE Full text] [doi: [10.1177/2325958218759209](https://doi.org/10.1177/2325958218759209)] [Medline: [29473483](https://pubmed.ncbi.nlm.nih.gov/29473483/)]
41. Portz JD, Bayliss EA, Bull S, Boxer RS, Bekelman DB, Gleason K, et al. Using the technology acceptance model to explore user experience, intent to use, and use behavior of a patient portal among older adults with multiple chronic conditions: descriptive qualitative study. *J Med Internet Res* 2019 Apr 08;21(4):e11604 [FREE Full text] [doi: [10.2196/11604](https://doi.org/10.2196/11604)] [Medline: [30958272](https://pubmed.ncbi.nlm.nih.gov/30958272/)]
42. Shimada SL, Zocchi MS, Hogan TP, Kertesz SG, Rotondi AJ, Butler JM, et al. Impact of patient-clinical team secure messaging on communication patterns and patient experience: randomized encouragement design trial. *J Med Internet Res* 2020 Nov 18;22(11):e22307 [FREE Full text] [doi: [10.2196/22307](https://doi.org/10.2196/22307)] [Medline: [33206052](https://pubmed.ncbi.nlm.nih.gov/33206052/)]
43. Martinengo L, Spinazze P, Car J. Mobile messaging with patients. *Br Med J* 2020 Mar 16;368:m884. [doi: [10.1136/bmj.m884](https://doi.org/10.1136/bmj.m884)] [Medline: [32179548](https://pubmed.ncbi.nlm.nih.gov/32179548/)]
44. Storck LG. Policy statement: texting in health care. *Online J Nurs Inform*. 2017. URL: <https://www.himss.org/resources/policy-statement-texting-healthcare> [accessed 2021-04-29]
45. Freundlich RE, Freundlich KL, Drolet BC. Pagers, Smartphones, and HIPAA: finding the best solution for electronic communication of protected health information. *J Med Syst* 2017 Nov 25;42(1):9. [doi: [10.1007/s10916-017-0870-9](https://doi.org/10.1007/s10916-017-0870-9)] [Medline: [29177600](https://pubmed.ncbi.nlm.nih.gov/29177600/)]
46. Safaeinili N, Brown-Johnson C, Shaw JG, Mahoney M, Winget M. CFIR simplified: pragmatic application of and adaptations to the Consolidated Framework for Implementation Research (CFIR) for evaluation of a patient-centered care transformation within a learning health system. *Learn Health Syst* 2020 Sep 26;4(1):e10201 [FREE Full text] [doi: [10.1002/lrh2.10201](https://doi.org/10.1002/lrh2.10201)] [Medline: [31989028](https://pubmed.ncbi.nlm.nih.gov/31989028/)]
47. Kruse C, Fohn J, Wilson N, Patlan EN, Zipp S, Mileski M. Utilization barriers and medical outcomes commensurate with the use of telehealth among older adults: systematic review. *JMIR Med Inform* 2020 Aug 12;8(8):e20359 [FREE Full text] [doi: [10.2196/20359](https://doi.org/10.2196/20359)] [Medline: [32784177](https://pubmed.ncbi.nlm.nih.gov/32784177/)]
48. Hsiang E, Offer C, Prescott M, Rodriguez A, Behar E, Matheson T, et al. Bridging the digital divide among racial and ethnic minority men who have sex with men to reduce substance use and HIV risk: mixed methods feasibility study. *JMIR Mhealth Uhealth* 2020 Apr 29;8(4):e15282 [FREE Full text] [doi: [10.2196/15282](https://doi.org/10.2196/15282)] [Medline: [32347807](https://pubmed.ncbi.nlm.nih.gov/32347807/)]
49. Mishori R, Antono B. Telehealth, rural America, and the digital divide. *J Ambul Care Manage* 2020;43(4):319-322. [doi: [10.1097/JAC.0000000000000348](https://doi.org/10.1097/JAC.0000000000000348)] [Medline: [32858734](https://pubmed.ncbi.nlm.nih.gov/32858734/)]

Abbreviations

CFIR: Consolidated Framework for Implementation Research
HIPAA: Health Insurance Portability and Accountability Act
mHealth: mobile health

Edited by G Eysenbach; submitted 24.07.20; peer-reviewed by A(Kassavou, YH Kwan; comments to author 04.10.20; revised version received 29.11.20; accepted 13.04.21; published 27.05.21.

Please cite as:

Fonner VA, Kennedy S, Desai R, Eichberg C, Martin L, Meissner EG

Patient-Provider Text Messaging and Video Calling Among Case-Managed Patients Living With HIV: Formative Acceptability and Feasibility Study

JMIR Form Res 2021;5(5):e22513

URL: <https://formative.jmir.org/2021/5/e22513>

doi: [10.2196/22513](https://doi.org/10.2196/22513)

PMID: [34042596](https://pubmed.ncbi.nlm.nih.gov/34042596/)

©Virginia A Fonner, Samuel Kennedy, Rohan Desai, Christie Eichberg, Lisa Martin, Eric G Meissner. Originally published in JMIR Formative Research (<https://formative.jmir.org/>), 27.05.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly

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

Original Paper

Recorded Mental Health Recovery Narratives as a Resource for People Affected by Mental Health Problems: Development of the Narrative Experiences Online (NEON) Intervention

Mike Slade^{1*}, PhD; Stefan Rennick-Egglestone^{1*}, PhD; Joy Llewellyn-Beardsley¹, MA; Caroline Yeo¹, PhD; James Roe², PhD; Sylvia Bailey³; Roger Andrew Smith³, BSc; Susie Booth³; Julian Harrison³, MA; Adarsh Bhogal³; Patricia Penas Morán⁴, PhD; Ada Hui¹, PhD; Dania Quadri⁵, BSc; Clare Robinson⁶, PhD; Melanie Smuk⁷, PhD; Marianne Farkas⁸, PhD; Larry Davidson⁹, PhD; Lian van der Krieke¹⁰, PhD; Emily Slade¹¹; Carmel Bond¹², MSc; Joe Nicholson¹³, PhD; Andrew Grundy¹, MPhil; Ashleigh Charles¹, MA; Laurie Hare-Duke¹, PhD; Kristian Pollock¹⁴, PhD; Fiona Ng¹, PhD

¹School of Health Sciences, Institute of Mental Health, University of Nottingham, Nottingham, United Kingdom

²National Institute for Health Research, ARC East Midlands, University of Nottingham, Nottingham, United Kingdom

³NEON Lived Experience Advisory Panel, Nottingham, United Kingdom

⁴Department of Personality, Assessment and Psychological Treatment, University of Deusto, Bilbao, Spain

⁵GKT School of Medical Education, King's College London, London, United Kingdom

⁶Centre for Primary Care & Public Health, Pragmatic Clinical Trials Unit, Queen Mary University of London, London, United Kingdom

⁷Department of Medical Statistics, London School of Hygiene and Tropical Medicine, London, United Kingdom

⁸College of Health and Rehabilitation Sciences, Boston University, Boston, MA, United States

⁹Yale School of Medicine, Yale University, New Haven, CT, United States

¹⁰University Medical Center Groningen, University Center of Psychiatry, University of Groningen, Groningen, Netherlands

¹¹Department of Computer Science, University of Oxford, Oxford, United Kingdom

¹²Nottingham University Business School, University of Nottingham, Nottingham, United Kingdom

¹³School of Humanities, University of Nottingham, Nottingham, United Kingdom

¹⁴School of Health Sciences, University of Nottingham, Nottingham, United Kingdom

*these authors contributed equally

Corresponding Author:

Stefan Rennick-Egglestone, PhD

School of Health Sciences

Institute of Mental Health

University of Nottingham

University of Nottingham Innovation Park

Triumph Road

Nottingham, NG7 2TU

United Kingdom

Phone: 44 115 409326

Email: stefan.egglestone@nottingham.ac.uk

Abstract

Background: The internet enables sharing of narratives about health concerns on a substantial scale, and some digital health narratives have been integrated into digital health interventions. Narratives describing *recovery* from health problems are a focus of research, including those presented in *recorded* (eg, invariant) form. No clinical trial has been conducted on a web-based intervention providing access to a collection of Recorded Recovery Narratives (RRNs).

Objective: This study presents knowledge produced through the development of the Narrative Experiences Online (NEON) Intervention, a web-based intervention incorporating the algorithmic recommendation of RRNs.

Methods: Knowledge was gathered through knowledge integration (KI) activities. KI1 synthesized previous studies to produce the NEON Impact Model describing how accessing RRNs produces health-related outcomes. KI2 developed curation principles

for the NEON Collection of RRNs through consultation with the NEON Lived Experience Advisory Panel and the curation of a preliminary collection. KI3 identified harm minimization strategies for the NEON Intervention through consultation with the NEON International Advisory Board and Lived Experience Advisory Panel. The NEON Intervention was finalized through 2 research studies (RS). In RS1, mental health service users (N=40) rated the immediate impact of randomly presented narratives to validate narrative feedback questions used to inform the recommendation algorithm. In RS2, mental health service users (n=25) were interviewed about their immediate response to a prototype of the NEON Intervention and trial procedures and then were interviewed again after 1 month of use. The usability and acceptability of the prototype and trial procedures were evaluated and refinements were made.

Results: KI1 produced the NEON Impact Model, which identifies moderators (recipient and context), mechanisms of connection (reflection, comparison, learning, and empathy), processes (identification of change from narrative structure or content and internalization of observed change), and outcomes (helpful and unhelpful). KI2 identified 22 curation principles, including a mission to build a large, heterogeneous collection to maximize opportunities for connection. KI3 identified seven harm minimization strategies, including content warnings, proactive and reactive blocking of narratives, and providing resources for the self-management of emotional distress. RS1 found variation in the impact of narratives on different participants, indicating that participant-level feedback on individual narratives is needed to inform a recommender system. The order of presentation did not predict narrative feedback. RS2 identified amendments to web-based trial procedures and the NEON Intervention. Participants accessed some narratives multiple times, use reduced over the 4-week period, and narrative feedback was provided for 31.8% (105/330) of narrative accesses.

Conclusions: RRNs can be integrated into web-based interventions. Evaluating the NEON Intervention in a clinical trial is feasible. The mixed methods design for developing the NEON Intervention can guide its extension to other clinical populations, the design of other web-based mental health interventions, and the development of narrative-based interventions in mental health.

(*JMIR Form Res* 2021;5(5):e24417) doi:[10.2196/24417](https://doi.org/10.2196/24417)

KEYWORDS

narratives; storytelling; intervention development; mental health; online intervention; patient involvement; narrative medicine; internet; recovery; mobile phone

Introduction

Background

The growing use of social media platforms has enabled the sharing of digital narratives about health and health problems on a substantial scale [1]. The sharing of digital health narratives has been examined across a range of domains, including cancer [2], chronic pain [3], and mental health [4]. Digital health narratives have also been incorporated into complex health interventions, where they can be used to structure therapeutic discussions [5], support engagement with an intervention [6], address health inequalities by giving voice to underrepresented populations [7], and reach nonmajority cultural groups or hard-to-reach populations [6-8]. Access to digital health narratives might also normalize experiences of health conditions and facilitate attitudinal change in health care staff when used as part of clinical education programs [9]. Although some interventions have integrated bespoke digital narratives curated by intervention development teams [10], others draw on the vast range of available public narratives, sometimes using algorithms to select those narratives that might provide the most therapeutic benefit for a recipient [11].

One specific focus of research has been on the subset of health narratives that describe *recovery* from health problems [12,13]. The definitions of recovery narratives vary across health domains. A systematic review of mental health defined a recovery narrative as a first-person lived experience account, which refers to events or actions over a period and which includes elements of adversity or struggle as well as self-defined

strengths, successes, or survival [13]. Narratives matching this definition are henceforth referred to as recovery narratives.

Recovery narratives can be shared as part of synchronous interactions between people, and sharing a recovery narrative is a core component of the work of peer specialists [14], an established effective intervention [15]. Recovery narratives can also be shared asynchronously in invariant forms such as text (eg, prose or poetry), audio, video, and other media, including visual artwork [16]. Narratives shared asynchronously are referred to as recorded recovery narratives in the remainder of this paper. Early examples of recorded recovery narratives can be found in the 1957 book *The Plea for the Silent* [17], which was published with the intent of addressing stigma about mental health problems. Recorded recovery narratives have been shared by health service units through booklets collating stories produced by their clients [18] and in printed autobiographies [19]. When shared on the web, recovery narratives can be presented individually [20-23] or in curated collections [24]. Recovery narratives have been incorporated into collections produced by antistigma campaigns such as Time to Change in the United Kingdom [25] and into websites intended to support the recovery process, such as Here to Help in Canada [26].

Sharing a recovery narrative can provide substantial benefits for the narrator [27], and developing new stories about one's experiences is central to the work of the hearing voices groups [28] and supports posttraumatic growth [29]. Qualitative research suggests that the elements of narratives describing recovery can also provide specific benefits to recipients, such as reducing isolation or providing hope for the future [30]. Although randomized controlled trials have been conducted on

interventions incorporating digital health narratives in relation to weight loss [11] and cancer [31], no randomized controlled trial has been conducted on the use of recorded mental health recovery narratives to benefit recipients. The best interventional evidence to date is derived from a study that investigated the use of a bespoke website presenting recovery narratives, among other mental health material. This identified benefits in three domains: being inspired, knowing I am not alone, and believing recovery is possible [10]. Narratives could be received in private, but meetings with support workers were available.

NEON Program

Overview

This paper describes the development of a web-based intervention that presents recorded recovery narratives. Research has been conducted as a part of Narrative Experiences Online (NEON), a 5-year program of work funded through the Programme Grants for Applied Research scheme of the National Institute of Health Research from 2017 to 2022. The aim of NEON is to investigate whether accessing recorded recovery narratives can improve health-related quality of life for people affected by mental health problems [32]. The NEON program comprises three stages: theory studies, which are completed and summarized below; intervention development studies, which are reported in this paper and integrate knowledge produced by theory studies; and 3 randomized controlled trials to evaluate the intervention. The primary outcome measure for all 3 trials was the Manchester Short Assessment of Quality of Life [33].

NEON Theory Studies

A systematic review developed an empirically supported definition of a mental health recovery narrative [13]. The review also used a narrative synthesis of 45 included publications to develop a conceptual framework describing the characteristics of mental health recovery narratives, grouped into nine superordinate categories: genre, positioning, emotional tone, relationship with recovery, trajectory, use of turning points, narrative sequence, protagonists, and use of metaphors. This framework was validated and extended through the analysis of 77 recovery narrative interviews [34] to provide a finalized Recovery Narrative Conceptual Framework.

A systematic review was then conducted to develop a Narrative Impact Conceptual Framework, describing the forms of impact from accessing mental health recovery narratives [1]. This review included 5 articles. Narrative synthesis was conducted to identify five transdiagnostic benefits to recipients: connectedness, a better understanding of recovery, a reduction in stigmatic views (including self-stigma), the validation of difficult personal experiences, and potentially beneficial behavioral responses, such as the initiation of more meaningful interactions with support workers. Strong affective responses can be produced in recipients of recovery narratives. The review identified one harmful outcome: emulating harmful behaviors encountered in eating disorder recovery narratives. It identified that strategies are needed to support the processing of affective responses and to minimize harmful impact. The review concluded that interventions incorporating recorded recovery narratives might be particularly relevant in areas with low

population density, that is, where access to both mental health care and peers with experience of similar mental health problems may otherwise be difficult [35].

The Narrative Impact Conceptual Framework was then extended by a study that developed a long-term narrative impact model characterizing the longer-term impact of live and recorded mental health recovery narratives [36]. This study involved an iterative thematic analysis of 77 interviews in which participants told their own recovery narrative and talked about the impact of recovery narratives of others on them. Helpful changes were identified as perceptions of connectedness to the narrative or narrator (as the strongest mechanism), validation, hope, empowerment and appreciation, a reduction in stigma and self-stigma, and the initiation of a particular form of turning point [37] identified as a reference shift, where accessing narrative content leads to a rapid and radical change in how recipients view what is possible for them. Harmful changes were identified as perceptions of inadequacy (eg, if a narrative describes a recovery that the recipient thinks to be impossible), disconnection (eg, from narrators who appear to have experienced less distress than the recipient), pessimism (eg, how much recovery is possible for the recipient), and emotional burden caused by empathy with the parts of a recovery narrative that describes adversity or struggle. The model also identified factors that might moderate the impact of recovery narratives: the recipient is experiencing a crisis and the recipient perceives the recovery narrative as authentic or inauthentic. The long-term narrative impact model had two main findings. First, a universally helpful recovery narrative is unlikely to exist, as components of recovery narratives that create benefits for some, such as observing narrator achievements, can cause harm for others. Second, a careful selection of narratives for use in an intervention is not by itself a sufficient harm management strategy, and other approaches are also needed to minimize harm.

A short-term narrative impact model was then developed through an experimental study on the immediate effect of accessing recorded recovery narratives [38]. Current mental health service users (N=40) were shown a series of recovery narratives and asked for qualitative and quantitative feedback on their impact. The short-term narrative impact model was developed through thematic analysis of the qualitative data. In the model, change is initiated through a recipient reflecting on their own experiences and then forming a connection through three mechanisms: comparing oneself with the narrative and/or narrator, learning about other people's experiences, and experiencing empathy. The three mechanisms of connection lead to impact through the identification of change based on the narrative structure or the interpretation of change in the narrative content, both of which lead to the internalization of the interpretation by the individual. Factors moderating impact included clinical factors (eg, an inability to focus on the narrative due to symptomatology), personality (eg, long-term difficulty in connecting with others), and recipient preferences such as narrative modality.

The potential to create positive change means that using recorded recovery narratives as a mental health intervention is possible, but the possibility of negative impacts means that care

is needed. The use of recovery narratives by health services has also been widely questioned. The critical theorist and activist collective Recovery in the Bin [39,40] and others [41] have argued that only very particular types of narratives that are perceived to be successful or acceptable are promoted by services. This can narrow the range of recovery templates available to recipients and potentially cause harm by limiting or negating recipients' own ways of recovering [42]. Others have expressed concerns that recovery narratives might be co-opted, that is, used for purposes other than those intended by the narrator, and that this can sustain harmful structures such as poorly functioning health services [43] or deflect attention from systemic inequalities and social injustice [41]. These ethical issues must be carefully considered.

To inform the work of building an ethically defensible recovery narrative collection, a systematic review of decisions made in the curation of mental health recovery narrative collections was conducted [4]. The concept of curation draws on existing usage within the discipline of museum studies, where the work of curators has been extensively studied [44]. Curation is understood as both a purposeful and political act, with curators often engaging with artifacts or collections that are sensitive and challenging [45]. In total, 23 documents were identified, of which only one was a research publication. A significant knowledge gap was identified regarding curatorial decision making in relation to recorded mental health recovery narratives.

To address this gap, an interview study was conducted with 30 recovery narrative collection curators from 7 countries [24]. The qualitative analysis identified six categories of decisions made by curators comprising the VOICES (Values and motivations, Organization, Inclusion and exclusion, Control and collaboration, Ethics and legal, Safety and well-being) framework. It was concluded that collection curators have a great influence on how mental health and recovery issues are presented and understood and that recovery narrative collections can provide a mechanism for making collective rather than individual-level knowledge available.

NEON Intervention Development

This paper reports the process of developing the NEON Intervention, a new mental health intervention that provides access to recorded mental health recovery narratives. Digital health technologies [46] are increasingly being used in global mental health practice, motivated by challenges such as lengthy waiting lists for treatment [47], limited access to in-person mental health treatment in rural and remote communities [48-50], and the distress inherent in accessing in-person treatment for people experiencing social anxiety [51]. Digital health technologies are a crucial approach during the COVID-19 pandemic [52], when social connectedness is reduced [53].

The NEON Intervention provides access to the recorded mental health recovery narratives contained in the NEON Collection. The NEON Collection is a curated collection of recorded mental health recovery narratives. Narrators have given permission for their narratives to be used in the NEON Collection. Each narrative is characterized using a standardized inventory derived from the Recovery Narrative Conceptual Framework called the

Inventory of Characteristics of Recovery Stories (INCREASE) [54].

Users of the NEON Intervention can access narratives in four ways:

1. Recommended: they can request the automated recommendation of recovery narratives in the NEON Collection. Requests are served by a recommender system [55], a term encompassing a family of algorithms designed to match digital media items to users. Recommender systems are frequently used in web-based digital media hosting services such as Spotify [56]. The design of the recommender system used in the NEON Intervention draws on the findings presented in this paper. Recommendations are informed by the narrative feedback provided after each narrative is received. For example, if accessing a recovery narrative makes a recipient feel more hopeful, they will receive more recommendations for recovery narratives with similar characteristics.
2. They can directly browse narratives in the NEON Collection by selecting tags of interest.
3. They can choose to be shown a randomly selected narrative.
4. They can rerequest a previously shown narrative.

The NEON Intervention is being developed and evaluated with three target groups: individuals who experienced psychosis, individuals with other mental health problems, and informal carers.

1. People who experience psychosis regularly use digital technologies such as social networks [57], and a systematic review of digital interventions for psychosis incorporating web-based, social media, and mobile technologies concluded that these approaches are acceptable, are feasible, and have the potential to improve outcomes [58]. Messages that promote hope are known to be recovery-promoting in psychosis [59], and hope is known to mediate potential psychosis recovery indicators such as increases in structured activity [60]. Accessing recovery narratives can reduce self-stigma, and self-stigma predicts low adherence to psychosocial treatments in patients with schizophrenia [61]. The effectiveness of the NEON Intervention for people who experienced psychosis will be evaluated in the definitive NEON Trial [27] (ISRCTN11152837).
2. The evidence reviewed earlier indicates that the benefits of accessing recorded recovery narratives are primarily transdiagnostic; therefore, the NEON Intervention may also be effective for people with nonpsychotic mental health problems. The effectiveness of the NEON Intervention for people with mental health problems other than psychosis will be evaluated in the definitive NEON-O Trial [27] (ISRCTN63197153).
3. With less strong evidence, recovery narratives may be helpful for informal carers, that is, family or friends of people with mental health problems, both for supporting their well-being and informing their understanding of the experiences of the person they care for. The feasibility of using the NEON Intervention with informal carers will be evaluated in the NEON-C Trial [27] (ISRCTN76355273).

Aims and Objectives

The aim of this paper is to present selected items of knowledge developed during the intervention development process for the NEON Intervention and the study procedures for the 3 planned NEON trials. A detailed description of the final version of the NEON Intervention and of our chosen trial procedures is provided in our trial protocol [27]. In combination with our trial protocol, the knowledge presented in this paper will support the replication of the NEON Intervention, enable studies with other populations, and inform the development of new interventions using recorded recovery narratives. Some knowledge products (such as the curation principles described below) will have broader relevance to the design of interventions, integrating digital narratives about health and health problems.

The objectives of this study in the context of the development of the NEON Intervention by the NEON study are as follows:

1. Study objective 1: to develop the NEON Impact Model, a change model describing the impact of the NEON Intervention on recipients
2. Study objective 2: to identify appropriate curation principles for the NEON Collection of recorded recovery narratives used by the NEON Intervention
3. Study objective 3: to identify appropriate strategies for minimizing harm from the NEON Intervention
4. Study objective 4: to finalize narrative feedback questions for use by the recommender system
5. Study objective 5: to evaluate the acceptability and usability of an initial prototype of the NEON Intervention and associated trial procedures
6. Study objective 6: to identify features of intervention use relevant to trial planning

Study objectives 1-3 were addressed through three knowledge integration activities. These were methodical activities that drew on existing study expertise and knowledge to develop knowledge products underpinning the final version of the NEON Intervention evaluated in the NEON trials. Study objective 4 was initially addressed through research study 1, an experimental study in which quantitative feedback was provided by mental health service users who were shown items selected from the NEON Collection of recovery narratives. Study objectives 4-6 were addressed through research study 2, a feasibility evaluation of a prototype implementation of the NEON Intervention.

Knowledge produced through the three knowledge integration activities and 2 research studies is interlinked and is included in a single paper to provide a thorough account of the NEON Intervention development work. In selecting these five activities for inclusion, we focused on describing how human-computer interaction issues were successfully addressed in developing the NEON Intervention and trial procedures. This is in keeping with an accepted definition of human-computer interaction research as considering the broad personal and sociological context of technology usage [62]. Technical aspects of the development of the recommender system used in the NEON Intervention were also informed by the knowledge presented in this paper. These will be reported elsewhere.

The research reported here was conducted between September 2017 and October 2019. Ethical approval was received from the London - West London and Gene Therapy Advisory Committee Research Ethics Committee in advance (18/LO/0991). All participants provided informed consent, either in written or audio-recorded form. Appropriate consent was also collected for the inclusion of all narratives in the NEON Collection.

The methods and results for each of the five selected activities are described. As a result of research studies 1 and 2, refinements to the prototype implementation of the NEON Intervention used in research study 2 were made. These refinements are described in the Results section. We reflect on the broader implications of our findings in the Discussion section.

Methods

Knowledge Integration Activity 1: Development of the NEON Impact Model

This knowledge integration activity developed the NEON Impact Model, describing how recorded recovery narratives improve outcomes. The modeling of change is recommended in the UK Medical Research Council guidance on the development and evaluation of complex evaluations [63].

The 3 completed NEON studies on impact [1,36,38] were synthesized by the NEON research team to produce a change model comprehensively describing how health-related outcomes might be produced by receiving recorded recovery narratives. The short-term narrative impact model was selected as the theoretical foundation of the integrated change model because it was produced by an experimental study considering only recorded recovery narratives. To incorporate longer-term impacts, relevant components from the Narrative Impact Conceptual Framework and the long-term narrative impact model were integrated if they described health-related change. Informed by the biomedical principle of nonmaleficence [64], negative outcomes identified in other eHealth studies [65] were added if they might be produced through use of the NEON Intervention, to inform the design of harm minimization strategies or trial procedures. *Moderators* and *forms of learning* were present across all underpinning studies; hence, comprehensive lists were synthesized by merging similar items and grouping them into superordinate categories.

Knowledge Integration Activity 2: Identification of Curation Principles

This knowledge integration activity identified appropriate curation principles for the NEON Collection of recovery narratives. Implementing these principles ensured that the NEON Collection addresses all safeguarding, ethical, legal, clinical, and technological challenges associated with storing and using recovery narratives.

Measures

The INCREASE is a 77-item, researcher-rated standardized tool to identify manifest and latent characteristics of recorded mental health recovery narratives [54]. Categories in INCREASE

comprise narrative eligibility, narrative mode, narrator characteristics, narrative characteristics, and narrative content. In total, 71 items characterize manifest content (eg, narrator gender, narrator diagnosis, content requiring content warnings, types of turning point, specific topics such as family, education, or work), and 6 characterize latent content (stage of recovery, genre, positioning, tone, relationship with recovery, and trajectory). Specific attention was paid to the issue described earlier, relating to the importance of diverse recovery templates, so items and rating categories included *uses a nondiagnostic framework*, *recovery despite services*, *recovery outside services*, *narrator rejects the concept of recovery as used in mental health services*, and *circular trajectory*. Different rating scales are used for different sections, for example, present or mainly positive or mainly negative for specific narrative content and escape or endurance or endeavor or enlightenment for genre. No summary scores were produced.

Procedures

All elements of the design of the NEON Collection were explored using the VOICES typology of curatorial decisions as an organizing framework [24]. The research team sought advice on specialist issues from relevant experts, for example, legal advice was obtained in relation to intellectual property and references to third parties in recorded recovery narratives. To ensure that the curation principles for the NEON Collection were informed by a lived experience perspective, rather than reflecting solely research and clinical priorities, three consultation workshops with the NEON Lived Experience Advisory Panel (LEAP) were held between October 2017 and June 2018. The NEON LEAP comprises an independent chair and 10 members with an interest in recovery narratives and personal experience of mental health problems and services either directly or as family members. Workshop agendas were co-produced between a subgroup of the NEON LEAP and the research team, who met at least two weeks before each workshop to finalize the agenda and precirculate preparatory materials in line with best practices [66]. Workshops addressed issues presenting complex ethical challenges, such as inclusion and exclusion criteria, appropriate approaches to anonymization, withdrawal of narratives, and procedures for processing third-party requests for narrative withdrawal. Each workshop was attended by approximately 15 participants, comprising all NEON LEAP members and several research team members. Each workshop involved facilitated whole-group and subgroup discussions to systematically explore curatorial issues from a lived experience perspective. The meetings were minuted in detail. LEAP members were paid GB £150 (US \$210) plus expenses per workshop. This work informed the ethics application enabling intervention development work. This was approved by the UK Health Research Authority to govern the long-term management of the NEON Collection.

A key challenge identified through work with LEAP was the selection of appropriate inclusion and exclusion criteria for narratives. To refine these criteria and to ground them in practical challenges, a preliminary collection of 100 recovery narratives was assembled from publicly accessible resources and with the permission of the collection curator or narrator as required. The NEON Collection Steering Group (CSG)

comprised 4 NEON LEAP members and 2 NEON researchers. Each narrative was coded by at least one researcher against the proposed inclusion and exclusion criteria and classified as either clearly included, clearly excluded, or not clear. Narratives coded as not clear were referred to the NEON CSG for a final decision. The CSG decided on inclusion and identified any necessary clarifications regarding the wording and interpretation of the inclusion or exclusion criteria. All decisions were minuted, and the current inclusion and exclusion criteria were updated on the NEON Collection web page [67].

Curation involves a careful review of narratives, which can create a significant emotional burden [36]. Furthermore, narratives that met the inclusion criteria were coded using INCREASE before inclusion in the NEON Collection, which involved a detailed assessment of a large number of narratives. In an unplanned add-on study, a feedback questionnaire was completed by INCREASE coders, covering the emotional impact of characterizing narratives and personal strategies for well-being [54]. These questionnaires were analyzed to identify general principles to support curator well-being. A facilitated discussion was then held involving all CSG members or NEON researchers who had coded at least 30 narratives. Participants reflected on personal well-being strategies, considered the questionnaire analysis, and identified the principles of well-being. The consensus recommendations were minuted.

Analysis

To synthesize knowledge on the most critical decisions taken during this process, a document analysis [68] was conducted on the minutes of NEON LEAP consultation meetings (n=3), the approved research protocol, minuted recommendations produced by CSG, the NEON Collection page on the NEON website, and minuted recommendations from the workshop on coder well-being. The documents were imported into NVivo (QSR International; version 11). The text describing a decision or the rationale for the decision was identified and summarized. Decisions were then organized into six predefined categories of the VOICES framework of narrative curation.

Knowledge Integration Activity 3: Identification of Harm Minimization Strategies

This knowledge integration activity identified appropriate strategies for minimizing harm from the NEON Intervention.

Participants

Strategies were developed through consultation between the NEON research team, the NEON LEAP as described in knowledge integration activity 2, and the NEON International Advisory Board, which comprises 7 experts in research and intervention development around recovery narratives, eHealth interventions, peer research, and mental health recovery.

Procedures

An iterative approach was used to identify, implement, and refine candidate harm minimization strategies. The NEON Impact Model was used as a foundation for shaping harm minimization strategies, and an evolving design rationale [69] explaining essential decisions and why they were made was created. This was expressed in the form of a draft protocol for

the 3 NEON trials. An early draft of the trial protocol was discussed with NEON International Advisory Board members, and the prototype developed for research study 2 was then discussed at two meetings between a subgroup of NEON LEAP and NEON researchers. These consultations led to protocol and prototype enhancements. The prototype was considered at a NEON LEAP meeting, and further enhancements were made.

One area in which efforts were made to be guided by research was in relation to content warnings, also known as trigger warnings, defined in an education context as “offering prior notification of an educational topic so that students may prepare for or avoid distress that is automatically evoked by that topic due to clinical mental health problems” [70]. The research team conducted a nonsystematic narrative review of research on content warnings [27]. The review found limited and conflicting evidence, primarily drawn from educational or trauma treatment settings [70-75]. No directly relevant evidence concerning the impact of content warning in relation to mental health recovery narratives was identified. NEON is a health service-funded study, so it was decided to follow the standard clinical practice of giving content warnings on the basis of the ethical principle of nonmaleficence.

Research Study 1: Narrative Feedback

The aim of research study 1 was to evaluate an initial set of narrative feedback questions designed for use in the recommender system. This contributed to the selection of a final set of narrative feedback questions, in keeping with study objective 4. The specific objectives of research study 1 were to identify floor and ceiling effects in narrative feedback, to describe variability among participants, and to examine whether the order of presentation affected response. The latter was included because it might indicate fatigue in the repeated provision of feedback. The analysis used quantitative narrative feedback data collected in the short-term narrative impact model study [38] but only qualitative data from this study have previously been reported.

Participants

Eligible participants were people with current mental health concerns, using statutory mental health services, aged above 18 years, able to provide informed consent, and fluent in English. Individuals who were experiencing a crisis or who were otherwise unable to participate in the research were excluded.

Setting

Participants were recruited from statutory mental health services within a health care trust in the East Midlands of England.

Measures

The Herth Hope Index (HHI) is a 12-item measure of hope adapted from the Herth Hope Scale with adequate psychometric properties [76]. The HHI score ranged from 12 (low hope) to 48 (high hope).

Procedures

A subset of 30 narratives were assembled from the NEON Collection by 2 researchers. Narratives were purposively selected to maximize the variation in modality, narrator

diversity, and length. Modality was chosen because multimedia use in educational settings has been shown to increase the depth of learning in students [77], suggesting that this may promote engagement and cater to different learning styles within individuals. Narratives with a substantial range of modalities are available in the public domain, and the use of multimedia may also promote the inclusiveness of individuals who experience disabilities or who may have difficulty comprehending a specific mode of media, for example, due to dyslexia. The selected narratives were diverse in narrator age, gender, and ethnicity, given the evidence from the NEON Impact Model that sociodemographic characteristics can influence connection. Finally, to vary the cognitive demands of participants, the chosen narratives were different in length. Text narratives ranged from half a page to 3 pages, video narratives ranged from 1 to 5 minutes, and audio narratives ranged from 2 to 3 minutes. On the basis of a pilot study protocol, it was estimated that, on average, participants would take no longer than 10 minutes to read, watch, or listen to a narrative. A total of 30 narratives were selected, comprising 15 texts (poems and prose text), 10 videos, and 5 audio-based narratives.

The study was promoted as an investigation of narrative impact through social media, advertisements within services (eg, posters and newsletters), and by clinicians and managers from Improving Access to Psychological Therapy Services, community forensic services, locality mental health teams, and recovery colleges. Both clinician referrals and self-referrals to the study were accepted. Potential participants were given a participant information sheet by their clinician or researchers. Interested participants then contacted the researchers or gave their clinician permission to pass on their contact details. The researchers assessed eligibility, and informed consent was obtained before the interview. Interviews took place in a university or clinical setting.

Each participant took part in a two-hour research session, for which they were offered GB £20 (US \$28) plus expenses. After providing informed consent, the researcher completed a participant characterization form from the verbal responses provided by the participant. The characteristics recorded included age, gender, preference for narrative modality, current diagnosis, and hope (HHI). Each participant then sequentially received randomly selected narratives, discontinuing when the participant indicated that they wished to stop. If individuals expressed a preference for narrative modality, then only narratives consistent with their preferences were considered. After accessing each narrative, the participants provided narrative feedback by rating three questions: How connected to the story did you feel? How connected to the narrator of the story did you feel? How hopeful did the story make you feel? Each response was rated on an 11-point scale (rating: 0-10). The labels for the two questions on connection were as follows: 0, extremely disconnected; 2, somewhat disconnected; 5, neither connected nor disconnected; 7, somewhat connected; and 10, extremely connected. The labels for the question on hopefulness were as follows: 0, extremely pessimistic; 3, somewhat pessimistic; 5, neither hopeful nor pessimistic; 7, somewhat hopeful; 10, extremely hopeful.

Analysis

Quantitative analysis was conducted using SPSS statistics (IBM; version 25) and STATA SE (version 16). The significance level was set at $P=.05$. Narrative feedback responses were grouped by question and then tested for nonnormality using the Kolmogorov-Smirnov test. No evidence of nonnormality was found; therefore, parametric tests were used.

Floor and ceiling effects and variability across participants were inspected by creating box-and-whisker plots. Outliers were defined as cases that fell 1.5 times above or below the IQR, and extreme outliers were defined as cases falling 3 times above or below the IQR. The impact of order of presentation was evaluated by creating a derived variable called presentation order (first, second, third, fourth, fifth, etc) and conducting a linear multilevel regression with each narrative feedback question as a dependent variable. This was because all participants accessed at least four narratives but not all accessed five or more.

Research Study 2: Feasibility Evaluation

The aim of research study 2 was to evaluate the feasibility of using the NEON Intervention and associated trial procedures with people with experience of mental health problems. The specific objectives were to finalize narrative feedback questions (study objective 4), to evaluate the acceptability and usability of an initial prototype of the NEON Intervention and associated trial procedures (study objective 5), and to identify features of intervention usage relevant to trial planning (study objective 6).

Participants

Eligible participants were people with current mental health concerns; using statutory mental health services; aged ≥ 18 years; having the capability, with support if needed, to interact with a web-based intervention; having access to a computer or smartphone with an internet connection at home, in a community venue, or through a health service venue; who are able to provide informed consent; and who are fluent in English. Individuals who were experiencing a crisis or who were otherwise unable to participate in the research were excluded.

Setting

Participants were recruited from statutory mental health services within a health care trust in the East Midlands of England.

Procedures

Research participants were recruited through a single health care trust in England, using three strategies. First, posters and leaflets were placed in health services and community venues. Second, participants were recruited through a direct approach by clinical support officers who attended health care clinics. Third, communication was sent by the research team to prior NEON research participants who provided ongoing consent to contact.

A prototype web-based NEON Intervention was developed as an interactive platform for implementing strategies and allowing the features to be used by others to inform refinements. The platform integrated the NEON Intervention with selected

web-based trial procedures, such as eligibility testing and web-based consent processes. All components of the NEON Intervention were implemented apart from the recommender system, and the baseline measures component was implemented in reduced form, that is, with fewer measures. Data usage was automatically logged, including the device type used to access the intervention, time between key trial processes (informed consent, baseline measure completion, randomization, first narrative presentation, etc), frequency, and context of usage of specific features (eg, "I'm upset" or "Get me out of here").

Participants attended a 2-hour baseline interview at a research site, for which they were offered GB £20 (US \$28) payment plus expenses. They worked through a series of tasks using a prototype. The tasks were to read the web-based participant information sheet; provide web-based consent through an informed consent form; register a NEON Intervention account using an email address; complete a web-based participant demographics form; complete the 12-item Manchester Short Assessment of Quality of Life [33], which is the primary outcome; complete the 10-item CORE-10 (Clinical Outcomes in Routine Evaluation) measure of problems, functioning, and risk [78], which is a secondary outcome; access an initial random narrative and provide narrative feedback; use the NEON Intervention to select an additional narrative and provide narrative feedback using the same three questions and rating scale as in study 3; and if time allowed, to explore the interface, including features such as the "About me," "I'm upset," and "Get me out of here" buttons.

During these tasks, participants were asked to verbalize their thoughts using a think-aloud protocol [79]. The researcher did not guide the participant, although occasionally the researcher assisted if the participant was stuck and could not progress despite substantial effort. The planned exception was that the researcher would proactively assist a participant who indicated that they normally needed support in using a computer system, but this did not occur in practice. After the tasks were completed, the participants were asked two trial procedure questions: Would you be happy to receive payment for trial participation through online vouchers? What might prompt you to use the intervention after your first use of it? The interview was recorded, and the researcher made field notes on the observations.

Following the baseline interview, participants were given unconstrained access to the web-based prototype for 28 days, during which time their usage was automatically logged. A follow-up interview was then conducted at a research site, and participants were paid GB £20 (US \$28) plus expenses. A summary of their recorded usage of the intervention was prepared, showing the number of narratives accessed and rated, number and date of log-ins, and proportion of log-ins on mobile devices (smartphones or tablets) or computers (laptops or desktops). At the follow-up interview, participants were shown their usage and asked to discuss notable features, such as periods of high or low usage. Data usage summary sheets as a form of data visualization are a standard approach to support reflection on computer system usage [80]. If a system or system feature is underused, the discussion of usage data can differentiate whether this was due to (1) the system creating immediate positive change requiring no further engagement, (2) periods

of planned technological disconnection such as holidays, or (3) malfunction, dislike, or poor usability of the system or feature [81].

The prototype was also discussed at a NEON LEAP meeting and minutes were taken. NEON LEAP members were subsequently given access to the prototype and provided written feedback.

Analysis

To meet study objective 5, a list of prototype features was created. For each feature, feedback from baseline and follow-up interviews and NEON LEAP feedback were synthesized. Short narrative summaries of responses to the two trial procedure questions were produced, and the baseline field notes were analyzed thematically. Each element of feedback was categorized as acceptable, unacceptable, usable, less usable, and other. Categories of unacceptable and less usable identified features that most needed improvements. Categories of

acceptable and usable were included to indicate variations in response across the cohort, for example, to identify whether there were features that were acceptable to some but unacceptable to others. Feedback in the other category was reviewed individually.

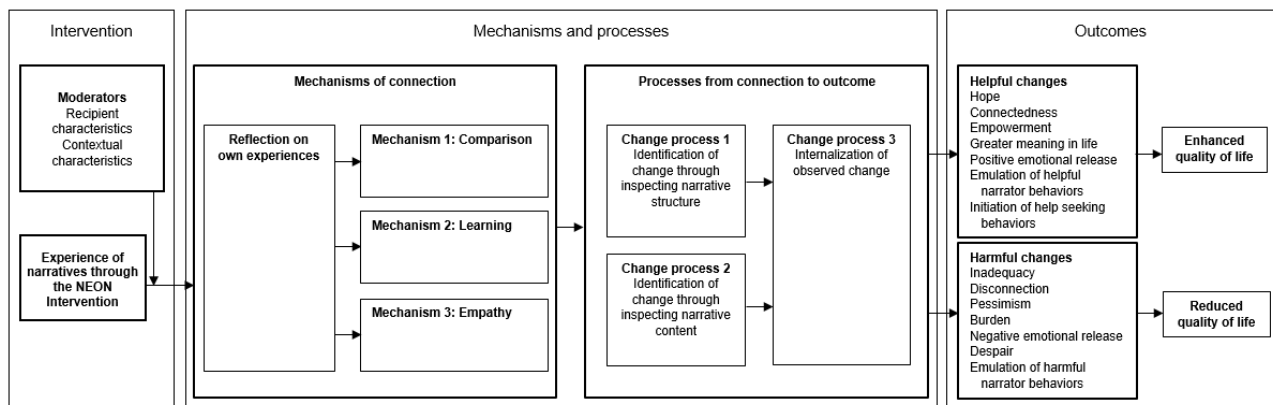
To meet study objectives 4 and 6, descriptive analyses were conducted to describe participant demographics, frequency and route of access, frequency of narrative feedback, and frequency and length of use of the NEON Intervention.

Results

Knowledge Integration Activity 1: NEON Impact Model

The NEON Impact Model, describing the processes by which engaging with a recorded recovery narrative can create change, is shown in Figure 1.

Figure 1. Narrative Experiences Online Impact Model linking experience of narratives to outcomes. NEON: Narrative Experiences Online.



In the NEON Impact Model, the impact of a recovery narrative is moderated by the characteristics of the recipient and the context in which they receive it. A complete list of recipient and contextual moderators synthesized from the source studies is shown in Table 1.

In the NEON Impact Model, one way for a narrative to make an impact is through the recipient learning something from the

narrative (mechanism 2 in Figure 1). The types of learning synthesized from the source studies are shown in Textbox 1.

The NEON Impact Model informed (1) the design of the NEON Intervention, including the narrative feedback questions and guidance about using the intervention, and (2) the trial process evaluation.

Table 1. Moderators of the impact of recovery narratives.

Moderator	Direction of influence
Recipient characteristics	
Recipient reports a long-term inability to connect with others.	Reduced impact
Recipient has experienced a recent event perceived as distressing.	Reduced impact
Recipient is experiencing a mental health crisis.	Reduced impact
Recipient has beliefs, values, or attitudes contradicting those of the narrator.	Reduced impact
Recipient is experiencing mental health problems that disrupts information processing (such as hearing voices).	Reduced impact
Recipient perceives the content of the narrative to be emotionally challenging.	Reduced impact
Recipient experiences difficulties in comprehending the form of the narrative (eg, if the narrative is presented as a poem).	Reduced impact
Recipient perceives the narrative or narrator to be inauthentic.	Reduced impact
Recipient perceives the narrative or narrator to be authentic.	Increased impact
Contextual characteristics	
Recipient has access to a private space to access challenging narratives.	Increased impact
Recipient has access to a mental health worker who supports processing.	Increased impact

Textbox 1. Types of learning from accessing recovery narratives.

<p>Learning About Mental Health</p> <ul style="list-style-type: none"> • How others experience a mental health condition • Alternative conceptualizations of mental health problems • The impact of mental health problems on others (eg, carers) • New coping strategies to enhance daily living <p>Learning About Recovery</p> <ul style="list-style-type: none"> • Recovery is possible • Specific recovery strategies that have helped others • Barriers to recovery that others have experienced • Differing beliefs and values that have supported recovery • How to manage treatment and make best use of services
--

Knowledge Integration Activity 2: Curation Principles

The NEON Collection curation principles are listed in [Textbox 2](#). These have been implemented in full to create, manage, and use the NEON Collection.

Textbox 2. Curation principles for the Narrative Experiences Online Collection.**Curation Principles**

VOICES (Values and motivations, Organization, Inclusion and exclusion, Control and collaboration, Ethics and legal, Safety and well-being) domain 1: values and motivations of the Narrative Experiences Online (NEON) Collection

- Purpose: the primary purpose of the NEON Collection is to provide benefits to recipients.
- Mission: the NEON Collection will seek for heterogeneity of narrative content, form, and narrator demographics and to be as large as possible. Greater heterogeneity and size increases the chance of a recipient finding someone like them or a story like theirs and hence experiencing the helpful outcomes outlined in [Figure 1](#). Insufficient heterogeneity and size risks a recipient failing to find someone like them and hence feeling more disconnected from others.
- Less hopeful narratives: the NEON Collection will include some narratives where adversity and struggle are the dominant themes. Although these narratives are less regularly used by health services, recipients experiencing profound distress may find it easier to connect with such narratives.

VOICES domain 2: organization of the NEON Collection

- Donation from existing collections: narratives can be donated to the NEON Collection by organizers of existing collections but only if the collection organizer confirms that appropriate consent has been obtained and only if the narrative is already public. Appropriate consent means either the collection organizer has previously collected consent to enable reuse or has obtained consent from individual narrators to donate their narrative to the NEON Collection. Details of the source collection will be retained and may be displayed to recipients to help them understand the context of the narrative.
- Donations from individuals: narratives can also be offered to the NEON Collection by individual narrators, even if they have not been published.
- Role of the curator: NEON researchers will assess inclusion of narratives, with involvement of the Collection Steering Group (CSG), the NEON Chief Investigator, or a legal expert where uncertainty exists around specific exclusion criteria. Diversity in the NEON Collection will be monitored to identify underrepresented groups to be targeted for narrative donation. Curators of the NEON Collection will not edit narratives, which will be displayed as close as possible to their original form.

VOICES domain 3: inclusion and exclusion of narratives in the NEON Collection

- Decision-making process: decisions on inclusion will be made with reference to formal current inclusion and exclusion criteria. These criteria will be publicly available to ensure transparency. Numbers of narratives considered, included, and excluded will be published for transparency. If stories link to external material, then the contents of this material should not be considered when deciding inclusion, as it may change. Donors (either individual narrators or collection organizers) will be informed if their narrative is not included, and there will be an appeal process.
- Inclusion criteria: a narrative is includable in the NEON Collection if all of the following criteria are met—(1) it includes elements of adversity or struggle that relate to mental health problems, broadly defined; (2) it includes descriptions of strength, success, or survival, as defined by the narrator or identifiable by a third party; (3) it refers to events or actions over a period (including either external events or internal mental events); (4) it is told by an individual with experience of mental health problems and recovery; (5) where language is used, the narrative is mainly in English or, if translated, the translation needs to be provided or approved by the narrator; (6) the story is provided in a digital file, the story is provided in a format that can easily be converted into a digital file, or the story is hosted on an existing webpage, the URL to the webpage is permanent, and the page does not contain links that would enable navigation to another page; or (7) consent to use the narrative in perpetuity (other than if the narrative is withdrawn) has been obtained from the narrator, from the owners of an existing collection who have previously collected consent from their donors that is broad enough to allow for reuse, or from the owners of an existing collection who have collected individual consent from their donors for usage in the NEON Collection.
- Exclusion criteria: a narrative is excluded from the NEON Collection if any of the following criteria are met—(1) it is presented as fictional; (2) it is told by anyone other than the individual experiencing mental health problems and recovery (such as a carer or journalist); (3) (for video and audio stories) the quality of recording is so low that the story is very difficult or not possible to understand; (4) it is split across multiple files or modalities or uses a multimedia approach that cannot easily be integrated into a single file; (5) it contains descriptions of potentially harmful behaviors in sufficient detail as to be likely to encourage imitation; (6) it indicates that the narrator has engaged in an undisclosed, serious criminal activity; (7) the narrator is a child or appears to be a child, unless it has been confirmed that the narrator is now an adult and has provided consent for a childhood story to be shared; (8) it contains hate speech; (9) it provides information about a third party that might reasonably lead to harm being caused to the third party such as providing directly identifying information about someone accused of abuse; (10) it includes sensitive personal information about individual third parties, unless the third party has already made this information public, for example, by publishing their own recovery story, or unless the third party is no longer alive. A story includes sensitive information about a third party if it clearly reveals their political or religious beliefs, mental or physical health conditions, sexual orientation or behaviors, or any offences committed or alleged to have been committed by them; (11) it reveals the adoption status of a third party, unless the third party has already made this information public; or (12) it raises any other unforeseen concerns, in which case this list of reasons for exclusion may be updated. Exclusion criterion 9 is included for predominantly legal reasons: the NEON Collection is hosted in the European Union (EU) and hence subject to the EU General Data Protection Regulations on personal information.
- Resolution of uncertainty: when making an assessment, there will be a bias toward the inclusion of a narrative. For example, inclusion criteria 2 is met if any rater can see strengths, successes, or survival in a narrative. If NEON researchers are uncertain whether a narrative meets all the inclusion criteria or exclusion criterion 1, 2, or 3, a final decision will be made by CSG. If the uncertainty is about exclusion criterion 4, 5, 7, 11, or 12, a final decision will be made by the NEON Chief Investigator. If the uncertainty is about exclusion criterion 6, 8, 9, or 10, an opinion will be sought from a legal representative approved by the study sponsor. Some forms of uncertainty can be resolved by asking the narrator for a short addendum to contextualize the narrative, but narrators are not required to submit this or may not be contactable, and hence final decisions may need to be made without it.

VOICES domain 4: control and collaboration around the NEON Collection

- Oversight: for the duration of the NEON program, the study sponsor will act as an auditor and may examine records relating to narrative consent. If use of the NEON Collection continues beyond the end of the study, an equivalent authority needs to be in place and approved by the study sponsor.
- Archiving and reinstatement: the NEON Collection can be temporarily archived (eg, at the end of the NEON program) and withdrawal requests cannot be met while it is archived. It can only be reinstated from the archive if a body with an equivalent status to a study sponsor is identified.
- Information about approvals: details of legal and ethical approvals for the NEON program and the NEON Collection will be displayed whenever narratives from the NEON Collection are used.

VOICES domain 5: ethical and legal considerations for the NEON Collection

- Documentation of consent: if a collection organizer wishes to offer narratives to the NEON Collection, they must confirm in writing that consent has been provided. This confirmation will be stored for audit purposes.
- Rights of collections: collection organizers have the right to withdraw any narratives that they have donated.
- Rights of the narrator: accepted narrators have a right to inclusion and publication of a short addendum. They might use this to illustrate how their life has changed since they created their narrative or to contextualize what was happening in their life at the time they wrote their narrative. All narrators have a right to withdraw a narrative. They can request withdrawal through a collection organizer if the narrative was donated from an existing collection or directly through the NEON Collection in all cases.
- Rights of third parties: third parties can request withdrawal, for example, if they assert that a narrator did not have capacity when they submitted a narrative, and each request will be individually assessed by the NEON CSG. To protect the right of narrators to have their story told, third parties do not have an automatic right to withdrawal.
- Processing of withdrawal requests: all narratives will be given a unique ID to aid withdrawal requests. Since some withdrawal requests may be malicious, such as an attempt by someone who is not the narrator to withdraw the narrative without due cause, in order to protect the rights of narrators and the existence of the NEON Collection, proof of identity may be required. Low-burden mechanisms will be provided to establish identity.
- Assertion of copyright breach: individuals can assert that a narrative has breached their copyright, and assertions of copyright theft will be processed in accordance with the European e-Commerce Directive 2000/31/EC [82].
- Expectations on recipients: to access narratives in the NEON Collection, a user must register an account and commit to not copying any material. This is because some individuals have donated narratives that are not published elsewhere.

VOICES domain 6: safety and well-being

- Safety of narrators: the NEON Collection will not edit or anonymize narratives that have been submitted because stories can be an economic and social resource for some narrators and because this may have intellectual property implications for the NEON program. If an individual donates their narrative to NEON, they will be provided with information about how the narrative will be used and encouraged to think about consequences of revealing their identity in a narrative, allowing them to make an informed choice about whether to be identifiable. They can submit identifying metadata (such as a story title that includes their name) if they wish.
- Safety of curators: when assessing narratives for inclusion, curators have the right to disengage from a narrative that distresses them, either temporarily or permanently, without providing a reason.

Knowledge Integration Activity 3: Harm Minimization Strategies

Several harm minimization strategies were identified and implemented in the NEON Intervention.

Strategy 1: Informed Consent

Potential participants were informed through a web-based participant information sheet about the potential harmful impacts listed in the NEON Impact Model (Figure 1). This allowed participants to make an informed choice about whether to participate in the study.

Strategy 2: Reflecting on Self-management

Before receiving the first narrative, participants received brief advice on how to handle difficult emotional responses to narratives. They were then asked to record self-management strategies that they find helpful to use when upset. This encouraged participants to apply their self-management strategies if needed when using the NEON Intervention. They

can change the recorded strategies in the “About me” section of the intervention.

Strategy 3: Dealing With Emotional Distress

Participants were encouraged to use the “I’m upset” button if they become upset while using the NEON Intervention. This button is prominently available on all intervention pages and opens a webpage, providing access to four resources:

- A reminder of any self-management strategies they previously recorded.
- Information about national helplines in England: Samaritans, Mind, The Mix, Elefriends, Big White Wall, Saneline, and Rethink Mental Illness.
- Evidence-based self-management resources organized into categories such as express yourself creatively, labeling your feelings, mindfulness, self-soothing, meditation, breaking up triggers, and distraction.

- Information about contacting a general practitioner, local mental health services, and the National Health Service (NHS) urgent and emergency care hotline.

Strategy 4: Content Warnings

INCREASE rates five sensitive issues: abuse or sexual violence, loss of life or endangerment to life, self-harm including eating disorders, violence or aggression and injustice, and prejudice and discrimination. As interpretation is needed in relation to some of these, compromising interrater reliability, INCREASE items relating to content warnings only will be independently rated a second time by a different rater. If either rater identifies a content warning as relevant, it will be included in the final INCREASE rating. If the INCREASE rating of a narrative indicates it deals with any of the five content warning issues, then all relevant content warnings are displayed before the presentation of the narrative. The participant must actively select to proceed to be presented with the narrative.

Strategy 5: Proactively Blocking Categories of Narratives

Participants have the ability to block categories of narratives. They can be blocked based on the modality or content. For modality, they have the option to block up to three of the four categories: text, audio, moving images, and static images. For content, they have the option to block any of the five INCREASE content warning categories: abuse or sexual violence, loss of

life or endangerment to life, self-harm including eating disorders, violence or aggression and injustice, and prejudice and discrimination. Narratives in each blocked category are not considered for future presentation. The “About me” section of the website contains the option to unblock previously blocked categories of narratives.

Strategy 6: Reactively Blocking Individual Narratives

If a participant finds a narrative distressing, they can block it during or after its presentation. The narrative will be immediately hidden from them and will not be considered for future presentation. The “About me” section contains the option to unblock previously blocked individual narratives.

Strategy 7: Easy Exit

A button labeled “Get me out of here” is prominently provided throughout the intervention, which when pressed goes to a neutral webpage. This can be used if a participant feels overwhelmingly distressed and wants to quickly leave the intervention or if the interface is being accessed in a public setting and a participant does not want others to know about their usage.

Research Study 1: Narrative Feedback

The clinical and sociodemographic characteristics of the 40 participants included in this study are shown in [Table 2](#).

Table 2. Characteristics of participants of study 3 (N=40).

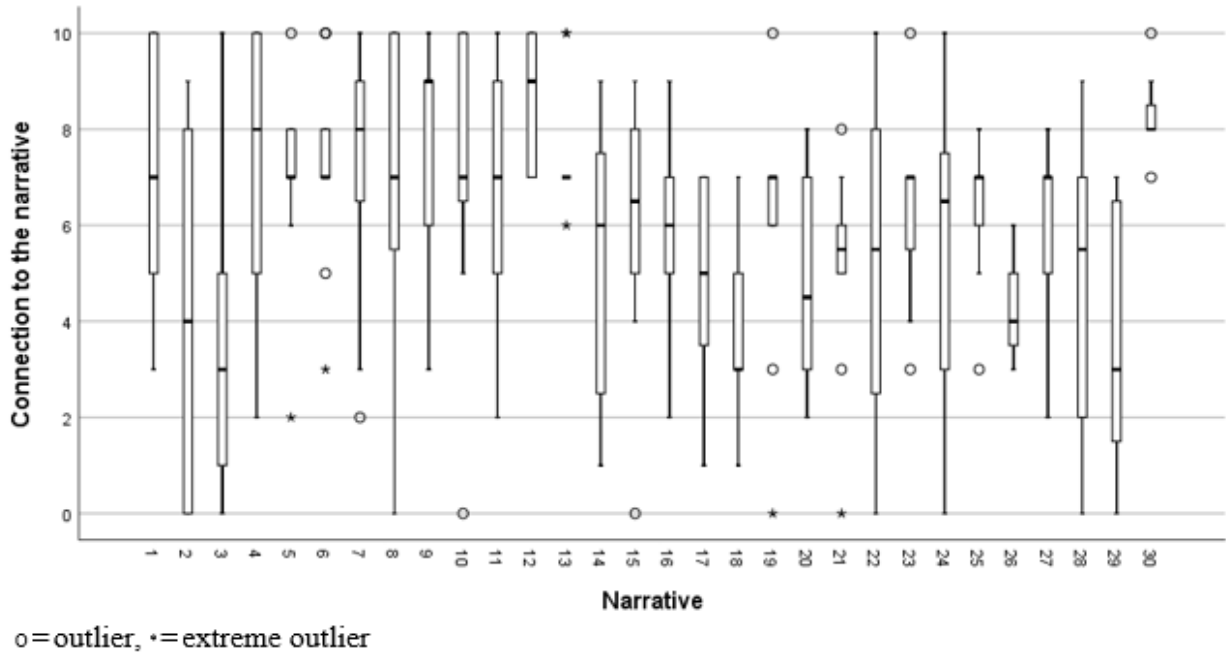
Characteristics	Participants
Gender (female participant), n (%)	24 (60)
Age (years), mean (SD)	44.4 (16.7)
Modality preference indicated, n (%)	28 (70)
Herth Hope Index, mean (SD)	31.1 (5.3)
Diagnosis, n (%)	
Mood disorder	15 (38)
Schizophrenia or other psychosis	9 (23)
Bipolar disorder	7 (18)
Personality disorder	8 (20)
Other	1 (3)
Self-rated recovery trajectory, n (%)	
I am recovered	1 (3)
I am living well	4 (10)
I am making progress	18 (45)
I am surviving day to day	17 (43)

Participants provided feedback on a total of 281 narratives, with a median of 7 randomly selected narratives (range 4-14) accessed per participant from a pool of 30 narratives.

For the 281 ratings of narratives, the mean ratings for connection to the narrative (mean 6.03, SD 2.77), connection to the narrator

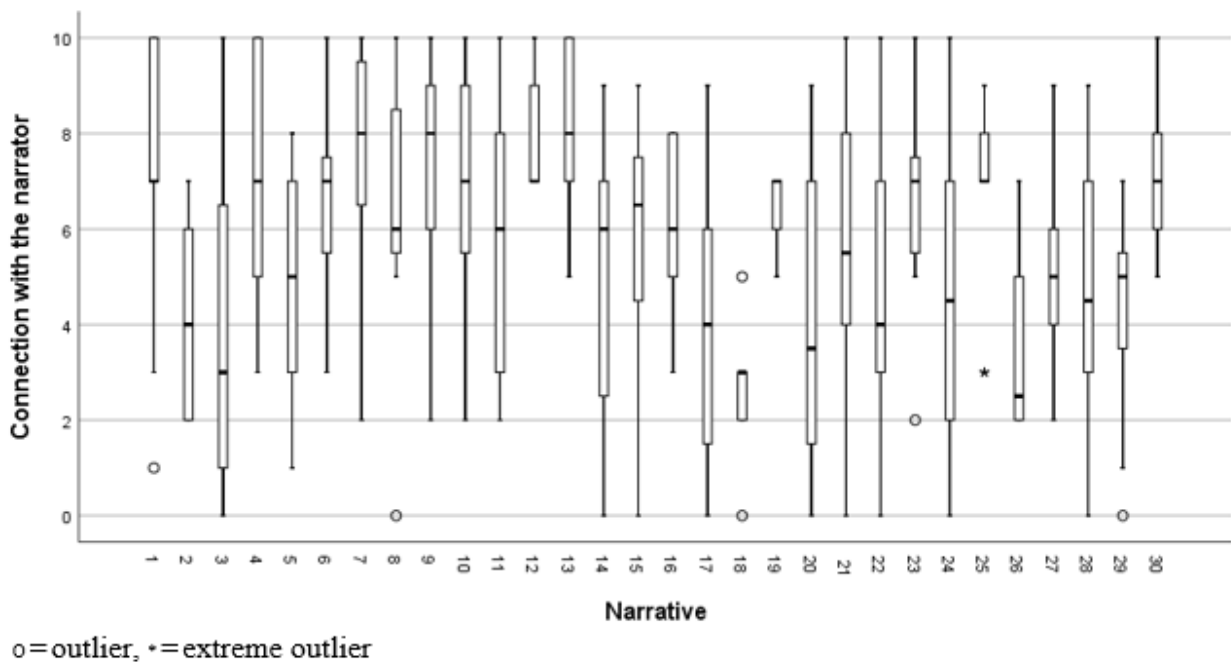
(mean 5.76, SD 2.80), and hope (mean 5.31, SD 2.63) indicated that the narratives had, on average, a neutral to small positive impact on participants in terms of connection and hope, as the chosen rating scales used a value of 5 to indicate a neutral impact. The distribution of narrative connection ratings for each of the 30 narratives is shown in [Figure 2](#).

Figure 2. Ratings by participants (N=40) of connection to the narrative.

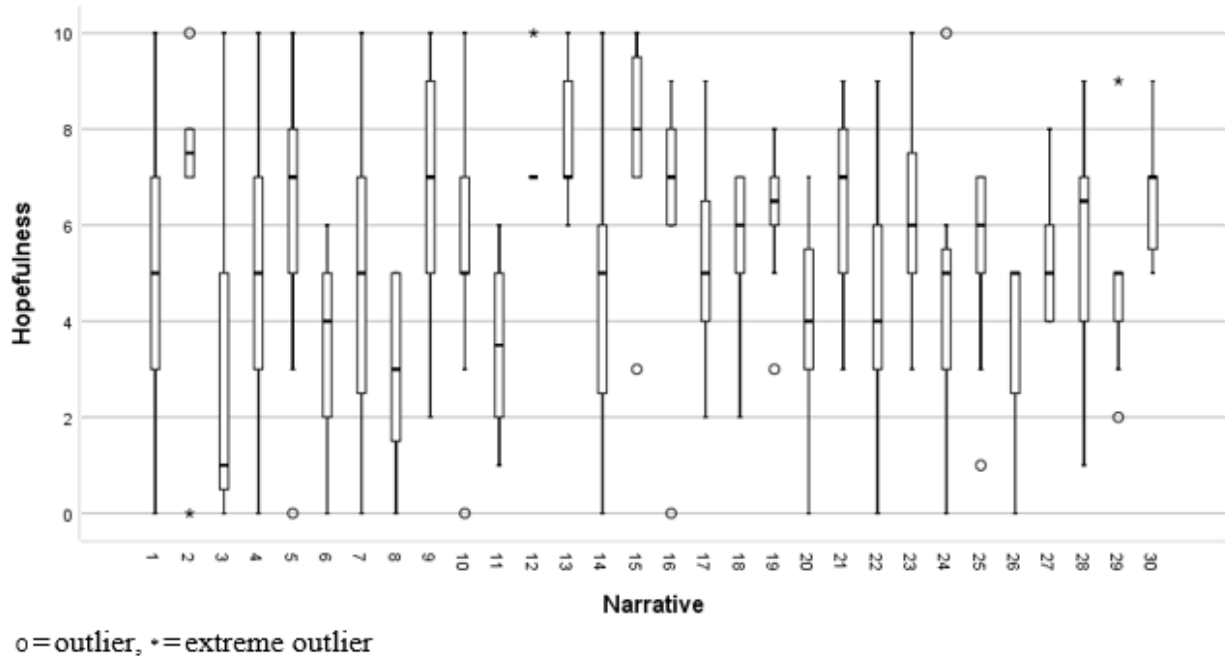


The distribution of narrator connection ratings for each narrative is shown in Figure 3.

Figure 3. Ratings by participants (N=40) of connection with the narrator.



The distribution of hopefulness ratings for each narrative is shown in Figure 4.

Figure 4. Ratings by participants (N=40) of hopefulness for each narrative.

There were no apparent floor or ceiling effects in narrative feedback. Narratives 12 (range 7-10) and 13 (range 6-10) were hope-promoting for all participants in the sample, and narrative 15 had the highest median hope rating of 8.0 (IQR 7-9.8). However, hope ratings were more widely distributed for the other 28 narratives, which were hope-promoting for some participants but hope-reducing for others. The order of presentation did not predict any of the three narrative feedback ratings, and hence, there was no evidence of fatigue effects.

Collectively, these results provide no rationale for discarding the candidate narrative feedback questions evaluated in this study, but they indicate that narrative feedback might be used to tailor recommendations of narratives to individuals, given substantial variability among participants on feedback provided.

Research Study 2: Feasibility Testing

The clinical and sociodemographic characteristics of the 25 participants included in this study are shown in [Table 3](#).

Table 3. Characteristics of participants of study 4 (n=25).

Characteristics	Participants
Age (years), mean (SD)	39.3 (11.6)
Ethnicity, n (%)	
Asian	4 (16)
White	21 (84)
Occupation, n (%)	
Sheltered employment	1 (4)
Employed	9 (36)
Unemployed	8 (32)
Retired	3 (12)
Training or education	4 (16)
Living situation, n (%)	
Alone	8 (32)
With others	17 (68)
Education, n (%)	
No qualification	2 (8)
O-levels, General Certificate of Secondary Education, or equivalent	3 (12)
A-levels, National Vocational Qualification, or equivalent	5 (20)
Degree-level qualification	11 (44)
Higher degree-level qualification	4 (16)
Service contact, n (%)	
No contact with any National Health Service	0 (0)
Contact with my general practitioner only (self-reported)	1 (4)
Contact with my general practitioner and with IAPT ^a	1 (4)
Contact with my general practitioner and a specialist mental health team	23 (92)
Duration of service support in years	
Mean (SD)	13.88 (9.45)
Median (IQR)	10 (7-20)
Has ever been an inpatient, n (%)	10 (40)
Current diagnosis, n (%)	
Schizophrenia or other psychosis	5 (19)
Bipolar disorder	9 (33)
Mood disorder	4 (15)
Other (included ADHD ^b , posttraumatic stress disorder, personality disorder, and autism)	9 (33)
MANSA^c score	
Mean (SD)	4.13 (1.01)
Median (IQR)	4.5 (3.58-4.67)
CORE^d-10 score	
Mean (SD)	18.84 (9.11)
Median (IQR)	20 (13-23)

^aIAPT: Improving Access to Psychological Therapies, a mental health treatment program provided by the National Health Service in England^bADHD: attention-deficit/hyperactivity disorder.

^cMANSA: Manchester Short Assessment of Quality of Life.

^dCORE: Clinical Outcomes in Routine Evaluation.

Follow-up interviews were conducted with 22 participants. One nonattending participant experienced a serious adverse event confirmed as unrelated to the study, and two were unavailable in the follow-up interview period due to significant life events.

Study Objective 5 (Acceptability and Usability)

Feedback during the baseline interview was broadly positive, and no participant indicated that trial procedures or the NEON

Intervention as a whole was unacceptable. Some participants commented that the design of the site used a large amount of blue and white, giving it the feel of a UK-based NHS website, which had negative associations for some.

A summary of the feedback specifically relating to trial procedures, which was rated as unacceptable or less usable, is shown in [Textbox 3](#).

Textbox 3. Identified issues relating to Narrative Experiences Online trial procedures.

Identified Issues
<ul style="list-style-type: none"> Participant Information Sheet (PIS): the PIS was too long. The focus on information about possible harms felt excessively negative and off-putting. A plain English summary would be helpful. Informed consent form (ICF): recording initials in the ICF created anxiety about confidentiality for one participant. Demographics form: the small number of categories for ethnicity was perceived as discriminatory. Some participants wanted to know whether the demographic information would be used in the recommender system. Manchester Short Assessment of Quality of Life (MANSA): some items (eg, “How satisfied are you with your sex life?”) were personal, embarrassing, or uncomfortable to fill in, especially with a researcher present. Submitting the form and then pressing the back button displayed a blank MANSA form, so it was not clear whether the data were successfully submitted. CORE-10: several participants felt distressed by a question on suicidality, and one suggested the “I’m upset” button should be available from this form onward. A suggestion was made that items be rated 1-10 instead of 1-5 and not to reverse-score some items.

A summary of the feedback relating to the NEON Intervention is shown in [Textbox 4](#).

Textbox 4. Identified issues relating to the Narrative Experiences Online Intervention.

Identified Issues
<ul style="list-style-type: none"> Welcome to NEON page: this should be preceded by a message thanking the participant for completing the enrollment process. Useful information page: this page includes the text, “It is normal to have a range of strong emotional responses to stories.” This made the stories sound scary and made the participant feel abnormal if they did not have a strong emotional response. Initial information page (collects information about the participant for use in the recommender system and allows proactive narrative blocking): more explanation was needed about what this was for. It was not clear what role information would have in matching people to stories. It was difficult to understand the difference between this and the demographics form. Some participants felt they had already provided this information. The titles of content warnings were considered to be very blunt and might cause distress. To get to the Initial Information page, participants have to navigate a substantial number of prior pages without seeing any recovery narratives, some showing warnings about harm. This might cause them to doubt the value of recovery narratives. First story page (shows an initial recovery story selected to have no content warnings): it felt quite abrupt to suddenly encounter this story after so many pages of information and data entry. More information was needed about the narrator—there was no explanation as to why this particular story had been chosen. Narrative feedback questions: six participants found it difficult to separate the meaning of the two narrative feedback questions on the connection to the narrator or to the story. Several participants did not understand what the purpose of the narrative feedback questions was. “Get me out of here” button: in total, 15 out of 25 participants explicitly indicated that they liked the inclusion of a button with the named “Get me out of here”; 8 felt that the name was too dramatic and that a more neutral name such as “Quick exit” would be better. “I’m upset” page: a third option was missing—how to access informal peer support for someone who wants to talk to others but does not want to contact a formal service. Give information about the local crisis team. Content warnings: in total, 15 participants felt that these were a good idea to include, but 1 participant suggested all stories should have content warnings and 2 participants suggested that the current content warnings do not sufficiently capture eating disorders.

Regarding the acceptability of web-based voucher payments for trial participation, 15 participants indicated that this was acceptable, one indicated that web-based vouchers restricted choice, and one mistakenly thought that vouchers had to be spent on the web (when in fact web-based vouchers can be

redeemed in a range of shops). In relation to usage prompts, answers included weekly reminders to check in, notifications of new stories, messages sent if the user had not logged in for a while, and messages thanking them for using the system. Messages could be sent by email or SMS text messages.

Feedback in follow-up interviews about using the NEON Intervention was broadly positive, and some participants indicated that access to the interface had helped them with their mental health problems. Some participants wanted direct access to what they described as *inspirational* stories to help lift their mood when they were feeling low. A few participants lost their log-in details, which was a sufficient barrier that they stopped using the system despite the availability of a password reset option. Some reported enjoying access to stories but failing to find someone like them, indicating a need for more diversity in the NEON Collection.

Study Objective 4 (Narrative Feedback) and Study Objective 6 (NEON Intervention Usage)

Participants accessed 330 narratives, of which 253 (76.6%) were selected by the participants and 77 (23.3%) were accessed by the participants asking for a randomly selected narrative. Narrative feedback was provided for 31.8% (105/330) of narratives accessed. The mean length of usage (ie, days from first to last log-in) was 9.8 (SD 9.4), with a median of 9 (IQR 1-16) and range of 1-28. At least one narrative was accessed by 25 participants in week 1, 6 participants in week 2, 4 participants in week 3, and 2 participants in week 4. The mean number of sessions was 2.7 (SD 1.9), with a median of 2 (IQR 2-3) and range of 1-8. The logged data on narrative access are summarized in [Table 4](#).

Table 4. Narrative accesses calculated from logged data about prototype usage by participants (n=25).

Parameter	Weeks 1-4	Week 1	Week 2	Week 3	Week 4
Total narratives accessed per participant					
Mean (SD)	13.2 (10.7)	10.1 (7.6)	6.6 (3.8)	7.3 (7.6)	4 (1.4)
Median (IQR)	9 (5-18)	7 (4-14)	6 (4-10)	5 (2-13)	4 (3-5)
Range	3-42	2-30	2-12	1-18	3-5

The mean number of unique narratives accessed was 9.2 (SD 6.3), with a median of 7 (IQR 4-13), which shows that participants accessed some narratives multiple times. Other key findings include that the function to select a random narrative was used, that usage reduces over the 4-week period, that no participant accessed all available narratives, and that narrative feedback was only provided for a third of the narratives accessed.

Modifications Made Following Research Study 2

Eight modifications to the trial procedures were implemented in light of the feedback collected through research study 2:

1. To reduce the length of the participant information sheet, some items were restructured to present summary text only, with the option to expand to complete the text if desired.
2. To allow anonymity, the informed consent form was modified to require yes or no responses rather than initials and to indicate that the person providing consent can use an email address that does not include their name if they wish. These changes are in keeping with the UK Health Research Authority guidance on seeking consent by electronic means [83].
3. To reduce usage barriers arising from losing log-in details, a participant can request that all necessary access details (ie, website address, email log-in, and chosen password) be sent to their email address as soon as they have given consent.
4. Short messages of a maximum of two paragraphs were added to explain the purpose of each trial procedure.
5. To clarify the role of baseline data collection, a message is shown before the presentation of the forms, which states that the information is being collected to support trial evaluation only and will not be used for matching narratives to participants.

6. To reduce participant discomfort and allow the participant to complete web-based forms in private if they wish, the same message also indicates that some questions might be perceived as sensitive.
7. To address concerns about ethnicity, subcategories were added, consistent with ethnicity guidance from the United Kingdom Office for National Statistics [84].
8. To support understanding of the data being collected in relation to trial procedures, a title is provided to describe the information that will be collected (eg, *Information about you* for the trial demographics form) and a subtitle is provided to describe the role of this information in the trial (eg, *This information will help us understand who is taking part in our trials*).

To produce a final version of the NEON Intervention, the following six modifications were implemented based on the findings presented in this paper. These modifications were included in the NEON Intervention version deployed in the NEON trials [27].

1. Response rates to the narrative feedback questions were lower than anticipated. To reduce burden and hence potentially increase completion of narrative feedback data, the two questions on connection described in research study 1 were made optional and the interface was adjusted accordingly. The rating scale for all feedback was simplified to a 4-point scale, for example, hope was changed to -1 (less hopeful than before), 0 (no change), 1 (a bit more hopeful), and 2 (much more hopeful).
2. To reflect the NEON Impact Model, two optional narrative feedback questions on learning and empathy were added.
3. For participants who wanted support without involving formal mental health services, the "I'm upset" page was extended with information about online peer support services. Services were selected by the research team based

- on features provided to support user safety, such as moderation of discussions and reporting mechanisms in the event of receiving abusive or inappropriate messages.
4. To address concerns about the blunt nature of the content warning systems and to clarify which content warning related to eating disorders, the titles of content warnings were updated. Final content warnings comprised the following: abuse or sexual violence; loss of life or endangerment to life; self-harm including eating disorders; violence or aggression; and injustice, prejudice, and discrimination.
 5. To address the negative perceptions of some participants that the NEON Intervention looked like a UK NHS website, a study logo was produced using a color scheme that is distinctively different from that employed by the UK NHS and was integrated into the NEON Intervention. The color scheme adopted by the NEON Intervention was updated to match the logo.
 6. To improve accessibility, the “Welcome to NEON” and “Useful information” pages were redrafted in collaboration with the NEON LEAP.

Discussion

Principal Findings

The NEON Impact Model developed through knowledge integration activity 1 guided the development of the NEON Intervention, including through informing the selection of narrative feedback questions and through the adoption of a mission to build a large, heterogeneous collection, given that this might increase opportunities for comparison, learning, and empathy for a diverse user base. The Impact Model is relevant to interventions that integrate recovery narratives. It might also guide the selection of recovery narratives by clinicians who integrate these into their practice, for example, by structuring discussions with clients as to which narratives might have the most positive impact.

The curation principles developed through knowledge integration activity 2 are grounded in the practical experience of building a substantial preliminary collection and reporting them supports the transparent management of a collection. This study demonstrates that the VOICES typology can be used as a guideline for collection reporting [24]. These principles can be adopted by other collections.

The experimental evaluation of preliminary narrative feedback questions presented in research study 1 provides evidence that these questions can be used as a mechanism to collect feedback on the impact of narratives, and findings on response rates collected in research study 2 have allowed their design to be refined by categorizing questions as optional or mandatory. The final set of questions may be used in other interventions that require narrative feedback. Research study 1 identified a small number of narratives that were hope-promoting for all in the sample but mostly confirmed our prior findings that narratives are not universally hope-promoting [36]. The study also identified substantial participant variability in response to most narratives, indicating the need to tailor narratives to the needs of participants.

The formal evaluation of a prototype of the NEON Intervention in research study 2 provides evidence that it is feasible to use this intervention in a clinical trial and has allowed for the refinement of intervention and trial procedures. Knowledge about the acceptability of web-based trial procedures designed for the NEON trials can inform the design of web-based trial procedures for other web-based interventions, particularly in relation to the selection of measures where we found evidence that measures designed for delivery on paper were associated with distress for some participants. Safety measures developed for the NEON Intervention in knowledge integration activity 3 might also be more generally applicable to other web-based interventions, particularly those presenting challenging materials.

The immediate clinical relevance of these findings is in informing the NEON randomized controlled trials [27]. The NEON Intervention has been finalized and is now (2020-2022) being evaluated in 3 randomized controlled trials that run in parallel and share the same digital infrastructure. Each trial is designed for participants who may or may not use mental health services and who may or may not choose to receive support from mental health workers in using the intervention.

The long-term generalizability of the findings includes the extension of the NEON Intervention from a mental health focus to (1) other clinical populations, including chronic disorders and palliative and end-of-life care, (2) other marginalized communities beyond health care who may benefit from access to narratives from their community, and (3) other languages and cultures. For any extension, the content of the NEON Collection will need to be widened, and such a program is currently underway to develop a multilanguage repository of Indian mental health recovery narratives, called NEON Collection India.

Strengths and Limitations

The strengths of this study are the mixed methods design to systematically address design issues, the use of multiple groups of participants with diverse mental health problems and experiences of service use, the consistent involvement of people with lived experience of mental health problems at every stage (including the research team and the NEON LEAP), and the use of web-based prototyping to bring concepts to life to obtain ecologically valid feedback.

Limitations of this study include the use of a single regional site for recruitment, the absence of a structured design approach such as the Double Diamond methodology of the Design Council, and the absence of testing of the final version of the trial procedures and NEON Intervention. This last limitation is being addressed through an internal pilot in the NEON Trial.

Comparison With Prior Work

Given that sharing a recovery narrative is a core component of the work of peer specialists [14], the NEON Impact Model might be compared with models created to describe the impact of peer workers. Gillard et al [85] have identified that change comes about through a peer worker (1) building trusting relationships based on shared lived experience, (2) role modeling individual recovery and living well with mental health problems,

and (3) engaging service users with mental health services and the community. These are more relational mechanisms than our own mechanism of *reflecting on personal experience*, presumably due to the recorded nature of recovery narratives that are being received. Synthesizing partial models presented in prior NEON studies [1,36,38] has enabled us to develop a more comprehensive understanding of what can be learned from recovery narratives and what might moderate the impact of narratives.

The variation in feedback received against individual narratives, with some participants finding the same narrative hope-promoting and others finding it pessimism-promoting, suggests a need for the tailoring of narrative selection to the needs of individuals and validates the choice of a recommender

system. The need to tailor digital interventions to recipients has long been recognized in health research, including in studies on stroke rehabilitation technologies [80]. Alankus et al [86] selected a target technology (rehabilitation gaming) and systematically demonstrated how to select properties of the technology that might be tailored. During the NEON trials, logging data will be collected from the recommender system, and an analysis of these data should reveal effective approaches to tailoring narrative selections to trial participants.

Conclusions

Recorded mental health recovery narratives can be integrated into web-based interventions, and it is feasible to conduct an evaluation of such interventions in a clinical trial.

Acknowledgments

This paper is independent research funded by the National Institute for Health Research (NIHR) under its Programme Grants for Applied Research Programme (Programme Grants for Applied Research, Personal Experience as a Recovery Resource in Psychosis: NEON Programme, RP-PG-0615-20016). MS acknowledges the support of the Center for Mental Health and Substance Abuse, University of South-Eastern Norway, and the NIHR Nottingham Biomedical Research Centre. JR acknowledges the support of the NIHR Collaboration for Leadership in Applied Health Research and Care East Midlands, now recommissioned as the NIHR Applied Research Collaboration East Midlands. CB and AC are supported by the Economic and Social Research Council (grant ES/P000711/1). The views expressed are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care.

Conflicts of Interest

None declared.

References

1. Rennick-Egglestone S, Morgan K, Llewellyn-Beardsley J, Ramsay A, McGranahan R, Gillard S, et al. Mental health recovery narratives and their impact on recipients: systematic review and narrative synthesis. *Can J Psychiatry* 2019 Oct;64(10):669-679 [FREE Full text] [doi: [10.1177/0706743719846108](https://doi.org/10.1177/0706743719846108)] [Medline: [31046432](https://pubmed.ncbi.nlm.nih.gov/31046432/)]
2. Chou WS, Hunt Y, Folkers A, Augustson E. Cancer survivorship in the age of YouTube and social media: a narrative analysis. *J Med Internet Res* 2011;13(1):e7 [FREE Full text] [doi: [10.2196/jmir.1569](https://doi.org/10.2196/jmir.1569)] [Medline: [21247864](https://pubmed.ncbi.nlm.nih.gov/21247864/)]
3. Ressler PK, Bradshaw YS, Gualtieri L, Chui KK. Communicating the experience of chronic pain and illness through blogging. *J Med Internet Res* 2012 Oct 23;14(5):e143 [FREE Full text] [doi: [10.2196/jmir.2002](https://doi.org/10.2196/jmir.2002)] [Medline: [23092747](https://pubmed.ncbi.nlm.nih.gov/23092747/)]
4. McGranahan R, Rennick-Egglestone S, Ramsay A, Llewellyn-Beardsley J, Bradstreet S, Callard F, et al. Curation of mental health recovery narrative collections: systematic review and qualitative synthesis. *JMIR Ment Health* 2019 Oct 04;6(10):e14233 [FREE Full text] [doi: [10.2196/14233](https://doi.org/10.2196/14233)] [Medline: [31588912](https://pubmed.ncbi.nlm.nih.gov/31588912/)]
5. Thomas N, Farhall J, Foley F, Leitan ND, Villagonzalo K, Ladd E, et al. Promoting personal recovery in people with persisting psychotic disorders: development and pilot study of a novel digital intervention. *Front Psychiatry* 2016;7:196 [FREE Full text] [doi: [10.3389/fpsy.2016.00196](https://doi.org/10.3389/fpsy.2016.00196)] [Medline: [28066271](https://pubmed.ncbi.nlm.nih.gov/28066271/)]
6. Heilemann MV, Martinez A, Soderlund PD. A mental health storytelling intervention using transmedia to engage Latinas: grounded theory analysis of participants' perceptions of the story's main character. *J Med Internet Res* 2018 May 02;20(5):e10028 [FREE Full text] [doi: [10.2196/10028](https://doi.org/10.2196/10028)] [Medline: [29720357](https://pubmed.ncbi.nlm.nih.gov/29720357/)]
7. Briant KJ, Halter A, Marchello N, Escareño M, Thompson B. The power of digital storytelling as a culturally relevant health promotion tool. *Health Promot Pract* 2016 Nov 10;17(6):793-801 [FREE Full text] [doi: [10.1177/1524839916658023](https://doi.org/10.1177/1524839916658023)] [Medline: [27402721](https://pubmed.ncbi.nlm.nih.gov/27402721/)]
8. Greenhalgh T, Collard A, Begum N. Sharing stories: complex intervention for diabetes education in minority ethnic groups who do not speak English. *Br Med J* 2005 Mar 19;330(7492):628 [FREE Full text] [doi: [10.1136/bmj.330.7492.628](https://doi.org/10.1136/bmj.330.7492.628)] [Medline: [15774990](https://pubmed.ncbi.nlm.nih.gov/15774990/)]
9. Roe J, Brown S, Yeo C, Rennick-Egglestone S, Repper J, Ng F, et al. Opportunities, enablers, and barriers to the use of recorded recovery narratives in clinical settings. *Front Psychiatry* 2020 Oct 30;11:589731 [FREE Full text] [doi: [10.3389/fpsy.2020.589731](https://doi.org/10.3389/fpsy.2020.589731)] [Medline: [33192738](https://pubmed.ncbi.nlm.nih.gov/33192738/)]

10. Williams A, Fossey E, Farhall J, Foley F, Thomas N. Recovery after psychosis: qualitative study of service user experiences of lived experience videos on a recovery-oriented website. *JMIR Ment Health* 2018 May 08;5(2):e37 [FREE Full text] [doi: [10.2196/mental.9934](https://doi.org/10.2196/mental.9934)] [Medline: [29739737](https://pubmed.ncbi.nlm.nih.gov/29739737/)]
11. Manuvinakurike R, Velicer WF, Bickmore TW. Automated indexing of internet stories for health behavior change: weight loss attitude pilot study. *J Med Internet Res* 2014;16(12):e285 [FREE Full text] [doi: [10.2196/jmir.3702](https://doi.org/10.2196/jmir.3702)] [Medline: [25491243](https://pubmed.ncbi.nlm.nih.gov/25491243/)]
12. Benish-Weisman M, Wu LM, Weinberger-Litman SL, Redd WH, Duhamel KN, Rini C. Healing stories: narrative characteristics in cancer survivorship narratives and psychological health among hematopoietic stem cell transplant survivors. *Palliat Support Care* 2014 Aug;12(4):261-267. [doi: [10.1017/S1478951513000205](https://doi.org/10.1017/S1478951513000205)] [Medline: [23942197](https://pubmed.ncbi.nlm.nih.gov/23942197/)]
13. Llewellyn-Beardsley J, Rennick-Egglestone S, Callard F, Crawford P, Farkas M, Hui A, et al. Characteristics of mental health recovery narratives: systematic review and narrative synthesis. *PLoS One* 2019 Mar 28;14(3):e0214678 [FREE Full text] [doi: [10.1371/journal.pone.0214678](https://doi.org/10.1371/journal.pone.0214678)] [Medline: [30921432](https://pubmed.ncbi.nlm.nih.gov/30921432/)]
14. Cronise R, Teixeira C, Rogers ES, Harrington S. The peer support workforce: results of a national survey. *Psychiatr Rehabil J* 2016 Sep;39(3):211-221. [doi: [10.1037/prj0000222](https://doi.org/10.1037/prj0000222)] [Medline: [27618458](https://pubmed.ncbi.nlm.nih.gov/27618458/)]
15. Pitt V, Lowe D, Hill S, Prictor M, Hetrick SE, Ryan R, et al. Consumer-providers of care for adult clients of statutory mental health services. *Cochrane Database Syst Rev* 2013 Mar 28(3):CD004807. [doi: [10.1002/14651858.CD004807.pub2](https://doi.org/10.1002/14651858.CD004807.pub2)] [Medline: [23543537](https://pubmed.ncbi.nlm.nih.gov/23543537/)]
16. Riding the storms. Recovery Devon, UK. URL: <https://library.recoverydevon.co.uk/document/riding-the-storms/> [accessed 2021-05-13]
17. Johnson MD, Dodds N. The plea for the silent. *Med J Aust* 1957 Oct;2(15):547. [doi: [10.5694/j.1326-5377.1957.tb60051.x](https://doi.org/10.5694/j.1326-5377.1957.tb60051.x)]
18. Patient stories: stories of recovery. South London and Maudsley NHS Foundation Trust. 2017. URL: <https://www.slam.nhs.uk/patients-and-carers/patient-information/patient-stories> [accessed 2021-05-13]
19. Smith R. Stop Paddling/Start Sailing: A Journey and Some Ideas. Baltimore, Maryland, United States: America Star Books; 2004:1-64.
20. The mental illness and bipolar video you need to watch. Chris Thompson Photography. 2017. URL: <https://www.youtube.com/watch?v=LuFbEKQME4A> [accessed 2019-10-24]
21. Aisha's schizophrenia story. UPMC Western Psychiatric Hospital. 2019. URL: <https://www.youtube.com/watch?v=uJNWQ7G7Eko> [accessed 2019-10-24]
22. Out of darkness: Joanna's story. Sunnybrook Hospital. 2018. URL: <https://www.youtube.com/watch?v=IIjNDpl2VBs> [accessed 2019-10-24]
23. Erculj N. 3 years after mental hospital. An Honest Chat About My Mental Health & Recovery. 2017. URL: <https://www.youtube.com/watch?v=pQZCcPyZMYE> [accessed 2019-10-24]
24. Yeo C, Hare-Duke L, Rennick-Egglestone S, Bradstreet S, Callard F, Hui A, et al. The VOICES typology of curatorial decisions in narrative collections of the lived experiences of mental health service use, recovery, or madness: qualitative study. *JMIR Ment Health* 2020 Sep 18;7(9):e16290 [FREE Full text] [doi: [10.2196/16290](https://doi.org/10.2196/16290)] [Medline: [32945771](https://pubmed.ncbi.nlm.nih.gov/32945771/)]
25. Telling your story - blogging or vlogging for mind. Mind. URL: <https://www.mind.org.uk/information-support/your-stories/about-the-mind-blog/telling-your-story-blogging-or-vlogging-for-mind/> [accessed 2019-02-22] [WebCite Cache ID [76NeTwf6L](https://www.webcitation.org/76NeTwf6L)]
26. Personal stories. Here to Help. URL: <https://www.heretohelp.bc.ca/personal-stories> [accessed 2020-08-27]
27. Rennick-Egglestone S, Elliott R, Smuk M, Robinson C, Bailey S, Smith R, et al. Impact of receiving recorded mental health recovery narratives on quality of life in people experiencing psychosis, people experiencing other mental health problems and for informal carers: Narrative Experiences Online (NEON) study protocol for three randomised controlled trials. *Trials* 2020 Jul 20;21(1):661 [FREE Full text] [doi: [10.1186/s13063-020-04428-6](https://doi.org/10.1186/s13063-020-04428-6)] [Medline: [32690105](https://pubmed.ncbi.nlm.nih.gov/32690105/)]
28. Slade M, Blackie L, Longden E. Personal growth in psychosis. *World Psychiatry* 2019 Feb;18(1):29-30 [FREE Full text] [doi: [10.1002/wps.20585](https://doi.org/10.1002/wps.20585)] [Medline: [30600621](https://pubmed.ncbi.nlm.nih.gov/30600621/)]
29. Slade M, Rennick-Egglestone S, Blackie L, Llewellyn-Beardsley J, Franklin D, Hui A, et al. Post-traumatic growth in mental health recovery: qualitative study of narratives. *BMJ Open* 2019 Jun 28;9(6):e029342 [FREE Full text] [doi: [10.1136/bmjopen-2019-029342](https://doi.org/10.1136/bmjopen-2019-029342)] [Medline: [31256037](https://pubmed.ncbi.nlm.nih.gov/31256037/)]
30. Shaw L, Homewood J. The effect of eating disorder memoirs in individuals with self-identified eating pathologies. *J Nerv Ment Dis* 2015 Aug;203(8):591-595. [doi: [10.1097/NMD.0000000000000347](https://doi.org/10.1097/NMD.0000000000000347)] [Medline: [26230647](https://pubmed.ncbi.nlm.nih.gov/26230647/)]
31. Larkey LK, McClain D, Roe DJ, Hector RD, Lopez AM, Sillanpaa B, et al. Randomized controlled trial of storytelling compared to a personal risk tool intervention on colorectal cancer screening in low-income patients. *Am J Health Promot* 2015;30(2):59-70. [doi: [10.4278/ajhp.131111-QUAN-572](https://doi.org/10.4278/ajhp.131111-QUAN-572)] [Medline: [25615708](https://pubmed.ncbi.nlm.nih.gov/25615708/)]
32. Guyatt GH, Feeny DH, Patrick DL. Measuring health-related quality of life. *Ann Intern Med* 1993 Apr 15;118(8):622-629. [doi: [10.7326/0003-4819-118-8-199304150-00009](https://doi.org/10.7326/0003-4819-118-8-199304150-00009)] [Medline: [8452328](https://pubmed.ncbi.nlm.nih.gov/8452328/)]
33. Priebe S, Huxley P, Knight S, Evans S. Application and results of the manchester short assessment of quality of life (Mansa). *Int J Soc Psychiatry* 2016 Jul 27;45(1):7-12. [doi: [10.1177/002076409904500102](https://doi.org/10.1177/002076409904500102)]
34. Llewellyn-Beardsley J, Rennick-Egglestone S, Bradstreet S, Davidson L, Franklin D, Hui A, et al. Not the story you want? Assessing the fit of a conceptual framework characterising mental health recovery narratives. *Soc Psychiatry Psychiatr Epidemiol* 2020 Mar;55(3):295-308. [doi: [10.1007/s00127-019-01791-x](https://doi.org/10.1007/s00127-019-01791-x)] [Medline: [31654089](https://pubmed.ncbi.nlm.nih.gov/31654089/)]

35. Serhal E, Crawford A, Cheng J, Kurdyak P. Implementation and utilisation of telepsychiatry in Ontario: a population-based study. *Can J Psychiatry* 2017 Oct;62(10):716-725 [FREE Full text] [doi: [10.1177/0706743717711171](https://doi.org/10.1177/0706743717711171)] [Medline: [28541753](https://pubmed.ncbi.nlm.nih.gov/28541753/)]
36. Rennick-Egglestone S, Ramsay A, McGranahan R, Llewellyn-Beardsley J, Hui A, Pollock K, et al. The impact of mental health recovery narratives on recipients experiencing mental health problems: qualitative analysis and change model. *PLoS One* 2019;14(12):e0226201 [FREE Full text] [doi: [10.1371/journal.pone.0226201](https://doi.org/10.1371/journal.pone.0226201)] [Medline: [31834902](https://pubmed.ncbi.nlm.nih.gov/31834902/)]
37. Kogstad RE, Ekeland T, Hummelvoll JK. In defence of a humanistic approach to mental health care: recovery processes investigated with the help of clients' narratives on turning points and processes of gradual change. *J Psychiatr Ment Health Nurs* 2011 Aug;18(6):479-486. [doi: [10.1111/j.1365-2850.2011.01695.x](https://doi.org/10.1111/j.1365-2850.2011.01695.x)] [Medline: [21749553](https://pubmed.ncbi.nlm.nih.gov/21749553/)]
38. Ng F, Charles A, Pollock K, Rennick-Egglestone S, Cuijpers P, Gillard S, et al. The mechanisms and processes of connection: developing a causal chain model capturing impacts of receiving recorded mental health recovery narratives. *BMC Psychiatry* 2019 Dec 21;19(1):413 [FREE Full text] [doi: [10.1186/s12888-019-2405-z](https://doi.org/10.1186/s12888-019-2405-z)] [Medline: [31864315](https://pubmed.ncbi.nlm.nih.gov/31864315/)]
39. RITB - key principles. Recovery in the Bin. URL: <https://recoveryinthebin.org/ritbkeyprinciples/> [accessed 2021-04-27]
40. Edwards BB, Burgess R, Thomas E. Neorecovery: a survivor led conceptualisation and critique. In: Proceedings of the 25th International Mental Health Nursing Research Conference. 2019 Presented at: 25th International Mental Health Nursing Research Conference; September 13, 2019; The Royal College of Nursing, London, UK.
41. Woods A, Hart A, Spandler H. The recovery narrative: politics and possibilities of a genre. *Cult Med Psychiatry* 2019 Mar 21. [doi: [10.1007/s11013-019-09623-y](https://doi.org/10.1007/s11013-019-09623-y)] [Medline: [30895516](https://pubmed.ncbi.nlm.nih.gov/30895516/)]
42. Fisher P, Lees J. Narrative approaches in mental health: preserving the emancipatory tradition. *Health (London)* 2016 Nov;20(6):599-615. [doi: [10.1177/1363459315600774](https://doi.org/10.1177/1363459315600774)] [Medline: [26304707](https://pubmed.ncbi.nlm.nih.gov/26304707/)]
43. Costa L, Voronka J, Landry D, Reid J, Mcfarlane B, Reville D, et al. "Recovering our Stories": a small act of resistance. *Stud Soc Just* 2012 Oct 16;6(1):85-101. [doi: [10.26522/ssj.v6i1.1070](https://doi.org/10.26522/ssj.v6i1.1070)]
44. Macdonald S. *A Companion to Museum Studies*. Hoboken, New Jersey, United States: Wiley-Blackwell; 2006:1-592.
45. Kreps C. Indigenous curation, museums, and intangible cultural heritage. In: *Intangible Heritage*. Milton, United Kingdom: Routledge; 2008:207-222.
46. Evidence standards framework for digital health technologies. National Institute for Health and Clinical Excellence. 2019. URL: <https://www.nice.org.uk/about/what-we-do/our-programmes/evidence-standards-framework-for-digital-health-technologies> [accessed 2021-04-27]
47. McClay C, Waters L, Schmidt U, Williams C. A survey of attitudes towards computerized self-help for eating disorders within a community-based sample. *Behav Cogn Psychother* 2016 Jan;44(1):65-78. [doi: [10.1017/S1352465814000484](https://doi.org/10.1017/S1352465814000484)] [Medline: [25430776](https://pubmed.ncbi.nlm.nih.gov/25430776/)]
48. Cassidy L. Online communities of practice to support collaborative mental health practice in rural areas. *Issues Ment Health Nurs* 2011;32(2):98-107. [doi: [10.3109/01612840.2010.535648](https://doi.org/10.3109/01612840.2010.535648)] [Medline: [21247275](https://pubmed.ncbi.nlm.nih.gov/21247275/)]
49. Griffiths KM, Christensen H. Internet-based mental health programs: a powerful tool in the rural medical kit. *Aust J Rural Health* 2007 Apr;15(2):81-87. [doi: [10.1111/j.1440-1584.2007.00859.x](https://doi.org/10.1111/j.1440-1584.2007.00859.x)] [Medline: [17441815](https://pubmed.ncbi.nlm.nih.gov/17441815/)]
50. Gibson K, Simms D, O'Donnell S, Molyneaux H. Clinicians' attitudes toward the use of information and communication technologies for mental health services in remote and rural areas. In: Proceedings of the Canadian Society of Telehealth Conference (CST 2009). 2009 Presented at: Canadian Society of Telehealth Conference (CST 2009); October 3-6, 2009; Vancouver, Canada p. 1-14 URL: <http://susanodo.ca/wp-content/uploads/2017/06/2009-CST-Gibson.pdf>
51. Andersson G, Titov N. Advantages and limitations of internet-based interventions for common mental disorders. *World Psychiatry* 2014 Feb;13(1):4-11 [FREE Full text] [doi: [10.1002/wps.20083](https://doi.org/10.1002/wps.20083)] [Medline: [24497236](https://pubmed.ncbi.nlm.nih.gov/24497236/)]
52. Holmes EA, O'Connor RC, Perry VH, Tracey I, Wessely S, Arseneault L, et al. Multidisciplinary research priorities for the COVID-19 pandemic: a call for action for mental health science. *Lancet Psychiatry* 2020 Apr 15 [FREE Full text] [doi: [10.1016/S2215-0366\(20\)30168-1](https://doi.org/10.1016/S2215-0366(20)30168-1)] [Medline: [32304649](https://pubmed.ncbi.nlm.nih.gov/32304649/)]
53. Hare-Duke L, Denning T, de Oliveira D, Milner K, Slade M. Conceptual framework for social connectedness in mental disorders: systematic review and narrative synthesis. *J Affect Disord* 2019 Feb 15;245:188-199. [doi: [10.1016/j.jad.2018.10.359](https://doi.org/10.1016/j.jad.2018.10.359)] [Medline: [30396057](https://pubmed.ncbi.nlm.nih.gov/30396057/)]
54. Llewellyn-Beardsley J, Barbic S, Rennick-Egglestone S, Ng F, Roe J, Hui A, et al. INCREASE: development of an inventory to characterize recorded mental health recovery narratives. *J Recov Ment Health* 2020;14(12):25-44 [FREE Full text]
55. Resnick P, Varian HR. Recommender systems. *Commun ACM* 1997 Mar;40(3):56-58. [doi: [10.1145/245108.245121](https://doi.org/10.1145/245108.245121)]
56. Giacaglia G. Spotify's recommendation engine. 2019. URL: <https://medium.com/datadriveninvestor/behind-spotify-recommendation-engine-a9b5a27a935> [accessed 2019-10-18]
57. Highton-Williamson E, Priebe S, Giacco D. Online social networking in people with psychosis: a systematic review. *Int J Soc Psychiatry* 2015 Feb;61(1):92-101. [doi: [10.1177/0020764014556392](https://doi.org/10.1177/0020764014556392)] [Medline: [25381145](https://pubmed.ncbi.nlm.nih.gov/25381145/)]
58. Alvarez-Jimenez M, Alcazar-Corcoles MA, González-Blanch C, Bendall S, McGorry PD, Gleeson JF. Online, social media and mobile technologies for psychosis treatment: a systematic review on novel user-led interventions. *Schizophr Res* 2014 Jun;156(1):96-106. [doi: [10.1016/j.schres.2014.03.021](https://doi.org/10.1016/j.schres.2014.03.021)] [Medline: [24746468](https://pubmed.ncbi.nlm.nih.gov/24746468/)]
59. Windell D, Norman RM. A qualitative analysis of influences on recovery following a first episode of psychosis. *Int J Soc Psychiatry* 2013 Aug;59(5):493-500. [doi: [10.1177/0020764012443751](https://doi.org/10.1177/0020764012443751)] [Medline: [22532125](https://pubmed.ncbi.nlm.nih.gov/22532125/)]

60. Hodgekins J, Fowler D. CBT and recovery from psychosis in the ISREP trial: mediating effects of hope and positive beliefs on activity. *Psychiatr Serv* 2010 Mar;61(3):321-324. [doi: [10.1176/ps.2010.61.3.321](https://doi.org/10.1176/ps.2010.61.3.321)] [Medline: [20194414](https://pubmed.ncbi.nlm.nih.gov/20194414/)]
61. Fung KM, Tsang HW, Corrigan PW. Self-stigma of people with schizophrenia as predictor of their adherence to psychosocial treatment. *Psychiatr Rehabil J* 2008;32(2):95-104. [doi: [10.2975/32.2.2008.95.104](https://doi.org/10.2975/32.2.2008.95.104)] [Medline: [18840563](https://pubmed.ncbi.nlm.nih.gov/18840563/)]
62. Bødker S. When second wave HCI meets third wave challenges. In: *Proceedings of the 4th Nordic Conference on Human-Computer Interaction: Changing Roles*. 2006 Presented at: *NORDICHI06: The 4th Nordic Conference on Human-Computer Interaction*; October, 2006; Oslo Norway p. 1-8. [doi: [10.1145/1182475.1182476](https://doi.org/10.1145/1182475.1182476)]
63. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M. Developing and evaluating complex interventions. Medical Research Council. 2019. URL: <https://mrc.ukri.org/documents/pdf/complex-interventions-guidance/> [accessed 2021-04-27]
64. Beauchamp TL, Childress JF. *Principles of Biomedical Ethics*. New York: Oxford University Press; 2001.
65. Bell J. Harmful or helpful? The role of the internet in self-harming and suicidal behaviour in young people. *Ment Health Rev J* 2014 Mar 05;19(1):61-71. [doi: [10.1108/MHRJ-05-2013-0019](https://doi.org/10.1108/MHRJ-05-2013-0019)]
66. Jennings H, Slade M, Bates P, Munday E, Toney R. Best practice framework for Patient and Public Involvement (PPI) in collaborative data analysis of qualitative mental health research: methodology development and refinement. *BMC Psychiatry* 2018 Jun 28;18(1):213 [FREE Full text] [doi: [10.1186/s12888-018-1794-8](https://doi.org/10.1186/s12888-018-1794-8)] [Medline: [29954373](https://pubmed.ncbi.nlm.nih.gov/29954373/)]
67. NEON collection web page. URL: <https://www.researchintorecovery.com/research/neon/neoncollection/> [accessed 2021-05-11]
68. Bowen GA. Document analysis as a qualitative research method. *Qual Res J* 2009 Aug 03;9(2):27-40. [doi: [10.3316/qrj0902027](https://doi.org/10.3316/qrj0902027)]
69. Moran TP, Carroll JM. *Design Rationale: Concepts, Techniques, and Use*. Boca Raton, Florida, United States: CRC Press; 1996:1-504.
70. Boysen GA. Evidence-based answers to questions about trigger warnings for clinically-based distress: a review for teachers. *Scholarship of Teaching and Learning in Psychology* 2017;3(2):163-177. [doi: [10.1037/st10000084](https://doi.org/10.1037/st10000084)]
71. Sanson M, Strange D, Garry M. Trigger warnings are trivially helpful at reducing negative affect, intrusive thoughts, and avoidance. *Clin Psychol Sci* 2019 Mar 04;7(4):778-793. [doi: [10.1177/2167702619827018](https://doi.org/10.1177/2167702619827018)]
72. Beverly EA, Díaz S, Kerr AM, Balbo JT, Prokopakis KE, Fredricks TR. Students' perceptions of trigger warnings in medical education. *Teach Learn Med* 2018 Jul 28;30(1):5-14. [doi: [10.1080/10401334.2017.1330690](https://doi.org/10.1080/10401334.2017.1330690)] [Medline: [28753041](https://pubmed.ncbi.nlm.nih.gov/28753041/)]
73. Gainsburg I, Earl A. Trigger warnings as an interpersonal emotion-regulation tool: avoidance, attention, and affect depend on beliefs. *J Experiment Soc Psychol* 2018 Nov;79:252-263. [doi: [10.1016/j.jesp.2018.08.006](https://doi.org/10.1016/j.jesp.2018.08.006)]
74. Kostouros P, Wenzel J. Depictions of suffering in the postsecondary classroom. *Traumatology* 2017 Sep;23(3):250-257. [doi: [10.1037/trm0000110](https://doi.org/10.1037/trm0000110)]
75. Shannon PJ, Simmelink-McCleary J, Im H, Becher E, Crook-Lyon RE. Experiences of stress in a trauma treatment course. *J Soc Work Edu* 2014 Oct;50(4):678-693. [doi: [10.1080/10437797.2014.947901](https://doi.org/10.1080/10437797.2014.947901)]
76. Herth K. Abbreviated instrument to measure hope: development and psychometric evaluation. *J Adv Nurs* 1992 Oct;17(10):1251-1259. [doi: [10.1111/j.1365-2648.1992.tb01843.x](https://doi.org/10.1111/j.1365-2648.1992.tb01843.x)] [Medline: [1430629](https://pubmed.ncbi.nlm.nih.gov/1430629/)]
77. Mayer RE. The promise of multimedia learning: using the same instructional design methods across different media. *Learn Instruct* 2003 Apr;13(2):125-139. [doi: [10.1016/S0959-4752\(02\)00016-6](https://doi.org/10.1016/S0959-4752(02)00016-6)]
78. Barkham M, Bewick B, Mullin T, Gilbody S, Connell J, Cahill J, et al. The CORE-10: a short measure of psychological distress for routine use in the psychological therapies. *Couns Psychother Res* 2013 Mar;13(1):3-13. [doi: [10.1080/14733145.2012.729069](https://doi.org/10.1080/14733145.2012.729069)]
79. Fonteyn ME, Kuipers B, Grobe SJ. A description of think aloud method and protocol analysis. *Qual Health Res* 2016 Jul 01;3(4):430-441. [doi: [10.1177/104973239300300403](https://doi.org/10.1177/104973239300300403)]
80. Rennick-Egglestone S. Tailoring technologies to the rehabilitational needs of stroke survivors. PhD thesis, University of Nottingham. 2014. URL: <http://eprints.nottingham.ac.uk/14153/> [accessed 2021-04-27]
81. Balaam M, Egglestone SR, Fitzpatrick G, Rodden T, Hughes AM, Wilkinson A, et al. Motivating mobility: designing for lived motivation in stroke rehabilitation. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 2011 Presented at: *CHI '11: CHI Conference on Human Factors in Computing Systems*; May, 2011; Vancouver BC Canada p. 3073-3082. [doi: [10.1145/1978942.1979397](https://doi.org/10.1145/1978942.1979397)]
82. European Commission. Directive 2000/31/EC of the European Parliament and of the Council of 8 June 2000 on certain legal aspects of information society services, in particular electronic commerce, in the internal market. The European Parliament and the Council of the European Union 2000:2-32. [doi: [10.5040/9781472559500.0004](https://doi.org/10.5040/9781472559500.0004)]
83. Joint statement on seeking consent by electronic methods. Health Research Authority and Medicines and Healthcare Products Regulatory Agency. 2018. URL: <https://www.hra.nhs.uk/documents/1588/hra-mhra-econsent-statement-sept-18.pdf> [accessed 2021-04-27]
84. National Cancer Institute. Ethnic group. NCI Thesaurus, Code C16564 2020. [doi: [10.32388/8t6czz](https://doi.org/10.32388/8t6czz)]
85. Gillard S, Gibson SL, Holley J, Lucock M. Developing a change model for peer worker interventions in mental health services: a qualitative research study. *Epidemiol Psychiatr Sci* 2015 Oct;24(5):435-445. [doi: [10.1017/S2045796014000407](https://doi.org/10.1017/S2045796014000407)] [Medline: [24992284](https://pubmed.ncbi.nlm.nih.gov/24992284/)]

86. Alankus G, Lazar A, May M, Kelleher C. Towards customizable games for stroke rehabilitation. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2010 Presented at: CHI '10: CHI Conference on Human Factors in Computing Systems; April, 2010; Atlanta Georgia USA p. 2113-2122 URL: <https://doi.org/10.1145/1753326.1753649> [doi: [10.1145/1753326.1753649](https://doi.org/10.1145/1753326.1753649)]

Abbreviations

CSG: Collection Steering Group

HHI: Herth Hope Index

INCREASE: Inventory of Characteristics of Recovery Stories

LEAP: Lived Experience Advisory Panel

NEON: Narrative Experiences Online

NHS: National Health Service

NIHR: National Institute for Health Research

VOICES: Values and motivations, Organization, Inclusion and exclusion, Control and collaboration, Ethics and legal, Safety and well-being

Edited by G Eysenbach; submitted 20.09.20; peer-reviewed by J Farhall, E Toki, O Ness; comments to author 24.10.20; revised version received 06.11.20; accepted 12.11.20; published 27.05.21.

Please cite as:

Slade M, Rennick-Egglestone S, Llewellyn-Beardsley J, Yeo C, Roe J, Bailey S, Smith RA, Booth S, Harrison J, Bhogal A, Penas Morán P, Hui A, Quadri D, Robinson C, Smuk M, Farkas M, Davidson L, van der Krieke L, Slade E, Bond C, Nicholson J, Grundy A, Charles A, Hare-Duke L, Pollock K, Ng F

Recorded Mental Health Recovery Narratives as a Resource for People Affected by Mental Health Problems: Development of the Narrative Experiences Online (NEON) Intervention

JMIR Form Res 2021;5(5):e24417

URL: <https://formative.jmir.org/2021/5/e24417>

doi: [10.2196/24417](https://doi.org/10.2196/24417)

PMID: [34042595](https://pubmed.ncbi.nlm.nih.gov/34042595/)

©Mike Slade, Stefan Rennick-Egglestone, Joy Llewellyn-Beardsley, Caroline Yeo, James Roe, Sylvia Bailey, Roger Andrew Smith, Susie Booth, Julian Harrison, Adarsh Bhogal, Patricia Penas Morán, Ada Hui, Dania Quadri, Clare Robinson, Melanie Smuk, Marianne Farkas, Larry Davidson, Lian van der Krieke, Emily Slade, Carmel Bond, Joe Nicholson, Andrew Grundy, Ashleigh Charles, Laurie Hare-Duke, Kristian Pollock, Fiona Ng. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 27.05.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

A Digital Platform for Facilitating Personalized Dementia Care in Nursing Homes: Formative Evaluation Study

Gubing Wang¹, MSc; Armagan Albayrak¹, PhD; Gerd Kortuem¹, Prof Dr; Tischa JM van der Cammen¹, Prof Dr, MD, PhD

Faculty of Industrial Design Engineering, Delft University of Technology, Delft, Netherlands

Corresponding Author:

Gubing Wang, MSc
Faculty of Industrial Design Engineering
Delft University of Technology
Landbergstraat 15
Delft, 2628 CE
Netherlands
Phone: 31 655548210
Email: g.wang-2@tudelft.nl

Abstract

Background: Care personalization is key to the well-being of people with dementia according to person-centered care. With the development of the internet of things, a large quantity of personal data can be collected securely and reliably, which has the potential to facilitate care personalization for people with dementia. Yet, there are limited assistive technologies developed for this purpose, and the user acceptance of assistive technologies is low in nursing homes. Therefore, through a data-enabled design approach, a digital platform was developed for helping the care team in a nursing home to personalize dementia care, specifically in the management of behavioral and psychological dementia symptoms.

Objective: This study aimed to evaluate the digital platform in a real-life context with potential users from the following two aspects: (1) to explore if the digital platform could help with generating insights on the current state of each person with dementia and (2) to gather feedback on the digital platform from the care team.

Methods: The digital platform was deployed in the nursing home for 7 weeks and the data collected were visualized and presented to the care team via the digital platform. The visualizations were analyzed by the researchers for pattern detection. Meanwhile, the care team was asked to examine the visualizations and were interviewed for the following: (1) if any insights and actions were generated from the examination, (2) the usefulness of the digital platform, and (3) the improvements they would like to see.

Results: The data collected on the digital platform demonstrated its potential for pattern detection. Insights were generated by the care team and categorized into “client level,” “ward level,” and “team level.” The corresponding actions taken by the care team were classified into “investigation” and “implementation.” User acceptance varied across the care team, and three aspects of improvement for the digital platform were identified.

Conclusions: By evaluating the digital platform, this study gained insights on applying data-enabled design for personalizing dementia care; besides, it offers future researchers some recommendations on how to integrate assistive technologies in the nursing home context.

(*JMIR Form Res* 2021;5(5):e25705) doi:[10.2196/25705](https://doi.org/10.2196/25705)

KEYWORDS

human-centered design; data visualizations; person-centered care; people with dementia; assistive technology; health care design; care management; internet of things; data-driven design; data-enabled design

Introduction

Background

Dementia has physical, psychological, and social impacts on affected people, their caregivers, and their families, and it is an economic burden to the society at large [1]. The “person-centered” perspective on dementia care was introduced in care practice by Kitwood [2], where “in addition to neurological impairments, the personality, past experiences, health, and other aspects of the person with dementia also influence how the person will behave.” Inspired by Kitwood’s theories, researchers in dementia care have been exploring how to manage Behavioral and Psychological Symptoms of Dementia (BPSD) in a person-centered manner [3-6].

Over 80% of people with dementia will develop BPSD during the course of their disease [7], and 90% of people with dementia living in nursing homes exhibit BPSD [8]. BPSD contribute to the most stressful, complex, and costly aspects of dementia care, which commonly result in poor health outcomes for people with dementia [9]. Personalization is a key element of the person-centered care approach [10], and most of the current practices accommodate the past experiences, hobbies, capabilities, and preferences of people with dementia in developing personalized care plans [11,12]. However, some of these capabilities, behaviors, personalities, and preferences of people with dementia could change over time as dementia progresses [13]. We propose that these changes should be monitored and considered for providing a holistic approach to personalized BPSD management. We argue a personalized care plan can only be helpful when it is up to date. Yet, it takes time and effort for care teams to notice and adapt to some of these changes given their high workload [14]. Recent research has been investigating the potential of data-driven assistive technologies in care personalization.

Related Work

Over the years, low-cost sensors, wearables, electronic health records, and artificial intelligence have been deployed in the health care setting to collect and analyze data about an individual in terms of his/her physical condition, living environment, lifestyle choices, etc [15]. Researchers hypothesize that these data could help health care professionals to make better and more timely decisions for each patient and hence allow the care received by the patient to be personalized and continuous rather than general and episodic [16]. Collectively, data-driven care personalization is made possible by assistive technologies that can integrate the data and best care practices for health service delivery [17]. A global agenda for personalized telehealth has been proposed [18].

Previous attempts have explored the application of data-driven assistive technologies in the field of cancer care [19], physiotherapy [20], cardiovascular disease [21], and elderly care at home [22]. Conversational agents [23] and data visualizations [19] are being developed for facilitating the interactions between the technologies and their users. However, few studies have investigated this type of assistive technology in personalizing dementia care. The adoption rate of assistive technologies for dementia care in nursing homes has been low.

A scenario-based survey study found that caregivers think assistive technologies have not been tailored to their needs and concerns [24]. Similarly, a systematic review identified that the acceptance of assistive technologies is low without users’ input during its development process [25]. Therefore, a main reason for the low adoption has been identified to be the lack of understanding regarding the interactions between the technology and the people involved.

Research Context

In our prior work, we involved the care team from a Dutch nursing home and developed a digital platform for personalizing BPSD management via a data-enabled design approach. Data-enabled design is about using quantitative data from sensors and qualitative data from users in the field as creative design material to inspire and inform the design process [26]. These data can be highly personalized for an individual. We, therefore, hypothesize that data-enabled design offers an approach to transform these personal data into valuable up-to-date insights about each person with dementia for the care team and hence can facilitate the care team in personalizing BPSD management holistically. The digital platform was then developed for visualizing and presenting the combination of quantitative and qualitative data to the care team and was evaluated in the ward for 15 days. A detailed description of this digital platform can be found in a previous report by Wang et al [27].

We positioned this digital platform as an assistive technology to facilitate the care team in personalizing BPSD management. Studies from the perspectives of care research [28,29], anthropology [30], and gerontology [31] have pointed out that understanding the interplay between the technology and the context (eg, users) to which it is introduced early on could help in the integration of the technology for regular use in the future. Therefore, this study was conducted to explore the effects of the digital platform as an assistive technology on the facilitation of personalized BPSD management in the daily care practice in a nursing home for a longer period of time. This is a qualitative study with the focus of researching the user experience of a care team in using the digital platform.

Aims

This study aimed to (1) explore if the data visualizations of the digital platform could help with generating insights into the current state of each person with dementia and (2) gather feedback on the digital platform from a care team after deployment in a nursing home for a longer period of time. In this way, it builds on our preceding work by investigating the interplay of the digital platform and the care team in a longer time frame to explore the opportunities and challenges for personalizing BPSD care with assistive technologies in a real-life context among potential users.

Methods

Study Setup

In our prior study, an Indoor Positioning System (IPS) was deployed in one of the nursing home sites of Zorggroep Elde Maasduinen in the Netherlands to collect continuous location

data from the residents and the professional caregivers (hereafter the caregivers). Zorggroep Elde Maasduinen is a large-scale organization for the care of older adults in the south of the Netherlands [32], with multiple sites throughout the south of the Netherlands. This study was performed in a nursing-home site in Boxtel, which has a special ward (the Oleander) for caring for residents exhibiting BPSD. This ward is a specific ward for assessing the behaviors of people in the later stages of dementia (from moderate to late stages), who exhibit BPSD, with the aim to find an intervention for these symptoms for each individual person with dementia so that, in time, the person with dementia can return to one of the regular dementia wards in the nursing home.

The Oleander was selected to be the development site of the digital platform since it offers a typical case of BPSD management in an institutional setting. The IPS was hence deployed in the Oleander with the consent of all the legal representatives of all the residents and all care team members. Contextual information concerning the culture of the nursing home, the ward ambiance of the Oleander, and the working structure of the care team is provided in Multimedia Appendix 1. There were 10 residents living in this ward, and the care team regarded each resident as a unique individual. In order to understand the user experiences of the residents and the care team with the use of the digital platform in-depth, the minimum sample size for this study was three people with dementia (given our budget limit).

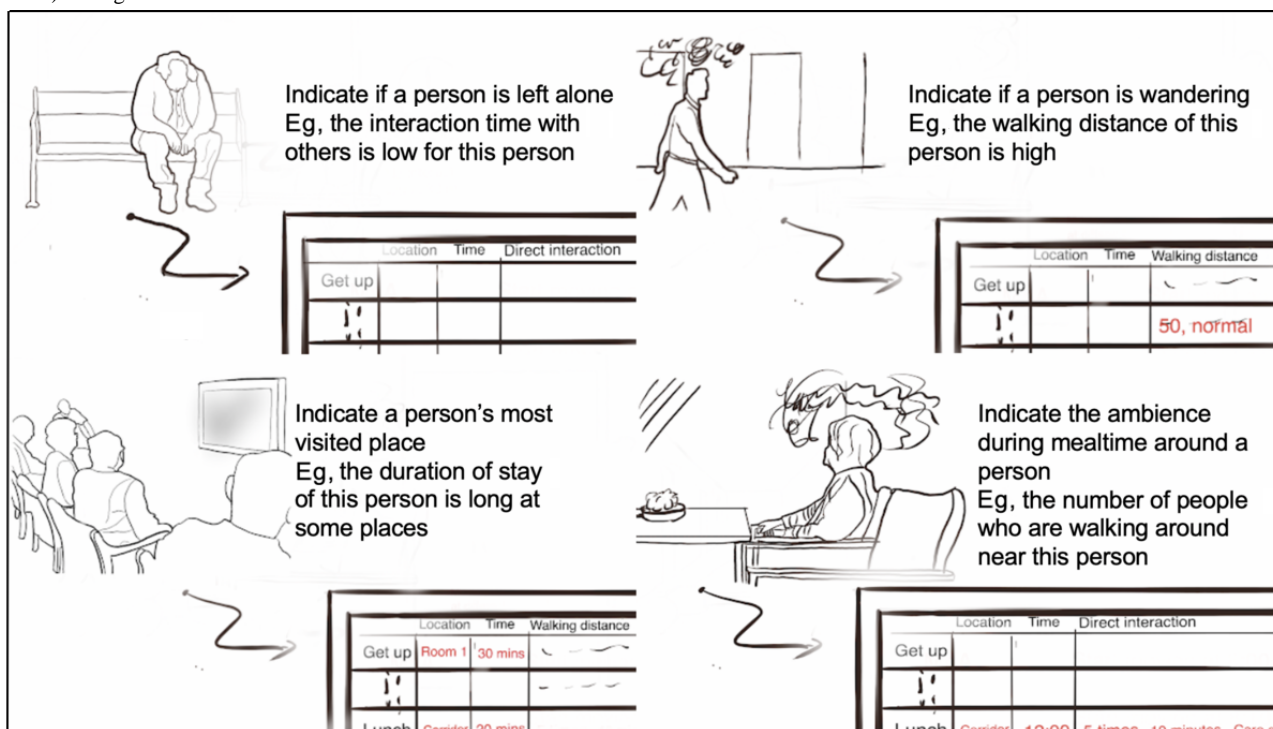
The study protocol was approved by the Human Research Ethics Committee of Delft University of Technology and the Board

of Directors of Zorggroep Elde Maasduinen (see Multimedia Appendix 2 for the ethics approval letters of Delft University of Technology and Zorggroep Elde Maasduinen). Since this study involved collecting location data from people with dementia and caregivers, written informed consent was obtained from the caregivers and from the legal representatives of the people with dementia.

Location data are collected because, among all types of sensor data, this type has been used for monitoring BPSD and is recognized as central to the context of BPSD management [33]. Specifically, not only movement patterns but also other relevant parameters (eg, traveled distance and interaction time with others) could be derived from location data. Both caregivers and residents who participated were given tags since the digital platform was designed with the intention to record the interaction times between caregivers and people with dementia. This interaction time was defined as when and for how long caregivers interacted with people with dementia per day. We programmed the IPS so that a physical distance between two tags of less than half a meter for more than 1 minute was registered as an interaction.

For each resident, several other parameters specific to the daily routine and behavior of the resident were derived from the location data collected (eg, traveled distance, traveled trajectory, and duration of stay). Meanwhile, qualitative data about each resident (eg, daily report) were collected to contextualize the location data. How might the digital platform help with personalized BPSD management is illustrated with a few envisioned scenarios in Figure 1.

Figure 1. Four envisioned scenarios where the digital platform could help with personalized BPSD (Behavioral and Psychological Symptoms of Dementia) management.

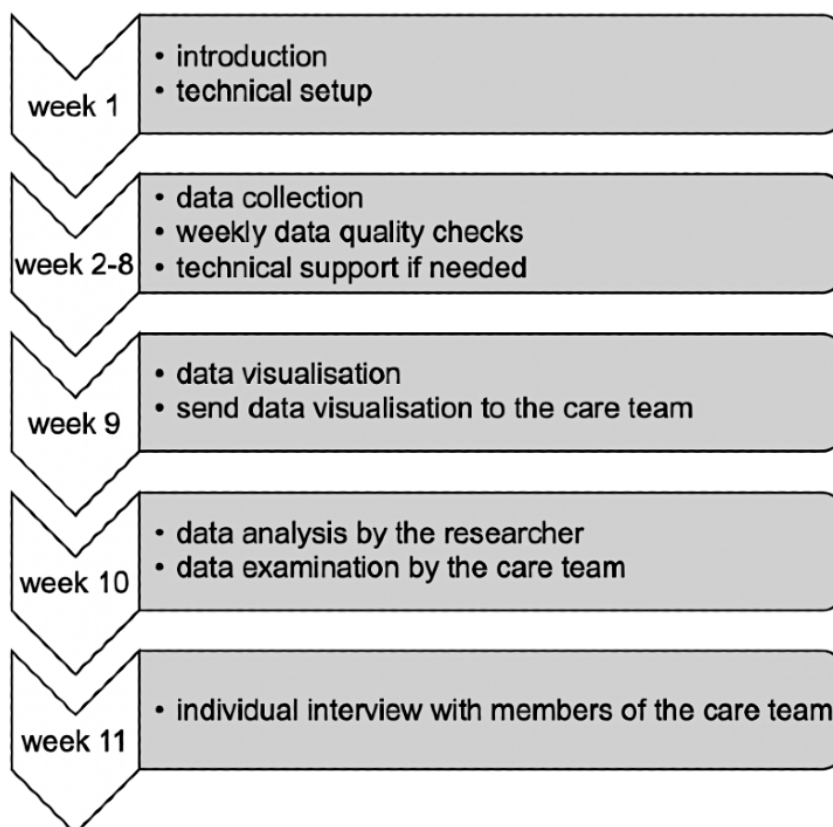


Study Design

In this study, a longitudinal design was employed. As demonstrated in the previous section, evaluating the digital platform for BPSD management in the Oleander constituted a macro case study. Within the macro case study were nested micro case studies, and each micro case study was about one person with dementia. The sampling strategy of this study was a combination of typical case sampling and criterion sampling. Regarding typical case sampling, a broad consensus about what is typical was achieved via a discussion involving the research team and care team. All 10 residents in the Oleander were identified as exhibiting typical BPSD. Within these residents, the legal representatives of eight residents with dementia signed the consent form. Three residents showed signs of dislike toward the tags, and they were excluded from the study. Two participants dropped out later because of signs of dislike toward the tags. Thus, three participants completed the study, resulting in three micro case studies. The three people with dementia were two females and one male. Moreover, one had vascular dementia, one had Lewy body dementia, and one had Alzheimer disease. Based on the information from the care team, all three

people with dementia were in the later stages of dementia, as supported by their Mini-Mental State Examination scores. Basic information about these three people with dementia can be found in [Multimedia Appendix 3](#). As for criterion sampling, having a direct influence on the care plan of each person with dementia was considered to be the inclusion criterion. For each person with dementia, the responsible caregiver of the person with dementia, the ward doctor (hereafter doctor), the ward psychologist (hereafter psychologist), and the ward dietitian (hereafter dietitian) had direct influences on the care plan. In addition, the ward manager (hereafter manager) was considered to have a direct influence on the care plans because of participation in multidisciplinary meetings for care plan updates. All members of the care team who met the criterion were approached, and all of them consented to participate in the study. Therefore, each micro case study included the person with dementia and his/her responsible caregiver, doctor, psychologist, dietitian, and manager. The sample size was considered to be adequate for this constructivist qualitative research, where a new, deep, and nuanced understanding was aimed to be gained on personalizing BPSD management with a data-enabled design [34]. The study procedure is illustrated in [Figure 2](#).

Figure 2. Study procedure.



Specifically, in week 1, an introduction was given to the participants, and the IPS was tested for its functionality. In weeks 2 to 8, the residents and caregivers wore tags for collecting their location data in the ward in everyday life. The caregivers wrote daily reports and gave a color code on the perceived stress levels of the people with dementia (ie, stress rating). For minimizing the workload of the care team, these tasks were designed to be as close as possible to the working routine of the care team. As for the daily report, it was part of

the working routine of the care team. The daily report usually recorded when and where the person with dementia got stressed, what did he/she do, and what did the caregiver do to reduce his/her stress. Once a report was written, caregivers could select a category for the report to be uploaded to the system (eg, physical health and mental health). These reports were shared with other members of the care team, which could sometimes result in additional remarks and editing. The stress rating was done with the Crisis Development model

(Crisisontwikkelingsmodel in Dutch) [35], which is a standardized assessment tool used in several Dutch nursing homes for rating the stress levels of the residents, and each member of the care team was trained to use it. This model divided the stress level of a person into green, yellow, orange, and red, indicating no stress (green) to high stress (red). The care team recorded and categorized the effective measures carried out to help a resident relax at different stress levels on a so-called signal plan. Each resident had a personalized signal plan. Based on the signal plan, the behaviors of the residents acted as signals for their stress levels. The care team could then react by adjusting the interaction style and environment for each resident according to the signal plan to reduce the stress for the resident. This way of working has been found to be helpful by the care team. Before the development of the digital platform, the care team did not explicitly record the stress of a resident over time. During this study, the caregivers were instructed to conduct the stress rating. The frequency of stress rating changed to every half an hour in this study (the previous frequency was once per day [27]) as the caregivers found that the stress levels of people with dementia usually change drastically over the day. The quality of data was checked weekly, and technical support was provided when needed.

In week 9, both quantitative and qualitative data were visualized on the digital platform. Since this study spans a longer period of time, new types of visualizations were created on the digital platform for conveying the large amount of collected data

effectively. Once the visualizations were done, a notification was sent to the caregivers to start the data examination. This is because the user scenario adopted was as follows: “the responsible caregiver examines the visualizations and discusses his/her findings with other team members in a care plan meeting” (refined based on a previous study [27]). Hence, the data visualizations were developed for caregivers to do the first round of data examination, and feedback on the digital platform from the whole care team was gathered. The term “examine” was used in referring to the caregiver’s and care team’s inspection of the data so as to reserve the term “analyze” for the researchers and algorithms. In week 10, both the researchers and care team studied the visualizations with different goals. While the goal of the care team was to uncover insights about people with dementia via a combination of quantitative and qualitative data, the goal of researchers was to investigate if behavioral patterns of people with dementia could be identified with the quantitative data (ie, the availability and utility of the data) and check with the care team about the identified patterns afterward. In week 11, individual interviews were carried out with the responsible caregiver of each person with dementia, doctor, psychologist, dietitian, and manager. The doctor, psychologist, dietitian, and manager were interviewed three times, and each time, they were interviewed specifically about each individual person with dementia. This led to 15 interviews (ie, five interviews per person with dementia). The participants involved in the study are shown in [Table 1](#).

Table 1. Participants involved in the study.

Study phase and participant type	Number of participants
Data collection	
People with dementia	3
Caregivers	12
Data examination	
Caregivers	3
Doctor	1
Psychologist	1
Dietitian	1
Manager	1

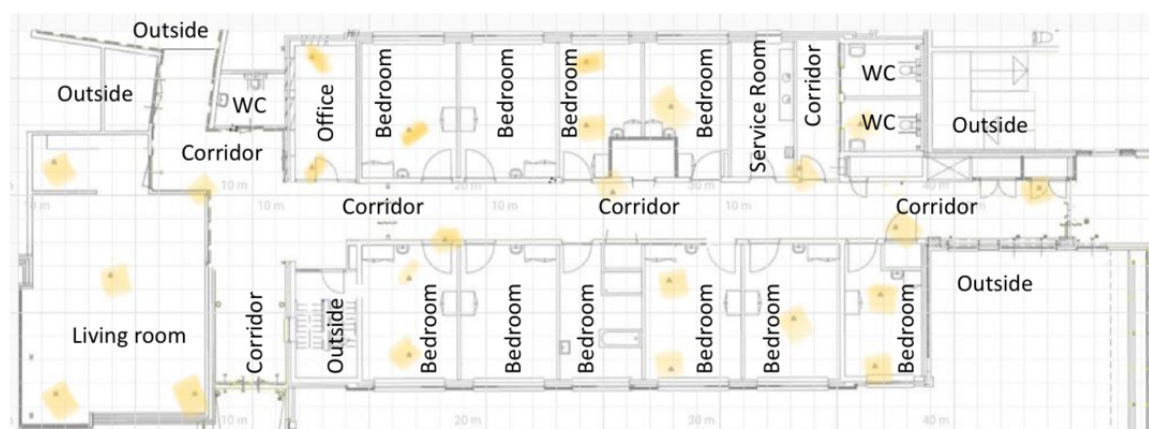
Data Collection

The categories of data collected and their collection times in this study are summarized in [Table 2](#). Precisely, some background information about each person with dementia was collected at the beginning of the study. The IPS data, daily

reports, and stress rating of each person with dementia were collected from weeks 2 to 8. Lastly, the interviews were conducted at the end of this study and transcribed verbatim. The setup of the location data collection in the ward is shown in [Figure 3](#).

Table 2. Details of collected data in the study.

Category of data	Collection time	Description
Background information	At the beginning of the study	Age, gender, basic clinical background, and typical behaviors of people with dementia
Indoor positioning system data	Continuous daytime collection from weeks 2 to 8	Location of people with dementia and caregivers
Daily reports	Once or a few times per day from weeks 2 to 8	Perception about when and where did the person with dementia get stressed, what did he/she do, and what did the caregiver do to reduce his/her stress
Stress rating	Every half an hour during the waking time of the residents from weeks 2 to 8	Perception of caregivers about the level of stress expressed by the person with dementia (color code used from low to high stress: green, yellow, orange, and red)
Feedback on the digital platform	At the end of the study	Semistructured interview on discovered insights and corresponding actions (if any), usefulness, and desired improvements of the digital platform

Figure 3. Set-up of location data collection in the ward (locations of the sensors are marked yellow on the map, and the data collected are sent to the server in the office via Wi-Fi gateways).

Data Analysis

Overview

Three analyses were performed to fulfill the research goals outlined above. The first analysis examined the availability and utility of the data collected by the digital platform. The second analysis investigated what types of insights and actions, if any, could be generated by the care team via using the digital platform. The third analysis studied the perceived usefulness and desired improvements regarding the digital platform.

Concerning the transferability of this study, we followed the strategies by Polit and Beck [36], which recommend that a sufficiently detailed description of the context is needed for the study results to be meaningful to other researchers. Transferability is defined as “the degree to which the results of qualitative research can be transferred to other contexts or settings with other respondents” [37]. Describing not only the study results but also their context could help the study results become meaningful to an outsider [37]. Hence, basic information about the nursing home where the research was conducted and a description of the three people with dementia are provided in

[Multimedia Appendix 1](#) and [Multimedia Appendix 3](#), respectively.

Analysis 1: Data Availability and Utility

The aim of this analysis was to determine whether the digital platform could collect adequate data for the intended purpose of identifying behavioral patterns. With “adequate,” we implied the amount of data that could enable behavioral patterns to be revealed via visual inspection. Visual inspection, sometimes referred to as visual analysis, is the most widely used and recommended method for interpreting single-subject data [38,39]. This method involves the researcher analyzing the data visually, which allows for a holistic evaluation for understanding the idiosyncrasies present in the data for each subject [40]. The amount of data collected depends both on the technical functioning of the IPS and the participants’ adherence (ie, keeping the tags charged and connected, and wearing the tags correctly).

The researchers examined the data availability by checking the proportion of the study period for which data were available for each person with dementia. Visual inspection was then applied to identify whether behavioral patterns could be revealed for each person with dementia. If behavioral patterns were found,

these findings would be validated with the care team. Three types of visualizations were created in the end, namely, tile plots (example in Figure 4), combined plots (example in Figure 5), and mapping plots (example in Figure 6). Only tile plots

were used for the purpose of visual inspection (the guide for visual inspection and the validation process with the care team can be found in Multimedia Appendix 4).

Figure 4. Movement distance in the corridor for participant 1 from April 1 to May 20, 2020, for each day in the daytime (tile plot).

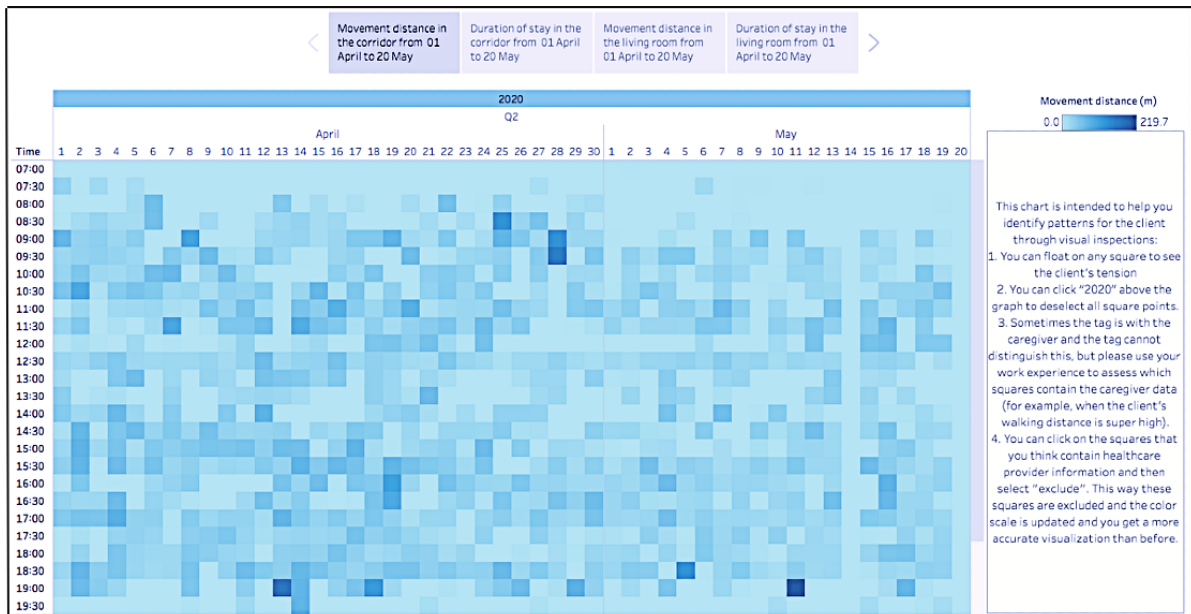


Figure 5. Duration of stay for participant 2 in all possible rooms in the ward per day in the daytime and corresponding daily report (combined plot).

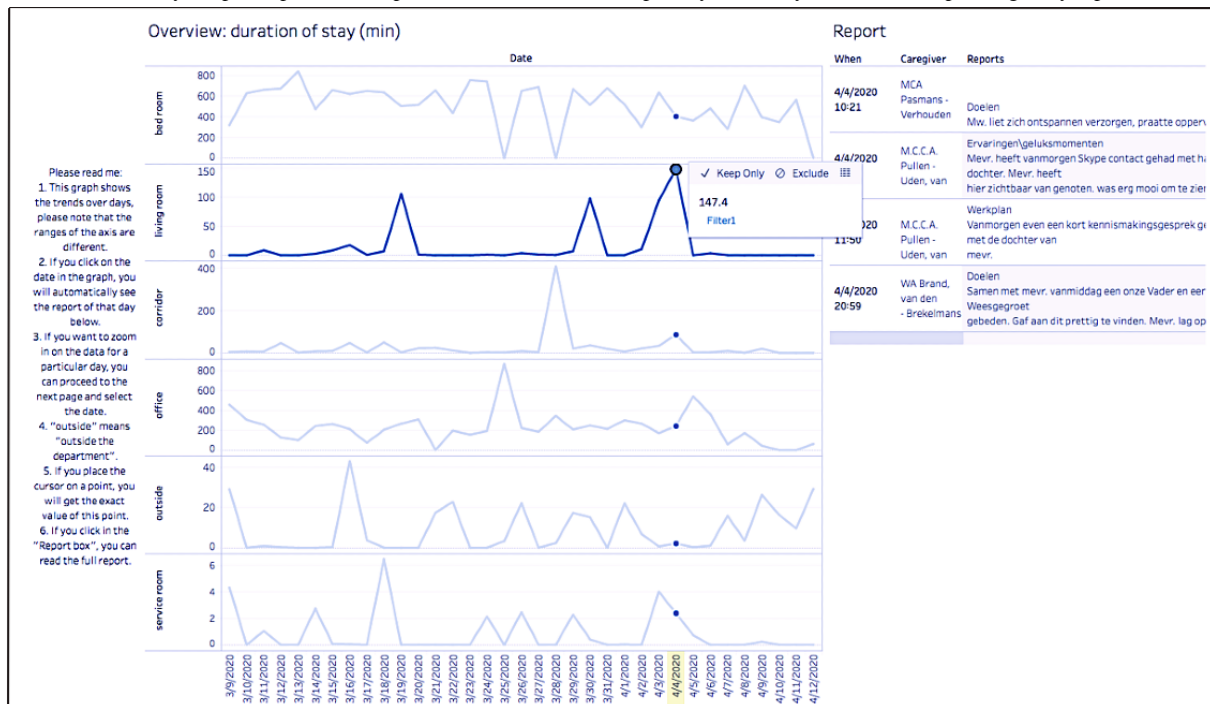
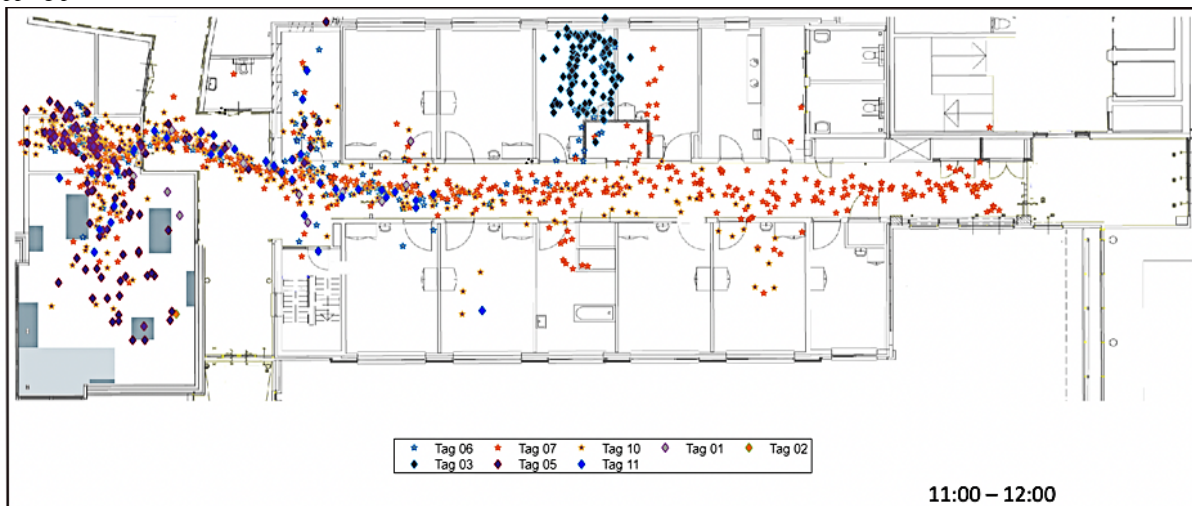


Figure 6. Movement trajectories of people with dementia (tags 1-3) and caregivers (tags 4-11) in the ward from 11:00 AM to 12:00 PM on April 16, 2020 (mapping plot).



Analysis 2: Types of Insights and Actions

The aim of this analysis was to explore what types of insights and actions, if any, could be generated by the care team from using the digital platform. To prepare the caregivers for the data examination, we sent them an email with a link to the three types of visualizations and instructions. The caregivers were asked to examine the data individually, and after that, they provided their insights at the care plan meetings. In this way, the insights were discussed within the care team. The caregivers were notified that they could contact the research team if any questions arose during the examination process. Being involved in developing the digital platform previously, the care team has gained experience in this type of data examination.

During the interview, each interviewee was asked if he/she had identified any insights from data examination and, if so, whether he/she would take any actions based on the insights. The interview guide can be found in [Multimedia Appendix 5](#). Based on thematic analysis of the interview transcripts, GW and AA categorized the types of insights and actions into themes according to the six-step guidance by Braun and Clarke [41].

Analysis 3: Usefulness and Future Improvement

The aim of the third analysis was to evaluate the perceived usefulness of the digital platform by the care team. During the interview, each interviewee was asked if this digital platform was useful to his/her work and what could be improved. The interview guide can be found in [Multimedia Appendix 5](#). Thematic analysis was conducted by GW and AA following the same guidelines as in Analysis 2.

We acknowledge that people with dementia were among the key stakeholders in this project. Since the people with dementia were in the later stages of dementia, we chose not to involve them in data examination and interviews. Instead, we involved care team members as they evaluated and monitored the current physical and emotional states of the residents daily and responded to their needs. All this information contributes to the personal care plan. For the people with dementia, their data were collected and studied, and the generated insights were shared with the care team and hence contributed to a timely

update of their personal care plans. Given that dementia is progressive, receiving up-to-date care is essential for the well-being of people with dementia. In this way, people with dementia can benefit from participating in this project.

Results

Data Availability and Utility

The researchers analyzed whether location data collected over a longer period of time could reveal any behavioral patterns of people with dementia, such as trends and fluctuations in measured parameters within a day or over days. After the improvement of the tag design, the amount of missing data was reduced in comparison to the preceding study. However, there were still cases where no data were collected for a whole day. For example, two days of data were missing over the 7-week period for participant 1, as shown in [Figure 4](#) (May 14 and May 20, 2020). Four days and three days of data were missing for participants 2 and 3, respectively. The missing data were mainly associated with the incorrect use of tags by the caregivers (ie, placing the tag on the wrong person or forgetting to charge or use the tags). Compared with a preceding study, the amount of missing data decreased [27]. Specifically, 3 days out of 7 weeks (on average) contained missing data in this study. In contrast, 7 days out of 15 days contained missing data in the preceding study [27].

It seems that the missed data did not inhibit the potential for identifying behavioral patterns via visual inspection. To illustrate, we present a tile plot of the location data for participant 1 from April 1 to May 20, 2020, in [Figure 4](#). Specifically, this figure visualizes the movement distance of this person with dementia in the corridor each day in the daytime, where the color intensity corresponds to the movement distance (ie, light color: short distance; dark color: long distance). From [Figure 4](#), participant 1 was found to have a longer movement distance in the corridor in April than in May and rarely moved in the corridor around noon (which corresponds to lunch time).

In general, this visualization in [Figure 4](#) is representative of all three micro case studies in terms of the availability and utility

of the location data (more tile plot examples can be found in [Multimedia Appendix 6](#)). The data collected over this period of time were adequate for behavioral patterns to be identified via visual inspection, despite some missing data. When discussing with the care team, we found that the daily patterns were usually within the expectations of the care team, mostly because they coincided with the daily routines of the participants (eg, the care team expected participant 1 to be in the living room at lunch time). In summary, the daily patterns uncovered about each person with dementia validated the experience and knowledge of the care team. Yet, the trends and periodic patterns over the days were usually new insights to the care team (eg, the care team did not know the total movement distance of participant 1 decreased over time).

In addition to tile plots, two more types of visualizations were also developed. The combined plot presented the daily reports and daily location data of the person with dementia side by side, as shown in [Figure 5](#). Specifically, [Figure 5](#) shows the duration of stay of participant 2 in all possible rooms in the ward each day, and clicking a data point on the graph provides the exact value of the duration of stay and the daily reports of this corresponding date. The main utility found for visualizing the duration of stay and daily reports together was to identify when the participants were under stress and what was done to reduce their stress. From previous research, we know that people with dementia might be sent back to their bedrooms if they are showing symptoms of stress, since there is a common belief within the care team that reducing the number of stimuli could help one to reduce stress, and the bedroom is regarded to have fewer stimuli than the rest of the ward. The daily reports normally recorded what happened and what was done around the stress moments.

The mapping plot presents the location data of all participants in their movement trajectories in the ward every hour in the daytime, as shown in [Figure 6](#). For this visualization, on selecting a specific participant's location trajectory, the

trajectory of the selected person is highlighted, with all other participants' trajectories visible in the background. The main utility found for this visualization was to identify if any person with dementia exhibits unique movement patterns spatially. From [Figure 6](#), participant 3 was found to walk in circles when in the bedroom alone, which the care team indicated as insightful. The care team wondered if participant 3 was looking for a way out. From past experience, they know that participant 3 tends to walk when under stress, and this visualization made them realize that participant 3 walked in the room for a while after being sent to the room for relieving stress. This type of visualization was created since the care team had expressed that they would like to know the exact locations (eg, where in a room) and movement trajectories of the residents to uncover detailed movement patterns [27].

Types of Insights and Actions

Types of Insights

The reported time spent on data examination ranged from 0.5 to 1.5 hours. Twenty-nine insights were generated by the caregivers, the doctor, the psychologist, and the dietitian based on their interpretation of the data visualizations. The manager did not find any insights. These insights were grouped into 13 subthemes that are under three themes. The first theme, client level (the care team refers to residents as their clients), includes (1) day structure according to care plan, (2) moments of unrest, (3) unusual movement trajectory, (4) behavior change over time, (5) physical activity over time, and (6) effect of medication. The second theme, ward level, includes (7) interaction with caregivers, (8) interaction with other people with dementia, and (9) dining environment. The third theme, team level, includes (10) more detailed reports, (11) tag usage, (12) behaviors of caregivers, and (13) workflow. These themes and subthemes of insights are presented in [Table 3](#), with each theme illustrated by an example quote. A detailed analysis of all the interviews can be found in [Multimedia Appendix 7](#).

Table 3. Types of insights generated from data examination (the care team refers to residents as their clients).

Theme and subtheme	Example quote ^a
Client level	
Day structure according to care plan	“He has a day structure, in which he goes to the toilet two times a day around 11 am and 3 pm. From the data, sometimes he goes to the toilet once, and sometimes he does not go to the toilet at all. Because of his agitation, he forgets to ask to go to the bathroom, and then we might forget about it too.” [Caregiver 1] (mapping plot + stress rating)
Moments of unrest	“He more often gets agitated in the afternoon than in the morning. I think maybe after lunch; he starts to think about the next step; he cannot wait to have something to eat. Sometimes he asks for food; sometimes, he asks what are we going to do next? He is a bit bored in the afternoon.” [Caregiver 1] (tile plot + daily report)
Unusual movement trajectory	“I see his usual pattern from his room to the living room and to the kitchen. There has been one time to the back of our ward. Why? I don't know, but that's not the route he usually takes. Yeah.” [Caregiver 3] (mapping plot + daily report)
Behavior change over time	“It seems that the connection (between restlessness and stress) is no longer there ... perhaps his stress manifests itself less in movement and more in shouting. So, the restlessness has moved from motor to verbal. That is something I know from experience.” [Doctor] (tile plot + stress rating)
Physical activity over time	“You can see that in the morning he is more active. And during the day his walking distance gets less.” [Caregiver 3] (combined plot)
Effect of medication	“He has medication at 8 am, 12 pm, 5 pm. When he is restless, he gets extra antipsychotic medicine. I sometimes noticed the medicine is effective and sometimes not.” [Caregiver 1] (combined plot)
Ward level	
Interaction with caregivers	“It strikes me that he is relatively alone when he goes back and forth (in the corridor). Except at 5-6 pm when there is care routine, and he is (with somebody and) keep moving back and forth (in the corridor).” [Doctor] (mapping plot)
Interaction with other people with dementia	“Most of the time, I would like to see my clients only (in the digital platform), but sometimes when my clients interact a lot with other clients, then it's also sometimes useful to know what other clients are doing.” [Dietitian] (mapping plot)
Dining environment	“I wonder what they have been doing during mealtime? Because some people I don't see in the living room (where the meal is served). It would also be nice if there is some quiet time around mealtime because when there's a lot of distraction, some people forget to eat or go walking.” [Dietitian] (mapping plot)
Team level	
More detailed reports	“Sometimes even when a high stress level is recorded for the client, there is no corresponding daily report to explain what happened.” [Psychologist] (combined plot)
Tag usage	“If I look at 3rd May, yeah, I think he had his tag put on only at 11:30. Sometimes, on 14th May, he doesn't have it (the tag) with him at all.” [Caregiver 3] (tile plot)
Behaviors of caregivers	“I am also impressed by the distance traveled by the care staff ... the staff is “more restless” than the residents. What's normal? I find it interesting to reflect on that with the team.” [Doctor] (mapping plot)
Workflow	“It also gives me insight about all our daily things. And that's I think that my colleagues they are...that they are more interested to see this.” [Caregiver 3] (mapping plot + daily report)

^aThe main data sources on which the insights were based are presented in parentheses after each example quote.

Types of Actions

Except for the manager, the rest of the care team participants wanted to take some actions based on the insights they uncovered. Their intended actions were grouped into six subthemes that are under two themes. The first theme, investigation, includes (1) discuss causes behind insights, (2)

monitor day structure, and (3) evaluate changes in care plan. The second theme, implementation, includes (4) change in care plan, (5) work practice, and (6) prediction. These themes and subthemes of actions are presented in [Table 4](#), with each theme illustrated by an example quote. A detailed analysis of all the interviews can be found in [Multimedia Appendix 7](#).

Table 4. Types of actions based on the insights generated.

Theme and subtheme	Example quote
Investigation	
Discuss causes behind insights	“I see this client moves a lot when he is in high stress level around 3 pm; we don't know what is going on, maybe because he wants to go to the bathroom, or he has nothing to do. It is a signal that things are not OK for him; hopefully, we can find reasons for this” [Psychologist]
Monitor day structure	“This allows us to see what has or has not already been offered in a day. This is easy to look back.” [Caregiver 2]
Evaluate changes in care plan	“We decided...to let him go to bed earlier. And I hope when we see a new view (visualization) then I can see a difference in that (stress rating). To get to know if it's helpful for him.” [Caregiver 3]
Implementation	
Change in care plan	“We should bring him to the bathroom twice a day; this could help him to relax. I will discuss with my colleagues and update the care plan on this.” [Caregiver 1]
Work practice	“There should always be a caregiver in the living room when he is in the living room. He doesn't like to be alone.” [Caregiver 1]
Prediction	“From the data, I know when the client is more likely to get tense. Previously, I only observe their behaviors to get to know if he is tensed or not. It is predictive.” [Caregiver 3]

Usefulness and Future Improvement

Usefulness

The feedback of the care team participants on the usefulness of the digital platform for their work was categorized according

to their professions. Their feedback is summarized in [Table 5](#) with example quotes. A detailed analysis of all the interviews can be found in [Multimedia Appendix 7](#).

Table 5. The perceived usefulness of the digital platform.

Profession	Perceived usefulness	Example quote
Caregiver (n=3)	Provide evidence for discussion and for confirming feelings	“It is nice that we use the data as the evidence when we discuss what we see with the doctor and psychologist.” [Caregiver 3] “The data are confirmations of our feelings towards the client.” [Caregiver 2] “I feel they are a confirmation about what I already knew.” [Caregiver 1]
Doctor (n=1)	The insights are useful, but more scientific evidence is needed	“I think it is getting better and better. The clinical relevance is still complicated; for each client, we can do something with that, but for a scientific basis, not only a feasibility study but also a real clinical study is needed.”
Psychologist (n=1)	Triangulate subjective report of the caregivers with collected data	“For me, since I am not in the ward myself, I normally talk with the caregivers; it is good to see how often he is in stress (from the visualizations).”
Dietitian (n=1)	Need more data related to food and dining	“It tells me where the person is, how long the person stays there. So, it gives some data for me. Yes. And stress. But it doesn't mean a lot for the dietitian. It doesn't say a lot about food; it is more about what's being done.”
Manager (n=1)	Digital platform is not helpful for my work	“Is it helpful for my work? not so much, because if I must put a conclusion, I have to have more data, more in an overview ... I couldn't draw any conclusions from this data.”

Future Improvement

The care team participants formulated 23 key areas for developing the digital platform further. These were grouped into 10 subthemes that are under three themes. The first theme, data collection, includes (1) more types of data, (2) define whose data to collect, and (3) reliability of stress rating. The second theme, data visualization, includes (4) personalized parameters,

(5) filter data by time, and (6) less clicking. Finally, the third theme, data examination, includes (7) reduce the examination time, (8) who should examine the data, (9) develop an examination workflow, and (10) automatic notifications. These themes and subthemes of areas of improvement are presented in [Table 6](#), with each theme illustrated by an example quote. A detailed analysis of all the interviews can be found in [Multimedia Appendix 7](#).

Table 6. Areas of improvement for the digital platform.

Theme and subtheme	Example quote
Data collection	
More types of data	“What I missed is the actions, the interventions that the team members have done and what the effect is on the behavior of the clients; for me, it's hard to find any conclusion about examining this data.” [Manager]
Define whose data to collect	“We have known him for a long time ... so we have done lots of analysis of his behaviors. I think these data will give much more information if it is someone new, who we don't know much about.” [Psychologist]
Reliability of stress rating	“Sometimes, stress-rating does not match the daily reports ... that is very unfortunate because you cannot see many things properly ... I don't know if there is any more convenient method for this (stress-rating).” [Doctor]
Data visualization	
Personalized parameters	“We only know the distance and how long he has been in the corridor. We would like to know how many times he moved back and forth in the corridor; this indicates his agitation.” [Caregiver 1]
Filter data by time	“It would be helpful to see what the person is doing at a time they should eat ... that would tell me where the person is at mealtime or is he walking around? And that's interesting. If I can select the time, then that would be nice.” [Dietitian]
Less clicking	“If I can see the report of the day when I hover on the data of that day, that would be good, I don't have to select the date for the report, and it will make the process faster.” [Psychologist]
Data examination	
Reduce the examination time	“This is new to me, I enjoyed doing it once. However, it is very difficult to look at the data when I have to care for clients.” [Caregiver 2]
Who should examine the data	“I noticed that caregivers have a lot of trouble in analyzing the data. And I think that's ... it has several reasons. But one of the most, I think the most important is that they are not used to analyzing data.” [Manager]
Develop an examination workflow	“We just talk about these in our regular meetings. It could be good if someone would first look at the data so that the data is not new for everybody ... and discuss with the team in the regular meetings. During the meeting, everyone can discuss if they see the same things and why or why not.” [Psychologist]
Automatic notifications	“It would be nice if the device can generate some insights automatically to help us with the examination over time; for example, it can tell us when the data deviates from the baseline.” [Caregiver 1]

Discussion

Principal Results

By evaluating the digital platform in a real-life context with potential users for 7 weeks, this study gained insights on applying data-enabled design for personalizing dementia care in a nursing home. The results demonstrated the potential for the data visualizations in the digital platform to reveal behavioral patterns despite the missing data. In addition, we identified three main types of insights generated from data analysis, two main types of corresponding actions, the perceived usefulness of the digital platform, and three areas for its improvement.

Implications of Analysis 1

The results showed that integrating the digital platform into the nursing home environment might offer an opportunity for the care team to uncover the behavior patterns of the residents and personalize care plans accordingly.

A study on adapting mobile and wearable technology to support and monitor rehabilitation for people with dementia in the home environment has discussed that the monitoring approach could replace traditional methods for behavior analysis (eg, questionnaires) with four advantages [42]. The findings from this study coincide with three of the four advantages identified, which are as follows: the behavioral patterns identified by monitoring could be of higher resolution, data quality is independent of human recall, and collaborative care could be facilitated by sharing the data visualizations. The last advantage

identified by this previous study is that the workload of the care team could be lowered by the monitoring approach [42]. In our case, the interplay of monitoring with assistive technology and the current practice of the care team was explored; hence, the time spent by the care team in managing the tags and examining the data was factored in. In this way, the overall workload increased in the short term, which was the main pain point mentioned by the care team.

The workload of the care team members might decrease in the long term as they gain experience with managing the tags and examining the visualizations. Meanwhile, the future development of the digital platform should focus on autonomizing some tasks for the care team. Moreover, BPSD management contributes to a high percentage of caring workload [43], and the insights generated from using the digital platform may offer novel and personalized ways for managing BPSD. Thus, using the digital platform would potentially enable BPSD management to be easier for the care team. In this study, most members of the care team identified the potential added value of this digital platform to their care practice; besides, they became more familiar with data collection and examination. Higher familiarity and having perceived needs have been recognized as two contributing factors for better user adherence [42]. Hence, a longer-term evaluation is needed to examine if the effort and time invested in using the digital platform could be outweighed by the effort and time saved from using it.

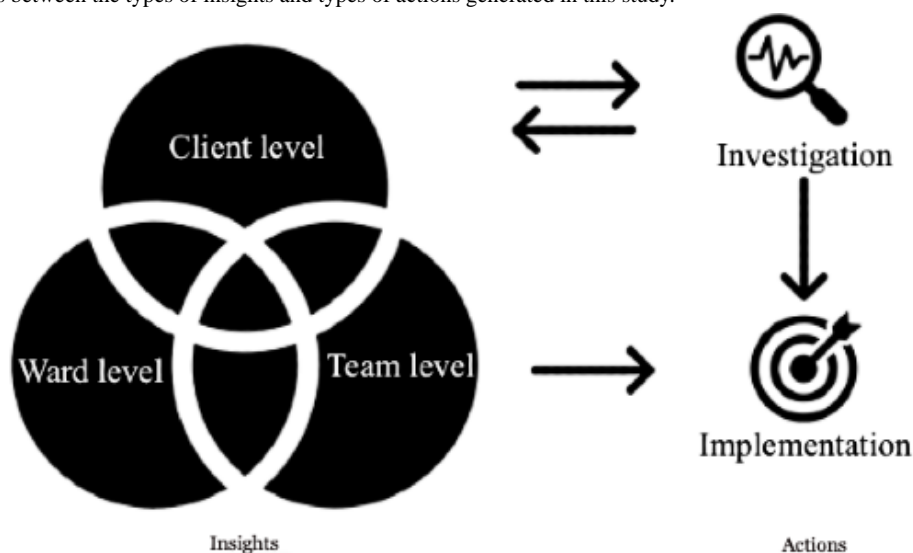
Implications of Analysis 2

The three main types of insights generated after data analysis were at the client level, ward level, and team level. Even though the initial aim of introducing the digital platform was to collect data that are unique to each individual, the care team generated insights that were not limited to the clients. This is because the context in which a person with dementia lives is also an important part to consider when personalizing dementia care, which reveals the importance of context in personalized dementia care. We recommend that future researchers and developers pay attention to this.

Although the types of insights and the types of actions were categorized in the Results section, they were interrelated in various ways. A tentative overview of the interrelations among them is shown in Figure 7. From the “Insight” section of this figure, some insights at the client level, ward level, and team

level mutually influence one another. For example, when some caregivers increase awareness about their behaviors (team level), their interaction approach with the residents might change (ward level), and this may lead to a reduction in the frequency of BPSD moments for the residents (client level). Moreover, due to the interrelations between these insights, uncovering one insight may lead to the discovery of other insights. For example, realizing that the day structure of a resident is usually disrupted at certain times (client level) may trigger the care team to recall if there are any differences in this resident’s environment at these times (ward level) or to reflect if there are any difficulties in the workflow of the caregivers around these times (team level). Despite these overlaps, we feel that all the themes and subthemes provide a set of specific types and subtypes of insights, respectively. Discussing the various separated themes and subthemes naturally results in making sense of their interconnections.

Figure 7. Tentative relations between the types of insights and types of actions generated in this study.



Zooming out from the “types of insights” in Figure 7, we postulate that these insights interact with the two types of actions identified, which are “investigation” and “implementation.” When an insight is in the form of a statement (eg, the client forgets to go to the bathroom when agitated), the care team will implement it in care practice (eg, bring the client to the bathroom even though he did not ask for it); however, when an insight is in the form of a question (eg, what has the client been doing during dinner time?), the care team will inspect to find the answer by “investigation” (eg, meeting, observation, and collect and examine more data), and this process will lead to more insights. When the care team members are certain about their findings after more investigation, they will then implement them in care practice.

These types of insights and actions together with their interactions could also be applied in evaluating assistive technologies in other contexts of personalizing dementia care in nursing homes. We hence encourage future researchers and developers to use Figure 7 as a conversation guide when evaluating these assistive technologies with a care team and be open-minded about more types of insights and actions generated.

Implications of Analysis 3

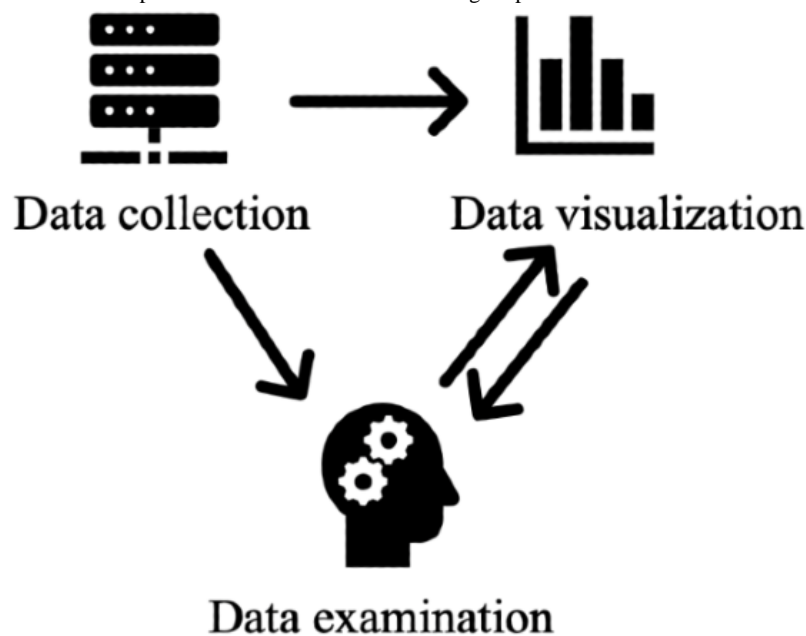
The people with dementia, caregivers, doctor, psychologist, dietitian, and manager were among the key stakeholders for the digital platform. Previous studies have found that caregivers’ acceptance of assistive technologies is vital for successful integration and usage of these technologies in nursing homes [24]. We would like to add that, in the case of managing BPSD with assistive technologies, all members of the care team should be involved. This is because a group effort is needed in incorporating insights from the data examination into the current care practice for BPSD management. Different stakeholders tend to have different interests in a project. In our previous study, we identified that the main value held by the care team is to provide better care to people with dementia. As the digital platform was implemented in the ward for a longer period of time, we noticed that, despite the main value, it is important to cater for the different interests among stakeholders. Specifically, most members of the care team identified the potential added value of this digital platform in their care practice. The manager was mainly interested in the efficiency of care in the ward and therefore saw fewer benefits from the digital platform for her work.

To manage the competing needs stemming from different stakeholders, we recommend future studies to apply a multicriteria analysis to prioritize the stakeholders involved [44]. For example, the manager might be interested in different aspects of care in comparison to other members of the care team and thus could be involved differently. Knowing the reasons behind the low acceptance of some stakeholders and prioritizing stakeholders early on could help with the development of future assistive technologies for dementia care in nursing homes.

We postulate that the three main areas of improvement identified for the digital platform, namely, data collection, data visualization, and data examination, are interconnected to each other. Their tentative connections are shown in Figure 8. First, a change in data collection would affect both data visualization and data examination. For example, introducing data about nonverbal behavior might lead to a change in previous data

visualizations to accommodate this new data type, and the examination process may have to be adapted accordingly. Second, an update in data visualization tends to affect data examination more than data collection. For instance, incorporating the hovering option for quick report reading can reduce the data examination time directly, but data collection will not be affected. Third, the data examination process, in turn, tends to affect data visualization more than data collection. Take the inclusion of automatic notifications as an example; while data collection could be the same, additional visualizations might be needed for presenting the data based on which the algorithmic analysis is made. This is to ensure the transparency of algorithms to the users, which has been found to be of paramount importance in decision-making [45]. The researchers and developers are advised to be aware of the mutual influences of these three areas when developing similar kinds of assistive technologies.

Figure 8. Tentative connections of the three improvement areas identified for the digital platform.



Reflection on Data-Enabled Design

Through the evaluation of the digital platform, we also learned about the value of data-enabled design in personalizing BPSD management. First, we experimented with the frequency and duration of collecting different types of data to explore the qualities and limitations of data as design material, as suggested by Bogers and van Kollenburg (who coined data-enabled design) [26]. In this way, the visualizations were perceived as relevant and meaningful to most members of the care team (eg, by increasing the frequency of stress rating). Second, after data had been collected over a longer period of time, the care team was able to gain insights from data visualizations that they otherwise would not have gained via traditional methods (eg, observation). The same has been found in previous data-enabled design studies [46]. Moreover, the care team generated insights at not only the client level but also the ward and team levels. This corresponds to the intention of data-enabled design as follows: “gain detailed and nuanced contextual, behavioral, and experiential insights” [26].

In addition, we adapted data-enabled design in several aspects to fit it in the dementia care context. First, even though one of our key stakeholder groups, people with dementia, was not involved in the data examination and interview sessions owing to cognitive impairments, we paid attention to their nonverbal behaviors when introducing the tags to them. Residents who showed signs of dislike toward the tags were excluded from the study, although their family members had signed the consent forms. We then included “design the tags to be more dementia friendly” on the to-do list. We found that more relevant design directions could be generated by combining data-enabled design with close observations of people with dementia (a way of co-designing with people having dementia [47]). In contrast, a previous data-enabled design study on developing a smart baby bottle only involved the parents in the design research [48]. Similar to people with dementia, babies can communicate their needs nonverbally, and more insights might be generated if the designers also involve the babies during the design process in the future.

Second, the data-enabled design approach was initially developed for designers to gain rich insights for their design projects. As stated by Bogers and van Kollenburg, “the aim of this approach is to, together with end-users, unravel the relevance, potential, and pitfalls of data in a specific context to design concepts that resonate” [49]. In our study, we identified that the care team could also gain insights about people with dementia and the contexts with this approach to improve their care practice. Since care team members are experts in understanding their clients and the care context, by involving them in the data examination, they can see things that designers cannot see. The designers can then learn from the insights generated by the care team and apply these insights in the design process. We, therefore, recommend the care team to be more involved in the data-enabled design process not only because the care team can gain more insights to improve their work, but also because the designers can learn more from working with the care team.

Finally, if one would like to apply data-enabled design in the dementia care context, an evaluation of the designed product (in our case, assistive technology) in the longer term is important. Both [Figures 7 and 8](#) indicate that there is an interplay between the technology introduced and the current BPSD care in the nursing home. There is a growing body of research on understanding how technology innovations change the dynamics between the people involved in health care and care practices, and these changes could, in turn, alter the role of the introduced technology in daily practice over time [50]. This study adds to this body of work by gaining an understanding of the drivers and barriers of integrating assistive technology via a data-enabled design approach in the institutional dementia care setting.

Limitations and Future Work

The main limitation of this study was that all the interviews were conducted in a one-to-one format. This is because finding a time slot suitable for all the participants was difficult given the fast pace and high uncertainty of the working environment in the nursing home. A potential shortcoming of this format is the lack of discussions within the care team, which might have stimulated more ideas during the interviews. The benefit of this format is that each participant can express his/her own views and interpretations without being influenced by other participants. The strengths of this study were that the digital platform was deployed in a real-life context and a wide range of stakeholders with different professional backgrounds was involved. Previous research has found that the use of different professionals' perspectives is valuable in developing assistive technologies for people with dementia [51-53].

For future work, the digital platform could be improved based on the feedback from the care team. For instance, regarding reducing the data examination time, the manual process of data visualization is planned to be automated by algorithms. In addition, further algorithms could be developed for pattern analysis (to replace visual inspection) and for making predictions. The results of the algorithmic analysis could then

be presented in the form of push notifications, with which relevant visualizations will be provided.

Moreover, it would be interesting to investigate if combining different types of visualizations could generate more insights. Although all three types of visualizations were sent to the care team at the same time, all participants discussed the visuals one by one, which indicates that combining visualizations might not have been explored during the data examination process. It could be that there was no explicit instruction for exploring visuals in a combined way, or all participants did not see added value when combining these visuals; therefore, these explorations were not mentioned in the interviews.

In this study, we presented an in-depth investigation with a care team that was insightful in understanding the value of the digital platform and the application of data-enabled design in developing assistive technologies for personalizing BPSD care. We, therefore, suggest researchers and developers consider this case study approach in the development of future assistive technologies for personalizing dementia care.

Implication for Dementia Care During a Pandemic

This study was carried out during the lockdown period in the Netherlands because of the COVID-19 pandemic. At the beginning of the project, the research team (GW, AA, and TvdC) went to the nursing home to introduce and discuss the study both at the management level and at the ward level. The principal researcher (GW) then spent a few months developing the digital platform with the care team. When the nursing home shut down because of COVID-19, the principal researcher stayed in touch with the care team via email, Skype, and Zoom. In this way, the research team collected all the data without going to the nursing home and conducted the interviews online with the care team. In hindsight, this digital platform has the potential to help care team members and researchers gain insights about each person with dementia remotely, which might be valuable in case of the current pandemic and future pandemics. In addition, the family members of people with dementia could benefit by using the digital platform to stay updated about the situation of their loved ones in the nursing homes remotely, which could decrease the already high communication workload of the care team during a pandemic.

Conclusions

By evaluating the digital platform developed, this study gained insights on applying data-enabled design for personalizing dementia care, specifically in BPSD management. The collected data demonstrated potential use for pattern detection. The types of insights and actions generated from using the digital platform were identified and found to be interconnected. The perceived usefulness of the digital platform was found to vary across the care team, and we uncovered three aspects of improvement for the digital platform. These findings could guide future researchers and developers in investigating similar assistive technologies for personalizing dementia care in nursing homes and beyond.

Acknowledgments

The authors would like to thank the employees and residents of Zorggroep Elde Maasduinen nursing home in the Netherlands for their participation and the company PinXact for providing the indoor positioning system. This study was funded by the Chinese Scholarship Council (201708060174).

Conflicts of Interest

None declared.

Multimedia Appendix 1

Contextual information about the study field.

[[DOCX File , 20 KB - formative_v5i5e25705_app1.docx](#)]

Multimedia Appendix 2

Ethics approval letters of Delft University of Technology and Zorggroep Elde Maasduinen.

[[PDF File \(Adobe PDF File\), 175 KB - formative_v5i5e25705_app2.pdf](#)]

Multimedia Appendix 3

Contextual information about the people with dementia who participated in the study.

[[DOCX File , 16 KB - formative_v5i5e25705_app3.docx](#)]

Multimedia Appendix 4

Guide for visual inspection.

[[DOCX File , 172 KB - formative_v5i5e25705_app4.docx](#)]

Multimedia Appendix 5

Interview guide.

[[DOCX File , 20 KB - formative_v5i5e25705_app5.docx](#)]

Multimedia Appendix 6

Examples of visual inspection with tile plots.

[[DOCX File , 1015 KB - formative_v5i5e25705_app6.docx](#)]

Multimedia Appendix 7

Detailed description of thematic analysis results.

[[DOCX File , 20 KB - formative_v5i5e25705_app7.docx](#)]

Multimedia Appendix 8

CONSORT-eHEALTH checklist (V 1.6.1).

[[PDF File \(Adobe PDF File\), 1151 KB - formative_v5i5e25705_app8.pdf](#)]

References

1. Dementia. World Health Organization. URL: <https://www.who.int/news-room/fact-sheets/detail/dementia> [accessed 2020-07-30]
2. Kitwood T. Dementia Reconsidered: The Person Comes First. Berkshire, United Kingdom: Open University Press; 1997.
3. Cohen-Mansfield J, Ray CA. Whose responsibility is it to make life worth living? Int. Psychogeriatr 2014 Jun 02;26(8):1231-1233. [doi: [10.1017/s1041610214000982](https://doi.org/10.1017/s1041610214000982)]
4. Cohen-Mansfield J, Thein K, Marx MS, Dakheel-Ali M. What are the barriers to performing nonpharmacological interventions for behavioral symptoms in the nursing home? J Am Med Dir Assoc 2012 May;13(4):400-405 [FREE Full text] [doi: [10.1016/j.jamda.2011.07.006](https://doi.org/10.1016/j.jamda.2011.07.006)] [Medline: [21872537](https://pubmed.ncbi.nlm.nih.gov/21872537/)]
5. Testad I, Corbett A, Aarstrand D, Lexow KO, Fossey J, Woods B, et al. The value of personalized psychosocial interventions to address behavioral and psychological symptoms in people with dementia living in care home settings: a systematic review. Int. Psychogeriatr 2014 Feb 24;26(7):1083-1098. [doi: [10.1017/s1041610214000131](https://doi.org/10.1017/s1041610214000131)]
6. van der Ploeg ES, Eppingstall B, Camp CJ, Runci SJ, Taffe J, O'Connor DW. A randomized crossover trial to study the effect of personalized, one-to-one interaction using Montessori-based activities on agitation, affect, and engagement in nursing home residents with Dementia. Int. Psychogeriatr 2012 Dec 14;25(4):565-575. [doi: [10.1017/s1041610212002128](https://doi.org/10.1017/s1041610212002128)]

7. Abraha I, Rimland JM, Trotta FM, Dell'Aquila G, Cruz-Jentoft A, Petrovic M, et al. Systematic review of systematic reviews of non-pharmacological interventions to treat behavioural disturbances in older patients with dementia. The SENATOR-OnTop series. *BMJ Open* 2017 Mar 16;7(3):e012759 [FREE Full text] [doi: [10.1136/bmjopen-2016-012759](https://doi.org/10.1136/bmjopen-2016-012759)] [Medline: [28302633](https://pubmed.ncbi.nlm.nih.gov/28302633/)]
8. Kales HC, Gitlin LN, Lyketsos CG, Detroit Expert Panel on Assessment and Management of Neuropsychiatric Symptoms of Dementia. Management of neuropsychiatric symptoms of dementia in clinical settings: recommendations from a multidisciplinary expert panel. *J Am Geriatr Soc* 2014 Apr 17;62(4):762-769 [FREE Full text] [doi: [10.1111/jgs.12730](https://doi.org/10.1111/jgs.12730)] [Medline: [24635665](https://pubmed.ncbi.nlm.nih.gov/24635665/)]
9. Kales HC, Gitlin LN, Lyketsos CG. Assessment and management of behavioral and psychological symptoms of dementia. *BMJ* 2015 Mar 02;350(mar02 7):h369-h369 [FREE Full text] [doi: [10.1136/bmj.h369](https://doi.org/10.1136/bmj.h369)] [Medline: [25731881](https://pubmed.ncbi.nlm.nih.gov/25731881/)]
10. Fazio S, Pace D, Flinner J, Kallmyer B. The Fundamentals of Person-Centered Care for Individuals With Dementia. *Gerontologist* 2018 Jan 18;58(suppl_1):S10-S19. [doi: [10.1093/geront/gnx122](https://doi.org/10.1093/geront/gnx122)] [Medline: [29361064](https://pubmed.ncbi.nlm.nih.gov/29361064/)]
11. BPSD Program of Care. URL: <http://bpsdcare.com/> [accessed 2019-01-17]
12. Nakanishi M, Endo K, Hirooka K, Nakashima T, Morimoto Y, Granvik E, et al. Dementia behaviour management programme at home: impact of a palliative care approach on care managers and professional caregivers of home care services. *Aging Ment Health* 2018 Aug 29;22(8):1057-1062. [doi: [10.1080/13607863.2017.1332160](https://doi.org/10.1080/13607863.2017.1332160)] [Medline: [28553880](https://pubmed.ncbi.nlm.nih.gov/28553880/)]
13. The progression and stages of dementia. Alzheimer's Society. URL: <https://www.alzheimers.org.uk/about-dementia/symptoms-and-diagnosis/how-dementia-progresses/progression-alzheimers-disease-dementia> [accessed 2021-02-03]
14. Griffiths AW, Kelley R, Garrod L, Perfect D, Robinson O, Shoesmith E, et al. Barriers and facilitators to implementing dementia care mapping in care homes: results from the DCM™ EPIC trial process evaluation. *BMC Geriatr* 2019 Feb 08;19(1):37 [FREE Full text] [doi: [10.1186/s12877-019-1045-y](https://doi.org/10.1186/s12877-019-1045-y)] [Medline: [30736748](https://pubmed.ncbi.nlm.nih.gov/30736748/)]
15. Spanakis EG, Santana S, Tsiknakis M, Marias K, Sakkalis V, Teixeira A, et al. Technology-Based Innovations to Foster Personalized Healthy Lifestyles and Well-Being: A Targeted Review. *J Med Internet Res* 2016 Jun 24;18(6):e128 [FREE Full text] [doi: [10.2196/jmir.4863](https://doi.org/10.2196/jmir.4863)] [Medline: [27342137](https://pubmed.ncbi.nlm.nih.gov/27342137/)]
16. Sheth A, Jaimini U, Thirunarayan K, Banerjee T. Augmented Personalized Health: How Smart Data with IoTs and AI is about to Change Healthcare. In: RTSI. 2017 Sep Presented at: 2017 IEEE 3rd International Forum on Research and Technologies for Society and Industry - Innovation to Shape the Future for Society and Industry (RTSI); September 11-13, 2017; Modena, Italy. [doi: [10.1109/RTSI.2017.8065963](https://doi.org/10.1109/RTSI.2017.8065963)]
17. Honka A, Kaipainen K, Hietala H, Saranummi N. Rethinking Health: ICT-Enabled Services to Empower People to Manage Their Health. *IEEE Rev. Biomed. Eng* 2011;4:119-139. [doi: [10.1109/rbme.2011.2174217](https://doi.org/10.1109/rbme.2011.2174217)]
18. Dinesen B, Nonnecke B, Lindeman D, Toft E, Kidholm K, Jethwani K, et al. Personalized Telehealth in the Future: A Global Research Agenda. *J Med Internet Res* 2016 Mar 01;18(3):e53 [FREE Full text] [doi: [10.2196/jmir.5257](https://doi.org/10.2196/jmir.5257)] [Medline: [26932229](https://pubmed.ncbi.nlm.nih.gov/26932229/)]
19. Backonja U, Haynes SC, Kim KK. Data Visualizations to Support Health Practitioners' Provision of Personalized Care for Patients With Cancer and Multiple Chronic Conditions: User-Centered Design Study. *JMIR Hum Factors* 2018 Oct 16;5(4):e11826 [FREE Full text] [doi: [10.2196/11826](https://doi.org/10.2196/11826)] [Medline: [30327290](https://pubmed.ncbi.nlm.nih.gov/30327290/)]
20. Korhonen O, Väyrynen K, Krautwald T, Bilby G, Broers HAT, Giunti G, et al. Data-Driven Personalization of a Physiotherapy Care Pathway: Case Study of Posture Scanning. *JMIR Rehabil Assist Technol* 2020 Sep 15;7(2):e18508 [FREE Full text] [doi: [10.2196/18508](https://doi.org/10.2196/18508)] [Medline: [32930667](https://pubmed.ncbi.nlm.nih.gov/32930667/)]
21. Yeh RW, Kramer DB. Decision Tools to Improve Personalized Care in Cardiovascular Disease. *Circulation* 2017 Mar 21;135(12):1097-1100. [doi: [10.1161/circulationaha.116.024247](https://doi.org/10.1161/circulationaha.116.024247)]
22. Wang H, Zhao Y, Yu L, Liu J, Zwetsloot IM, Cabrera J, et al. A Personalized Health Monitoring System for Community-Dwelling Elderly People in Hong Kong: Design, Implementation, and Evaluation Study. *J Med Internet Res* 2020 Sep 30;22(9):e19223 [FREE Full text] [doi: [10.2196/19223](https://doi.org/10.2196/19223)] [Medline: [32996887](https://pubmed.ncbi.nlm.nih.gov/32996887/)]
23. Kocaballi AB, Berkovsky S, Quiroz JC, Laranjo L, Tong HL, Rezazadegan D, et al. The Personalization of Conversational Agents in Health Care: Systematic Review. *J Med Internet Res* 2019 Nov 07;21(11):e15360 [FREE Full text] [doi: [10.2196/15360](https://doi.org/10.2196/15360)] [Medline: [31697237](https://pubmed.ncbi.nlm.nih.gov/31697237/)]
24. Offermann-van Heek J, Ziefle M. They Don't Care About Us! Care Personnel's Perspectives on Ambient Assisted Living Technology Usage: Scenario-Based Survey Study. *JMIR Rehabil Assist Technol* 2018 Sep 24;5(2):e10424 [FREE Full text] [doi: [10.2196/10424](https://doi.org/10.2196/10424)] [Medline: [30249592](https://pubmed.ncbi.nlm.nih.gov/30249592/)]
25. Evans J, Brown M, Coughlan T, Lawson G, Craven MP. A Systematic Review of Dementia Focused Assistive Technology. In: Kurosu M, editor. *Human-Computer Interaction: Interaction Technologies*. HCI 2015. Lecture Notes in Computer Science, vol 9170. Cham: Springer; 2015:406-417.
26. van Kollenburg, J, Bogers SJA. Data-enabled design : a situated design approach that uses data as creative material when designing for intelligent ecosystems. Eindhoven, the Netherlands: Technische Universiteit Eindhoven; 2019. URL: https://research.tue.nl/files/118827915/20190228_Bogers_Kollenburg.pdf
27. Wang G, Gong H, Albayrak A, van der Cammen TJM, Kortuem G. Personalising Management of Behavioural and Psychological Symptoms of Dementia in Nursing Homes: Exploring the Synergy of Quantitative and Qualitative Data. *Biomed Res Int* 2020 Jul 10;2020:3920284-3920216 [FREE Full text] [doi: [10.1155/2020/3920284](https://doi.org/10.1155/2020/3920284)] [Medline: [32695812](https://pubmed.ncbi.nlm.nih.gov/32695812/)]

28. Grisot M, Moltubakk Kempton A, Hagen L, Aanestad M. Data-work for personalized care: Examining nurses' practices in remote monitoring of chronic patients. *Health Informatics J* 2019 Sep 18;25(3):608-616 [FREE Full text] [doi: [10.1177/1460458219833110](https://doi.org/10.1177/1460458219833110)] [Medline: [30880551](https://pubmed.ncbi.nlm.nih.gov/30880551/)]
29. Stokke R. From "Plug" to "play": Making established technology innovations work in caring services. Norway: Høgskolen i Innlandet; 2018.
30. Repairing Innovation: A Study of Integrating AI in Clinical Care. *Data & Society*. URL: <https://datasociety.net/library/repairing-innovation/> [accessed 2020-12-22]
31. Chang F, Eriksson A, Östlund B. Discrepancies between Expected and Actual Implementation: The Process Evaluation of PERS Integration in Nursing Homes. *Int J Environ Res Public Health* 2020 Jun 14;17(12):4245 [FREE Full text] [doi: [10.3390/ijerph17124245](https://doi.org/10.3390/ijerph17124245)] [Medline: [32545871](https://pubmed.ncbi.nlm.nih.gov/32545871/)]
32. Zorggroep Elde Maasduinen. URL: <https://zgem.nl/> [accessed 2021-05-16]
33. Qassem T, Moore P, Tadros G, Xhafa F. Emerging Technologies for Monitoring Behavioural and Psychological Symptoms of Dementia. 2014 Presented at: 9th International Conference on P2P, Parallel, Grid, Cloud and Internet Computing; November 8-10, 2014; Guangdong, China. [doi: [10.1109/3pgcic.2014.82](https://doi.org/10.1109/3pgcic.2014.82)]
34. Boddy CR. Sample size for qualitative research. *QMR* 2016 Sep 12;19(4):426-432. [doi: [10.1108/qmr-06-2016-0053](https://doi.org/10.1108/qmr-06-2016-0053)]
35. Geelen R. Agressie en ontregeld gedrag: het Crisis Ontwikkelings Model (COM). In: *Dementiezorg in de praktijk - deel 2*. Houten, the Netherlands: Bohn Stafleu van Loghum; 2015:167-184.
36. Polit DF, Beck CT. Generalization in quantitative and qualitative research: myths and strategies. *Int J Nurs Stud* 2010 Nov;47(11):1451-1458. [doi: [10.1016/j.ijnurstu.2010.06.004](https://doi.org/10.1016/j.ijnurstu.2010.06.004)] [Medline: [20598692](https://pubmed.ncbi.nlm.nih.gov/20598692/)]
37. Korstjens I, Moser A. Series: Practical guidance to qualitative research. Part 4: Trustworthiness and publishing. *Eur J Gen Pract* 2018 Dec 05;24(1):120-124. [doi: [10.1080/13814788.2017.1375092](https://doi.org/10.1080/13814788.2017.1375092)] [Medline: [29202616](https://pubmed.ncbi.nlm.nih.gov/29202616/)]
38. Smith JD. Single-case experimental designs: a systematic review of published research and current standards. *Psychol Methods* 2012 Dec;17(4):510-550 [FREE Full text] [doi: [10.1037/a0029312](https://doi.org/10.1037/a0029312)] [Medline: [22845874](https://pubmed.ncbi.nlm.nih.gov/22845874/)]
39. Kratochwill TR, Hitchcock JH, Horner RH, Levin JR, Odom SL, Rindskopf DM, et al. Single-Case Intervention Research Design Standards. *Remedial and Special Education* 2012 Aug 15;34(1):26-38. [doi: [10.1177/0741932512452794](https://doi.org/10.1177/0741932512452794)]
40. Ninci J, Vannest KJ, Willson V, Zhang N. Interrater agreement between visual analysts of single-case data: a meta-analysis. *Behav Modif* 2015 Jul 14;39(4):510-541. [doi: [10.1177/0145445515581327](https://doi.org/10.1177/0145445515581327)] [Medline: [25878161](https://pubmed.ncbi.nlm.nih.gov/25878161/)]
41. Braun V, Clarke V. Using thematic analysis in psychology. *Qualitative Research in Psychology* 2006 Jan;3(2):77-101. [doi: [10.1191/1478088706qp063oa](https://doi.org/10.1191/1478088706qp063oa)]
42. Thorpe J, Forchhammer BH, Maier AM. Adapting Mobile and Wearable Technology to Provide Support and Monitoring in Rehabilitation for Dementia: Feasibility Case Series. *JMIR Form Res* 2019 Oct 17;3(4):e12346 [FREE Full text] [doi: [10.2196/12346](https://doi.org/10.2196/12346)] [Medline: [31625951](https://pubmed.ncbi.nlm.nih.gov/31625951/)]
43. Wang G, Albayrak A, van der Cammen TJM. A systematic review of non-pharmacological interventions for BPSD in nursing home residents with dementia: from a perspective of ergonomics. *Int. Psychogeriatr* 2018 Oct 18;31(08):1137-1149. [doi: [10.1017/s1041610218001679](https://doi.org/10.1017/s1041610218001679)]
44. Bendjenna H, Charre P, Eddine Zarour N. Using multi - criteria analysis to prioritize stakeholders. *J of Systems and Info Tech* 2012 Aug 10;14(3):264-280. [doi: [10.1108/13287261211255365](https://doi.org/10.1108/13287261211255365)]
45. Lysaght T, Lim HY, Xafis V, Ngiam KY. AI-Assisted Decision-making in Healthcare: The Application of an Ethics Framework for Big Data in Health and Research. *Asian Bioeth Rev* 2019 Sep 12;11(3):299-314 [FREE Full text] [doi: [10.1007/s41649-019-00096-0](https://doi.org/10.1007/s41649-019-00096-0)] [Medline: [33717318](https://pubmed.ncbi.nlm.nih.gov/33717318/)]
46. van Kollenburg J, Bogers S, Rutjes H, Deckers E, Frens J, Hummels C. Exploring the Value of Parent Tracked Baby Data in Interactions with Healthcare Professionals: A Data-Enabled Design Exploration. In: *CHI '18: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 2018 Presented at: 2018 CHI Conference on Human Factors in Computing Systems; April 21-26, 2018; Montréal, Canada. [doi: [10.1145/3173574.3173871](https://doi.org/10.1145/3173574.3173871)]
47. Wang G, Marradi C, Albayrak A, van der Cammen TJM. Co-designing with people with dementia: A scoping review of involving people with dementia in design research. *Maturitas* 2019 Sep;127:55-63 [FREE Full text] [doi: [10.1016/j.maturitas.2019.06.003](https://doi.org/10.1016/j.maturitas.2019.06.003)] [Medline: [31351521](https://pubmed.ncbi.nlm.nih.gov/31351521/)]
48. Bogers S, Frens J, van Kollenburg J, Deckers E, Hummels C. Connected Baby Bottle: A Design Case Study Towards a Framework for Data-Enabled Design. In: *DIS '16: Proceedings of the 2016 ACM Conference on Designing Interactive Systems*. 2016 Presented at: 2016 ACM Conference on Designing Interactive Systems; June 4-8, 2016; Brisbane, Australia. [doi: [10.1145/2901790.2901855](https://doi.org/10.1145/2901790.2901855)]
49. Data-enabled Design. Technische Universiteit Eindhoven. URL: <https://www.tue.nl/en/our-university/departments/industrial-design/innoveren/projects/systemic-change/data-enabled-design/> [accessed 2021-02-04]
50. Peine A, Neven L. From Intervention to Co-constitution: New Directions in Theorizing about Aging and Technology. *Gerontologist* 2019 Jan 09;59(1):15-21. [doi: [10.1093/geront/gny050](https://doi.org/10.1093/geront/gny050)] [Medline: [29850812](https://pubmed.ncbi.nlm.nih.gov/29850812/)]
51. Meiland FJM, Hattink BJJ, Overmars-Marx T, de Boer ME, Jedlitschka A, Ebben PWG, et al. Participation of end users in the design of assistive technology for people with mild to severe cognitive problems; the European Rosetta project. *Int. Psychogeriatr* 2014 Feb 10;26(5):769-779. [doi: [10.1017/s1041610214000088](https://doi.org/10.1017/s1041610214000088)]

52. Aloulou H, Mokhtari M, Tiberghien T, Biswas J, Phua C, Kenneth Lin JH, et al. Deployment of assistive living technology in a nursing home environment: methods and lessons learned. *BMC Med Inform Decis Mak* 2013 Apr 08;13(1):42 [FREE Full text] [doi: [10.1186/1472-6947-13-42](https://doi.org/10.1186/1472-6947-13-42)] [Medline: [23565984](https://pubmed.ncbi.nlm.nih.gov/23565984/)]
53. Wangmo T, Lipps M, Kressig RW, Ienca M. Ethical concerns with the use of intelligent assistive technology: findings from a qualitative study with professional stakeholders. *BMC Med Ethics* 2019 Dec 19;20(1):98 [FREE Full text] [doi: [10.1186/s12910-019-0437-z](https://doi.org/10.1186/s12910-019-0437-z)] [Medline: [31856798](https://pubmed.ncbi.nlm.nih.gov/31856798/)]

Abbreviations

BPSD: Behavioral and Psychological Symptoms of Dementia

IPS: Indoor Positioning System

Edited by G Eysenbach; submitted 12.11.20; peer-reviewed by B Östlund, H Mehdizadeh, K Yakubu, Y Tani; comments to author 18.01.21; revised version received 10.02.21; accepted 04.04.21; published 28.05.21.

Please cite as:

Wang G, Albayrak A, Kortuem G, van der Cammen TJM

A Digital Platform for Facilitating Personalized Dementia Care in Nursing Homes: Formative Evaluation Study

JMIR Form Res 2021;5(5):e25705

URL: <https://formative.jmir.org/2021/5/e25705>

doi: [10.2196/25705](https://doi.org/10.2196/25705)

PMID: [34047703](https://pubmed.ncbi.nlm.nih.gov/34047703/)

©Gubing Wang, Armagan Albayrak, Gerd Kortuem, Tischa JM van der Cammen. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 28.05.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

A Clinician-Controlled Just-in-time Adaptive Intervention System (CBT+) Designed to Promote Acquisition and Utilization of Cognitive Behavioral Therapy Skills in Bulimia Nervosa: Development and Preliminary Evaluation Study

Adrienne Juarascio¹, PhD; Paakhi Srivastava¹, PhD; Emily Presseller¹, BA; Kelsey Clark¹, MA; Stephanie Manasse¹, PhD; Evan Forman¹, PhD

Center for Weight, Eating and Lifestyle Science, Drexel University, Philadelphia, PA, United States

Corresponding Author:

Paakhi Srivastava, PhD
Center for Weight, Eating and Lifestyle Science
Drexel University
3141 Chestnut Street, Stratton Hall, Suite 232
Philadelphia, PA, 19104
United States
Phone: 1 2678156511
Email: ps887@drexel.edu

Related Article:

This is a corrected version. See correction statement: <https://formative.jmir.org/2021/7/e31964>

Abstract

Background: Cognitive behavioral therapy (CBT) for bulimia nervosa (BN) is most effective when patients demonstrate adequate skill utilization (ie, the frequency with which a patient practices or uses therapeutic skills) and skill acquisition (ie, the ability to successfully perform a skill learned in treatment). However, rates of utilization and acquisition of key treatment skills (eg, regular eating, urge management skills, and mood management skills) by the end of the treatment are frequently low; as a result, outcomes from CBT for BN are affected. Just-in-time adaptive interventions (JITAI) may improve skill acquisition and utilization by delivering real-time interventions during algorithm-identified opportunities for skill practice.

Objective: In this manuscript, we describe a newly developed JITAI system called CBT+ that is designed to facilitate the acquisition and utilization of CBT for BN treatment skills when used as a treatment augmentation. We also present feasibility, acceptability, and preliminary outcomes data from a small proof-of-concept pilot trial (n=5 patients and n=3 clinicians) designed to identify opportunities for iterative development of CBT+ ahead of a larger ongoing randomized controlled trial.

Methods: A total of 5 individuals with BN received 16 sessions of outpatient CBT for BN while using the CBT+ app. Data were collected from patients and clinicians to evaluate the feasibility (eg, app use and user adherence), acceptability (eg, qualitative patient and clinician feedback, including usefulness ratings of CBT+ on a 6-point Likert scale ranging from 1=extremely useless to 6=extremely useful), and preliminary outcomes (eg, improvements in skill utilization and acquisition and BN symptoms) of the CBT+ system.

Results: Patients reported that CBT+ was a relatively low burden (eg, quick and easy-to-use self-monitoring interface), and adherence to in-app self-monitoring was high (mean entries per day 3.13, SD 1.03). JITAI was perceived as useful by both patients (median rating 5/6) and clinicians (median rating 5/6) for encouraging the use of CBT skills. Large improvements in CBT skills and clinically significant reductions in BN symptoms were observed post treatment. Although preliminary findings indicated that the CBT+ system was acceptable to most patients and clinicians, the overall study dropout was relatively high (ie, 2/5, 40% patients), which could indicate some moderate concerns regarding feasibility.

Conclusions: CBT+, the first-ever JITAI system designed to facilitate the acquisition and utilization of CBT for BN treatment skills when used as a treatment augmentation, was shown to be feasible and acceptable. The results indicate that the CBT+ system

should be subjected to more rigorous evaluations with larger samples and should be considered for wider implementation if found effective. Areas for iterative improvement of the CBT+ system ahead of a randomized controlled trial are also discussed.

(*JMIR Form Res* 2021;5(5):e18261) doi:[10.2196/18261](https://doi.org/10.2196/18261)

KEYWORDS

eating disorders; telemedicine; mobile phone; smartphone; technology; cognitive behavioral therapy

Introduction

Background

Bulimia nervosa (BN) is characterized by recurrent episodes of binge eating (ie, eating a large amount of food within a discrete period accompanied by a sense of loss of control over eating) and compensatory behaviors such as purging (eg, self-induced vomiting and misuse of a laxative or diuretic), fasting, or driven exercise [1]. The leading evidence-based treatment for BN is cognitive behavioral therapy (CBT), a present-focused, active, skill-oriented treatment. Although CBT can be an effective treatment for BN, a recent meta-analysis found that nearly 70% of patients remain at least partially symptomatic at the end of the treatment [2].

One reason many patients may fail to achieve remission from CBT for BN is suboptimal levels of therapeutic skill acquisition (ie, the ability to successfully perform a skill learned in treatment) and skill utilization (ie, the frequency with which a patient practices or uses therapeutic skills) [3,4]. Acquisition and utilization of therapeutic skills designed to reduce dietary restraint (eg, regular eating) consistently predict treatment outcomes [4-7]. In addition, numerous studies have demonstrated that reduction of dietary restraint is a key mechanism of action in CBT for BN [2,8-11]. Although less well-studied, failure to respond adaptively to cues for binge eating (particularly a failure to regulate negative affect) is also strongly associated with the maintenance of BN symptoms [12,13]. A growing body of literature has identified that the use of therapeutic skills such as mood management techniques and the resulting improvements in the ability to manage negative affect during CBT for BN are associated with symptom improvement [12,14]. Existing data thus suggest that improvements in the acquisition and utilization of skills related to reducing dietary restraint and increasing adaptive responses to cues could strongly improve treatment outcomes for BN.

Mobile health (mHealth) technologies, specifically just-in-time adaptive interventions (JITAI), are a promising intervention paradigm that may be able to increase skill acquisition and utilization when used in conjunction with in-person treatment. JITAI monitor the temporal dynamics of an individual's state in real time and quickly adapt to the individual's current contextual state to provide individually tailored interventions at crucial times of need [15]. JITAI are increasingly being developed and used to augment treatment for several health care concerns [16-21] and generally show promise in the treatment of numerous physical [22] and mental health concerns [23]. Despite early successes in other fields, JITAI are still relatively rare, and few have been tested as an augmentation to a comprehensive therapy package. In addition, although JITAI are often described as providing tailored interventions, the

tailoring of intervention content is typically limited to a series of prewritten interventions that are then matched to the user's current contextual state (eg, providing a prewritten reminder to get up and walk when an app has identified that the user has not walked in a certain period). The ability of JITAI to (1) augment a comprehensive therapy package such as CBT for BN and (2) allow for more sophisticated tailoring of abilities by enabling a treating clinician to flexibly adjust the timing, content, and goals of a JITAI to match an individual patient remains largely untested.

This Study

As part of an ongoing National Institute of Mental Health (NIMH)-funded clinical trial (R34MH116021), our team recently developed CBT+, a JITAI system comprising a patient smartphone app and clinician portal designed to augment traditional in-person CBT for BN. In the development and design of the CBT+ system section below, we describe the core features of the CBT+ app with a particular focus on how we designed CBT+ to work as an augmentation to an ongoing treatment program by allowing the treating clinician to flexibly adjust the timing, content, and goals of CBT+ to match the changing needs of an individual patient. We also present data on feasibility (eg, app use and user adherence), acceptability (eg, qualitative patient and clinician feedback), and preliminary clinical outcomes from our initial wave of pilot patients (n=5) and therapists (n=3).

Development and Design of the CBT+ System

Identifying the Optimal Ways to Extend CBT for BN

As described above, CBT for BN strives to help patients acquire and use skills designed to (1) reduce dietary restraint and (2) increase adaptive responses to cues [3]. To reduce dietary restraint, CBT for BN encourages the development of 3 core skills: (1) scheduling eating episodes at regular intervals throughout the day, with the goal of eating 3 meals and 1-2 snacks per day; (2) eating a sufficient number of calories at each meal or snack to prevent acute hunger; and (3) eating a sufficient range of food, including foods the patient may fear eating (eg, desserts and carbohydrates), to reduce or prevent feelings of deprivation. The ultimate goal of these 3 skills is to reduce several types of dietary restraint (eg, caloric restriction, hedonic restriction, and adherence to rigid food rules) that increase vulnerability to binge-eating episodes. To increase adaptive responses to cues, CBT for BN also teaches 3 core skills: (1) how to recognize internal (eg, negative affect and urges) and external triggers (eg, presence of palatable foods) for BN symptoms, (2) how to modulate mood when experiencing negative emotions (without resorting to binge eating), and (3) how to manage urges to binge or use a compensatory behavior. These 3 skills are designed to help patients identify internal and

external triggers for binge eating before these triggers lead to a binge-eating episode and successfully use therapeutic skills to manage negative affect and urges [3].

We developed the CBT+ JITAI system as a CBT for BN treatment augmentation with the goal of increasing the utilization and acquisition of the 6 core skills described above. Accordingly, the full CBT+ JITAI system consists of 2 highly integrated subsystems: (1) the CBT+ patient smartphone app (CBT+ app) and (2) a CBT+ website that is accessible by the CBT therapist (CBT+ clinician portal). Below, we describe both the subsystems in more detail.

CBT+ App

CBT+ Electronic Self-Monitoring Forms

The CBT+ app (Figure 1) was designed to replace the traditional paper-based monitoring forms used in CBT for BN with

easy-to-use electronic self-monitoring forms (Figure 2). Individuals were instructed to record all of their eating episodes in real time to capture information about their eating pattern, including the type of meal (eg, breakfast, midmorning snack, and lunch), time of the meal and snack, and food consumed. At each meal entry, patients also reported the presence or absence of any disordered eating behaviors or urges and rated their current mood. Individuals were also instructed to complete an *other* entry any time they engaged in disordered eating behavior outside of meal entries (eg, laxative use and excessive exercise), experienced a strong urge to engage in a disordered eating behavior, or experienced a significant change in mood. An open text box is available at all entries for patients to add additional context and comments if desired. The data entered into the electronic monitoring form provide the momentary data needed to power the clinician-controlled JITAI algorithm (described in greater detail in *JITAI Algorithm* section below).

Figure 1. Screenshots of the home screen showing all the tabs CBT+ users can access.

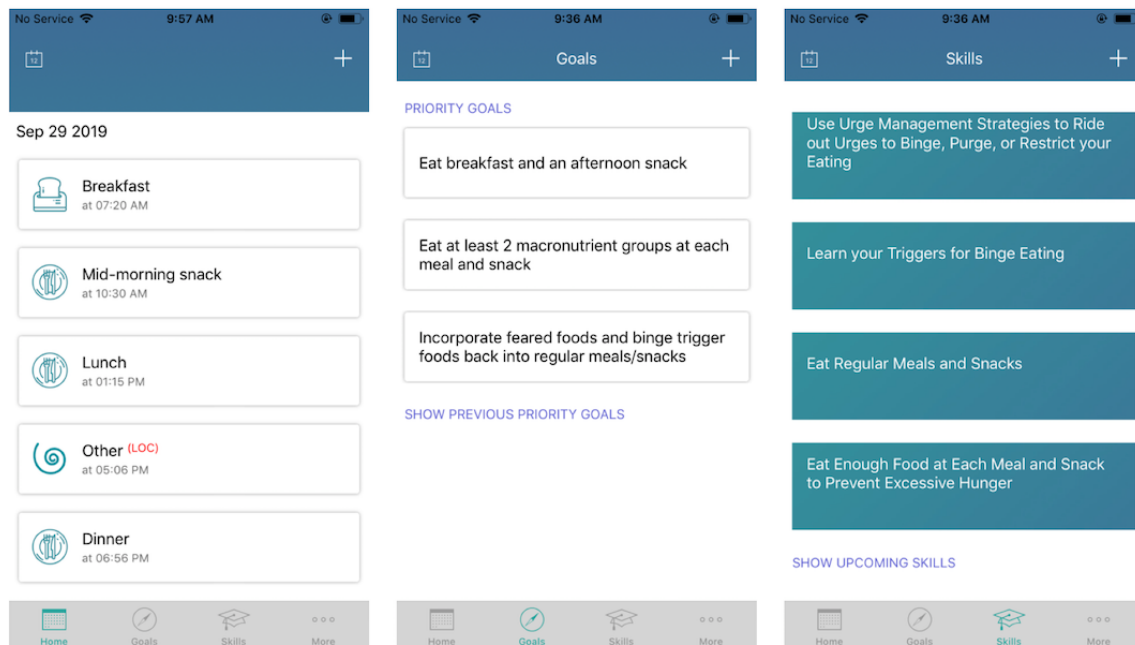


Figure 2. Screenshots of the electronic self-monitoring form of the CBT+ smartphone app.

Goals

The goals feature allows patients to view up to 3 weekly goals (called priority goals in the app) that they collaboratively set with their clinician during treatment sessions. Priority goals are a method by which the clinician can control the JITAI algorithm, as individuals will only receive momentary interventions that have been linked with a priority goal. Priority goals are typically (though not always) linked to 1 of the 6 CBT skills described above, but clinicians can scale up or scale down the goal to meet the needs of an individual patient. For example, although regular eating skills recommend patients eat 3 meals and 1-2 snacks per day, the priority goals for regular eating can be more modest at first (eg, eating 2 meals and 1 snack) for patients who need to slowly build up to 3 meals and 1-2 snacks per day over several weeks. Weekly goals can be viewed on the goals tab of the CBT+ app. Before attending each therapy session, patients were instructed to rate how successful they were in meeting their goals each week, which provides information to the treating clinician that can be used to inform goal progression.

JITAI Algorithm

Each time an individual enters data, an embedded algorithm determines whether there is an opportunity for the individual to practice a therapeutic skill. The starting algorithm is based on CBT theory [24] and triggers interventions when self-monitoring logs suggest that CBT skills practice would be beneficial. For example, if one of an individual's priority goals is to *practice urge management skills whenever you experience an urge to binge*, the CBT+ algorithm will check self-monitoring records at each entry and suggest an urge management skill to try whenever an urge is reported. As described above, the

starting algorithm can be flexibly adjusted by the clinician by altering the priority goals set for that week. Clinicians can also control how often the algorithm intervenes directly by scaling up or down the frequency of interventions. For example, the starting algorithm prompts users to regularly eat if 5 or more waking hours have passed without an eating episode. However, if a clinician determines that a specific patient would benefit from a more frequent reminder, they can adjust the algorithm timing. By allowing clinicians to flexibly adjust the embedded algorithm, CBT+ will only intervene when opportunities are identified to practice the specific skills that the clinician and patient are working on together in the therapy sessions. For example, in the earlier scenario where an urge management skill was recommended, the patient may have also reported feeling sad. However, because the clinician and patient have not yet begun working on mood management skills and no priority goal was set in this domain, the algorithm will not intervene on this practice opportunity.

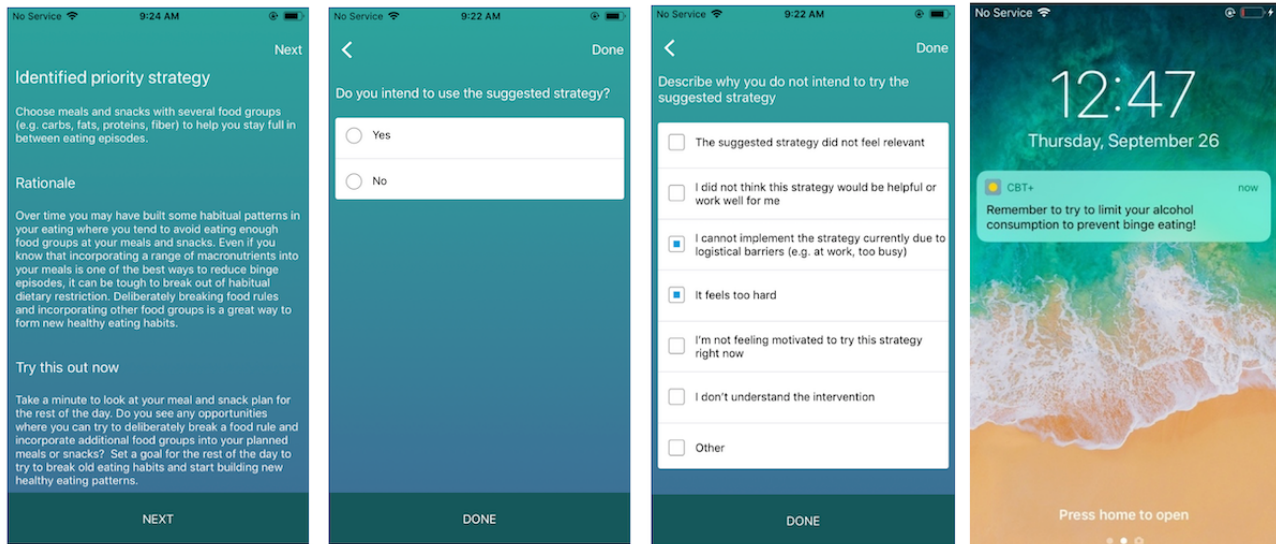
JITAI Interventions

All CBT+ interventions follow a specific structure that is designed to facilitate awareness of when, why, and how to use a specific skill at the moment that the skill is suggested. Each intervention has 3 key components: (1) the skill identified as most important to practice based on the data entered by the patient; (2) a brief rationale explaining why the patient would benefit from practicing that skill in the current moment; and (3) instructions to *try this out now*, which provide guidance on how to implement the suggested skill at the current time. Additional information about how clinicians can control the specific content of the interventions will be discussed in the *Clinician Portal* section below. Every time an intervention is delivered, patients

are asked if they intend to use the recommended skill. If patients report that they do not intend to try out the recommended skill, they can select a reason (eg, *the suggested strategy did not feel relevant and it feels too hard*) that will be shared with their treating clinician and can be used to refine future intervention content. In addition to the algorithm-generated interventions, patients can also receive push notification interventions at

scheduled times preset by their treating clinician. Push notification interventions will be discussed further in the *Clinician Portal* section below. A standard intervention example and a push notification example are shown in Figure 3, and 2 examples of patient self-monitoring log entries and a corresponding intervention can be found in Textbox 1.

Figure 3. Screenshots of the delivered just-in-time adaptive interventions and follow-up questions.



Textbox 1. Examples of events that would trigger just-in-time adaptive interventions (JITAs) and examples of corresponding JITAs.

- Patient X had a priority goal of *eat regular meals and snacks* on her self-monitoring form. She noted that it had been more than 5 hours since her last meal or snack. The algorithm identified this as an opportunity to reinforce using regular eating skills and delivered a just-in-time adaptive intervention (JITAI; for example, to address logistics or problem-solving issues using regular eating skills).

On the basis of your most recent data, CBT+ has identified a strategy you can try out now to build your skills:

 - Identified priority strategy: Stick to your planned meal and snack schedule, even if you are worried it is too much food or that you may gain weight.
 - Rationale: Getting back on track with your regular eating plan after you missed a meal or snack because of fears of weight gain often requires some problem solving.
 - Try this out now: Take a minute to think through how you will meet your regular eating goals for the rest of the day. Are there any logistical barriers you may experience? If so, try to identify some options for how you may be able to work around some of these barriers so that you can meet your goals.
- Patient Y had a priority goal of *learning to manage negative emotions skill*, and on her self-monitoring form, she gave a rating of 4 on the mood rating question (mood range 1 [good]-5 [bad]) and also checked off *no* box in response to the question “have you used a mood management strategy?” The CBT+ algorithm identified this as an opportunity to practice a mood management strategy and delivered a JITAI (eg, encouragement for using alternative activities strategies for mood management).

On the basis of your most recent data, CBT+ has identified a strategy you can try out now to build your skills:

 - Identified priority strategy: Engage in alternative activities to improve your mood.
 - Rationale: Breaking out of certain problematic habits (eg, engaging in an alternative activity instead of binge eating when feeling sad) requires tolerating uncomfortable thoughts, urges, and/or cravings. Learning to tolerate distress is key to long-term success.
 - Try this out now: Many people find reflecting on their long-term goals to be helpful in tolerating distress at the moment. How will engaging in an alternative activity right now make your life better in the long term, even if it feels uncomfortable in the short term? Remember that thoughts, feelings, urges, and cravings cannot physically hurt you, and engaging in an alternative activity will take you closer to reaching your long-term goals. Practice tolerating distress and engaging in an alternative activity from our list right now.

On-Demand CBT Skills

The CBT+ app also includes a repository of the skills that individuals have learned in therapy to date. Individuals can

always access a brief description of what each skill is and how to use that skill.

CBT+ Clinician Portal

Patient Dashboard

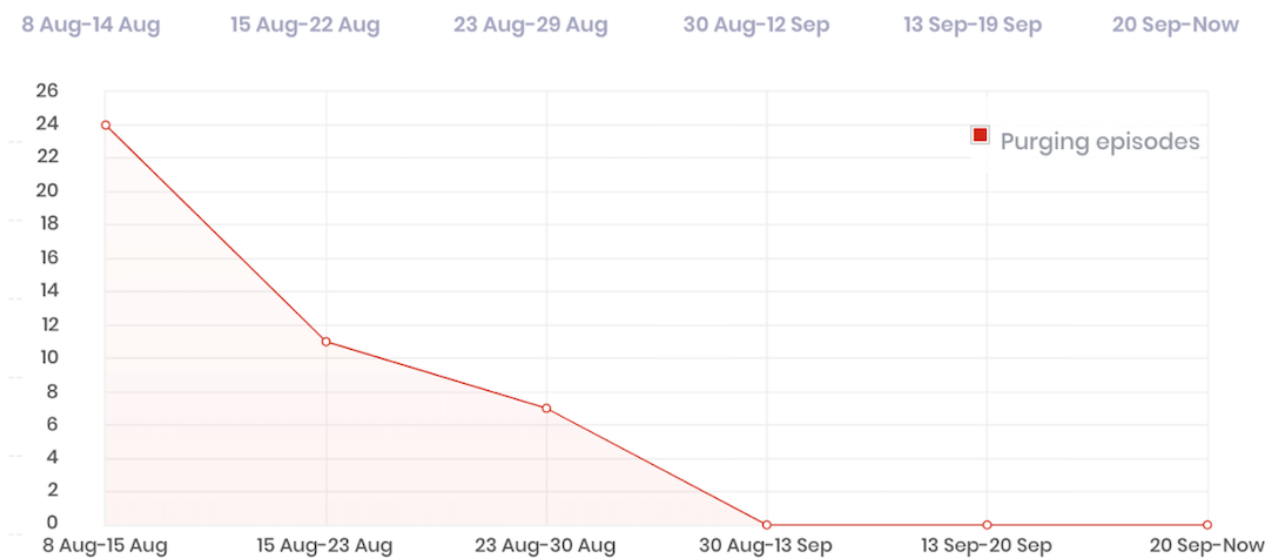
The patient dashboard quickly summarizes key data from the patients last week (eg, number of binge episodes and number

of self-monitoring entries). Clinicians can also view trends in their patients' disordered eating behaviors over time through auto-generated graphs by clicking on any dashboard summary. A screenshot of the patient dashboard and a trend graph are shown in [Figure 4](#).

Figure 4. Patient dashboard of the CBT+ clinicians' portal.

Summary Calendar Visits Skills Records Shown intervention popups Push notifications

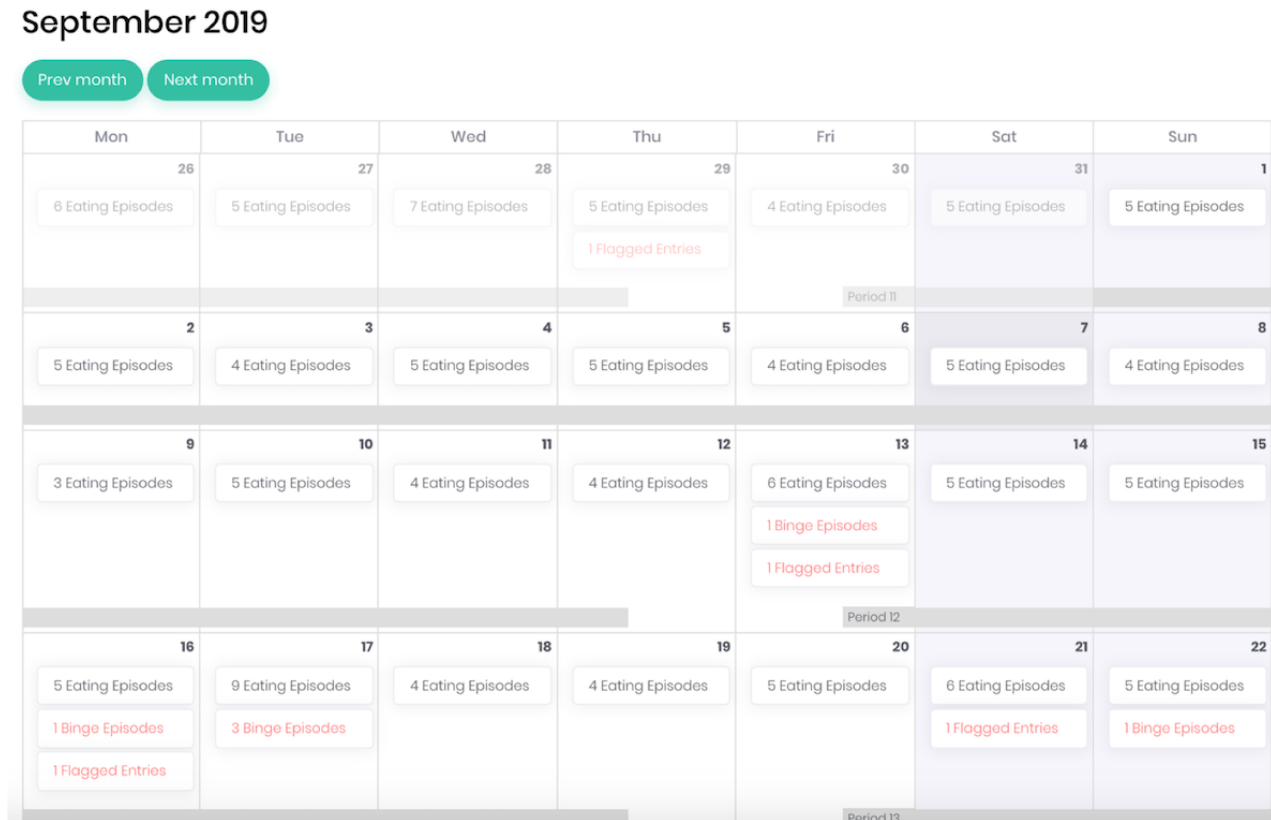
Progression of Treatment							
	8 Aug-14 Aug	15 Aug-22 Aug	23 Aug-29 Aug	30 Aug-12 Sep	13 Sep-19 Sep	20 Sep-Now	
Loss of control eating	1	3	0	0	5	1	
Purging episodes	24	11	7	0	0	0	
Laxatives	0	0	0	0	0	0	
4+ hours without eating	6	1	3	7	3	2	
Eating / day AVG	2.9	5.8	5.6	4.4	5.4	4.4	
Data entries / day AVG	5.9	5.9	5.6	4.4	5.4	4.6	
Days with 0-2 data entries	1	0	0	0	0	1	



Self-Monitoring Records

Clinicians can use a calendar feature in which daily self-monitoring records can be viewed. All eating disorder behaviors endorsed by the patient are flagged in red to allow

clinicians to quickly view the data from the past week and identify which days are most important to discuss during the therapy sessions. [Figure 5](#) shows an example of the calendar feature.

Figure 5. Calendar feature of the CBT+ clinicians' portal.

Skills Tab

Clinicians can use the skills tab to open new skills once they have been introduced into therapy. Once a skill has been introduced during a session, the skill becomes viewable in the CBT+ app skill section, and priority goals that intervene in this skill can be assigned.

Interventions Tab

In the interventions tab, the clinician can view which interventions their patients received in the past week. Clinicians can also view if their patients intend to implement the suggested strategy and the reasons for not implementing a recommended strategy. This tab can be used to understand which intervention the patient found helpful or provide guidance for adjusting intervention content over time to address the reported barriers to skill use.

Push Notification Intervention Tab

The push notifications tab on the patient dashboard shows a list of all push notifications that have been sent or are scheduled to be sent to the patient. There are 2 types of push notifications: self-monitoring push notifications that prompt patients to enter data if they have not done so for a set number of waking hours (default is set at 5 hours, but this is adjustable by clinicians) and custom push notifications. Custom push notifications allow clinicians to set a one-time or recurring reminder to practice a therapeutic skill. These notifications allow clinicians to preset JITAIs that are not dependent on self-monitoring log completion. Custom push notifications can be used to remind patients ahead of sessions to try out a specific homework at agreed-upon times, encourage skill use during known high-risk times (eg, right

before a scheduled weekly happy hour and on weekday evenings where unscheduled time can be a common trigger for binge-eating episodes), or send motivational messages.

Visits Tab

Clinicians use the visits tab to see the summary of goals that have been assigned to the patient at each session and record session notes. During each in-person therapy session, clinicians add a new visit and set priority goals for the period between the current visit and the next therapy session. As described above, priority goals are typically linked to 1 of 6 CBT skills. Linking to 1 of the 6 skills allows clinicians to view a list of prewritten interventions that can be delivered during algorithm-identified moments when that skill should be practiced. The prewritten interventions contain suggestions for specific CBT strategies that may facilitate skill use, and they are designed to address common barriers to skill implementation (ie, low accountability, poor awareness of when and how to use skills, habitual use of maladaptive coping skills, distress intolerance, poor problem-solving abilities, low motivation, and poor memory). Clinicians can choose any or all of the prewritten interventions that may be relevant for their patient, or if none of the prewritten options are relevant, they can create their own custom intervention. After assigning weekly goals and linking goals to interventions, clinicians save and close the visit, which triggers an automatic update to the individual patient's CBT+ algorithm.

Methods

Study Design

We conducted a small pilot trial (n=5 patients and n=3 clinicians) to assess the feasibility and acceptability of augmenting CBT for BN with the CBT+ JITAI system and identify opportunities for improvement to CBT+ ahead of a larger ongoing randomized controlled trial (RCT). In this study, we present data obtained from clinicians and patients that were used to inform iterative development.

Participants

Patients were eligible if they were aged between 18 and 70 years and met the criteria for a primary diagnosis of BN. Exclusion criteria included active severe psychiatric comorbidity that would limit the ability to participate in an outpatient clinical trial for BN (eg, psychosis, acute suicidality, and severe substance use disorder), inability to speak and write English, diagnosis of an intellectual disability that would impair the use of the CBT+ app, or having received a full trial of CBT for BN in the past. Although the CBT+ app is currently only compatible with Apple iPhone iOS, participants were not excluded based on ownership of an iPhone; participants without iPhones were given loaner iPhones to use for the duration of the study. Patients' (n=5, 4 females and 1 male) average age was 35.6 (SD 6.8, range 25-42) years and average BMI was 31.2 (SD 4.43, range 25.8-37.1) kg/m². A total of 4 patients self-reported as Caucasian or White, 1 identified as multiracial, and 2 identified as Latino or Hispanic.

Clinicians were either masters (n=2) or doctoral-level clinicians (n=1) with a minimum of 2 years of experience in delivering CBT for BN. All 3 clinicians previously completed the Centre for Research on Eating Disorders at Oxford web-based training program developed by Fairburn et al [25] and received weekly supervision from a licensed psychologist during the course of the treatment trial.

Treatment

The treatment consisted of 16 sessions of CBT for eating disorders based on the treatment approach developed by Fairburn [24]. In addition to in-person treatment, all patients

were instructed to use the CBT+ app daily for the duration of the clinical trial.

Assessments

Full assessments occurred at baseline, midtreatment (week 8), and posttreatment (week 16). Assessments included the Eating Disorder Examination (EDE, version 17.0) [26]; a series of self-report measures, including the Quality of Life Inventory (QOLI) [27]; and behavioral tasks. Patients completed phone interviews after 4, 8, 12, and 16 weeks of treatment to provide specific feedback on the CBT+ app. Patients also completed weekly pre-session questionnaires in which they rated different components of the CBT+ app, reported on their skill use in the past week, and provided qualitative feedback about the app. To measure dietary restraint and emotion regulation, patients completed the EDE questionnaire restraint subscale items [28] and 4 individual items selected from the Difficulties in Emotion Regulation Scale [29] in the weekly pre-session questionnaires. Clinicians completed post-session questionnaires after every session and rated the usefulness of the CBT+ system. To quantify CBT+ app use, the number of days the app was used (ie, the number of days when at least one entry was completed) and the number of entries completed per day (ie, adherence) were calculated. For this aim, we only included days during which the patient was still enrolled in the trial to allow for the determination of app use during the period when the patient intended to use the CBT+ system. Thus, for the 2 patients who discontinued treatment, we only included use data up to the date that they discontinued participation in the study. Data from the app were used to examine skill use during each treatment period. To examine the use of the skill *eating enough at each meal and snack to prevent excessive hunger* throughout the treatment, we calculated the weekly percentage of regular eating episodes (ie, excluding those accompanied by loss-of-control over eating) where participant endorsed that they "ate enough food to prevent excessive hunger before [their] next planned eating episode" and/or that they "ate a range of macronutrients (eg, protein, fat, carbs)" (Figure 6). The percentage of records in which the participant endorsed using an urge management skill when relevant (ie, when the participant also experienced an urge to binge or use a compensatory behavior) was calculated for each week in treatment to identify changes in the use of urge management skills throughout treatment (Figure 7).

Figure 6. Change in frequency with which patients endorsed eating enough at regular eating episodes throughout treatment.

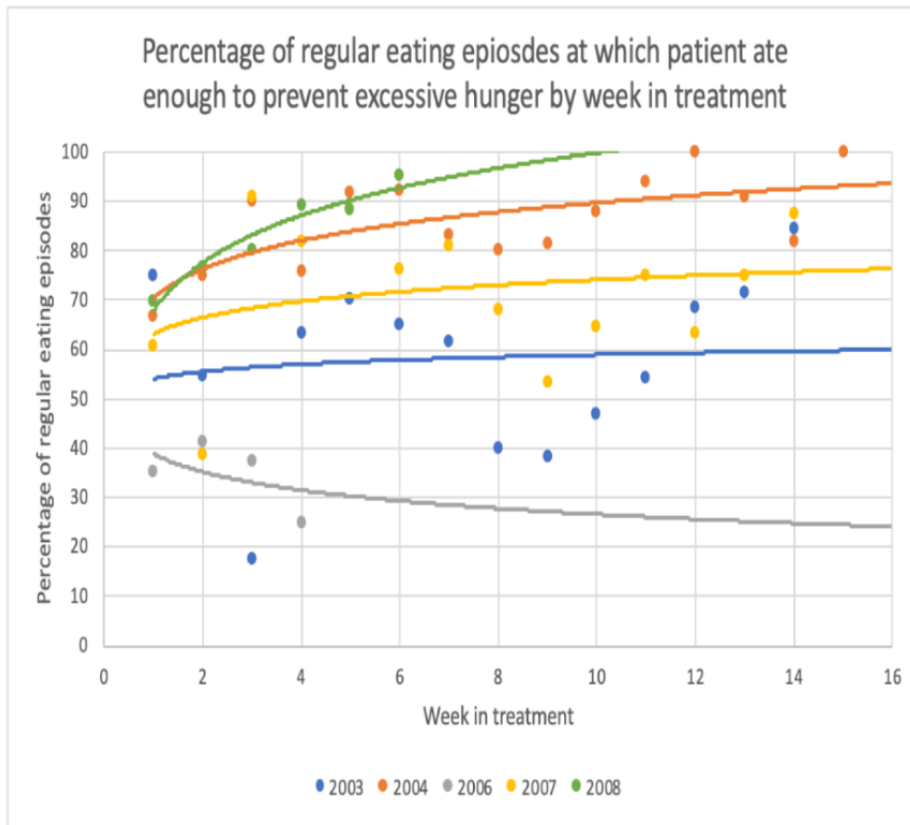
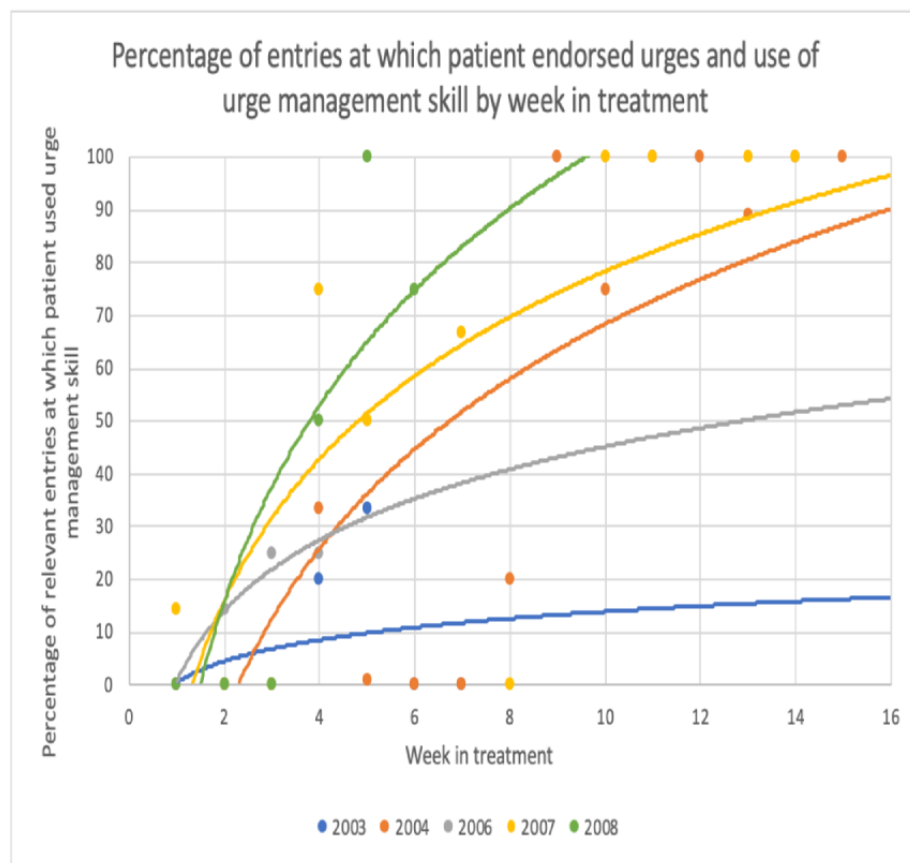


Figure 7. Change in frequency of reported urge management skill use throughout treatment.



Results

Study Feasibility

Of the 7 patients who were eligible to participate in the study after the baseline assessment, 5 chose to enroll in the study. Two patients who began treatment dropped out of the study: one after session 4 and the other after session 7. Both stated that their reason for dropping out was being too busy to continue treatment. The remaining patients (n=3) completed all 16 treatment sessions and all assessments.

Feasibility of the CBT+ System

In weekly presession questionnaires, patients reported a few minor technical issues with the CBT+ app (eg, occasional app crashing and occasional occurrences of an entry not saving), but no other major problems with the app were reported. Patients also noted occasional minor issues using the CBT+ system that were not because of technical issues with the system itself but still prevented app use (eg, forgetting login information and poor Wi-Fi connectivity). Technical issues that limited the use of the CBT+ app for at least one entry were reported in 5% (3/58) of weekly patient questionnaires. Clinicians reported only 1 bug with the clinician portal: eating episodes were occasionally mistakenly flagged as binge episodes on the calendar page. This bug was reported in 3% (2/58) of the weekly clinician questionnaires. Given that the technical issues with the CBT+ system were minor and did not interfere with treatment delivery, no live changes in the system were made during this study. As described above, CBT+ app use was evaluated during the periods when a patient was active in the study (eg, for the 2 users who discontinued in-person treatment, data on their app use were included until the time they discontinued study participation). Although we considered reporting adherence across 16 weeks for all participants, regardless of whether they completed treatment or not, we ultimately decided that understanding app adherence during active treatment would be more valuable for future iterations than reporting app adherence overall, which would mix the results of those who intentionally used the app with those who intentionally discontinued participation and thus limit interpretability. On average, patients used the app on 86.1% of

days (range 74.6%-92.9%; SD 7.5%) and made 3.13 entries per day (range 2.53-4.34; SD 0.74) while they were actively in treatment. User adherence declined somewhat throughout the course of the treatment but remained high among patients who continued their study participation. All 5 patients used the app on 100% of days and completed 3.59 entries per day (range 2.67-4.88; SD 0.82) during the first week of treatment. By week 8 (n=3), this number had declined to 89.4% of days (range 68%-100%; SD 18.4%) and 2.77 entries per day (range 2.20-3.88; SD 0.96). By the final week of app use, patients (n=3) used the app on 85.5% of days (range 78%-90%; SD 6.78%) and made an average of 3.19 entries per day (range 2.80-3.44; SD 0.34).

Acceptability of the Overall CBT+ System

Data from all patients were included in acceptability analyses and included the final acceptability interviews completed with the 2 patients who discontinued treatment. Patients' median rating of the overall usefulness of CBT+ was 5 (out of 6; IQR 1) on a 6-point Likert scale ranging from 1=extremely useless to 6=extremely useful. During qualitative feedback interviews, all patients reported that the app was easy to use and that they liked that the app kept them on track with their weekly goals and reinforced the use of strategies. [Textbox 2](#) shows additional qualitative feedback from patients. Key themes that emerged from qualitative interviews with patients included positive feedback regarding the ability of CBT+ to provide added accountability and awareness of the patient's own patterns, ability to track progress, and reinforcement of skills and strategies. Negative feedback primarily centered on the redundancy of interventions and concerns regarding the burden associated with completing daily records. Patients reported that self-monitoring via the CBT+ app could be burdensome, partially because tracking food intake is in itself burdensome, regardless of the format used, but also partially because of the number of questions asked in the CBT+ app at each entry. Patients also reported that they were less likely to thoroughly read the CBT+ intervention texts as treatment progressed because they had already read a similar version of the intervention before and found that they did not need to reread the intervention content.

Textbox 2. Qualitative feedback from patients on the CBT+ app.

Accountability and Identifying Patterns

- “Tracking my food intake allows me to check in with myself throughout the day.”
- “I like being prompted to reflect on how I’m feeling.”
- “It’s a good record to look back on and see what was happening if I fall off with my goals.”
- “It helps us to have more honest discussions and to validate what [I] think with actual data.”
- “It reminds you and when you reflect back...it helps to put things into perspective and see why things are happening.”
- “The fact that the app says, ‘Did this include two...or three [food] groups?’ helps you to pay attention to what you’re eating.”
- “It was super helpful. It made me think about [eating enough] more and made me be more cognizant of what I was putting in my meals.”
- “I could start to see patterns by recording everything.”
- “It helps you to understand why you’re eating so much in a certain moment. I always use the box at the end to make notes about whether something is happening and where I am. Writing it down and reflecting on it helps.”
- “It is helpful since I have to fill out the form and I have to reflect on everything that happened when I ate the meal. It definitely helps me realize why I ate what I did or why I ate as much as I did. It helps me realize why I was so hungry, if I didn’t eat for the last 5 hours.”

Tracking Progress

- “I like that it provides structure and feedback on my progress. It keeps me on track with my goals.”
- “[The app] helped by reminding me to record and being able to see my progress prompted me to maintain that progress.”

Reinforcement of Skills and Strategies

- “I like how the app reinforces the use of strategies.”
- “I think it helps to keep you aware and reflective about what you discuss in your therapy sessions. Like it keeps it in the back of your mind. What you’re eating and putting into the app is very purposeful and connected to what you discuss in therapy.”
- “I find [the identified priority strategies] somewhat helpful in changing my behavior.”
- “[The app] helps because it reminds you of [urge management strategies].”
- “I’m not the best at making entries in the moment, I think [the app] would be more useful if I actually did that. Otherwise, it’s less useful for me if I make an entry after. But overall, having it there and having to make the entry later certainly makes you more mindful [of strategies].”
- “I’d say the app and the therapy are just as helpful as each other with [incorporating feared foods]. Because again it keeps it in the back of my mind when I have to enter the stuff in.”

Redundancy and Burden

- “Over the last four weeks [of treatment] it was kind of redundant and burdensome because it was the same stuff that I had been doing...It just felt kind of taxing to do it.”
- “It was a little cumbersome to record everything I ate.”
- “It’s just really burdensome, the same questions over and over again. I almost wish it was all on one page, because I feel like [clicking] next, next, next just makes it feel longer.”
- “A lot of times the same strategies kept coming up so I started ignoring them after a while.”
- “I find [the identified priority strategies] helpful but also a little repetitive. But at the same time, you want to be reminding of the strategies.”
- “I liked that at the beginning. I feel like later on [in treatment], I wouldn’t read them because I had read them a bunch of times and I’d seen them and I knew what they were talking about.”
- “I definitely find [the reminder push notifications] helpful, but a lot of times the push notifications just reinforce to me how burdensome it is. It’s almost overwhelming. It’s just another thing on my to-do list.”

Clinicians’ median rating of the overall usefulness of CBT+ was 5 (out of 6; IQR 0), using the same Likert scale described above. Clinician interviews were overall positive, but a few suggestions for improvements to the clinician portal were noted, including changing the layout for viewing self-monitoring data and flagging compensatory behaviors on the calendar page in addition to binge episodes.

Acceptability of JITAI

The frequency of JITAIs was evaluated during the periods when a patient was active in the study (eg, for the 2 users who discontinued in-person treatment, data on JITAI frequency were included until the time they discontinued study participation). Patients received an average of 2.14 push notification reminders to enter data (range 1.76-2.69, SD 0.47) and 0.78 interventions (range 0.41-1.70, SD 0.53) per day. [Table 1](#) shows the average

percentage of total interventions associated with each skill. Significantly more interventions were received for the 3 skills related to reducing dietary restraint (1461/1590, 91.89% of all interventions) than for the 3 skills related to increasing adaptive response to cues (65/1590, 4.09% of all interventions). This imbalance was primarily driven by regular eating interventions (1288/190, 81.01% of all interventions), which may be partially due to the fact that it is (1) the first of the 6 skills introduced in

CBT for BN and (2) considered an essential skill to prioritize in CBT for BN; thus, many clinicians maintained regular eating as one of the key priority goals for the majority of treatment. A secondary factor that may explain the imbalance is that 2 of the 5 patients in the pilot trial dropped out of the treatment before they were introduced to several adaptive responses to cues skills, which were introduced in the latter half of the treatment.

Table 1. The average percentage of total interventions associated with each skill.

Cognitive behavioral therapy skills	Interventions (%)		
	Median	25th percentile	75th percentile
Eat regular meals and snacks	86.0	57.0	96.3
Eat enough food at each meal and snack	9.5	0.2	24.4
Incorporate feared foods and binge-trigger foods	0.0	0.0	1.4
Use urge management strategies to cope with urges	3.3	0.9	5.9
Learn your triggers for binge eating	0.0	0.0	0.7
Learn to manage negative emotions	0.0	0.0	1.5

Clinicians created custom push notifications during 16% (9/56; range 0/4, 0%-5/6, 83%) of sessions. The most common type of custom push notification was a personalized reminder notification to complete self-monitoring entries as close to their eating episodes as possible or to make an *other* entry as soon as they engaged in any disordered eating behavior. The second and third most common custom push notifications were reminders to adhere to treatment recommendations (eg, “refrain from weighing yourself at home”) and interventions designed to address common treatment-interfering behaviors (eg, “try to avoid drinking too much alcohol on the weekend to prevent binge eating”), respectively.

Overall, patients reported that they intended to use the suggested strategies 97.1% (372/383) of the time when identified priority

strategies were delivered (range 84/88, 95.5%-13/13 and 91/91, 100%). During qualitative interviews, 100% (5/5) of patients reported that the JITAIs provided situationally relevant interventions, although 40% (2/5) of patients described the interventions as *repetitive* and reported that they frequently received interventions related to the same types of problematic behaviors.

Descriptive statistics were calculated for the patients’ and clinicians’ ratings of the app’s usefulness. Patients’ ratings of the usefulness of JITAIs for each of the 6 skills are shown in [Table 2](#). Patients’ median usefulness ratings for the 6 skills ranged from 4 to 5 on a 6-point Likert scale (ranging from 1=extremely useless to 6=extremely useful).

Table 2. Median patient ratings of the usefulness of CBT+ for the acquisition of each cognitive behavioral therapy skill.

Cognitive behavioral therapy skills	Participant rating		
	Median	25th percentile	75th percentile
Eat regular meals and snacks	5	5	5
Use urge management strategies to cope with urges	4	4	5
Eat enough food at each meal and snack	4	4	5
Learn your triggers for binge eating	4	4	5
Learn to manage negative emotions	4	4	4
Incorporate feared foods and binge-trigger foods	4	4	5

Preliminary Clinical Outcomes

To examine changes in CBT skills during the treatment period, we graphed data collected via the CBT+ app that reflects several CBT skills. Of note, not all skills were measured weekly using the app. For these skills, we used data from the weekly

self-report questionnaires described in the *Assessment* section. [Figures 8](#) and [9](#) show 2 examples from the weekly questionnaires. All skills examined via app or weekly questionnaires showed a visual pattern of improvement throughout the treatment.

Figure 8. Change in weekly Eating Disorder Examination Questionnaire restraint subscale score throughout treatment.

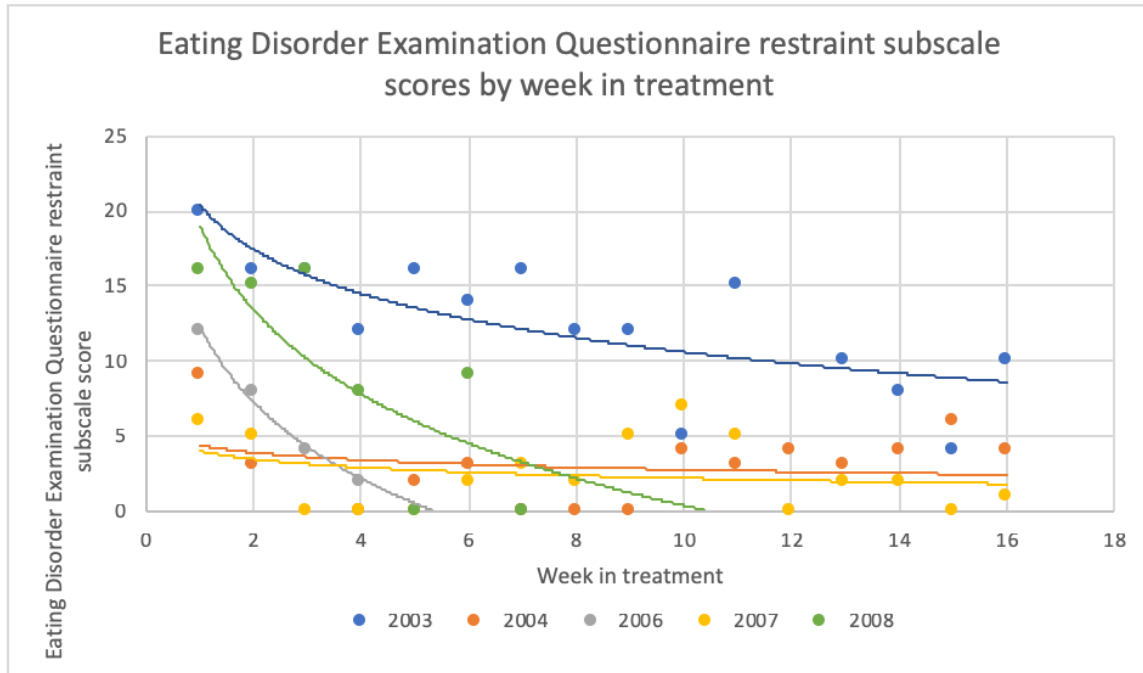
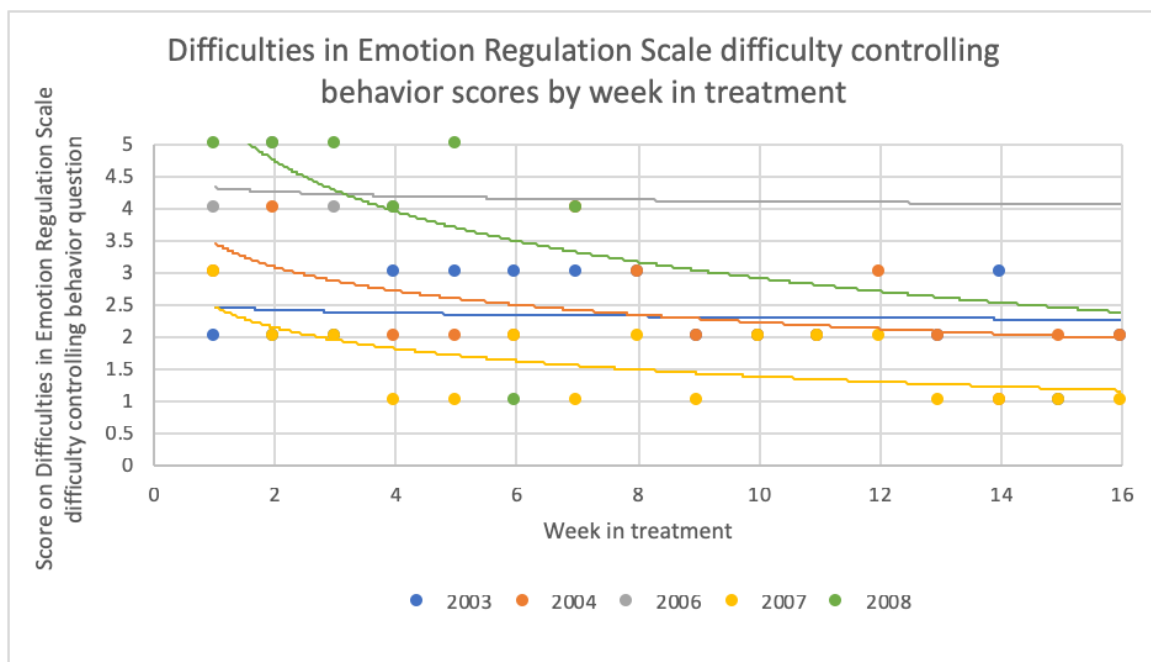


Figure 9. Change in score on the Difficulties in Emotion Regulation Scale item, “When I was upset, I had difficulty controlling my behaviors,” throughout treatment.



For clinical outcomes, we only included data from treatment completers, as both patients who dropped out discontinued their participation in the study before the midtreatment assessment point, and thus, we only had baseline EDE and QOLI scores. From pre- to posttreatment, patients who completed treatment demonstrated large decreases in the frequency of binge eating and compensatory behaviors and in EDE global scores. The detailed results are presented in Table 3. At posttreatment, the

mean EDE global score for treatment completers fell within 1 SD of community norms. Although paired samples *t* tests from pre- to posttreatment did not demonstrate statistically significant changes, the effect sizes were large for eating disorder symptoms. Patients generally demonstrated slight improvements in the QOLI total score from pre- to posttreatment with a small effect size (Table 3).

Table 3. Change in eating disorder symptoms and quality of life from pre- to posttreatment.

Eating disorder symptoms and quality of life	Pretreatment, median (range)	Posttreatment, median (range)	<i>t</i> test (<i>df</i>)	<i>P</i> value	Cohen <i>d</i>
Past month objective binge episodes	4 (2-14)	0 (0-1)	1.65 (2)	.24	0.95
Past month total loss of control eating episodes	8 (4-14)	0 (0-1)	2.62 (2)	.12	1.52
Past month total compensatory behaviors	10 (2-10)	1 (0-4)	1.47 (2)	.28	0.85
Global Eating Disorder Examination score	2.59 (2.18-3.88)	1.73 (1.64-1.86)	2.06 (2)	.18	1.18
Quality of Life Inventory total score	15.28 (14.78-17.28)	16.75 (14.22-18.00)	-0.62 (2)	.60	0.36

Discussion

As part of an ongoing NIMH-funded clinical trial (R34MH116021), our team developed CBT+, a JITAI system consisting of a patient smartphone app and clinician portal designed to augment traditional in-person CBT for BN. CBT+ is the first JITAI system designed as an augmentation to CBT for BN with the specific goal of promoting skills acquisition and utilization.

Principal Findings

Our preliminary evaluation demonstrates that it is broadly feasible and acceptable to integrate a JITAI system with CBT for BN, although some concerns were noted with 2 out of 5 patients discontinuing participation in the study. Among treatment completers, outcomes were positive in all 3 patients, showing substantial decreases in objectively large binge episodes and compensatory behaviors, and average posttreatment EDE global scores were within 1 SD of community norms, a metric commonly used to define remission [30,31].

We found that user adherence to app use and data entry on the CBT+ app were high throughout the study period for patients who remained in the study. Although there was a modest fall-off in adherence rates over time, patients used the app on 87.2% (470/539) of the days, even during the last weeks of prescribed app use in treatment, which is greater than that reported in other studies (eg, the overall percentage of the days when smartphone app was used for breast cancer=45% [32] and for depression=66% [33]) and clinical reviews of mHealth apps developed to augment treatments [34]. Devising a system that facilitates continued adherence during many months is critical for JITAIs, as JITAI algorithms rely on these data to inform the timing and content of intervention delivery. High user adherence may have been because of 3 key strengths in the design of the CBT+ app. First, the user interface of CBT+ was specifically designed to be low burden, that is, patients only needed to enter a few key variables at each data entry, which allowed us to keep entry time between 1 and 2 minutes per entry. Second, the self-monitoring data were directly relevant to the in-person treatment component, and therapists routinely referred to and made use of app-collected information during treatment, which allowed therapists to reinforce frequent use of the CBT+ app. Third, patients reported finding the JITAIs helpful, which may have increased the likelihood of completing data entry because of the perceived usefulness of the JITAI system.

Qualitative feedback and weekly questionnaire results suggested that both patients and clinicians believed that the CBT+ system was easy to use and helped them keep track of their goals. With regard to the acceptability of JITAIs, CBT+ algorithms produced slightly less than 1 intervention per day. Patients perceived the JITAIs to be accurate and appropriate to the entries they completed and found the JITAIs to be helpful in promoting the use and acquisition of therapeutic skills. JITAIs designed to reduce dietary restraint comprised the majority of interventions that were delivered. This finding suggests that most patients needed continued support to practice skills designed to reduce dietary restraint (such as eating regular meals and snacks and eating enough food at each meal and snack) compared with strategies designed to reduce unhelpful responses to cues to binge eating (such as learning to manage triggers for binge eating and learning to manage negative emotions). As the reduction of dietary restraint is the most well-established mechanism of action in CBT for BN [2,8], JITAIs designed to improve skills related to reducing dietary restraint may be the most impactful.

One reason that clinicians and patients may have found CBT+ beneficial was the high customizability and personalization allowed by the JITAI system. Previous research conducted by our team found that both clinicians and users value a system that allows personalized interventions [35]. The CBT+ clinician portal offered multiple options to clinicians to personalize the content, timing, and frequency of interventions to their patients' needs. For example, clinicians could craft personalized push notification interventions and used this feature in approximately 1 out of 5 sessions. These personalized push notifications were frequently used to target adherence to treatment recommendations (eg, reminder delivered every morning to remind participants to avoid at-home weighing) and to address treatment-interfering behaviors (eg, push notifications scheduled to remind participants every weekend to avoid high alcohol use to prevent the risk of binge eating).

Patients who completed treatment experienced improvements in CBT skills and significant reductions in binge eating and compensatory behaviors over the course of 16 weeks of treatment, and the average posttreatment EDE scores were in the remission range by the end of treatment. Although the additive efficacy of CBT+ cannot be directly assessed because of the lack of a control group, preliminary outcomes are promising and support the need for an RCT assessing whether CBT+ enhances outcomes when added to CBT for BN.

Limitations and Future Work

Several limitations of this study must be acknowledged. First, we conducted an open trial, which limits our ability to draw conclusions about the effectiveness of CBT+ in improving skills acquisition and utilization above and beyond CBT alone. Second, our study had a small number of patients, as the primary goal of this pilot study was to evaluate the feasibility and acceptability of CBT+ ahead of an ongoing RCT. Our team is currently conducting a randomized control trial comparing CBT for BN with electronic self-monitoring via the CBT+ app without the JITAI system to CBT for BN with electronic self-monitoring via the CBT+ app with the JITAI system (ie, full CBT+ system) to better understand the additive value of CBT+. Third, although user adherence in our study was better than that observed in other mHealth app studies, we observed a modest decrease in user adherence in the latter half of treatment. Fourth, the dropout rate was relatively high. Although this rate is comparable with that reported in other treatment studies for individuals with BN (eg, 15%-65% across a range of other treatment studies [36-40]), these results suggest that there may be some participants for whom the CBT+ system is not feasible or acceptable. Fifth, as is typically the case in eating disorder studies, most participants were female; thus, app adherence and outcomes may not generalize to men. Sixth,

similar to standard in-person CBT for BN [41], completing self-monitoring forms on the CBT+ app was perceived to be burdensome. Finally, although the interventions were generally found to be helpful, they were considered less useful later in treatment, likely because of their repetitive content.

Future iterations of the CBT+ app could better tailor JITAI to improve their usefulness throughout the treatment. Similarly, the CBT+ app could incorporate and use more passive channels of data collection, such as continuous glucose monitoring, which could help increase the accuracy of the algorithm (eg, provide a risk alert to a participant when they have not eaten for more than 4 hours and are at a risk of binge eating) [42].

Conclusions

In summary, our preliminary results suggest that the CBT+ system is a feasible and acceptable treatment augmentation for both patients and clinicians. If the results are confirmed in a larger RCT, the CBT+ system could prove to be a valuable addition to treatment to facilitate skill use and acquisition and ultimately reduce bulimic symptoms. Further research is necessary to determine the extent to which CBT+ is responsible for improving skills acquisition and utilization and whether improvements in skills acquisition and utilization mediate symptom reductions.

Acknowledgments

Funding for this study was awarded to AJ by the NIMH (grant R34MH116021).

Conflicts of Interest

None declared.

References

1. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. 5th Ed. Washington, DC: American Psychiatric Association Publishing; 2013:1-991.
2. Linardon J. Meta-analysis of the effects of cognitive-behavioral therapy on the core eating disorder maintaining mechanisms: implications for mechanisms of therapeutic change. *Cogn Behav Ther* 2018 Mar 30;47(2):107-125. [doi: [10.1080/16506073.2018.1427785](https://doi.org/10.1080/16506073.2018.1427785)] [Medline: [29378481](https://pubmed.ncbi.nlm.nih.gov/29378481/)]
3. Juarascio AS, Parker MN, Lagacey MA, Godfrey KM. Just-in-time adaptive interventions: a novel approach for enhancing skill utilization and acquisition in cognitive behavioral therapy for eating disorders. *Int J Eat Disord* 2018 Aug 26;51(8):826-830 [FREE Full text] [doi: [10.1002/eat.22924](https://doi.org/10.1002/eat.22924)] [Medline: [30051495](https://pubmed.ncbi.nlm.nih.gov/30051495/)]
4. Zendegui EA, West JA, Zandberg LJ. Binge eating frequency and regular eating adherence: the role of eating pattern in cognitive behavioral guided self-help. *Eat Behav* 2014 Apr;15(2):241-243. [doi: [10.1016/j.eatbeh.2014.03.002](https://doi.org/10.1016/j.eatbeh.2014.03.002)] [Medline: [24854811](https://pubmed.ncbi.nlm.nih.gov/24854811/)]
5. Barakat S, Maguire S, Surgenor L, Donnelly B, Miceska B, Fromholtz K, et al. The role of regular eating and self-monitoring in the treatment of bulimia nervosa: a pilot study of an online guided self-help CBT program. *Behav Sci (Basel)* 2017 Jun 26;7(3):39 [FREE Full text] [doi: [10.3390/bs7030039](https://doi.org/10.3390/bs7030039)] [Medline: [28672851](https://pubmed.ncbi.nlm.nih.gov/28672851/)]
6. Shah N, Passi V, Bryson S, Agras WS. Patterns of eating and abstinence in women treated for bulimia nervosa. *Int J Eat Disord* 2005 Dec;38(4):330-334. [doi: [10.1002/eat.20204](https://doi.org/10.1002/eat.20204)] [Medline: [16231339](https://pubmed.ncbi.nlm.nih.gov/16231339/)]
7. Ellison JM, Simonich HK, Wonderlich SA, Crosby RD, Cao L, Mitchell JE, et al. Meal patterning in the treatment of bulimia nervosa. *Eat Behav* 2016 Jan;20:39-42 [FREE Full text] [doi: [10.1016/j.eatbeh.2015.11.008](https://doi.org/10.1016/j.eatbeh.2015.11.008)] [Medline: [26630618](https://pubmed.ncbi.nlm.nih.gov/26630618/)]
8. Wilson GT, Fairburn CC, Agras WS, Walsh BT, Kraemer H. Cognitive-behavioral therapy for bulimia nervosa: time course and mechanisms of change. *J Consult Clin Psychol* 2002;70(2):267-274. [doi: [10.1037/0022-006x.70.2.267](https://doi.org/10.1037/0022-006x.70.2.267)]
9. Towell D, Woodford S, Reid S, Rooney B, Towell A. Compliance and outcome in treatment-resistant anorexia and bulimia: a retrospective study. *Br J Clin Psychol* 2001 Jun;40(2):189-195. [doi: [10.1348/014466501163634](https://doi.org/10.1348/014466501163634)] [Medline: [11446240](https://pubmed.ncbi.nlm.nih.gov/11446240/)]
10. Thiels C, Schmidt U, Troop N, Treasure J, Garthe R. Compliance with a self-care manual in guided self-change for bulimia nervosa. *Eur Eat Disorders Rev* 2001 Jul;9(4):255-262. [doi: [10.1002/erv.395](https://doi.org/10.1002/erv.395)]

11. Troop N, Schmidt U, Tiller J, Todd G, Keilen M, Treasure J. Compliance with a self-care manual for bulimia nervosa: predictors and outcome. *Br J Clin Psychol* 1996 Sep;35(3):435-438. [doi: [10.1111/j.2044-8260.1996.tb01197.x](https://doi.org/10.1111/j.2044-8260.1996.tb01197.x)] [Medline: [8889084](https://pubmed.ncbi.nlm.nih.gov/8889084/)]
12. Lavender JM, Wonderlich SA, Engel SG, Gordon KH, Kaye WH, Mitchell JE. Dimensions of emotion dysregulation in anorexia nervosa and bulimia nervosa: a conceptual review of the empirical literature. *Clin Psychol Rev* 2015 Aug;40:111-122 [FREE Full text] [doi: [10.1016/j.cpr.2015.05.010](https://doi.org/10.1016/j.cpr.2015.05.010)] [Medline: [26112760](https://pubmed.ncbi.nlm.nih.gov/26112760/)]
13. Fischer S, Peterson CM, McCarthy D. A prospective test of the influence of negative urgency and expectancies on binge eating and purging. *Psychol Addict Behav* 2013 Mar;27(1):294-300. [doi: [10.1037/a0029323](https://doi.org/10.1037/a0029323)] [Medline: [22823545](https://pubmed.ncbi.nlm.nih.gov/22823545/)]
14. Peterson CB, Berg KC, Crosby RD, Lavender JM, Accurso EC, Ciao AC, et al. The effects of psychotherapy treatment on outcome in bulimia nervosa: examining indirect effects through emotion regulation, self-directed behavior, and self-discrepancy within the mediation model. *Int J Eat Disord* 2017 Jun 24;50(6):636-647 [FREE Full text] [doi: [10.1002/eat.22669](https://doi.org/10.1002/eat.22669)] [Medline: [28117906](https://pubmed.ncbi.nlm.nih.gov/28117906/)]
15. Nahum-Shani I, Smith SN, Spring BJ, Collins LM, Witkiewitz K, Tewari A, et al. Just-in-time adaptive interventions (JITAI): an organizing framework for ongoing health behavior support. University Park, PA: The Methodology Center, Penn State (Technical Report No. 14-126). 2014. URL: <http://static1.1.sqspcdn.com/static/f/1067572/26291673/1433513949823/jitai-techreport.pdf?token=9kLD43wvHAgjsdhvMUnY2STJ3gA%3D> [accessed 2021-05-17]
16. Adams MA, Sallis JF, Norman GJ, Hovell MF, Hekler EB, Perata E. An adaptive physical activity intervention for overweight adults: a randomized controlled trial. *PLoS One* 2013 Dec 9;8(12):e82901 [FREE Full text] [doi: [10.1371/journal.pone.0082901](https://doi.org/10.1371/journal.pone.0082901)] [Medline: [24349392](https://pubmed.ncbi.nlm.nih.gov/24349392/)]
17. Thomas JG, Bond DS. Behavioral response to a just-in-time adaptive intervention (JITAI) to reduce sedentary behavior in obese adults: Implications for JITAI optimization. *Health Psychol* 2015 Dec;34S(Suppl):1261-1267 [FREE Full text] [doi: [10.1037/hea0000304](https://doi.org/10.1037/hea0000304)] [Medline: [26651467](https://pubmed.ncbi.nlm.nih.gov/26651467/)]
18. Kumar S, Abowd GD, Abraham WT, al'Absi M, Beck JG, Chau DH, et al. Center of excellence for mobile sensor data-to-knowledge (MD2K). *J Am Med Inform Assoc* 2015 Nov;22(6):1137-1142. [doi: [10.1093/jamia/ocv056](https://doi.org/10.1093/jamia/ocv056)] [Medline: [26555017](https://pubmed.ncbi.nlm.nih.gov/26555017/)]
19. Haapala I, Barengo NC, Biggs S, Surakka L, Manninen P. Weight loss by mobile phone: a 1-year effectiveness study. *Public Health Nutr* 2009 Mar 27;12(12):2382-2391. [doi: [10.1017/s1368980009005230](https://doi.org/10.1017/s1368980009005230)]
20. Burke LE, Styn MA, Glanz K, Ewing LJ, Elci OU, Conroy MB, et al. SMART trial: a randomized clinical trial of self-monitoring in behavioral weight management—design and baseline findings. *Contemp Clin Trials* 2009 Nov;30(6):540-551 [FREE Full text] [doi: [10.1016/j.cct.2009.07.003](https://doi.org/10.1016/j.cct.2009.07.003)] [Medline: [19665588](https://pubmed.ncbi.nlm.nih.gov/19665588/)]
21. Beasley JM, Riley WT, Davis A, Singh J. Evaluation of a PDA-based dietary assessment and intervention program: a randomized controlled trial. *J Am Coll Nutr* 2008 Apr;27(2):280-286. [doi: [10.1080/07315724.2008.10719701](https://doi.org/10.1080/07315724.2008.10719701)] [Medline: [18689560](https://pubmed.ncbi.nlm.nih.gov/18689560/)]
22. 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 Feb 24;17(2):e52 [FREE Full text] [doi: [10.2196/jmir.3951](https://doi.org/10.2196/jmir.3951)] [Medline: [25803266](https://pubmed.ncbi.nlm.nih.gov/25803266/)]
23. Versluis A, Verkuil B, Spinhoven P, van der Ploeg MM, Brosschot JF. Changing mental health and positive psychological well-being using ecological momentary interventions: a systematic review and meta-analysis. *J Med Internet Res* 2016 Jun 27;18(6):e152 [FREE Full text] [doi: [10.2196/jmir.5642](https://doi.org/10.2196/jmir.5642)] [Medline: [27349305](https://pubmed.ncbi.nlm.nih.gov/27349305/)]
24. Fairburn C. *Cognitive Behavior Therapy and Eating Disorders*. New York, United States: The Guilford Press; 2008:1-324.
25. Fairburn CG, Cooper Z, Shafran R. Cognitive behaviour therapy for eating disorders: a "transdiagnostic" theory and treatment. *Behav Res Ther* 2003 May;41(5):509-528. [doi: [10.1016/s0005-7967\(02\)00088-8](https://doi.org/10.1016/s0005-7967(02)00088-8)] [Medline: [12711261](https://pubmed.ncbi.nlm.nih.gov/12711261/)]
26. Cooper Z, Fairburn C. The eating disorder examination: a semi-structured interview for the assessment of the specific psychopathology of eating disorders. *Int J Eat Disord* 1987 Jan;6(1):1-8. [doi: [10.1002/1098-108x\(198701\)6:1<1::aid-eat2260060102>3.0.co;2-9](https://doi.org/10.1002/1098-108x(198701)6:1<1::aid-eat2260060102>3.0.co;2-9)]
27. Frisch MB. Quality of Life Inventory® (QOLI®). National Computer Systems. 1994. URL: <https://eprovide.mapi-trust.org/instruments/quality-of-life-inventory-r> [accessed 2021-05-17]
28. Fairburn CG, Beglin SJ. Assessment of eating disorders: interview or self-report questionnaire? *Int J Eat Disord* 1994 Dec;16(4):363-370. [Medline: [7866415](https://pubmed.ncbi.nlm.nih.gov/7866415/)]
29. Gratz KL, Roemer L. Multidimensional assessment of emotion regulation and dysregulation: development, factor structure, and initial validation of the difficulties in emotion regulation scale. *J Psychopathol Behav Assess* 2008 Oct 17;30(4):315. [doi: [10.1007/s10862-008-9102-4](https://doi.org/10.1007/s10862-008-9102-4)]
30. Kordy H, Krämer B, Palmer RL, Papezova H, Pellet J, Richard M, et al. Remission, recovery, relapse, and recurrence in eating disorders: conceptualization and illustration of a validation strategy. *J Clin Psychol* 2002 Jul;58(7):833-846. [doi: [10.1002/jclp.2013](https://doi.org/10.1002/jclp.2013)] [Medline: [12205723](https://pubmed.ncbi.nlm.nih.gov/12205723/)]
31. Le Grange D, Lock J, Agras WS, Moye A, Bryson SW, Jo B, et al. Moderators and mediators of remission in family-based treatment and adolescent focused therapy for anorexia nervosa. *Behav Res Ther* 2012 Feb;50(2):85-92 [FREE Full text] [doi: [10.1016/j.brat.2011.11.003](https://doi.org/10.1016/j.brat.2011.11.003)] [Medline: [22172564](https://pubmed.ncbi.nlm.nih.gov/22172564/)]

32. Min YH, Lee JW, Shin Y, Jo M, Sohn G, Lee J, et al. Daily collection of self-reporting sleep disturbance data via a smartphone app in breast cancer patients receiving chemotherapy: a feasibility study. *J Med Internet Res* 2014 May 23;16(5):e135 [FREE Full text] [doi: [10.2196/jmir.3421](https://doi.org/10.2196/jmir.3421)] [Medline: [24860070](https://pubmed.ncbi.nlm.nih.gov/24860070/)]
33. Arean PA, Hallgren KA, Jordan JT, Gazzaley A, Atkins DC, Heagerty PJ, et al. The use and effectiveness of mobile apps for depression: results from a fully remote clinical trial. *J Med Internet Res* 2016 Dec 20;18(12):e330 [FREE Full text] [doi: [10.2196/jmir.6482](https://doi.org/10.2196/jmir.6482)] [Medline: [27998876](https://pubmed.ncbi.nlm.nih.gov/27998876/)]
34. Torous J, Nicholas J, Larsen ME, Firth J, Christensen H. Clinical review of user engagement with mental health smartphone apps: evidence, theory and improvements. *Evid Based Ment Health* 2018 Aug 05;21(3):116-119. [doi: [10.1136/eb-2018-102891](https://doi.org/10.1136/eb-2018-102891)] [Medline: [29871870](https://pubmed.ncbi.nlm.nih.gov/29871870/)]
35. Juarascio AS, Goldstein SP, Manasse SM, Forman EM, Butryn ML. Perceptions of the feasibility and acceptability of a smartphone application for the treatment of binge eating disorders: qualitative feedback from a user population and clinicians. *Int J Med Inform* 2015 Oct;84(10):808-816 [FREE Full text] [doi: [10.1016/j.ijmedinf.2015.06.004](https://doi.org/10.1016/j.ijmedinf.2015.06.004)] [Medline: [26113461](https://pubmed.ncbi.nlm.nih.gov/26113461/)]
36. Agras W, Crow SJ, Halmi KA, Mitchell JE, Wilson GT, Kraemer HC. Outcome predictors for the cognitive behavior treatment of bulimia nervosa: data from a multisite study. *Am J Psychiatry* 2000 Aug;157(8):1302-1308. [doi: [10.1176/appi.ajp.157.8.1302](https://doi.org/10.1176/appi.ajp.157.8.1302)] [Medline: [10910795](https://pubmed.ncbi.nlm.nih.gov/10910795/)]
37. Fairburn CG, Jones R, Peveler RC, Hope RA, O'Connor M. Psychotherapy and bulimia nervosa. Longer-term effects of interpersonal psychotherapy, behavior therapy, and cognitive behavior therapy. *Arch Gen Psychiatry* 1993 Jun 01;50(6):419-428. [doi: [10.1001/archpsyc.1993.01820180009001](https://doi.org/10.1001/archpsyc.1993.01820180009001)] [Medline: [8498876](https://pubmed.ncbi.nlm.nih.gov/8498876/)]
38. Hoste RR, Zaitsoff S, Hewell K, le Grange D. What can dropouts teach us about retention in eating disorder treatment studies? *Int J Eat Disord* 2007 Nov;40(7):668-671. [doi: [10.1002/eat.20421](https://doi.org/10.1002/eat.20421)] [Medline: [17607714](https://pubmed.ncbi.nlm.nih.gov/17607714/)]
39. Schnicker K, Hiller W, Legenbauer T. Drop-out and treatment outcome of outpatient cognitive-behavioral therapy for anorexia nervosa and bulimia nervosa. *Compr Psychiatry* 2013 Oct;54(7):812-823. [doi: [10.1016/j.comppsy.2013.02.007](https://doi.org/10.1016/j.comppsy.2013.02.007)] [Medline: [23587528](https://pubmed.ncbi.nlm.nih.gov/23587528/)]
40. Walsh BT, Wilson GT, Loeb KL, Devlin MJ, Pike KM, Roose SP, et al. Medication and psychotherapy in the treatment of bulimia nervosa. *Am J Psychiatry* 1997 Apr;154(4):523-531. [doi: [10.1176/ajp.154.4.523](https://doi.org/10.1176/ajp.154.4.523)] [Medline: [9090340](https://pubmed.ncbi.nlm.nih.gov/9090340/)]
41. Wilson GT, Vitousek KM. Self-monitoring in the assessment of eating disorders. *Psychol Assess* 1999;11(4):480-489. [doi: [10.1037/1040-3590.11.4.480](https://doi.org/10.1037/1040-3590.11.4.480)]
42. Presseller EK, Parker MN, Lin M, Weimer J, Juarascio AS. The application of continuous glucose monitoring technology to eating disorders research: an idea worth researching. *Int J Eat Disord* 2020 Dec 07;53(12):1901-1905. [doi: [10.1002/eat.23404](https://doi.org/10.1002/eat.23404)] [Medline: [33159708](https://pubmed.ncbi.nlm.nih.gov/33159708/)]

Abbreviations

- BN:** bulimia nervosa
CBT: cognitive behavioral therapy
EDE: Eating Disorder Examination
JITAI: just-in-time adaptive intervention
NIMH: National Institute of Mental Health
RCT: randomized controlled trial
QOLI: Quality of Life Inventory

Edited by G Eysenbach; submitted 15.02.20; peer-reviewed by E Caiani, A Tsanas, N Zhao; comments to author 06.09.20; revised version received 01.11.20; accepted 11.04.21; published 31.05.21.

Please cite as:

Juarascio A, Srivastava P, Presseller E, Clark K, Manasse S, Forman E
A Clinician-Controlled Just-in-time Adaptive Intervention System (CBT+) Designed to Promote Acquisition and Utilization of Cognitive Behavioral Therapy Skills in Bulimia Nervosa: Development and Preliminary Evaluation Study
JMIR Form Res 2021;5(5):e18261
URL: <https://formative.jmir.org/2021/5/e18261>
doi: [10.2196/18261](https://doi.org/10.2196/18261)
PMID: [34057416](https://pubmed.ncbi.nlm.nih.gov/34057416/)

©Adrienne Juarascio, Paakhi Srivastava, Emily Presseller, Kelsey Clark, Stephanie Manasse, Evan Forman. Originally published in *JMIR Formative Research* (<https://formative.jmir.org>), 31.05.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Formative Research*, is properly

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

Original Paper

Displayed Depression Symptoms on Facebook at Two Time Points: Content Analysis

Megan A Moreno¹, MD, MEd, MPH; Quintin Gaus¹, BS; Megan Wilt², BS, RN; Alina Arseniev-Koehler³, MA; Adrienne Ton⁴, NP; Molly Adrian⁵, PhD; Ann VanderStoep⁵, PhD

¹Department of Pediatrics, University of Wisconsin-Madison, Madison, WI, United States

²Department of Nursing, St Catherine University, Minneapolis, MN, United States

³Department of Sociology, University of California, Los Angeles, Los Angeles, CA, United States

⁴Department of Nursing, Columbia University, New York, NY, United States

⁵Department of Psychiatry, University of Washington, Seattle, WA, United States

Corresponding Author:

Megan A Moreno, MD, MEd, MPH

Department of Pediatrics

University of Wisconsin-Madison

2870 University Ave, Suite 200

Madison, WI, 53705

United States

Phone: 1 (608) 262 4440

Email: mamoreno@pediatrics.wisc.edu

Abstract

Background: Depression is a prevalent and problematic mental disorder that often has an onset in adolescence. Previous studies have illustrated that depression disclosures on social media are common and may be linked to an individual's experiences of depression. However, most studies have examined depression displays on social media at a single time point.

Objective: This study aims to investigate displayed depression symptoms on Facebook at 2 developmental time points based on symptom type and gender.

Methods: Participants were recruited from an ongoing longitudinal cohort study. The content analysis of text-based Facebook data over 1 year was conducted at 2 time points: time 1 (adolescence; age 17-18 years) and time 2 (young adulthood; ages 20-22 years). Diagnostic criteria for depression were applied to each post to identify the displayed depression symptoms. Data were extracted verbatim. The analysis included nonparametric tests for comparisons.

Results: A total of 78 participants' Facebook profiles were examined, of which 40 (51%) were male. At time 1, 62% (48/78) of the adolescents had a Facebook profile, and 54% (26/78) displayed depression symptom references with an average of 9.4 (SD 3.1) references and 3.3 (SD 2.3) symptom types. Of the 78 participants, 15 (19%) females and 12 (15%) males displayed depression symptom references; these prevalence estimates were not significantly different by gender ($P=.59$). At time 2, 35 young adults displayed symptoms of depression with an average of 4.6 (SD 2.3) references and 2.4 (SD 1.3) symptom types. There were no differences in the prevalence of symptoms of depression displayed between males ($n=19$) and females ($n=16$; $P=.63$).

Conclusions: This content analysis study within an ongoing cohort study illustrates the differences in depression displays on Facebook by developmental stage and symptom. This study contributes to a growing body of literature by showing that using social media to observe and understand depression during the emerging adult developmental period may be a valuable approach.

(*JMIR Form Res* 2021;5(5):e20179) doi:[10.2196/20179](https://doi.org/10.2196/20179)

KEYWORDS

adolescents; content analysis; depression; Facebook; social media

Introduction

Background

Depression often has an onset during adolescence and young adulthood [1-3]. Social networking sites (SNSs) may present new opportunities to understand emerging adults' experiences of depression, as over 90% of youth and 75% of adults use one or more SNS [4,5]. Previous work has found that displayed depression symptoms on SNSs are common; a study of college students found that these displays were present on up to a quarter of participants' profiles [6]. Furthermore, displayed depression symptoms on SNSs were positively associated with self-reported depression symptoms using clinical screening tools [7]. Given the nearly ubiquitous use of SNSs among adolescents and young adults, the disclosure of depression on SNSs may present an innovative opportunity to evaluate how depression displays on social media may change during the developmental transitions of emerging adulthood. Thus, SNSs may allow researchers to understand unique aspects of the depression experience during adolescence and young adulthood, such as gender differences or changes in the types of symptoms displayed over time. Research approaches incorporating SNSs may enhance our understanding of symptom burden, symptoms over time, and differences by developmental stage and gender.

Depression During the Emerging Adult Years

The emergence of depression is a major health concern during adolescence; the prevalence of depression increases between the ages of 10 and 18 years [1-3]. The most common form of depression in adolescents and young adults is major depressive disorder, with a yearly incidence of approximately 8% [8,9]. An additional 22% of adolescents and young adults have *subdiagnostic* levels of depression symptoms, which is an important and understudied group [10].

Patients with a major depression diagnosis and those with *subdiagnostic* levels of depression symptoms experience impaired functioning and morbidity. It is well established that depression is associated with numerous negative health and social outcomes. These outcomes can include poor academic achievement, increased rates of substance use, comorbid psychiatric conditions, and, most severely, suicide [11-15].

Depression and Gender

Gender is another critical factor in the depression experience; previous work suggests that beginning in adolescence, females have higher rates of depression compared with males [16-18]. It has been proposed that gender differences reported in the literature may reflect differences in depression prevalence and depression symptom types or differences in diagnosis and treatment by providers [19]. A recent meta-analysis found that gender differences in depression diagnoses were seen at the young age of 12 years, peaked in adolescence, and declined in adulthood [20].

A New Lens to Understand Depression Experiences: SNSs

SNSs may present new opportunities to investigate depression, particularly among adolescents and young adults. The vast

majority of adolescents and young adults use an SNS, such as Instagram, Snapchat, or Facebook [5,21]. SNSs allow users to build a personal profile, communicate with others, and build a social network. Thus, SNSs provide unique tapestry for curating personal content and sharing with viewers.

Through content displayed on an SNS profile, profile owners may provide researchers insight into experiences that are not always apparent in their life offline. SNSs, such as Facebook, allow individuals to create personal profiles that are public representations of themselves. Facebook users have the ability to change their web-based identity on a daily or even minute-to-minute basis. These disclosures may include references to depression symptoms in the form of *status updates*, that is, personally written text describing the profile owner's current state of mind. Thus, SNSs allow users to capture their lived feelings, behaviors, and experiences at any time or place.

Examining content on an SNS as a representation of one's self has theoretical support. Even before the advent of using social media as a venue for personal expression, research has shown that human-computer interaction use fostered self-disclosure and uninhibited personal expression [22,23]. Social media may foster the self-expression of depression symptoms through their unique affordances. Affordances are typically described as properties of artifacts that can be recognized by users and contribute to their function [24]. An affordance approach can be useful because it provides a way to identify functionalities that are similar and different across these platforms [25-28]. Social media affordances can include social affordances, cognitive affordances, or functional affordances. Social affordances provided by social media may enhance the likelihood of disclosures of depression symptoms in social media venues and include a sense of belonging to a group, such as a group focused on a particular interest, experience, social group, or religion [29]. This sense of belonging may lead to social media being a safe space in which depression disclosures can occur. Another social affordance is its capacity to promote network-informed associations, such as when Facebook suggests friends for a user based on their friends' friends. This allows users to see how friends are connected to other people and their interests [30]. This feature may again contribute to the sense that Facebook is a safe and connected space for disclosure. These affordances may inform mechanisms by determining which types of communication are common on particular platforms.

Affordances are also a useful lens through which to view social media platforms. As platforms can change in popularity or features within an individual platform change, researchers can focus on underlying affordances to interpret observations and predict whether similar observations would be seen on platforms with similar affordances [31].

Evidence Supporting Depression Symptom Disclosures on SNSs

Several areas of previous research support the investigation of depression using SNSs. First, previous work illustrates that from the inception of social media, adolescents have used SNSs to disclose personal information and seek support. One study from the early period of social media found that adolescents reported

that they often disclosed more about themselves on SNSs than they did in person [32]. Adolescents with depression have also described using SNSs to express themselves and seek social support [33]. Peers may respond to these disclosures with social support. Previous work supports that peers often view SNS depression symptom disclosures as a *cry for help* and respond by providing support on the web and offline [34].

Second, depression symptom displays are common across SNSs. Previous studies have described depression symptom displays across various SNSs, including Facebook, Twitter, and Tumblr [6,35-37]. Previous studies have reported that between 25% and 33% of Facebook profiles of adolescents and young adults include references to depression symptoms consistent with the symptom criteria for a major depressive episode (MDE) [6].

Third, studies support links between displayed depression symptoms on SNSs and self-reported depression symptoms. One study found a positive association among young adults between posting displayed references to depression on Facebook and self-reporting depression symptoms [7]. Another study found that adults who used Facebook to post about health concerns were also more likely to have depression via a medical record review [38]. To date, these studies have focused on displayed depression symptoms in young adults or adults on SNSs.

Finally, previous literature has explored different approaches to examine the relationship between social media and depression, such as through studies examining the interaction between individual users and social media [39,40]. Other studies have examined social media language and compared it with the language present in medical records [41]. These studies complement approaches that examine social media directly and highlight the potential knowledge to be gained from understanding how depression is represented on social media and the user experience.

The Role of Development

Adolescence is often the time when depression symptoms arise; thus, adolescents may be faced with these symptoms during a developmental stage in which their abstract thinking and ability to conceptualize the future is still developing [1-3]. As a result, adolescents may post their experiences on social media as an outlet for expression or to seek peer support. Developing and relying on peer support is a major task in adolescence. Given the social affordances of social media, these platforms may represent safe spaces in which adolescents feel comfortable expressing their new experiences. Alternatively, adolescents may be reluctant to share depression-related experiences on social media, as the fear of stigma among their peer groups may limit their comfort in expressing depression symptoms. As most studies to date of displayed depression symptoms on social media have focused on young adults, it remains unclear how and how often adolescents choose to represent depression symptoms on social media.

Young adulthood is a time in which youth achieve new levels of independence from parents and experience additional responsibilities for work or education. For young adults with depression, the developmental changes that occur in early

adulthood can cause increased stressors [42]. Therefore, early adults may use social media to express these stressors to a known peer network where they can seek support. Previous studies have shown that young adult peers frequently see displayed depression symptoms by their peers and often interpret these displays as requests for support [34].

The developmental stage, which combines adolescence with young adulthood, has been proposed as a contributor to the symptom experiences of adolescents and young adults with depression [42]. A previous study argued that although the core experiences of depression were similar for adolescents and young adults, experiences related to stressors and depression triggers differed by developmental stage [43]. Other studies have found an improvement in depression symptoms in early adulthood compared with adolescence, suggesting that an increased level of experience, cognitive development, and psychosocial maturity may be key factors in improving an individual's self-management of depression [44,45].

Gaps in the Literature

At present, there are important gaps in the literature. First, this study is unique in its use of displayed depression symptoms on SNS samples at 2 developmentally distinct time points. There has been a call for studies on media about the adolescent developmental period [46]. Previous approaches using limited longitudinal study designs using SNSs have been successful. A previous study evaluated displayed references on SNSs regarding religion and sexual behavior over a one-year period [47]. A second study followed students during their first year of college to understand the emergence of alcohol displays on Facebook [48]. However, this study provides a unique contribution of examining 2 distinct time points, years apart, to understand differences by developmental period within a longitudinal cohort study.

Second, this study evaluates depression displays on SNSs at the symptom level. To date, most previous studies have focused on undifferentiated and aggregate displayed depression on SNSs. Depression symptoms can include sleep issues, which may be nonspecific to depression, and depressed mood, which is a more specific concern. The investigation of specific depression symptom displays on SNSs remains to be understudied. Furthermore, whether there are changes in displayed depression symptom frequency or types of depression symptoms displayed in adolescence compared with early adulthood remains to be unknown.

Third, studies provide conflicting findings regarding whether males or females are more likely to display depression references on SNSs. One previous study of college students found that 40% of females were depression symptom displayers (DSDs), compared with 25% of males [7]. Another study of college students found that females were more likely to display references to both stress and depression on SNSs [49]. Most of these studies have focused on college students, so it remains to be unclear how these results apply to adolescents or populations that are more general. A previous qualitative study focused on college students found that students perceived no differences in the frequency of seeing depression symptom disclosures on SNSs from male versus female peers [34].

To summarize, there is incomplete knowledge of depression displays by developmental stage, depression symptom displays, and gender, considering the unique time points of adolescence and young adulthood.

Study Purpose

The aim of this content analysis study was to evaluate displayed depression symptoms on Facebook posts among a cohort of participants at 2 time points: adolescence and young adulthood. Our study objectives are to determine (1) the prevalence of displayed depression symptoms at 2 developmental time points: adolescence and young adulthood; (2) the most common depression symptoms displayed on Facebook; and (3) the prevalence of displayed depression symptoms by gender. As we conducted this study in the context of an ongoing longitudinal study, we also aim to illustrate our methodological process and its strengths and limitations so that future studies could consider such approaches in the context of their studies.

Methods

Overview

This study used a cohort of participants enrolled in the Developmental Pathways Project (DPP). This ongoing parent study was ideal for the purposes of this study for 3 reasons. First, the DPP study design was longitudinal and collected data from participants in adolescence (ages 17-18 years) and young adulthood (ages 21-22 years). Second, the study design of recruiting students in adolescence allowed us to leverage a more general population of adolescents, rather than focusing on college student populations, as in many previous studies. Third, recruitment was designed to oversample participants with early signs of depression, to increase the likelihood of having depression symptom displays present on SNSs.

The detailed screening and sampling procedures have been described previously [50,51]. Informed consent was obtained from adolescents and their parents at the start of the study. All study procedures were approved by the University of Washington Human Subjects Division.

Participants

Recruitment began with mailing an introduction letter from the DPP principal investigators with a fact sheet for this study. This sheet explained that the purpose of the study was to learn about how young adults use social media. The participants were informed that we would be contacting them via email to request their Facebook profile name and URL to send a Facebook *friend request*. We created a Facebook profile for this study that participants would link their profile with for the duration of the study. We informed participants that if they accepted our *friend request* via Facebook, we would view their Facebook profile to get an idea of how often they used the profile and what they posted on it. They were assured that we would post no comments

or public messages to their profile and that all information we gathered would be kept confidential. Participants were told that accepting the Facebook friend request served as their voluntary, informed consent to participate in the Facebook study. Participants received US \$10 for participation.

One week after mailing the letter, we sent each eligible participant an email that contained the same letter and fact sheet and asked for their Facebook profile name and URL. For those who did not respond, we searched for their Facebook profile using their full name or email address. To verify that the participant's Facebook profile belonged to the actual participant, one of the following data-matching criteria had to be met by finding the information on the Facebook profile and matching it to the information on file for that participant: their email, birthday, city and high school, city and high school graduation year, or high school and high school graduation year.

One week later, we sent friend requests to DPP participants who met the eligibility criteria for whom we had Facebook profile names. The friend requests included a Facebook message explaining that there was a new DPP-related study opportunity and that if they wished to participate, they would need to accept the friend request within 3 months. If we were unable to find the Facebook account, we sent a follow-up email requesting account information. If the participant had not accepted the friend request one week later, we sent a text message reminding them to check their email and Facebook accounts. If the participant still had not friended the Facebook account after 3 months, the friend request was rescinded.

Depression Symptom Display Codebook Development

Using previous studies, we created a depression codebook containing keywords used to evaluate Facebook profiles and Twitter feeds for references to depression [6,35,36]. References to depression symptoms were defined using the symptoms outlined in the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV) criteria for an MDE (Table 1) [52]. We chose the established clinical criteria as the basis of our research codebook to provide the most clinically relevant, objective framework for profile evaluation. The criteria for MDE based on the version of the DSM used in this study included depressed mood, loss of interest or pleasure in activities, weight or appetite changes, sleep problems such as insomnia or hypersomnia, slowing down of thought and movement (referred to as psychomotor retardation in the DSM-IV), energy loss, feelings of worthlessness or guilt, decreased concentration, or suicidal ideation [52]. Some of these criteria in the DSM-IV are further described as requiring observation by a physician to be considered valid, such as changes in appetite or weight. However, as this study involved evaluating self-disclosed feelings, thoughts, and experiences by participants, all symptoms were evaluated at face value, as reported by the participant.

Table 1. Depression symptom definitions applied to Facebook profile status updates.

DSM-IV ^a criteria	Example key words or phrases	Exclusions
Depressed mood	Sad, empty, crying, tearful, alone, lonely, distressed, down (unless context clarifies otherwise)	"I had a bad day": reference to a negative event without expressing emotion, FML (f*ck my life)
Decreased interest or pleasure in activities or anhedonia	Not having fun, don't feel like doing anything, giving up, lack of purpose, not caring, not giving a "f*ck"	"I'm bored": references to being bored or decreased interest in something specific such as a class or project
Changes in weight or appetite	No appetite, don't feel like eating, can't stop eating, eating everything in sight	"I ate too much McDonalds this weekend": references to poor eating habits rather than change in appetite or references to single episodes of eating too much, such as at Thanksgiving
Insomnia or hypersomnia	Tired, exhausted, sleepy, need a nap, easily tired, not sleeping well, restless sleep, can't sleep, can't get to sleep, falling asleep in an unusual place, insomnia	"I took a nap after finals": references to specific fatigue-related events
Agitation or slowing down of movement	Inability to sit still or feeling slow	"So excited for this concert that I can't sit still!": references to energy increases linked to specific events
Fatigue or loss of energy	Can't get anything done, can't get motivated due to fatigue	"I'm tapped out after that soccer game": references to specific events or activity leading to low energy
Feelings of worthlessness or guilt	Feel guilty or worthless, "I am stupid," "I'm not cool," "I'm average," "I'm crazy" or "I'm insane," regretting something, being a failure, or failing	"I'm feeling so guilty after binge watching the whole fifth season": references to single episodes of guilt linked to one-time activities
Difficulty concentrating or indecisiveness	Can't study, can't finish work, can't concentrate because of emotion, can't decide on something, don't feel like deciding, can't make up your mind, not knowing what to do	"I can't concentrate on this homework because I am on Facebook": references to not wanting to concentrate, or can't concentrate because of activity
Recurrent thoughts of death	Thinking of ways to commit suicide, references to jumping, referencing death of self, thinking about death	^b

^aDSM-IV: Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition.

^bNot available. No data fit this category.

Text-based Facebook displays were coded as depression references if they fit one of the described depression criteria based on keywords or synonyms. For example, one symptom of major depression is depressed mood; therefore, a status update stating "I feel depressed" would be coded as a depression symptom reference. Feeling down is often considered synonymous with feeling sad; therefore, a status update disclosing "I feel really down" would be coded as a depression symptom reference. Status updates or comments that referenced the common experience of having a bad day, such as "in a bad mood today because of my math exam," would not meet the criteria for depression symptoms.

For this study, we began with the previously established codebook and engaged coders in testing and revising it for this study to ensure that symptom categories were well-developed and represented the current language on SNS profiles. A total of 4 rounds of pilot testing were conducted with participant profiles that were not included in the study data set. Modifications and revisions to the codebook were made to ensure that specific diagnostic criteria were represented clearly and as objectively as possible. Table 1 presents the revised codebook with examples of the symptom types.

Depression Symptom Display Coding Procedure

Coding Depression Symptom Displays by Time Point

To evaluate depression symptom disclosures, coders reviewed each profile's history of status updates and comments beginning one day before the date of interview at time 1 (ages 17-18 years) and time 2 (ages 21-22 years) and coded in reverse chronological order back to the same date, a year ago. Therefore, each period was tailored to the individual participant linked to the interview date but covered an identical range of time of 1 year. Each coding period was thus separated by 3 to 4 years.

Coding Depression Symptom Displays: Exclusions

For each social media post examined, if the post did not include any reference to depression symptoms, it was not coded. Status updates or comments that referenced a person other than the profile owner (ie, "Matt is sitting next to me in class and he looks bummed") were not considered personal references and were excluded. Only text was evaluated in this study; thus, photos, links, and videos were excluded from the evaluation.

Coding Depression Symptom Displays: Process

For profiles in which a status update or comment disclosed a depression symptom, investigators categorized the type of

depression symptom into one of the specific symptom categories. The unit of coding was a statement matching the depression symptoms. Therefore, if a post included multiple statements describing multiple symptoms, then each statement was counted as an individual depression symptom. For any displayed depression symptoms that met the criteria for representing suicidal ideation, researchers used an algorithm from previous studies to determine whether a response to that post was warranted [36]. Responses could include providing mental health resources or contacting participants to check if the concerning post was recent.

Interrater Agreement

Interrater agreement across depression symptom coding categories was compared across all the coders. The interrater agreement for depression symptom type ranged from a low of 88% for anhedonia to 98% for sleep concerns.

Analysis

Qualitative assessment of the data was performed via content analysis. All coders went through a training phase of coding using pilot data not included in this study and began coding for this study once interrater agreement goals of >90% were met. When coding posts that met the criteria as a depression symptom, coders evaluated symptom types and placed them into appropriate symptom categories based on the codebook. Weekly meetings of all coders allowed the discussion of questionable codes and decisions were made by consensus.

For quantitative data, demographic variables and displayed depression references on Facebook were evaluated using descriptive statistics. We used Wilcoxon rank sum tests to test for associations between gender and the number of Facebook depression displays at each time point. Chi-square tests were used to test for an association between the presence of depression display and gender at each time point. All *P* values were two-sided, and $P < .05$ was used to indicate statistical significance. Statistical analyses were conducted using Stata 12 (StataCorp LP).

Results

Overview

This study was funded in October 2014 and received institutional review board approval from the University of Washington in February 2015. Participants from the larger DPP study were recruited for this study between February and April 2015. Retrospective data collection using Facebook profiles was conducted between April 2015 and August 2016. Statistical

analyses and data interpretation were conducted subsequent to that time.

Participants

Of 394 participants, 212 (53.8% of the DPP sample) participants were eligible for recruitment on the basis of the inclusion criteria. Among these 212 eligible DPP participants, 199 (93.9%) participants provided information about their Facebook profile, and 89 (41.9%) participants consented to participate in this study. A total of 78 out of 89 participants accepted our friend request and had data available at the time 2 data point.

At time 1, the sample of participants who had a Facebook profile as an adolescent included 48 participants aged between 17 and 18 years, of whom 48% (23/48) were male. The time 2 sample of participants who had a Facebook profile as a young adult included 78 participants, aged between 21 and 22 years, of whom 51% (40/78) were male.

Depression Symptom Displays by Time Point

At time 1, 62% (48/78) of the participants had Facebook profiles and were included in the analyses. Of these 48 individuals whose Facebook profiles were coded, 27 (56%) displayed one or more depression symptoms on Facebook. We will refer to these participants as *DSDs*. The total number of depression displays over the year among all DSD profiles was 269. DSDs had an average of 5.3 (SD 9.2) depressive symptom displays and a median of 1 depressive symptom display in the one-year period. At time 1, the Facebook profile often represented status updates in the third-person language. A typical status update would read as “[Profile owner name] is....” Thus, many status updates were third person, such as “is feeling very sad.” Common types of language associated with these status updates included the use of all capital letters for emphasis (eg, “can’t stop EATING”) and unique spellings in shorthand (eg, “sleep d-prived”).

At time 2, all 78 participants had Facebook profiles and were included in the analyses. Of these 78 participants, 35 (45%) were DSDs. The total number of depression symptom displays by these DSDs at time 2 was 172. At time 2, DSDs had an average of 4.6 (SD 3.9) and a median of 3 displays. While by this point on Facebook, the third-person status update was no longer the norm, we saw the continued use of all capital letters for emphasis (eg, “just ate SO MUCH”) and seemingly purposeful misspellings, such as “I don’t wanna.”

Depression Symptom Types by Time Point

Depression symptom displays were present across both time points (Table 2).

Table 2. Displayed depression symptom type on Facebook profiles by time point.

Depression symptom type	Time 1 (17-18 years; 2008-2010)		Time 2 (21-22 years; 2011-2014)	
	Posts (n=255), n (%)	Examples	Posts (n=167), n (%)	Examples
Depressed mood	36 (14.1)	“is kind of sad,” “is terrible, thanks for asking,” “is really, truly depressed”	32 (19.2)	“feeling the worst right now, just wanting to cry...,” “can’t depend on anyone, so sad”
Decreased interest or pleasure in activities or anhedonia	43 (16.9)	“doesn’t care anymore,” “too late to pull together the value of my life”	27 (16.1)	“and everything continues to suck...,” “what is this life? other than a slow and painful death”
Changes in weight or appetite	3 (1.2)	“fasting the rest of the week,” “can’t stop... EATING”	4 (2.4)	“craving every sort of food,” “another day of overeating”
Insomnia or hypersomnia	85 (33.3)	“5 naps in a day, phew,” “sleep deprived”	69 (41.3)	“just another night of insomnia,” “i don’t want to wake up”
Agitation or slowing down of movement	11 (4.3)	“so restless,” “feels paralyzed”	0 (0)	N/A ^a
Fatigue or loss of energy	3 (1.2)	“says get the heck up!” “feeling a bit listless...”	4 (2.4)	“don’t wanna do a thing,” “no motivation”
Feelings of worthlessness or guilt	34 (13.3)	“no one likes me,” “tired of not being good enough,” “wishing she was beautiful”	18 (10.8)	“officially a weekend loser... fail,” “ahhh i struggle with life”
Difficulty concentrating or indecisiveness	25 (9.8)	“NEEDS to learn how to focus!” “what am i doing with life...”	6 (3.6)	“don’t know how to feel...,” “so clueless and lost, i am”
Recurrent thoughts of death	15 (5.9)	“going to die,” “I am terminal,” “hard times call for drastic measures- got a gun,” “dad I’m still here, wishing I was dead with you.”	7 (4.2)	“have a party at my funeral,” “if I died tomorrow, what would you say to me,” “i wanna kill myself”

^aN/A: not applicable.

At time 1, depression symptom displays included an average of 3.3 symptom types (SD 2.3; median 2) per participant. A common symptom type was sleep issues; 85 posts referenced sleep challenges, such as “5 naps in a day, phew” and “tired, I don’t want to wake up.” Another common type of depression symptom display was decreased interest or pleasure in activities (43 posts referenced this), such as “doesn’t care anymore” and “not having fun ever.”

At time 2, status updates referencing depression included an average of 2.4 (SD 1.3; median 2) depression symptoms among 9 possible symptom types. Similar to time 1, commonly displayed symptoms at time 2 included sleep issues (69/167, 41.3% of depression symptom posts), depressed mood (32/167, 19.2% of depression symptom posts), and decreased interest or pleasure (27/167, 16.2%). Table 2 shows examples of depression posts by symptoms.

Prevalence of Displayed Depression Symptoms by Gender

Depression symptoms were observed in both females and males at both time points (Table 3).

At time 1, 52% (12/23) of males and 60% (15/25) of females displayed depression symptom references on Facebook; the prevalence estimates were not significantly different ($P=.59$). The median number of depression symptom displays for females and males was not significantly different ($P=.28$). The median number of displays for both genders was 1.

Across each of the 9 depression symptom categories except suicidal ideation, females consistently had a high proportion of displays; however, these differences were not statistically significant ($P=.31$ and $P=.72$ with Yates correction).

At time 2, displayed depression symptoms were present in 48% (19/40) of male profiles and 42% (16/38) of female profiles ($P=.63$); there were no significant differences by gender in the number of displayed depression symptoms ($P=.94$).

While comparing females and males across types of depression symptoms, males had higher rates of decreased interest or pleasure and loss of energy compared with females; however, these differences were not statistically significant ($P=.11$ and $P=.40$ with Yates correction).

Table 3. Facebook depression symptom displays by gender at time 1 and time 2.

Time point	Gender	Participant, n	Facebook depression symptom display, n (%)	Depression symptom displays on Facebook profiles	
				Value, mean (SD)	Value, median (range)
Time 1 (n=48)					
	Males	23	12 (52)	2.7 (5.7)	1 (0-25)
	Females	25	15 (60)	7.6 (11.2)	1 (0-33)
Time 2 (n=78)					
	Males	40	19 (48)	1.6 (2.8)	0 (0-11)
	Females	38	16 (42)	2.4 (3.9)	0 (0-15)

Discussion

Principal Findings

This study applied content analysis to evaluate displayed depression symptoms on Facebook within an ongoing longitudinal cohort study of adolescents. The findings contribute to the state of the science regarding how symptoms of depression are displayed on SNSs at 2 time points across a critical developmental period of emerging adulthood: by time, symptom, and gender. We identified that DSDs were common at both time points, with a variation in prevalence and symptom type among adolescents and young adults. We did not identify any gender differences in the prevalence of depression symptom displays at either time point.

Types and Timing of Displayed Depression Symptoms on Facebook

The first finding is the high prevalence of displayed depression symptoms at both time 1 and time 2. This finding is different from that previously reported in the literature. Previous studies have reported that the prevalence of displays of depression symptoms on Facebook ranges between 25% and 33% [6,7]. It is likely that this study sample, which was oversampled for early signs of depression as part of the prospective cohort design, was a major contributing factor to the higher prevalence of displayed depression symptoms in this study. However, it is also important to note that most previous studies have focused on college students. This study focused on adolescents, not all of whom were in college at the young adult time point. As this is the first study to examine displayed depression disclosures by a population of adolescents, our study findings support the importance of future work focused on the adolescent population to better understand how displayed depression symptoms may present opportunities to better understand, detect, or provide early intervention for depression.

Furthermore, we found that 53% (41/78) of our participants chose to represent depression symptom displays on Facebook as adolescents and about 44% (34/78) chose to represent depression symptom displays on Facebook as young adults. Another interesting finding is that at time 1, our 38 participants generated 269 depression symptom displays, whereas at time 2, our larger number of 78 participants posted a smaller number of 172 depression displays. These findings align with previous longitudinal studies of the emerging adolescence period, and one study found decreases in depression symptoms over 7 years

between adolescence and early adulthood [44]. These results support early adulthood as a time in which many early adults experience improved psychological well-being and psychosocial maturity, which may manifest via decreased depression disclosures on social media.

A second finding was that both adolescents and young adults represented their depression symptoms using all 9 symptom types, including mood symptoms, physical symptoms, such as sleep, and even suicidal ideation. Furthermore, participants expressed these symptoms in informal ways, such as slang, shortened text, and all capitals for emphasis. This language may be more commonly attributed to the adolescent developmental stage but was observed across both time points in our study. In addition, both adolescents and young adults were willing to share information that may be seen as quite deep and personal information about how they were feeling, including clear descriptions of depression symptoms, such as hopelessness. We also found descriptions of poor self-worth, sleep challenges, and appetite changes. It is possible that having a network of peers on SNSs allows emerging adults to share developmental concerns around self-worth, sleep difficulties, and changing appetites that are fundamental to the developmental transition from adolescence to young adulthood.

The rich descriptions that emerging adults were willing to share may be startling as representations of publicly shared text. However, these disclosures may reflect an emerging cultural norm around what types of content are acceptable to share, both in a web-based public format and among a new generation of emerging adults who may view mental health as less stigmatizing compared with previous generations. It is also possible that some of these disclosures were not perceived by participants as depression symptoms, such as disclosures related to sleep problems or weight changes. Previous studies have illustrated that emerging adults often receive peer support for these disclosures, which may include social media comments in response, or offers to get together offline [6,53]. Thus, some adolescents and young adults may display depression symptoms in an effort to reach out for help or support from their peer network. Peers may be the first to observe and respond to depression symptom disclosures, and symptoms suggesting hopelessness or suicidal ideation may be the most critical for peer response. Future studies exploring the motivations and mechanisms of peer support in response to displayed depression symptoms are warranted.

No Differences in Depression Symptom Displays on SNSs by Gender

A third finding was that no differences in the prevalence of displays or types of symptoms displayed between females and males were found in this study. This finding may seem surprising, given that the literature suggests that females more often display personal or relational content and mention more psychological and social processes, such as emotions and feelings [3,6]. There are several possible explanations for the absence of gender differences in our study. It is possible that some aspects of the initial DPP study design, such as the sampling procedures for the original DPP study, might have influenced SNS disclosure patterns for both females and males in ways that we cannot understand. It is also possible that SNSs allow males a safe or culturally appropriate place to share depression symptoms or experiences. Previous research focused on early computer use communication found that human-computer interaction can lead to increased self-disclosure and less inhibited personal expression [54,55]. It is possible that this expression in males allows them the opportunity to express emotions in ways that may be more intimidating in offline communication. Thus, SNSs may be a venue in which males feel comfortable expressing depression symptoms.

It is interesting to consider another perspective that it is possible that our coding approach may have underestimated the frequency of male depression disclosure. It has been suggested that males are more likely to endorse other symptoms of distress, such as aggression, substance abuse, and risk-taking behavior, whereas women are more likely to endorse symptoms of sadness, stress, irritability, sleep problems, and loss of interest [56]. Our coding approach was based on the DSM-IV, and the symptoms of distress commonly noted by males were not among the criteria. However, the DSM-IV criteria does include the symptoms most commonly reported by women. Therefore, it is possible that some males' disclosures related to underlying depression may have been missed based on our coding criteria, leading to an underestimation of mental health issues of males displayed on social media.

Limitations

Our study had several limitations. First, although the overall sample size of the cohort study from which we recruited for this study was large, the number of participants who met our inclusion criteria and were recruited was relatively small. The smaller sample and cohort study design allowed us to thoroughly evaluate participants' displayed symptoms at 2 time points, yielding rich participant-generated data that allowed both quantitative and qualitative analyses. However, ongoing studies that integrate social media approaches as additional data collection may improve their participant numbers through approaches such as phone calls to participants or increased stipends. Another limitation of our study recruitment approach is potential sampling bias, as the participants who consented may have had different Facebook display patterns compared with those who declined participation. Furthermore, this study's purpose was to evaluate user-generated text describing depression symptoms applying the DSM-IV criteria, which are also in writing and not represented via images. Thus, multimedia

content such as photos, videos, or links to articles was not evaluated, although there are no current criteria or standards defined to evaluate such content. We evaluated all disclosures through the lens of the DSM-IV but did not factor in additional requirements for some criteria to be observed by a clinician, given our study's focus on social media displays by participants. In taking all public web-based disclosures at face value, we cannot make claims about the motivation or context for participants' posts. Another limitation is that we evaluated displayed depression symptoms on only one site, Facebook. As noted repeatedly in this paper, Facebook as a platform is not a static entity but is dynamic and changes over time. These changes were well illustrated within our data; at time 1, status updates were in the third-person language, and by time 2, this was no longer the case. Therefore, aspects or patterns of display at the earlier or later data collection times in our study may have changed. This is a limitation of any social media study and emphasizes the importance of reflecting on overall aspects of communication and culture around social media, such as through an affordance approach, rather than particular nuances of a single platform. Finally, because this study was retrospective within an existing cohort, data are not recent and may not reflect current trends in depression symptom displays on social media.

Conclusions and Implications

Despite these limitations, our study provides novel findings that contribute toward understanding how adolescents and young adults represent depression symptoms on Facebook during the critical developmental period of emerging adulthood. An important research implication of our study findings is that our findings can contribute to other studies in this important area. First, our study may inform text mining or big data studies of depression disclosure. It is important to note that many of the depression symptoms coded in this study that met the criteria within our codebook would not likely be identified through big data approaches tied to keywords. Adolescents used slang and deliberate misspellings, such as "don't wanna do anything," to represent constructs such as anhedonia. Studies using keywords or big data approaches to mine text-based social media data may find it useful to consider unexpected yet common word combinations to identify content that may signify depression symptoms. A previous systematic review noted that assembling large data sets of mental health symptoms on social media is problematic because of limitations and biases with data collection methods [57]. Combining *thick data*, such as what was collected in this study through human coders, with big data approaches using automated detection may allow for a broader yet more nuanced understanding of depression.

A second implication for future studies is to use our findings to design future studies more narrowly focused on specific or critical depression symptoms. Our study took the approach of understanding the ways in which all types of depression symptoms may be represented on Facebook. One finding was that depression symptoms around sleep concerns were common, although this particular symptom type could be interpreted as nonspecific for depression. Thus, future studies may elect to narrow the focus on sleep-related coding to only code posts that specifically mention sleep issues and depression. Alternatively,

future studies could consider focusing on coding symptoms that are more specific to depression, such as depressed mood.

Furthermore, our study provides methodological contributions on how existing studies may incorporate SNS evaluations into the study design. We found that using an existing study sample from a longitudinal cohort study was feasible and provided a novel approach to the evaluation of depression symptom displays on Facebook. We found that from our larger

longitudinal sample, very few participants had data at the 2 time points evaluated. This finding may be applicable to other studies, as social media patterns and platforms change over time. Finally, given the frequency of depression symptom displays, SNSs may be an innovative avenue for raising self-awareness and combating stigma surrounding mental health conditions. As social media continues to evolve, further studies to evaluate these types of displays over time or across platforms are warranted.

Acknowledgments

The authors would like to thank Nika Sulakvelidze for his contributions to data collection and Dr Elizabeth McCauley for her contributions to the study design.

Conflicts of Interest

None declared.

References

1. Anttila KI, Anttila MJ, Kurki MH, Välimäki MA. Social relationships among adolescents as described in an electronic diary: a mixed methods study. *Patient Prefer Adherence* 2017;11:343-352 [FREE Full text] [doi: [10.2147/PPA.S126327](https://doi.org/10.2147/PPA.S126327)] [Medline: [28280307](https://pubmed.ncbi.nlm.nih.gov/28280307/)]
2. Waloszek JM, Schwartz O, Simmons JG, Blake M, Blake L, Murray G, et al. The SENSE Study (Sleep and Education: learning New Skills Early): a community cognitive-behavioural therapy and mindfulness-based sleep intervention to prevent depression and improve cardiac health in adolescence. *BMC Psychol* 2015 Nov 04;3:39 [FREE Full text] [doi: [10.1186/s40359-015-0096-x](https://doi.org/10.1186/s40359-015-0096-x)] [Medline: [26537175](https://pubmed.ncbi.nlm.nih.gov/26537175/)]
3. Hankin BL, Abramson LY, Moffitt TE, Silva PA, McGee R, Angell KE. Development of depression from preadolescence to young adulthood: emerging gender differences in a 10-year longitudinal study. *J Abnorm Psychol* 1998 Feb;107(1):128-140. [doi: [10.1037//0021-843x.107.1.128](https://doi.org/10.1037//0021-843x.107.1.128)] [Medline: [9505045](https://pubmed.ncbi.nlm.nih.gov/9505045/)]
4. Duggan M, Ellison NB, Lampe C, Lenhart A, Madden M. Social media update 2014. Pew Research Center. 2015. URL: <https://www.pewresearch.org/internet/2015/01/09/social-media-update-2014/> [accessed 2020-01-09]
5. Anderson M, Jiang J. Pew Research Center. 2018. URL: <https://www.pewresearch.org/internet/2018/05/31/teens-social-media-technology-2018/> [accessed 2020-01-09]
6. Moreno MA, Jelenchick LA, Egan KG, Cox E, Young H, Gannon KE, et al. Feeling bad on Facebook: depression disclosures by college students on a social networking site. *Depress Anxiety* 2011 Jun;28(6):447-455 [FREE Full text] [doi: [10.1002/da.20805](https://doi.org/10.1002/da.20805)] [Medline: [21400639](https://pubmed.ncbi.nlm.nih.gov/21400639/)]
7. Moreno MA, Christakis DA, Egan KG, Jelenchick LA, Cox E, Young H, et al. A pilot evaluation of associations between displayed depression references on Facebook and self-reported depression using a clinical scale. *J Behav Health Serv Res* 2012 Jul;39(3):295-304 [FREE Full text] [doi: [10.1007/s11414-011-9258-7](https://doi.org/10.1007/s11414-011-9258-7)] [Medline: [21863354](https://pubmed.ncbi.nlm.nih.gov/21863354/)]
8. Eisenberg D, Golberstein E, Gollust SE. Help-seeking and access to mental health care in a university student population. *Med Care* 2007 Jul;45(7):594-601. [doi: [10.1097/MLR.0b013e31803bb4c1](https://doi.org/10.1097/MLR.0b013e31803bb4c1)] [Medline: [17571007](https://pubmed.ncbi.nlm.nih.gov/17571007/)]
9. Mendoza JA, Zimmerman FJ, Christakis DA. Television viewing, computer use, obesity, and adiposity in US preschool children. *Int J Behav Nutr Phys Act* 2007 Sep 25;4:44 [FREE Full text] [doi: [10.1186/1479-5868-4-44](https://doi.org/10.1186/1479-5868-4-44)] [Medline: [17894878](https://pubmed.ncbi.nlm.nih.gov/17894878/)]
10. Yang J, Yao S, Zhu X, Zhang C, Ling Y, Abela JRZ, et al. The impact of stress on depressive symptoms is moderated by social support in Chinese adolescents with subthreshold depression: a multi-wave longitudinal study. *J Affect Disord* 2010 Dec;127(1-3):113-121. [doi: [10.1016/j.jad.2010.04.023](https://doi.org/10.1016/j.jad.2010.04.023)] [Medline: [20554013](https://pubmed.ncbi.nlm.nih.gov/20554013/)]
11. Rao U. Links between depression and substance abuse in adolescents: neurobiological mechanisms. *Am J Prev Med* 2006 Dec;31(6 Suppl 1):161-174. [doi: [10.1016/j.amepre.2006.07.002](https://doi.org/10.1016/j.amepre.2006.07.002)] [Medline: [17175411](https://pubmed.ncbi.nlm.nih.gov/17175411/)]
12. Garlow SJ, Rosenberg J, Moore JD, Haas AP, Koestner B, Hendin H, et al. Depression, desperation, and suicidal ideation in college students: results from the American Foundation for Suicide Prevention College Screening Project at Emory University. *Depress Anxiety* 2008;25(6):482-488. [doi: [10.1002/da.20321](https://doi.org/10.1002/da.20321)] [Medline: [17559087](https://pubmed.ncbi.nlm.nih.gov/17559087/)]
13. Rao U, Chen L. Characteristics, correlates, and outcomes of childhood and adolescent depressive disorders. *Dialogues Clin Neurosci* 2009;11(1):45-62 [FREE Full text] [Medline: [19432387](https://pubmed.ncbi.nlm.nih.gov/19432387/)]
14. Deas D, Brown ES. Adolescent substance abuse and psychiatric comorbidities. *J Clin Psychiatry* 2006 Jul;67(7):e02. [doi: [10.4088/jcp.0706e02](https://doi.org/10.4088/jcp.0706e02)] [Medline: [17107227](https://pubmed.ncbi.nlm.nih.gov/17107227/)]
15. Kessler RC, Foster CL, Saunders WB, Stang PE. Social consequences of psychiatric disorders, I: educational attainment. *Am J Psychiatry* 1995 Jul;152(7):1026-1032. [doi: [10.1176/ajp.152.7.1026](https://doi.org/10.1176/ajp.152.7.1026)] [Medline: [7793438](https://pubmed.ncbi.nlm.nih.gov/7793438/)]

16. Zender R, Olshansky E. Women's mental health: depression and anxiety. *Nurs Clin North Am* 2009 Sep;44(3):355-364. [doi: [10.1016/j.cnur.2009.06.002](https://doi.org/10.1016/j.cnur.2009.06.002)] [Medline: [19683096](https://pubmed.ncbi.nlm.nih.gov/19683096/)]
17. Kessler RC. Epidemiology of women and depression. *J Affect Disord* 2003 Mar;74(1):5-13. [doi: [10.1016/s0165-0327\(02\)00426-3](https://doi.org/10.1016/s0165-0327(02)00426-3)]
18. Kornstein SG, Schatzberg AF, Thase ME, Yonkers KA, McCullough JP, Keitner GI, et al. Gender differences in chronic major and double depression. *J Affect Disord* 2000 Oct;60(1):1-11. [doi: [10.1016/s0165-0327\(99\)00158-5](https://doi.org/10.1016/s0165-0327(99)00158-5)] [Medline: [10940442](https://pubmed.ncbi.nlm.nih.gov/10940442/)]
19. Albert PR. Why is depression more prevalent in women? *J Psychiatry Neurosci* 2015 Jul;40(4):219-221 [FREE Full text] [doi: [10.1503/jpn.150205](https://doi.org/10.1503/jpn.150205)] [Medline: [26107348](https://pubmed.ncbi.nlm.nih.gov/26107348/)]
20. Salk RH, Hyde JS, Abramson LY. Gender differences in depression in representative national samples: meta-analyses of diagnoses and symptoms. *Psychol Bull* 2017 Aug;143(8):783-822 [FREE Full text] [doi: [10.1037/bul0000102](https://doi.org/10.1037/bul0000102)] [Medline: [28447828](https://pubmed.ncbi.nlm.nih.gov/28447828/)]
21. Brenner J, Smith A. 72% of Online adults are social networking site users. Pew Research Center. 2013. URL: <https://www.pewresearch.org/internet/2013/08/05/72-of-online-adults-are-social-networking-site-users/> [accessed 2020-01-09]
22. Fleming PJ. Software and sympathy; therapeutic interaction with the computer. In: Fish SL, editor. *Talking to Strangers: Mediated Therapeutic Communication*. Norwood, CA: Ablex; 1990.
23. Walther JB, Parks MR. Cues filtered out, cues filtered in: computer mediated communication and relationships. In: Miller GR, editor. *The Handbook of Interpersonal Communication*. Thousand Oaks, CA: SAGE Publications; 2002.
24. Zhao Y, Liu J, Tang J, Zhu Q. Conceptualizing perceived affordances in social media interaction design. *Aslib Proc* 2013 Mar;65(3):289-303. [doi: [10.1108/00012531311330656](https://doi.org/10.1108/00012531311330656)]
25. Choi M, Toma CL. Social sharing through interpersonal media: patterns and effects on emotional well-being. *Comput Hum Behav* 2014 Jul;36:530-541. [doi: [10.1016/j.chb.2014.04.026](https://doi.org/10.1016/j.chb.2014.04.026)]
26. Halpern D, Gibbs J. Social media as a catalyst for online deliberation? Exploring the affordances of Facebook and YouTube for political expression. *Comput Hum Behav* 2013 May;29(3):1159-1168. [doi: [10.1016/j.chb.2012.10.008](https://doi.org/10.1016/j.chb.2012.10.008)]
27. Steinfield C, Ellison NB, Lampe C. Social capital, self-esteem, and use of online social network sites: a longitudinal analysis. *J Appl Dev Psychol* 2008 Nov;29(6):434-445. [doi: [10.1016/j.appdev.2008.07.002](https://doi.org/10.1016/j.appdev.2008.07.002)]
28. Bazarova NN. Public intimacy: disclosure interpretation and social judgments on Facebook. *J Commun* 2012 Aug 27;62(5):815-832. [doi: [10.1111/j.1460-2466.2012.01664.x](https://doi.org/10.1111/j.1460-2466.2012.01664.x)]
29. Lundby K. Patterns of belonging in online/offline interfaces of religion. *Inf Commun Soc* 2011 Dec;14(8):1219-1235. [doi: [10.1080/1369118x.2011.594077](https://doi.org/10.1080/1369118x.2011.594077)]
30. Majchrzak A, Faraj S, Kane GC, Azad B. The contradictory influence of social media affordances on online communal knowledge sharing. *J Comput-Mediat Comm* 2013 Oct 18;19(1):38-55. [doi: [10.1111/jcc4.12030](https://doi.org/10.1111/jcc4.12030)]
31. Moreno MA, Uhls YT. Applying an affordances approach and a developmental lens to approach adolescent social media use. *Digit Health* 2019;5:- [FREE Full text] [doi: [10.1177/2055207619826678](https://doi.org/10.1177/2055207619826678)] [Medline: [30746154](https://pubmed.ncbi.nlm.nih.gov/30746154/)]
32. Christofides E, Muise A, Desmarais S. Information disclosure and control on Facebook: are they two sides of the same coin or two different processes? *Cyberpsychol Behav* 2009 Jun;12(3):341-345. [doi: [10.1089/cpb.2008.0226](https://doi.org/10.1089/cpb.2008.0226)] [Medline: [19250020](https://pubmed.ncbi.nlm.nih.gov/19250020/)]
33. Tran TB, Uebelacker L, Wenze SJ, Collins C, Broughton MK. Adaptive and maladaptive means of using facebook: a qualitative pilot study to inform suggestions for development of a future intervention for depression. *J Psychiatr Pract* 2015 Nov;21(6):458-473 [FREE Full text] [doi: [10.1097/PRA.000000000000109](https://doi.org/10.1097/PRA.000000000000109)] [Medline: [26554330](https://pubmed.ncbi.nlm.nih.gov/26554330/)]
34. Egan KG, Koff RN, Moreno MA. College students' responses to mental health status updates on Facebook. *Issues Ment Health Nurs* 2013 Jan;34(1):46-51 [FREE Full text] [doi: [10.3109/01612840.2012.718044](https://doi.org/10.3109/01612840.2012.718044)] [Medline: [23301569](https://pubmed.ncbi.nlm.nih.gov/23301569/)]
35. Pumper MB, Kelleher E, Whitehill JM, Moreno MA. Identifying depression and suicide displays on Twitter: key search terms and characteristics. *J Psychiatry Behav Sci* 2014:1-7 [FREE Full text] [doi: [10.1037/e488702008-003](https://doi.org/10.1037/e488702008-003)]
36. Moreno MA, Kelleher E, Pumper M. Evaluating displayed depression symptoms on social media sites. *Soc Network J* 2013;02(04):185-192. [doi: [10.4236/sn.2013.24018](https://doi.org/10.4236/sn.2013.24018)]
37. Kelleher E, Moreno M, Wilt MP. Recruitment of participants and delivery of online mental health resources for depressed individuals using tumblr: pilot randomized control trial. *JMIR Res Protoc* 2018 Apr 12;7(4):e95 [FREE Full text] [doi: [10.2196/resprot.9421](https://doi.org/10.2196/resprot.9421)] [Medline: [29650507](https://pubmed.ncbi.nlm.nih.gov/29650507/)]
38. Smith RJ, Crutchley P, Schwartz HA, Ungar L, Shofer F, Padrez KA, et al. Variations in Facebook posting patterns across validated patient health conditions: a prospective cohort study. *J Med Internet Res* 2017 Jan 06;19(1):e7 [FREE Full text] [doi: [10.2196/jmir.6486](https://doi.org/10.2196/jmir.6486)] [Medline: [28062392](https://pubmed.ncbi.nlm.nih.gov/28062392/)]
39. Keles B, McCrae N, Grealish A. A systematic review: the influence of social media on depression, anxiety and psychological distress in adolescents. *Int J Adolesc Youth* 2019 Mar 21;25(1):79-93. [doi: [10.1080/02673843.2019.1590851](https://doi.org/10.1080/02673843.2019.1590851)]
40. Odgers CL, Jensen MR. Annual research review: adolescent mental health in the digital age: facts, fears, and future directions. *J Child Psychol Psychiatry* 2020 Mar;61(3):336-348. [doi: [10.1111/jcpp.13190](https://doi.org/10.1111/jcpp.13190)] [Medline: [31951670](https://pubmed.ncbi.nlm.nih.gov/31951670/)]
41. Eichstaedt JC, Smith RJ, Merchant RM, Ungar LH, Crutchley P, Preotiu-Pietro D, et al. Facebook language predicts depression in medical records. *Proc Natl Acad Sci U S A* 2018 Oct 30;115(44):11203-11208 [FREE Full text] [doi: [10.1073/pnas.1802331115](https://doi.org/10.1073/pnas.1802331115)] [Medline: [30322910](https://pubmed.ncbi.nlm.nih.gov/30322910/)]

42. Keenan K, Hipwell A, Babinski D, Bortner J, Henneberger A, Hinze A, et al. Examining the developmental interface of cortisol and depression symptoms in young adolescent girls. *Psychoneuroendocrinology* 2013 Oct;38(10):2291-2299 [FREE Full text] [doi: [10.1016/j.psyneuen.2013.04.017](https://doi.org/10.1016/j.psyneuen.2013.04.017)] [Medline: [23726646](https://pubmed.ncbi.nlm.nih.gov/23726646/)]
43. Szabó M. The short version of the Depression Anxiety Stress Scales (DASS-21): factor structure in a young adolescent sample. *J Adolesc* 2010 Feb;33(1):1-8. [doi: [10.1016/j.adolescence.2009.05.014](https://doi.org/10.1016/j.adolescence.2009.05.014)] [Medline: [19560196](https://pubmed.ncbi.nlm.nih.gov/19560196/)]
44. Galambos NL, Barker ET, Krahn HJ. Depression, self-esteem, and anger in emerging adulthood: seven-year trajectories. *Dev Psychol* 2006 Mar;42(2):350-365. [doi: [10.1037/0012-1649.42.2.350](https://doi.org/10.1037/0012-1649.42.2.350)] [Medline: [16569173](https://pubmed.ncbi.nlm.nih.gov/16569173/)]
45. Merikangas KR, Zhang H, Avenevoli S, Acharyya S, Neuenschwander M, Angst J, Zurich Cohort Study. Longitudinal trajectories of depression and anxiety in a prospective community study: the Zurich Cohort Study. *Arch Gen Psychiatry* 2003 Oct;60(10):993-1000. [doi: [10.1001/archpsyc.60.9.993](https://doi.org/10.1001/archpsyc.60.9.993)] [Medline: [14557144](https://pubmed.ncbi.nlm.nih.gov/14557144/)]
46. James C, Davis K, Charmaraman L, Konrath S, Slovak P, Weinstein E, et al. Digital life and youth well-being, social connectedness, empathy, and narcissism. *Pediatrics* 2017 Nov 01;140(Supplement 2):71-75. [doi: [10.1542/peds.2016-1758f](https://doi.org/10.1542/peds.2016-1758f)]
47. Gannon KE, Becker T, Moreno MA. Religion and sex among college freshmen. *J Adolesc Res* 2012 Nov 16;28(5):535-556. [doi: [10.1177/0743558412464521](https://doi.org/10.1177/0743558412464521)]
48. Moreno MA, D'Angelo J, Kacvinsky LE, Kerr B, Zhang C, Eickhoff J. Emergence and predictors of alcohol reference displays on Facebook during the first year of college. *Comput Human Behav* 2014 Jan;30:- [FREE Full text] [doi: [10.1016/j.chb.2013.07.060](https://doi.org/10.1016/j.chb.2013.07.060)] [Medline: [24415846](https://pubmed.ncbi.nlm.nih.gov/24415846/)]
49. Egan KG, Moreno MA. Prevalence of stress references on college freshmen Facebook profiles. *Comput Inform Nurs* 2011 Oct;29(10):586-592 [FREE Full text] [doi: [10.1097/NCN.0b013e3182160663](https://doi.org/10.1097/NCN.0b013e3182160663)] [Medline: [21436681](https://pubmed.ncbi.nlm.nih.gov/21436681/)]
50. Adrian M, Miller AB, McCauley E, Vander Stoep A. Suicidal ideation in early to middle adolescence: sex-specific trajectories and predictors. *J Child Psychol Psychiatry* 2016 May;57(5):645-653 [FREE Full text] [doi: [10.1111/jcpp.12484](https://doi.org/10.1111/jcpp.12484)] [Medline: [26610726](https://pubmed.ncbi.nlm.nih.gov/26610726/)]
51. Rhew IC, Fleming CB, Vander Stoep A, Nicodimos S, Zheng C, McCauley E. Examination of cumulative effects of early adolescent depression on cannabis and alcohol use disorder in late adolescence in a community-based cohort. *Addiction* 2017 Nov;112(11):1952-1960 [FREE Full text] [doi: [10.1111/add.13907](https://doi.org/10.1111/add.13907)] [Medline: [28600897](https://pubmed.ncbi.nlm.nih.gov/28600897/)]
52. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders (DSM-5®)*, Fifth Edition. Washington, D.C: American Psychiatric Association; 2013:1-991.
53. Gritton J, Rushing SC, Stephens D, Ghost Dog T, Kerr B, Moreno MA. Responding to concerning posts on social media: insights and solutions from American Indian and Alaska Native youth. *Am Indian Alsk Native Ment Health Res* 2017;24(3):63-87. [doi: [10.5820/aian.2403.2017.63](https://doi.org/10.5820/aian.2403.2017.63)] [Medline: [29161455](https://pubmed.ncbi.nlm.nih.gov/29161455/)]
54. Wallace P, Linke S, Murray E, McCambridge J, Thompson S. A randomized controlled trial of an interactive web-based intervention for reducing alcohol consumption. *J Telemed Telecare* 2006 Dec 02;12 Suppl 1(1_suppl):52-54. [doi: [10.1258/13576330677978452](https://doi.org/10.1258/13576330677978452)] [Medline: [16884581](https://pubmed.ncbi.nlm.nih.gov/16884581/)]
55. Walther JB, Parks MR. Cues filtered in, cues filtered out: computer-mediated communication relationships. In: *The Handbook of Communication Science*. Thousand Oaks, California, United States: SAGE Publications; 2009:529-563.
56. Martin LA, Neighbors HW, Griffith DM. The experience of symptoms of depression in men vs women: analysis of the National Comorbidity Survey Replication. *JAMA Psychiatry* 2013 Oct;70(10):1100-1106. [doi: [10.1001/jamapsychiatry.2013.1985](https://doi.org/10.1001/jamapsychiatry.2013.1985)] [Medline: [23986338](https://pubmed.ncbi.nlm.nih.gov/23986338/)]
57. Wongkoblap A, Vadillo MA, Curcin V. Researching mental health disorders in the era of social media: systematic review. *J Med Internet Res* 2017 Jun 29;19(6):e228 [FREE Full text] [doi: [10.2196/jmir.7215](https://doi.org/10.2196/jmir.7215)] [Medline: [28663166](https://pubmed.ncbi.nlm.nih.gov/28663166/)]

Abbreviations

DPP: Developmental Pathways Project

DSD: depression symptom displayer

DSM-IV: Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition

MDE: major depressive episode

SNS: social networking site

Edited by G Eysenbach; submitted 15.05.20; peer-reviewed by F Palmieri, J Li; comments to author 16.07.20; revised version received 29.09.20; accepted 16.03.21; published 31.05.21.

Please cite as:

Moreno MA, Gaus Q, Wilt M, Arseniev-Koehler A, Ton A, Adrian M, VanderStoep A

Displayed Depression Symptoms on Facebook at Two Time Points: Content Analysis

JMIR Form Res 2021;5(5):e20179

URL: <https://formative.jmir.org/2021/5/e20179>

doi: [10.2196/20179](https://doi.org/10.2196/20179)

PMID: [34057422](https://pubmed.ncbi.nlm.nih.gov/34057422/)

©Megan A Moreno, Quintin Gaus, Megan Wilt, Alina Arseniev-Koehler, Adrienne Ton, Molly Adrian, Ann VanderStoep. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 31.05.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Using Social Media for Qualitative Health Research in Danish Women of Reproductive Age: Online Focus Group Study on Facebook

Camilla Gry Temmesen^{1,2,3,4*}, MSc; Henriette Svarre Nielsen^{5,6*}, MD, DMSc; Heidi Lene Myglegård Andersen^{3*}, MA, MPP, PhD; Kathrine Birch Petersen^{7*}, MD, MGP, PhD; Jane Clemensen^{1,2,4*}, RN, MScN, PhD

¹HCA Research, Hans Christian Andersen Children's Hospital, Odense University Hospital, Odense, Denmark

²Centre for Innovative Medical Technology, Odense University Hospital, Odense, Denmark

³Department of Nursing, University College Absalon, Roskilde, Denmark

⁴Institute of Clinical Research, Faculty of Health Sciences, University of Southern Denmark, Odense, Denmark

⁵Department of Obstetrics and Gynecology, Copenhagen University Hospital, Hvidovre, Denmark

⁶Department of Clinical Medicine, University of Copenhagen, Copenhagen, Denmark

⁷TFP Stork Fertility, The Fertility Partnership DK, Copenhagen, Denmark

* all authors contributed equally

Corresponding Author:

Camilla Gry Temmesen, MSc

Institute of Clinical Research, Faculty of Health Sciences

University of Southern Denmark

Winsløwparken 19, 3

Odense, 5000

Denmark

Phone: 45 26215135

Email: ctemmesen@health.sdu.dk

Abstract

Background: Social media platforms provide new possibilities within health research. With Facebook being the largest social network in the world, it constitutes a potential platform for recruitment and data collection from women of reproductive age. Women in Denmark and in other Western countries postpone motherhood and risk infertility due to their advanced age when they try to conceive. To date, no study has explored Danish women's reflections on the timing of motherhood within a social media setting.

Objective: The aim of this study was to explore the challenges and opportunities of using Facebook as a platform for qualitative health research in Danish women of reproductive age.

Methods: This study was a qualitative study based on 3 online focus groups on Facebook with 26 Danish women of reproductive age discussing the timing of motherhood in January 2020.

Results: Conducting online focus groups on Facebook was successful in this study as the web-based approach was found suitable for developing qualitative data with women of reproductive age and made recruitment easy and free of charge. All participants found participating in an online focus group to be a positive experience. More than half of the women participating in the online focus groups found it advantageous to meet on Facebook instead of meeting face-to-face.

Conclusions: Conducting online focus groups on Facebook is a suitable method to access qualitative data from women of reproductive age. Participants were positive toward being a part of an online focus group. Online focus groups on social media have the potential to give women of reproductive age a voice in the debate of motherhood.

(*JMIR Form Res* 2021;5(5):e24108) doi:[10.2196/24108](https://doi.org/10.2196/24108)

KEYWORDS

internet; social media; Facebook; online focus groups; women; reproduction; reproductive age; motherhood; participatory design

Introduction

Background of This Study

Within the last decade, social media platforms have gained a prominent role in the way people communicate and interact with one another and have enabled people to be connected and accessible to a greater extent. Social media offer multiple possibilities such as chat forums, blogs, virtual worlds, and social networks such as Facebook, Twitter, Instagram, and SnapChat. Facebook is currently the largest and most popular social network in the world with an average of 1.73 billion daily active users and an average of 2.6 billion monthly active users worldwide, as reported in 2020 [1]. In Denmark too, Facebook constitutes the most popular social media with 77% of all Danes having a Facebook profile [2]. In 2018, 95.3% of Danes aged 19-34 years had a Facebook profile followed by 86% of Danes aged 35-54 years [2], with 72% using Facebook daily [3]. Social media networks are popular among women of reproductive age. On Facebook, women can engage in different web-based groups, such as pregnancy, due date, and baby groups with the possibility of sharing experiences, photos, life stories, and to be a part of a virtual community with other future parents. However, web-based groups for women who have not yet had children are not common neither are the possibilities for discussing the future with or without children online with other women of reproductive age. Within health care, social media has increasingly become a way of informing, communicating, interacting, and engaging with people. Previous studies have explored the use of social media within health research, for example, a study that explores mommy bloggers disseminating breast cancer risk information [4], a study that uses Facebook to deliver smoking cessation treatment to young adults [5], or a study using online support groups for women with endometriosis [6].

Web-Based Recruitment and Web-Based Data Collection

Within qualitative research, focus groups are a popular data collection method. However, this method is often associated with challenges to recruit enough eligible participants, and inclusion of participants across the country can be time-consuming and economically and practically challenging [7]. Web-based recruitment is a well-known strategy in health research, especially pointed at “hard-to-reach” populations [8] such as marginalized groups or people who do not have access to the internet. Choosing traditional recruitment approaches such as posters, brochures, or personal approaching in health clinics when recruiting women of reproductive age who have not yet had children can be challenging, as most of these women do not have the need for consulting health professionals or have a natural web-based community. When planning this study, Facebook seemed to have the potential for recruitment and collection of data, since Danish women in the age group of 18-45 years constitute the most active users on Facebook with 86.5% using Facebook several times a day or almost every day compared to 77.9% of men of the same age group (Index Danmark, Gallup, Social media: use of Facebook by age and sex, unpublished data, 2019). Women tend to engage on

Facebook by liking and sharing content [9], which can contribute to a faster recruitment process. Online focus groups, also referred to as virtual focus groups, are known from social science research, but online focus groups have become an interesting methodological approach within health science too. Williams et al [10] defines online focus groups as “A selected group of individuals who have volunteered to participate in a moderated, structured online discussion in order to explore a particular topic for the purpose of research.” Previously, online focus groups have been used in a study exploring DNA paternity testing with single mothers with young babies, for whom it was difficult to participate in face-to-face focus groups [11,12] and in another study as an approach to explore breastfeeding women’s use of social media [13].

Timing of Motherhood

Danish women’s age at the birth of their first child has risen significantly throughout the last 50 years. In 1969, Danish women were 23.3 years old when giving birth to their first child, which has risen to 29.5 years of age in 2019 [14]. Postponing motherhood increases the risk of infertility and pregnancy complications such as miscarriage [15,16], chromosomal abnormalities [15,17], hypertensive complications, fetal growth restrictions, fetal death [18], and birth complications such as cesarean section [19]. The higher the age also means an increased risk of other risk factors, including overweight, sexually transmitted diseases, cancer, endometriosis, or environmental exposures, for example, hormone disruptive chemicals that may affect female fertility [15]. Postponement of motherhood is not unique to Denmark but has been reported in other Western and Northern European countries too [20-24].

Aim and Purpose of This Study

This paper presents an exploration of a qualitative methodological approach using Facebook to collect data from women of reproductive age, discussing their thoughts on the timing of motherhood. The aim of this study was to explore Facebook as a platform for qualitative health research within the reproductive field by recruiting women who had not yet had children into an online focus group discussing the timing of motherhood as a way to explore the challenges and opportunities of using Facebook to collect qualitative data. The online focus group was a part of a Participatory Design research-inspired study, which included a literature study and individual interviews with women of reproductive age. To the best of our knowledge, no previous research has addressed the approach of conducting online focus groups with a group of women of reproductive age in a social media setting. This paper focuses solely upon the methodological approaches of conducting online focus groups. In-depth empirical data from the online focus groups will be presented elsewhere.

Methods

Study Design

This study was conducted as an asynchronous text-based online focus group study on Facebook. In previous literature, focus groups within computer-mediated settings were also known as virtual focus groups. In this study, we prefer using the term

“online focus groups” describing focus group activities connected to the internet [25], as we see the term “virtual” referring to a 3D image or environment that can be interacted with in a seemingly real or physical way by a person using special electronic equipment, as known from virtual reality [26]. Facebook was chosen as a web-based platform for the online focus groups because of its popularity, user friendliness, and the possibility to create private online groups. Prior to this study, a pilot study in 2016 with 14 Danish women tested Facebook as a platform for discussing the timing of motherhood in a social media setting, and Facebook was then found to be a suitable platform for this target group [27].

Participant Recruitment

We used a purposive recruitment strategy where participants were invited to the online focus groups if they met the following inclusion criteria: women of reproductive age defined as 18-45 years, with no children, regardless of whether they want to have children in the future, single or in a relationship, capable of

speaking and writing Danish, and having a personal Facebook profile. Participants were recruited using the snowball sampling method on Facebook in the beginning of January 2020, initially through a recruitment post shared within the main author’s social network. The recruitment post, containing information about the project, was shared 287 times, reaching potential participants in all 5 regions of Denmark. None of the participants knew the facilitator prior to participation but were recruited through mutual connections within the social network. Participants who fulfilled the inclusion criteria contacted the facilitator (main author) by mail or *Messenger* (Textbox 1) [28] and answered a short web-based demographic questionnaire with a written consent to participate. We aimed to have 8-10 participants for each group, which was achieved for 2 out of 3 groups. Recruitment lasted for 20 days and finished when the desired number of participants were reached. Two potential participants withdrew from the study without explanation prior to data collection.

Textbox 1. Social media dictionary.

Post: A piece of writing, image, or other item of content published online, typically on a blog or social media website.

Tagging: Linking a person to a comment or a post by adding his/her profile name.

Notifications: A message, email, icon or another symbol that appears when an app wants you to pay attention.

Emoji: A small digital image or icon used to express emotions.

Reaction symbols: A series of 6 animated “emoji” reactions that customers can add when responding to a post. Reaction symbols are Facebook’s way of facilitating an emotional conversation online.

Messenger: Messenger is a free mobile messaging app developed by Facebook, used for instant messaging, sharing photos, videos, audio recordings, and group chats.

Participant Characteristics

A total of 26 women agreed to participate in the online focus groups lasting for 4 days (January 27-30, 2020). Participants were sent a personal invite linked to their private Facebook profiles to join a private Facebook group that was only visible to the facilitator and the participants. Participants were divided into 3 age groups with a private Facebook group created for each group. The participants were divided into different age groups with the intention to obtain insight into the potential similarities or differences among (1) younger women aged 18-24 years (n=8), (2) women aged 25-34 years (n=11), who according to their age are most likely to consider having children, and (3) women aged 35-45 years (n=7), who represent being at advanced maternal age defined as 35 years and older.

Online Focus Groups

The online focus groups were initiated by an elaboration of how the focus groups would proceed. As recommended by Abrams and Gaiser [29], participants were encouraged to make a short presentation of themselves as a way of enabling trust and making them feel comfortable. Since the participants used their personal

Facebook profile, their identities were inevitably visible to other participants in the online focus groups, which is why we encouraged confidentiality between the participants. Participants were briefed that they could expect a daily topic to be uploaded and discussed within the group and that the estimated time each participant was expected to spend in the group was approximately 15 minutes per day. An interview guide inspired by traditional focus groups was prepared, as recommended by Malterud [30] and Skelton et al [13] in their study of breastfeeding mothers’ use of social media. The interview guide focused on 4 broad main themes: (1) *considerations regarding when to have children*, (2) *the influence of others*, (3) *age and women’s fertility*, and (4) *fertility knowledge*. Each theme represented a daily topic and was visually presented with a banner (photo or illustration) to make it easier to navigate between the various discussion threads in the group. The facilitator uploaded a new topic with support questions to be discussed each of the 4 days the online focus groups lasted, which addressed the 4 main themes of the interview guide. A range of open-ended support questions was designed to facilitate reflections on the topic and to promote a dialogue between the participants (Textbox 2).

Textbox 2. An example of a topic for discussion in the online focus groups on the second day.

Today I would ask you to discuss the following issue: Do you feel that the outside world influences your thoughts about having children?

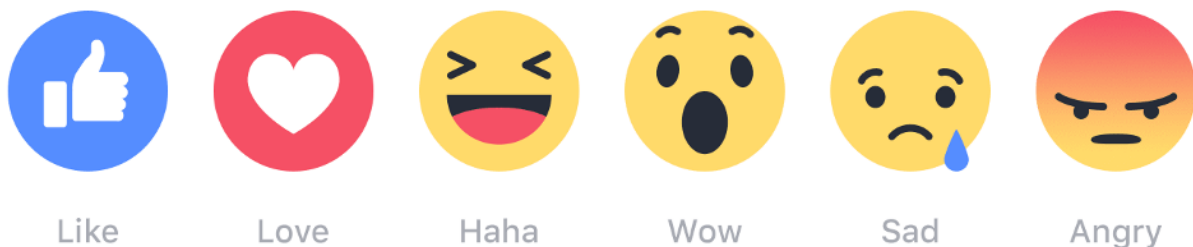
Supporting questions:

- Feel free to define what you mean by the outside world (is it your possible partner, family, parents, siblings, grandparents), friends/girlfriends, politicians, media, community or otherwise?)
- Do you find that the outside world positively affects you to have children?
- Do you find that the outside world negatively affects you to have children?
- Does that make you want to have children earlier or later?
- Have you told your outside world about your considerations or choices regarding your thoughts on having children?
- Who are your loved ones that you talk to about your thoughts on having children? (eg. partner, parent, sibling, friends/girlfriends, healthcare professionals or others?)
- Is it hard to talk to anyone else about these thoughts? Why/why not?

Apart from text-written communication, participants had the possibility of using Facebook's *reaction symbols* [31] (Figure 1) [32] to acknowledge comments or posts from coparticipants. Reaction symbols are Facebook's way of facilitating an emotional conversation online [31,33]. The online focus groups were asynchronous, meaning that participants could access the online focus groups at a time that suited them best during the day rather than real-time online focus groups where participants access the online focus groups at a predefined time [34]. Three days after the online focus groups were completed, participants received a web-based debriefing questionnaire to anonymously evaluate how they experienced participating in an online focus

group. A debriefing can give the researcher insight into group dynamics and how the participants experienced being a part of the focus group [35,36]. A total of 464 comments or 97 full pages of transcripts were generated from the 3 online focus groups and were copied as written by the participants into a text-based document, including other forms of *interactions*, for example, uploaded photos or videos, *emojis* [37], and punctuation. Identifiable profile names were replaced with pseudonyms and profile pictures were excluded from the analysis. As the native language of the participants was Danish, data were translated to English after analysis to reduce the risk of important pointers being lost in translation.

Figure 1. Facebook reaction symbols.



Ethical Considerations

This study was notified to the Regional Ethical Committee as well as the Danish Data Protection Agency, but in Denmark, interview studies no longer require approval [38]. This study followed the principles of the Helsinki Declaration [39] and The Danish Code of Conduct for Research Integrity [40], in which participants received written information with the purpose of the study, just as they were informed about the confidentiality and the right to withdraw from the study. As Facebook users, all participants had accepted the Facebook Terms of Service and Facebook Data Policy prior to participation in this study but were reminded of these policies when invited to the study. The online focus groups were created as private Facebook groups, meaning that all content within the groups was only visible to the facilitator and the participants. Only the facilitator could invite participants to the group. All groups were deleted from Facebook at 1 month after participation.

Results

Participant Characteristics

The age range of the participants was 18-43 years, which was considered as a representative age span for women of reproductive age. Approximately 35% (9/26) of the women were single and 65% (17/26) were in a relationship, of which 9% (2/26) were married, 39% (10/26) were in a relationship and cohabiting, and 19% (5/26) were in a relationship but not cohabiting. All 5 regions of Denmark were represented with 58% (15/26) of the women living in urban areas and 42% (11/26) living in provincial areas. The majority of the women (21/26, 81%) had a high educational level with 42% (11/26) having a medium higher education or 39% (10/26) having a higher education. One participant discovered she was pregnant during the data collection but was not excluded from the group since she was at a very early stage of pregnancy. These demographical data are shown in Table 1.

Table 1. Demographics of the women of reproductive age included in the online focus groups in this study.

Participant ID	Pseudonym	Age (years)	Region of living	Area of living	Educational level	Relationship status
W1	Anna	39	Southern Denmark	Rural	Medium higher education (3-4.5 years)	Single
W2	Amelia	31	Southern Denmark	Urban	Medium higher education (3-4.5 years)	Single
W3	Ava	31	Zealand	Rural	Medium higher education (3-4.5 years)	Single
W4	Beatrice	32	Capital	Urban	Higher education (5-6 years)	Single
W5	Catherine	30	Northern Denmark	Rural	Higher education (5-6 years)	In a relationship, cohabiting
W6	Charlotte	28	Capital	Rural	Higher education (5-6 years)	In a relationship, not cohabiting
W7	Emily	43	Capital	Urban	Medium higher education (3-4.5 years)	Single
W8	Hannah	32	Capital	Urban	Medium higher education (3-4.5 years)	In a relationship, cohabiting
W9	Helen	35	Zealand	Rural	Higher education (5-6 years)	In a relationship, not cohabiting
W10	Julia	23	Capital	Urban	Medium higher education (3-4.5 years)	In a relationship, not cohabiting
W11	Josephine	32	Capital	Urban	Medium higher education (3-4.5 years)	In a relationship, not cohabiting
W12	Lauren	38	Central Denmark	Urban	Higher education (5-6 years)	Single
W13	Lily	27	Capital	Urban	Higher education (5-6 years)	In a relationship, cohabiting
W14	Megan	23	Zealand	Rural	Medium higher education (3-4.5 years)	In a relationship, cohabiting
W15	Mia	25	Zealand	Rural	Medium higher education (3-4.5 years)	In a relationship, cohabiting
W16	Olivia	35	Capital	Rural	Shorter higher education (2-3 years)	In a relationship, cohabiting
W17	Sarah	25	Central Denmark	Urban	Other	Single
W18	Sophie	33	Southern Denmark	Rural	Higher education (5-6 years)	In a relationship, cohabiting
W19	Susan	19	Zealand	Rural	Primary school	In a relationship, cohabiting
W20	Tessa	29	Southern Denmark	Urban	Medium higher education (3-4.5 years)	Single
W21	Tori	36	Southern Denmark	Urban	Higher education (5-6 years)	Single
W22	Tracy	38	Southern Denmark	Urban	Higher education (5-6 years)	In a relationship, cohabiting
W23	Vanessa	26	Capital	Urban	Higher education (5-6 years)	Married
W24	Vera	18	Southern Denmark	Urban	High school	In a relationship, not cohabiting
W25	Veronica	24	Capital	Urban	Medium higher education (3-4.5 years)	Married

Participant ID	Pseudonym	Age (years)	Region of living	Area of living	Educational level	Relationship status
W26	Victoria	20	Capital	Rural	High school	In a relationship, cohabiting

Participant Interactions

Facebook provided a *notification* every time a new comment or a reaction was added in the group, making it possible for the participants and the facilitator to keep track on the activity in the groups. Where the majority of the participants appeared to be the most active within the groups during day and evening time, often being active online at the same time, which at times made the online focus groups synchronous, a few participants were more likely to contribute with their comments late at night. All posts uploaded by the facilitator were seen by all the participants, but not every participant contributed to the discussion. Facebook offers a feature that indicates whether the participants have seen the posts or not, which make it visible to the facilitator if some participants were not active in the

group. Participants were considered as low responders if they contributed less than 5 comments or interactions (n=3), medium responders if they contributed 5-15 comments or interactions (n=15), and as high responders if they contributed 15 comments or interactions and more (n=8) (Table 2). The number of comments ranged from 2 comments or interactions from 1 participant to 45 comments or interactions from another participant. None of the participants left the online groups or withdrew their wish to participate in the study during data collection. Overall, there was a high engagement level in all 3 online focus groups fostering rich discussions on the topics with an average of 142 comments per group (Table 2). Participants naturally applied the different emojis and reaction symbols as a way of interacting nonverbally with each other (Textbox 3).

Table 2. Interaction among the women of reproductive age in online focus groups categorized by age and type of responder.

Responses	18-25 years age group (n=8)	26-34 years age group (n=11)	35-45 years age group (n=7)
Low responders ^a (n=3)	1	1	1
Medium responders ^b (n=15)	7	5	3
High responders ^c (n=8)	0	5	3
Comments per group	75	195	157
Comments per person (median)	6.5	15	15

^a<5 comments.

^b5-15 comments.

^c>15 comments.

In all groups, there was a friendly informal tone among the participants. Between the age groups, there were some differences in how often and in which way participants interacted with each other. While participants in age group 26-34 years and age group 35-45 years often addressed specific questions or comments to other participants by *tagging* [41] (Textbox 3), participants in the age group of 18-25 years

primarily answered the questions from their own point of view and did not interact as much with each other as the other groups did, which led to markedly fewer comments and less lively discussions in this age group. Two participants did not contribute to the discussions; they only contributed with a presentation of themselves, thereby being merely passive participants.

Textbox 3. Examples of text interactions among Tori, Anna, and Lauren (age group 35-45 years) in the online focus group on Facebook.

Tori (W21): Anna, *I always say that there will be some divorced men with cute children, where you can become a bonus mother when the marital crises hit* [👍]. *Then you can experience couple relationships and the mother role too, if you have the needs. Then they never quite know what to answer and the conversation changes topic* [👍]

Anna (W1): Tori, *hey good tactics! I like it* [👍]

Lauren (W12): *Yes, I have also met people who say, "you will regret" or "in 5 years we'll see you with a baby on the arm". Very condescending, I think...*

Facilitator Role

The facilitator applied a low management role to the discussions, which, apart from addressing the daily topic and supporting questions consisted of having a more observant role in the

discussions. However, the facilitator ongoingly addressed specific questions either as general questions within the groups if the interview topics did not appear naturally within the discussions among the participants or asked individual questions to get a participant to elaborate on her answer (Textbox 4).

Textbox 4. An example of a communication thread between the facilitator and a participant.

Susan (W19): (...) *I think everyone has an opinion on when one should or should not have children. Whether it is one or the other they think, then it is completely natural for me to listen to the arguments they come up with. I personally want to be a young mother, and there are many who have an attitude to it. And I can be negatively affected in relation to my attitude to have children at a young age, as there are so many who do not think it is a good idea - and since it is one's primary circle, it's hard not to be influenced by their opinion (...)*

Facilitator: Susan, *can you say more about how you experience the negative attitude towards your desire to become a mother at a young age?*

Susan (W19): Facilitator, *I often think that people rule out the possibility that you can have experiences after having children (...)*

Responses in the Debriefing Questionnaire

Out of 26 women, 89% (23/26) answered the debriefing questionnaire. When asked why they decided to participate in the study, 65% (15/23) of the women stated that they felt obligated to participate in the research aiming at them, 61% (14/23) of the women participated because they found it exciting to be a part of an online focus group, and 48% (11/23) of the women participated in the study because the topic motherhood was found to be interesting (multiple answers were possible). When asked on a 5-point Likert Scale (very high degree, some degree, low degree, very low degree, don't know), all participants agreed that participating in the online focus groups was a positive experience, with 70% (16/23) agreeing to a high degree and 30% (7/23) agreeing to some degree. Overall, 61% (14/23) of the women felt that participating in an online focus group made them feel as a part of a community with 4% (1/23) who felt to a high degree and 57% (13/23) who felt to some degree that they were a part of a community, whereas 26% (6/23) felt to a low degree and 4% (1/23) felt to a very low degree that they were a part of a community. Approximately 17% (4/23) of the participants felt to a high degree and 43% (10/23) felt to some degree that it was an advantage to meet on Facebook instead of meeting face-to-face, whereas 13% (3/23) of the participants felt to a low degree and 9% (2/23) of the participants felt to a very low degree that it was advantageous to meet on Facebook. One participant stated that "I think that it would have supported the internal dialogue between participants having been physically together, since communication is so much more than words (...) However, I probably wouldn't have had the opportunity to attend if it required physical attendance, so I think this is a great solution!" The majority of the women spent more time than the estimated 15 minutes per day participating in online focus groups, with 35% (8/23) of the participants spending 5-15 minutes daily, 39% (9/23) of the participants spending more than 15 minutes, and 26% (6/23) of the participants spending more than 30 minutes daily. One participant claimed that "It was a little more time-consuming than I expected, and because I was lacking time, I attended less than I had hoped for."

Discussion

Overview of This Study

The aim of this study was to explore the challenges and opportunities of using Facebook as a platform for qualitative health research in women of reproductive age. This study represents 26 Danish women from 3 different age groups reflecting their thoughts upon the timing of motherhood. Rather than aiming to conduct more online focus groups, we decided

to include participants who were different in terms of age, relationship status, and area of living (urban or rural), and from the 3 online focus groups, we obtained rich data on reflections on the timing of motherhood.

Facebook as a Platform

Using Facebook as a platform for conducting online focus groups with women of reproductive age was beneficial for this study, as it made recruitment and data collection free of charge and inclusion of women across the country without the need for covering travelling and venue expenses. Another benefit was that through Facebook, we gained access to participants and data by creating a temporary online space for women of reproductive age, which could otherwise have been difficult, as this group of women do not have a natural web-based meeting space and can be hard to reach using traditional recruitment methods. It was an advantage that the recruitment as well as the data collection took place on the same platform, since Facebook has user-friendly features and was well-known to the participants. When conducting online focus groups with women of reproductive age on Facebook, we met participants on a media platform that they were already familiar with [42], which makes recruitment and adaptation to the research tool easier. Transcriptions were easy, as participants generated the transcripts themselves and the text-based data were copy-pasted for analysis, which increases the accuracy of the transcripts and removes the potential for error [10,43]. However, the anonymization of the participants and the use of emojis, reaction symbols, punctuation, photos, and videos uploaded by participants required some preparation of data prior to analysis.

Asynchronous Versus Real-Time Online Focus Groups

The asynchronous interaction in the online focus groups meant that participants had the freedom to participate at a time that suited them best, thereby enhancing flexibility and convenience. Another advantage was that participants had the ability to return to a previously asked question and elaborate on their answers or contribute with their points of view later on if they did not have the opportunity to answer immediately. This gave participants more time for reflection before submitting a response, thereby enabling richer collection of data. This finding is supported by Lijadi and van Schalkwyk [34] in their study of Third Culture Kids. A disadvantage of asynchronous online focus groups was that when a topic was being discussed, participants had to wait for a response from the other participants before the discussion could start properly. Compared to real-time chat-based interviews or traditional focus groups, this asynchronous approach may cause participants not to respond as spontaneously because they consider their answers [36]. The asynchronous approach required that participants checked updates in the Facebook group several times a day to keep up

with the discussions, with the risk of overlooking some comments. The danger of Facebook being a platform that can be accessed anytime and anywhere is that participants can easily be distracted and forget to return to a discussion thread. However, Facebook has a helpful tool in the form of *notifications* [44] that reminded participants about new comments or posts in the group. Some participants were online simultaneously, which from time to time unintentionally caused the online focus groups to have the character of being synchronous, because participants responded and interacted in real time. Being online at the same time made the discussions lively and caused participants to interact more. However, in-depth comments often required participants to reflect upon their answers, which made the asynchronous approach more suitable, but a combination of both synchronous and asynchronous interactions seems to be ideal. This finding is supported by Graffigna and Bosio [45] in their study of young people and HIV/AIDS, who state that a combination of synchronous and asynchronous approaches to online-focused discussions maximizes the richness of data collected by fostering both immediate interaction and considered responses. However, Bloor et al [46] state that when doing synchronous online focus groups, there is a risk of the discussions going too fast, which requires greater attention from participants and the facilitator, when multiple discussion threads need to be controlled simultaneously. One participant elaborated on participating in an online focus group on Facebook compared to a face-to-face focus group and stated that she would probably not have participated if she were to meet physically. This illustrates that some participants who would otherwise have declined to participate in a face-to-face focus group owing to time restraints agree to participate when the research takes place online. Lijadi and van Schalkwyk [34] suggested that online focus group participants tend to contribute with shorter comments than participants in face-to-face focus groups. This contradicts with our findings since the majority of the comments in this study tended to be relatively long and reflective. In our experience, being a part of an online focus group makes it possible for participants to reflect upon thoughts and elaborate on them, which in our study contributed to richness in data. Williams et al [10] argue that written language can enhance the communication for people who feel more comfortable in expressing their feelings in text with the ability to express rich feelings and detailed reflections rather than expressing themselves face-to-face. In their study of online focus group discussions within the setting of pediatric oncology, Tates et al [43] argued that owing to the nature of being online without the physical presence, online focus groups can increase self-disclosure and lead to a higher level of interaction among participants.

Facilitator Role

When doing research on social media, as a researcher, you need to adapt to the platform being used by being familiar with the formal codes of conduct (in this case the Facebook Policy) and the informal codes of conduct (netiquette, ie, a set of rules for acceptable online behavior) and know the jargons of the web-based platform to help participants navigate on the platform and to appear as a natural member of the group. Halkier [36]

emphasizes that when planning online focus groups, it is important to consider the level of management needed. Online discussions can easily get off-track, which is why we considered it necessary to define a daily topic for participants to discuss, thereby incorporating it in the interview guide, supported by a number of support questions and associated banners to visually support the topic. Having too strictly formulated questions can cause participants to only respond to what they are asked, which will cause the discussions to be less nuanced [36]. In this study, this did not seem to be the case, because when participants occasionally moved away from the original question, new and interesting angles emerged, which gave rise to even richer discussions. Where moderate management was applied to the questions asked, the management of the discussions themselves was very limited, in order to allow the participants to reflect on the topic in question as naturally as possible and to reduce potential interviewer effect. This low form of managing the discussions was a deliberate choice to get participants to interact with each other rather than with the facilitator. In the study of Bloor et al [46], this finding is supported by Murray [47], one of the pioneers within online focus groups, who found that too high level of interactions from the facilitator can lead to participants answering the facilitator rather than stimulate discussions among participants [46].

Participant Interactions

In traditional face-to-face focus groups, various techniques can be used to support group processes, promote participant interaction, and focus the discussion of the topic (eg, so-called prompts, which are actions that help to make the conversation flow; in the form of affirmative body language such as nod, mimic, and encouraging expressions or probes that encourage participants to elaborate on their answers such as “And what happened then?” or “Tell me more about this”) [48]. In the online focus groups, where the use of body language was excluded, we adapted traditional interviewing techniques to the web-based universe by using Facebook's *reaction symbols* [31] to nonverbally acknowledge a comment. As in traditional focus groups, online focus groups may also face challenges with some participants retaining and not interacting naturally with the other participants. This was also the case in this study, where some participants or groups did not participate as actively as others who were very active in the discussions and interacted mutually with each other. Surprisingly, in this study, the youngest group (age group of 18-25 years) had remarkably less interactions than the other groups. A cautious suggestion can be that this age group felt it harder to relate to the topic of timing of motherhood, as they are not at a time in their lives where motherhood seems as relevant yet, as it can be in the other age groups. However, more research is needed to explore how younger women reflect upon motherhood.

In this study, we discovered different behavioral patterns among the participants. The majority of the participants responded to all the questions uploaded by the facilitator and interacted vividly with one another by using text, photos, videos, reaction symbols, and emojis to comment on each other's posts. Often, an individual opinion was developed through a discussion with other participants, which in a constructivist perspective can be seen as data are being socially constructed among participants

and framed during the interpersonal exchange of opinions and different perspectives [45]. In addition, some participants made updates on other thoughts of motherhood at their own initiatives, which emphasizes that there was a great deal of interest in discussing the topic of motherhood while providing good group dynamics and demanding discussions within the group, without getting too far off the original topic. Clemensen et al [49] define a participant who starts a communication thread as an updater, and in this study, we identified a couple of updaters within each age group, who helped stimulating group discussions and fostered interaction with less active participants. In contrast, we also identified a few passive participants, whom according to Bloor et al [46] could be considered as so-called lurkers who only contributed with a presentation of themselves on the first day and did not contribute to the discussions at all. Remarkably, they read all the posts uploaded by the facilitator, were online daily, responded to the debriefing questionnaire, and did not withdraw from the study. This indicates that even though passive participants do not necessarily interact or contribute with comments, they can still have a sense of participation, which shows how participation can take place at different levels. Although Bloor et al [46] implied that it is easier for the researcher to encourage participation from those who read, but not respond, when doing asynchronous online focus groups, we found it difficult to point out the low responders, as it was hard to figure out whether the low responders were simply not online at the same time as other participants or whether they did not want to comment. It is the facilitator's role to try to involve these lesser active participants, but in this study, we found it problematic to address questions directly to a specific participant without the participant feeling designated, with the risk that some participants were hiding and just reading the comments of others without participating in the discussions themselves.

Use of Emojis and Reaction Symbols

In the written language, there is a lack of nonverbal expressions such as facial expressions, body language, and emphasis on words. When participants communicated with each other in the online focus groups, emojis or reaction symbols were used several times, typically when participants needed to denote a feeling, express humor or irony, or to emphasize a point. The use of emojis and reaction symbols are examples of how verbal communication in a face-to-face focus group can be translated into web-based language, showing how it is possible to express nonverbal communication without the physical presence. Since all participants in this study were familiar with Facebook, the use of emojis and reaction symbols appeared natural and did not cause obvious misunderstandings between the participants. The use of emojis, reaction symbols, and punctuation was incorporated in transcripts and contributed to richness in data and gave a greater understanding for group dynamics, for example, the use of abbreviations (eg, OMG for Oh my God) to express surprise or dismay, or punctuation (eg, '!!!'), when something needed to be emphasized.

When using reaction symbols (eg, interacting with a heart), it seemed as if participants were emotionally closer than what online media usually foster. For example, when Catherine discloses her family history with mental problems, which causes her to consider whether she should become a mother or not

herself, other participants responded to her by conveying care online with the use of reaction symbols. This indicates that web-based communication can elicit a different range of feelings, for example, receiving care from other online participants. This finding is supported in the study of Bloor et al [46] by Stone (1995) who argues that technology can convey more than just words, in the sense of smell, touch, sight, etc, and that interactions in cyberspace are social in character.

Reflections on Future Motherhood

More than half of the women stated that participating in online focus groups made them feel as a part of a community. This indicates that there is a possibility that women of reproductive age have unmet needs for discussing thoughts on future motherhood, including timing of motherhood, before they decide whether and when to start a family. Online communities have the potential to give women of reproductive age a feeling of belonging and a voice in the debate of motherhood. Prefertility web-based forums can be a solution to address these needs, but further research needs to be done to explore this. Two participants stated that they spent more time than estimated to engage in the discussions. When studying a sensitive topic as the timing of motherhood, it can be necessary to allow more time for reflection than the estimated 15–20 minutes suggested in this study. Lack of time can explain why some participants did not respond as much as others.

Strengths and Limitations

This study has several strengths, including the insight in a rather unexplored field discussing the timing of motherhood among a group of women of reproductive age who had not yet had children—a topic that many people can consider as personal and sensitive. Another strength is the innovative approach of this study using online focus groups to collect qualitative data on social media with women of reproductive age. Facebook was found to be a suitable platform for collecting qualitative health data, as it was easy for recruitment, cost-effective, well-known, and easy to use by participants and researchers. However, despite these strengths, there are some limitations that need to be considered. A limitation of this study was that we only included women who had a Facebook profile, which meant having access to internet was a prerequisite for participation. In Denmark, where 98% of the population has internet access in their homes, access to participation was not a challenge, but in other settings, internet access should be considered. Another limitation addresses the absence of complete anonymization. When participants used their private Facebook profile, they, to some extent, accepted not being anonymous to other participants, depending on the degree to which they had limited their privacy settings on Facebook. This differs from other types of web-based research, where you typically guarantee participant anonymization, which is not possible to the same extent when doing research on Facebook [50] and why confidentiality in the processing of data is essential, especially in vulnerable or stigmatized groups. Given the reservations about not being anonymous, Lijadi and van Schalkwyk [34] describe that the advantage of participants using their own identity (being nononymous) allows the researcher to verify the participant's authenticity and member identity. A

concern associated with conducting online focus groups on Facebook is that research data were subject to Facebook's privacy policy [51], in which Facebook claims to collect data on communication on the platform, leaving the researcher with less control over the data. All participants had accepted these terms as they were Facebook users prior to participation in this study. As the majority of participants had a high educational level (medium higher or higher education), women with lower education were not represented to the same degree in this study. However, where we often see a connection between less educated people and the development of various diseases, we see the opposite in the reproductive context, based on the fact that it is often highly educated women who have children at advanced maternal age. Conducting online focus groups on Facebook proved to be particularly beneficial during the COVID-19 pandemic, where data collection was compromised in relation to arranging physical meetings. Thus, web-based research on social media makes it possible to connect with participants even under special circumstances. By showing the advantages as well as the disadvantages of conducting online focus groups in a social media setting, we want to stress that

online focus groups is not a replacement for traditional focus groups but shall be considered as an independent alternative method for researchers to use when the topic of interest is suitable for web-based research.

Conclusion

The results of our study show that Facebook is an eligible platform to access qualitative data from women of reproductive age, as we succeeded in recruiting women for this study and collecting qualitative data. Conducting online focus groups on Facebook is an eligible method to access qualitative data from women of reproductive age within health research, especially when participants have access to the internet and are familiar with the platform. Overall, participants were positive toward being a part of an online focus group, and the majority of the participants considered it an advantage to meet on Facebook instead of a physical meeting. Online focus groups have the potential to give women of reproductive age a voice in the debate of motherhood. Further research must be done to explore the impact of conducting qualitative health research on web-based platforms.

Acknowledgments

We thank the women who participated in this study. This study was funded by The Fertility Department, University Hospital Zealand, Region Zealand, Denmark.

Authors' Contributions

CGT and JC were responsible for the design of the study. CGT conducted the online focus groups and was responsible for organizing the analysis and drafting the manuscript. All authors contributed to the manuscript.

Conflicts of Interest

None declared.

References

1. Tankovska H. Number of daily active Facebook users worldwide as of 4th quarter 2020. Statista. URL: <https://www.statista.com/statistics/346167/facebook-global-dau/> [accessed 2021-05-11]
2. Internetbrug og enheder (The use of internet and units). Kulturministeriet (Ministry of Culture Denmark). 2019. URL: https://mediernesudvikling.slks.dk/fileadmin/user_upload/dokumenter/medier/Mediernes_udvikling/2019/Internetbrug_og_enheder/PDF-filer_Internetbrug_og_enheder/Internetbrug_2019.pdf [accessed 2020-05-05]
3. Brug af sociale medier i 2018 (The use of social media 2018). Kulturministeriet (Ministry of Culture Denmark). URL: <https://mediernesudvikling.slks.dk/2018/kort-nyt/brug-af-sociale-medier-i-2018/> [accessed 2020-05-05]
4. Wright K, Fisher C, Rising C, Burke-Garcia A, Afanaseva D, Cai X. Metadata Correction: Partnering With Mommy Bloggers to Disseminate Breast Cancer Risk Information: Social Media Intervention. *J Med Internet Res* 2019 Jun 17;21(6):e14158 [FREE Full text] [doi: 10.2196/14158] [Medline: 31210137]
5. Thurl J, Klein AB, Ramo DE. Smoking Cessation Intervention on Facebook: Which Content Generates the Best Engagement? *J Med Internet Res* 2015 Nov 11;17(11):e244 [FREE Full text] [doi: 10.2196/jmir.4575] [Medline: 26561529]
6. Shoebbotham A, Coulson NS. Therapeutic Affordances of Online Support Group Use in Women With Endometriosis. *J Med Internet Res* 2016 May 09;18(5):e109 [FREE Full text] [doi: 10.2196/jmir.5548] [Medline: 27160641]
7. Green J, Thorogood N. Group interviews. In: *Qualitative Methods for Health Research*, 3rd edition. Los Angeles: SAGE Publications; 2014:126-150.
8. Arigo D, Pagoto S, Carter-Harris L, Lillie SE, Nebeker C. Using social media for health research: Methodological and ethical considerations for recruitment and intervention delivery. *Digit Health* 2018;4:2055207618771757 [FREE Full text] [doi: 10.1177/2055207618771757] [Medline: 29942634]
9. Mellema V. Do men use Facebook differently than women? *Social Media Today*. URL: <https://www.socialmediatoday.com/content/do-men-use-facebook-differently-women> [accessed 2020-05-11]

10. Williams S, Clausen MG, Robertson A, Peacock S, McPherson K. Methodological Reflections on the Use of Asynchronous Online Focus Groups in Health Research. *International Journal of Qualitative Methods* 2012 Sep;11(4):368-383. [doi: [10.1177/160940691201100405](https://doi.org/10.1177/160940691201100405)]
11. Turney L, Pocknee C. Virtual Focus Groups: New Frontiers in Research. *International Journal of Qualitative Methods* 2016 Nov 29;4(2):32-43. [doi: [10.1177/160940690500400203](https://doi.org/10.1177/160940690500400203)]
12. Liamputtong P. Virtual focus groups. In: *Focus Group Methodology: Principles and Practice*. London: SAGE Publications Ltd; 2015:149-164.
13. Skelton K, Evans R, LaChenaye J. Hidden Communities of Practice in Social Media Groups: Mixed Methods Study. *JMIR Pediatr Parent* 2020 Mar 24;3(1):e14355 [FREE Full text] [doi: [10.2196/14355](https://doi.org/10.2196/14355)] [Medline: [32207693](https://pubmed.ncbi.nlm.nih.gov/32207693/)]
14. Befolkning og valg Gennemsnitsalder for fødende kvinder og nybagte fædre (Population and elections average age of birth for women giving birth and newborn fathers). Danmarks Statistik (Statistics Denmark). URL: <https://www.statbank.dk/statbank5a/SelectVarVal/saveselections.asp> [accessed 2020-02-27]
15. Nielsen HS, Schmidt L, Andersen AN, Petersen KB, Gyrd-Hansen D, Jensen TK. Forebyggelse af nedsat frugtbarhed (Prevention of reduced fertility). In: *Vidensråd for forebyggelse*. Copenhagen: Vidensråd; 2016:1-152.
16. Nybo Andersen AM, Wohlfahrt J, Christens P, Olsen J, Melbye M. Maternal age and fetal loss: population based register linkage study. *BMJ* 2000 Jun 24;320(7251):1708-1712 [FREE Full text] [doi: [10.1136/bmj.320.7251.1708](https://doi.org/10.1136/bmj.320.7251.1708)] [Medline: [10864550](https://pubmed.ncbi.nlm.nih.gov/10864550/)]
17. Liu KE, Case A. No. 346-Advanced Reproductive Age and Fertility. *J Obstet Gynaecol Can* 2017 Aug;39(8):685-695. [doi: [10.1016/j.jogc.2016.12.004](https://doi.org/10.1016/j.jogc.2016.12.004)] [Medline: [28549563](https://pubmed.ncbi.nlm.nih.gov/28549563/)]
18. Lean SC, Derricott H, Jones RL, Heazell AEP. Advanced maternal age and adverse pregnancy outcomes: A systematic review and meta-analysis. *PLoS One* 2017;12(10):e0186287 [FREE Full text] [doi: [10.1371/journal.pone.0186287](https://doi.org/10.1371/journal.pone.0186287)] [Medline: [29040334](https://pubmed.ncbi.nlm.nih.gov/29040334/)]
19. Rydahl E, Declercq E, Juhl M, Maimburg RD. Cesarean section on a rise-Does advanced maternal age explain the increase? A population register-based study. *PLoS One* 2019;14(1):e0210655 [FREE Full text] [doi: [10.1371/journal.pone.0210655](https://doi.org/10.1371/journal.pone.0210655)] [Medline: [30677047](https://pubmed.ncbi.nlm.nih.gov/30677047/)]
20. Cooke A, Mills TA, Lavender T. 'Informed and uninformed decision making'--women's reasoning, experiences and perceptions with regard to advanced maternal age and delayed childbearing: a meta-synthesis. *Int J Nurs Stud* 2010 Oct;47(10):1317-1329. [doi: [10.1016/j.ijnurstu.2010.06.001](https://doi.org/10.1016/j.ijnurstu.2010.06.001)] [Medline: [20580363](https://pubmed.ncbi.nlm.nih.gov/20580363/)]
21. Lavender T, Logan J, Cooke A, Lavender R, Mills TA. "Nature makes you blind to the risks": An exploration of womens' views surrounding decisions on the timing of childbearing in contemporary society. *Sexual & Reproductive Healthcare* 2015 Oct;6(3):157-163. [doi: [10.1016/j.srhc.2015.04.006](https://doi.org/10.1016/j.srhc.2015.04.006)]
22. Soderberg M, Lundgren I, Olsson P, Christensson K. A burden and a blessing--young Swedish women's experience of fertility. A study among women lacking experience of pregnancy and parenthood. *Health Care Women Int* 2011 May;32(5):402-419. [doi: [10.1080/07399332.2010.530725](https://doi.org/10.1080/07399332.2010.530725)] [Medline: [21476160](https://pubmed.ncbi.nlm.nih.gov/21476160/)]
23. Sylvest R, Koert E, Vittrup I, Birch Petersen K, Nyboe Andersen A, Pinborg A, et al. Status one year after fertility assessment and counselling in women of reproductive age-a qualitative study. *Ups J Med Sci* 2018 Dec;123(4):264-270 [FREE Full text] [doi: [10.1080/03009734.2018.1546243](https://doi.org/10.1080/03009734.2018.1546243)] [Medline: [30539672](https://pubmed.ncbi.nlm.nih.gov/30539672/)]
24. Birch Petersen K, Sylvest R, Nyboe Andersen A, Pinborg A, Westring Hvidman H, Schmidt L. Attitudes towards family formation in cohabiting and single childless women in their mid- to late thirties. *Hum Fertil (Camb)* 2016 Apr;19(1):48-55. [doi: [10.3109/14647273.2016.1156171](https://doi.org/10.3109/14647273.2016.1156171)] [Medline: [27006139](https://pubmed.ncbi.nlm.nih.gov/27006139/)]
25. Definition of online in English. Lexico. URL: <https://www.lexico.com/definition/online> [accessed 2020-06-17]
26. Definition of virtual reality in English. Lexico. URL: https://www.lexico.com/definition/virtual_reality [accessed 2020-06-17]
27. Temmesen C. Få de børn du ønsker, mens du kan. Afdækning af kvinders overvejelser om fertilitet og reproduktion ved brug af Participatory Design (Get The Children You Want While You Can. Identification of Women's Reflections Regarding Fertility and Reproduction Using Participatory Design). Odense, Denmark: Faculty of Health, University of Southern Denmark; 2016.
28. What is messenger? Webwise. URL: <https://www.webwise.ie/parents/explained-what-is-messenger/> [accessed 2020-07-07]
29. Abrams K, Gaiser T. Online focus groups. In: *The SAGE Handbook of Online Research Methods*. London: SAGE Publications Ltd; 2017:435-449.
30. Malterud K. Planlegging. In: *Fokusgrupper som Forskningsmetode for Medisin og Helsefag*. Oslo: Universitetsforlaget; 2012:29-36.
31. Facebook reactions. Emojipedia. URL: <https://emojipedia.org/facebook/> [accessed 2020-06-29]
32. Definition of post in English. Lexico. URL: <https://www.lexico.com/en/definition/post> [accessed 2020-06-17]
33. How to use emoji in marketing to drive engagement. Sprout Social. URL: <https://sproutsocial.com/insights/emoji-marketing/> [accessed 2020-06-30]
34. Lijadi AA, van Schalkwyk GJ. Online Facebook Focus Group Research of Hard-to-Reach Participants. *International Journal of Qualitative Methods* 2015 Dec 11;14(5):1609406915621383. [doi: [10.1177/1609406915621383](https://doi.org/10.1177/1609406915621383)]
35. Malterud K. Fokusgruppeinterviewet. In: *Fokusgrupper som Forskningsmetode for Medisin og Helsefag*. Oslo: Universitetsforlaget; 2012:68-80.

36. Halkier B. Fokusgrupper. Denmark: Samfundslitteratur; 2018:1-126.
37. Definition of emoji in English. Lexico. URL: <https://www.lexico.com/en/definition/emoji> [accessed 2020-05-05]
38. Research and statistics - Especially regarding health care. Datatilsynet (The Danish Data Protection Agency). URL: <https://www.datatilsynet.dk/emner/forskning-og-statistik/saerligt-om-sundhedsomraadet> [accessed 2020-07-07]
39. Declaration of Helsinki-Ethical principles for medical research involving human subjects. World Medical Association. 2018 Jul 19. URL: <https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/> [accessed 2021-05-11]
40. Den danske kodeks for integritet i forskning (The Danish Code of Conduct for Research Integrity). Uddannelses-og forskningsministeriet (Ministry of Higher Education Science). URL: <https://ufm.dk/publikationer/2015/filer/file> [accessed 2021-05-11]
41. What is tagging and how does it work? Facebook Help Center. URL: <https://www.facebook.com/help/124970597582337> [accessed 2021-05-10]
42. Medley-Rath S. Using Facebook Secret Groups for Qualitative Data Collection. TQR 2019 Jul 22;24(7):1765-1777. [doi: [10.46743/2160-3715/2019.3963](https://doi.org/10.46743/2160-3715/2019.3963)]
43. Tate K, Zwaanswijk M, Otten R, van Dulmen S, Hoogerbrugge PM, Kamps WA, et al. Online focus groups as a tool to collect data in hard-to-include populations: examples from paediatric oncology. BMC Med Res Methodol 2009 Mar 03;9:15 [FREE Full text] [doi: [10.1186/1471-2288-9-15](https://doi.org/10.1186/1471-2288-9-15)] [Medline: [19257883](https://pubmed.ncbi.nlm.nih.gov/19257883/)]
44. What is a notification? Bobology. URL: <https://www.bobology.com/public/What-is-a-Notification.cfm> [accessed 2020-06-30]
45. Graffigna G, Bosio AC. The Influence of Setting on Findings Produced in Qualitative Health Research: A Comparison between Face-to-Face and Online Discussion Groups about HIV/AIDS. International Journal of Qualitative Methods 2016 Nov 29;5(3):55-76. [doi: [10.1177/160940690600500307](https://doi.org/10.1177/160940690600500307)]
46. Bloor M, Frankland J, Thomas M, Robson K. Virtual focus groups. In: Focus Groups in Social Research. London: SAGE Publications Ltd; 2011:74-88.
47. Murray PJ. Using Virtual Focus Groups in Qualitative Research. Qual Health Res 2016 Jul 01;7(4):542-549. [doi: [10.1177/104973239700700408](https://doi.org/10.1177/104973239700700408)]
48. Green J, Thorogood N. In-depth interviews. In: Qualitative Methods for Health Research. Los Angeles: SAGE; 2014:95-125.
49. Clemensen J, Syse M, Coxon I. The rise of patient 3.0: The impact of social media. 2016 Presented at: 8th International Conference e-Health; July 2016; Madeira, Portugal p. 139-148.
50. Kamp K, Herbell K, Magginis WH, Berry D, Given B. Facebook Recruitment and the Protection of Human Subjects. West J Nurs Res 2019 Sep;41(9):1270-1281 [FREE Full text] [doi: [10.1177/0193945919828108](https://doi.org/10.1177/0193945919828108)] [Medline: [30729866](https://pubmed.ncbi.nlm.nih.gov/30729866/)]
51. Data policy. Facebook. URL: <https://www.facebook.com/policy.php> [accessed 2020-06-30]

Edited by G Eysenbach; submitted 04.09.20; peer-reviewed by L Chang, SL Lee; comments to author 18.12.20; revised version received 25.01.21; accepted 13.04.21; published 31.05.21.

Please cite as:

Temmesen CG, Nielsen HS, Andersen HLM, Birch Petersen K, Clemensen J

Using Social Media for Qualitative Health Research in Danish Women of Reproductive Age: Online Focus Group Study on Facebook
JMIR Form Res 2021;5(5):e24108

URL: <https://formative.jmir.org/2021/5/e24108>

doi: [10.2196/24108](https://doi.org/10.2196/24108)

PMID: [34057418](https://pubmed.ncbi.nlm.nih.gov/34057418/)

©Camilla Gry Temmesen, Henriette Svarre Nielsen, Heidi Lene Myglegård Andersen, Kathrine Birch Petersen, Jane Clemensen. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 31.05.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Use of a Self-guided Computerized Cognitive Behavioral Tool During COVID-19: Evaluation Study

Isadora Detweiler Guarino¹, BS; Devin R Cowan¹, BS; Abigail M Fellows¹, MA; Jay C Buckey¹, MD

Space Medicine Innovations Laboratory, Geisel School of Medicine at Dartmouth, Lebanon, NH, United States

Corresponding Author:

Jay C Buckey, MD
Space Medicine Innovations Laboratory
Geisel School of Medicine at Dartmouth
One Medical Center Drive
Lebanon, NH, 03756
United States
Phone: 1 603 650 6012
Email: Jay.C.Buckey.Jr@dartmouth.edu

Abstract

Background: Internet-based programs can help provide accessible and inexpensive behavioral health care to those in need; however, the evaluation of these interventions has been mostly limited to controlled trials. Data regarding patterns of use and effectiveness of self-referred, open-access online interventions are lacking. We evaluated an online-based treatment designed to address stress, depression, and conflict management, the Dartmouth PATH Program, in a freely available and self-guided format during the COVID-19 pandemic.

Objective: The primary aim is to determine users' levels of stress and depression, and the nature of problems and triggers they reported during the COVID-19 pandemic. A secondary objective is to assess the acceptability and usability of the PATH content and determine whether such a program would be useful as a stand-alone open-access resource. The final objective is understanding the high dropout rates associated with online behavioral programs by contrasting the use pattern and program efficacy of individuals who completed session one and did not return to the program with those who came back to complete more sessions.

Methods: Cumulative anonymous data from 562 individuals were analyzed. Stress triggers, stress responses, and reported problems were analyzed using qualitative analysis techniques. Scores on usability and acceptability questionnaires were evaluated using the sign test and Wilcoxon signed rank test. Mixed-effects linear modeling was used to evaluate changes in stress and depression over time.

Results: A total of 2484 users registered from April through October 2020, most of whom created an account without initiating a module. A total of 562 individuals started the program and were considered in the data analysis. The most common stress triggers individuals reported involved either conflicts with family or spouses and work or workload. The most common problems addressed in the mood module were worry, anxiousness, or stress and difficulty concentrating or procrastination. The attrition rate was high with 13% (21/156) completing the conflict module, 17% (50/289) completing session one of the mood module, and 14% (16/117) completing session one of the stress module. Usability and acceptability scores for the mood and stress modules were significantly better than average. In those who returned to complete sessions, symptoms of stress showed a significant improvement over time ($P=.03$), and there was a significant decrease in depressive symptoms over all time points ($P=.01$). Depression severity decreased on average by 20% (SD 35.2%; $P=.60$) between sessions one and two.

Conclusions: Conflicts with others, worry, and difficulty concentrating were some of the most common problems people used the programs to address. Individuals who completed the modules indicated improvements in self-reported stress and depression symptoms. Users also found the modules to be effective and rated the program highly for usability and acceptability. Nevertheless, the attrition rate was very high, as has been found with other freely available online-based interventions.

Trial Registration: ClinicalTrials.gov NCT02726061; <https://clinicaltrials.gov/ct2/show/NCT02726061>

(*JMIR Form Res* 2021;5(5):e26989) doi:[10.2196/26989](https://doi.org/10.2196/26989)

KEYWORDS

computerized cognitive behavioral therapy; interactive media; COVID-19; computer-based therapy; usability; acceptability; cognitive behavioral therapy; therapy; effectiveness; digital health; depression; stress

Introduction

Community epidemiological surveys estimate that as many as 30% of adults in the United States are affected by a mental disorder, yet less than half see a physician, and only a quarter are treated properly [1,2]. A number of barriers limit access to cognitive behavioral therapy (CBT). It is often not widely available, in part due to a lack of adequately trained CBT professionals [3], high cost, potential stigma, inconvenient hours, demands of attending in-person treatments, and concerns over privacy [4].

Technology can increase access to care by providing secure, inexpensive, and easily accessible treatment tools. Internet-delivered CBT (ICBT) has existed for 20 years, and a number of controlled trials and meta-analytic reviews have demonstrated the effectiveness of this approach [5-10]. Although studies have shown ICBT can be as effective as conventional face-to-face therapy [7,8], dropout rates remain exceptionally high. For example, a study has found that as few as 1% of total users completed a full course of an open-access nontracked online program, and fewer than 25% of participants completed programs in a research trial setting [11]. Meta-analytic studies have shown that self-guided web-based interventions (defined as interventions that patients work through on their own without support or guidance) exhibit less promising results than guided web-based interventions (defined as interventions that are delivered with support from a therapist or coach) [6,12-16]. However, the evaluation of web-based programs has been mostly limited to controlled trials, rather than open-access interventions [17], restricting the interpretations of the efficacy and feasibility of these programs in an unstructured format.

Our study evaluated a web-based program in a freely available format. The Dartmouth PATH Program is a multimedia-based computerized CBT tool designed to address stress, depression, and conflict management. The program was developed as a psychosocial training and treatment resource for the National Aeronautics and Space Administration (NASA) with the aim of addressing psychological challenges endured by astronauts on long duration spaceflights [18]. The operational demands of living in such isolated, confined environments can induce conflict, stress, and depression [19,20]. The PATH program has already been tested in extreme environments, such as the Hawaii Space Exploration Analog and Simulation (HI-SEAS) Mars analog and Australian Antarctic stations, and was shown to be acceptable, usable, and valuable [18,21].

Comparably, the COVID-19 pandemic has also been associated with mental health challenges related directly to the virus' morbidity and mortality, and indirectly by the impact of physical distancing and stay-at-home orders [22]. According to the Centers for Disease Control and Prevention, symptoms of anxiety and depressive disorders increased considerably in the United States during the period of April through October 2020 compared to the same period in 2019 [23,24]. Although the

PATH program was made available to the public initially in June 2016, it grew in popularity during the pandemic through news outlets and social media coverage, which increased website traffic. This popularity was largely due to the psychological challenges that were present during that time, leading to an increase in interest of news stories focusing on how people could address their psychological problems on their own. This provided a unique opportunity to determine the type of problems driving people to self-help tools and to assess the uptake, completion, and effectiveness of this ICBT resource in a self-referred, open-access fashion in comparison to its previous evaluations in controlled clinical trials.

The program was freely available as part of NCT02726061. Participants needed to agree to participate using an online consent form. We evaluated responses to the PATH program during the COVID-19 pandemic from April through October 2020. The data collected were fully anonymous and self-reported. Our objectives were to determine the levels of stress and depression as well as the nature of problems affecting individuals during the pandemic, assess the acceptability and usability of the PATH content and determine whether such a program would be useful as a stand-alone open-access resource, and understand the high dropout rates associated with online behavioral programs by contrasting the use pattern and program efficacy of individuals who completed session one and did not return to the program with those who came back to complete more sessions.

Methods

The PATH Program

The PATH program is an interactive, media-intensive CBT-based program that interacts with users in real time and delivers individualized feedback based on self-reported responses (Multimedia Appendix 1). In addition to self-assessment questionnaires and manuals, the program contains three primary modules: conflict management, stress management, and depression treatment. The user can complete the program by following each module's guided instructions session by session in sequential order. Alternatively, the participant can use the self-assessments, which consist of questionnaires that identify the main problems affecting participants and redirects them to the appropriate content module.

Conflict Module

The conflict module teaches participants how to approach conflict and reach effective solutions using CBT principles. It includes a conflict briefing, an interactive conflict simulation, a cognitive restructuring exercise, and a training module on interest-based negotiation [18]. An evaluation study in the isolated and confined HI-SEAS III expedition found this module to be useful, valuable, and interesting [18].

At the end of each one of the four sections in the conflict module, users received a survey on how valuable, feasible, and realistic the simulation was on a scale of 0 to 4 ranging from “Strongly disagree” to “Strongly agree.” The questions covered whether the activity had too much or too little information and if it was interesting or valuable.

Stress Module

Like the conflict module, the stress module also uses CBT principles and focuses on stress management and resilience training. The module consists of six approximately 1-hour sessions. A randomized controlled trial of a version of the stress module reported significant reductions in perceived stress and increases in perceived control over stress [25]. Each session focuses on teaching participants different methods on how to deal with thoughts, feelings, and actions associated with stress. The sessions contained a mixture of activities from three major domains of feelings, thoughts, and actions. “Feelings” activities included guided muscle relaxation and focused breathing. “Thoughts” activities included compartmentalization and weighing evidence, aiming to educate the user on cognitive flexibility. Compartmentalization required the user to imagine a stressful scenario and proceed to shift their attention to perform a task quickly and accurately without being distracted by the previous stressful image. Weighing evidence used cognitive restructuring to help participants identify and dispute the validity of an automatic negative thought by weighing the evidence for and against that thought with the goal of reaching a rational conclusion. Finally, “Action” activities included effective communication, strategic problem solving, and resilience through writing. Effective communication taught assertive communication strategies, strategic problem solving involved problem-solving therapy, and resilience through writing consisted of a journaling activity [25]. At the end of each session, participants received a printout summary that included a stress profile based on the selected stressful triggers, thoughts, physical feelings, emotional feelings, current actions, and selected resilience strategies. It also contained a strategic problem-solving action plan to address the selected stressful trigger and a resilience practice plan consisting of exercises learned in that session to be practiced in the upcoming week.

Users’ progress was tracked at the beginning of each session through a self-reported survey regarding satisfaction with their progress since the last session on a binary scale (“It went well,” “It went not so well”). Participants were also asked how often they practiced each skill learned in the previous session on a scale ranging from 0 to 3 (0 “none,” 1 “once or twice,” 2 “every other day,” and 3 “daily”).

Participants completed the Perceived Stress Scale-14 (PSS) questionnaire at the beginning of each session to assess the degree to which events were perceived as stressful since their last visit. The questions in the PSS were meant to convey feelings and thoughts experienced in the last month on a 0 to 4 scale ranging from “never” to “very often.” The questionnaire contains 14 questions, and the final score is obtained by reversing the scores on seven of the positively stated items and then summing across all 14 questions. The PSS has established

adequate internal and test-retest reliability, and correlates with life event scores [26].

Depression Module

The depression module uses problem-solving treatment and consists of six sessions lasting 30 to 60 minutes. In these sessions, a mentor guides users through a step-by-step problem-solving therapy tool. Participants were first asked to identify and clarify a problem, establish an achievable goal, brainstorm solutions to the problem, evaluate pros and cons of each solution, develop an action plan to implement the selected solutions, and finally schedule enjoyable activities they will do during the next week. At the end of each session, a summary printout was available containing the problem selected, the action plan developed in that session, and a list of the selected scheduled enjoyable activities to be completed before the following session [27]. The program provides tailored feedback through branching algorithms based on user choices in their problem-solving efforts and their scores on the depression questionnaire. A randomized clinical trial on an earlier version of the module showed significant improvements in depression outcomes when using this program compared to a no treatment control group [28].

Participant progress was tracked through surveys regarding satisfaction with the amount of effort spent trying to solve the problem on a 1 to 10 scale ranging from “not satisfied at all” to “extremely satisfied” at the beginning of each session.

Users completed the Patient Health Questionnaire-9 (PHQ-9) Item Depression Scale at the beginning of each session to assess depression symptoms they had experienced for the past 2 weeks. The questionnaire had participants rating nine questions concerning their depression symptoms on a 0 to 3 scale ranging from “not at all” to “nearly every day.” The final score was calculated by adding up the total scores from each question, which ranges from 0 to 27 (0-4: no depression; 5-9: mild depression; 10-14: moderate depression; 15-19: moderate to severe depression; ≥ 20 : severe depression). The questionnaire has demonstrated strong test-retest reliability and internal consistency [29].

Participants

The study was approved by the Dartmouth Committee for the Protection of Human Subjects. The online program was made freely available to anyone interested in participating in the study. Users were required to accept an online consent form and then create a username and password of their choice. The data were acquired during the period of the COVID-19 pandemic from April through October 2020.

Procedures

Users were able to browse the program’s website freely and to choose to go through the cognitive behavioral modules they selected at their own pace. The main page contains all the potential choices for users to choose from, which includes the three primary modules of depression, stress, and conflict management; a guide on how to use the program; self-assessments to guide users on finding relevant content based on individual needs; and other resources and publications.

Demographic Questionnaire

Participant age and gender were collected upon signing up in the study.

Usability and Acceptability Measures

The Post-Study System Usability Questionnaire (PSSUQ) is a 19-item self-report questionnaire used to assess user satisfaction with system usability at the end of a study.

The items are scored on a 7-point scale (1-7) on the strength of agreement with each statement (eg, "It was simple to use this system." "The interface of this system was pleasant."). The scale ranges from "Strongly agree" (1) to "Strongly disagree" (7), and a "Not applicable" indicates answers outside the scale. The overall score is obtained by summing across scores for all questions (1-19), and the evaluation is further subdivided into three subscales: system usefulness (questions 1-8), information quality (questions 9-15), and interface quality (questions 16-18). The overall scale and its subscales have shown adequate levels of reliability, validity, and sensitivity [30,31]. Initially, the program was programmed to present the PSSUQ after the third session in the stress or mood modules. When it became apparent that few people were returning for three sessions, the usability questionnaire was moved to after the first, fourth, and sixth sessions. This meant PSSUQ data were not collected for participants who visited before this change was made.

The Acceptability of Self-Guided Treatment (AST) is a 16-item self-report questionnaire developed in previous research on an earlier version of the depression module as a stand-alone treatment for depression [32]. It comprises of 16 statements (eg, "Computer programs can help with emotional problems such as depression." "I would feel comfortable using this program without a clinician's supervision.") scored on a 7-point scale (1-7) ranging from "Strongly disagree" (1) to "Strongly agree" (7). This question was presented after the first, fourth, and sixth sessions in the mood and stress modules.

Statistical Analysis

To evaluate for statistically significant improvements in PSS (stress module) and PHQ-9 (mood module) scores, we performed the nonparametric Wilcoxon signed rank test using MATLAB (MathWorks). Since the dropout rate of sessions three to six were too high, this test was only performed to assess the median of differences of paired samples for both PSS and PHQ-9 questionnaires from sessions one and two at a 5% significance level (MATLAB's default value). The PSSUQ and AST analyses were also performed using the sign test to evaluate whether the given scores were statistically significant against the neutral mean value of each questionnaire's scale. A right-sided test for the median of the AST scores and a left-sided test for the median of the PSSUQ scores were performed at a 5% significance level. This tested the hypothesis that the median scores of the AST questionnaires were higher than neutral (more acceptable) and that the median scores of the PSSUQ questionnaires were lower than neutral (more usable). To determine whether stress and depression would show statistically significant improvement over time, we conducted linear mixed-effects modeling to account for the time-varying natures

of the variables and missing data. We entered time as the fixed effect and the intercept of subjects as random effects into the model. ATLAS.ti software (ATLAS.ti Scientific Software Development GmbH) was used to synthesize the qualitative data for both the stress and mood modules (stress triggers, emotional response to perceived stressors, and problems selected in the mood module). The numerical values are presented in this study with mean (SD), median (range), or both. The statistical analyses were conducted using MATLAB R2020a (v9.8.0).

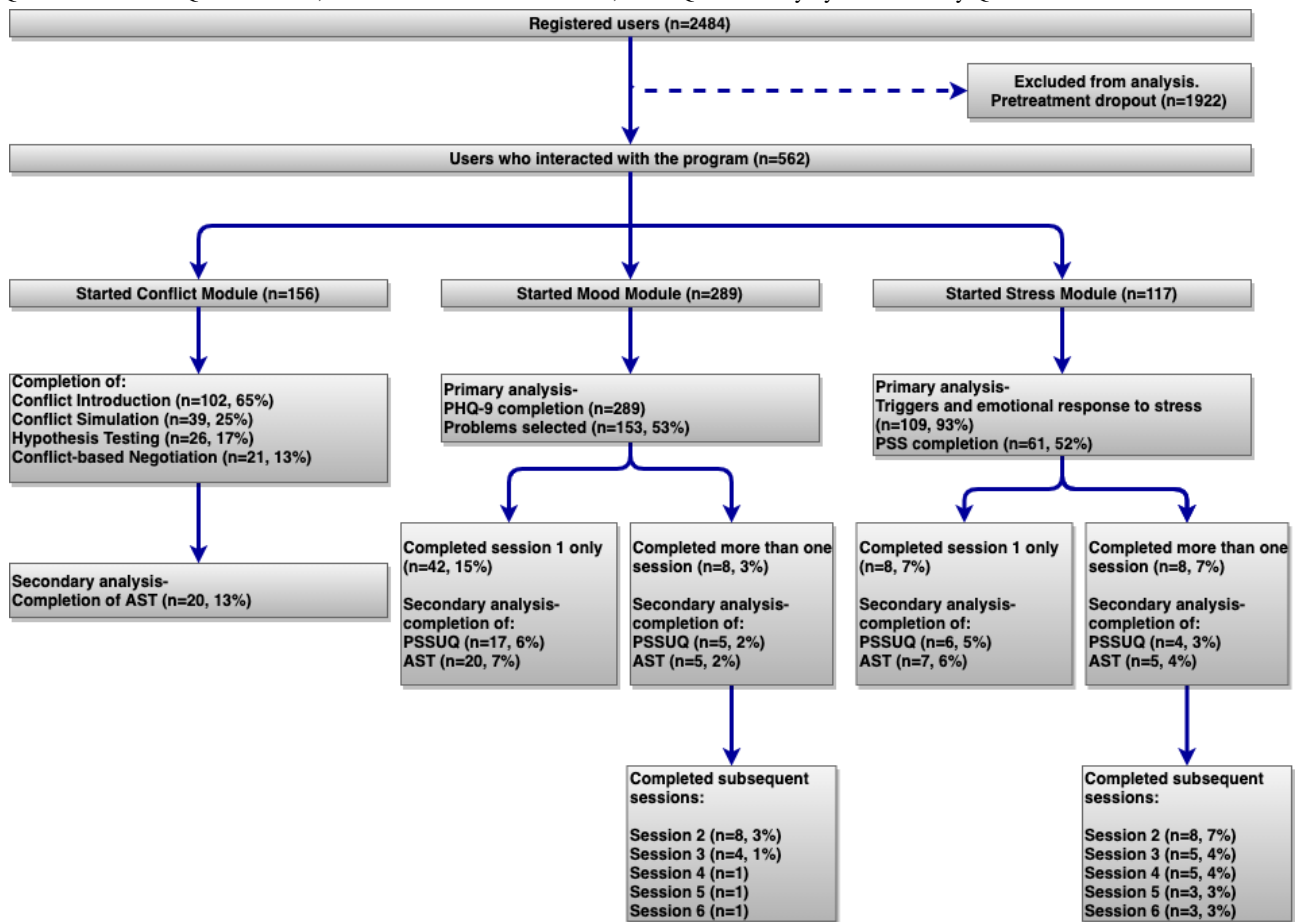
Results

Demographics and Dropout Rate

Between April and October 2020, a total of 2484 users registered with the program website in response to media coverage of the program during the COVID-19 pandemic. Of those who registered, 1321 (53.2%) self-identified as females and 1042 (41.9%) as males with a mean age for the entire group of 44 (SD 15.2) years. The majority of those who registered did not interact with the program and therefore were not included in the treatment dropout rate. The dropout rate involves leaving treatment before its completion [33] and may occur at any point throughout the treatment. For example, a user may withdraw from the program before interacting with any of the module sessions (pretreatment dropout), prior to completion of treatment sessions at any point once treatment had started (treatment dropout), or prior to completing follow-up assessments (follow-up dropout). The majority of registered users created an account but did not proceed with any of the program sessions (pretreatment dropout) [34].

A total of 562 participants interacted with the program modules and were included in the data analysis. At baseline, 156 users interacted with the conflict module, 289 with the mood module, and 117 with the stress module. The conflict module had four subsections and was completed by 21 users (Figure 1). Session one of the mood module was completed by 50 individuals, 8 of which went on to complete subsequent sessions. The PSSUQ and AST questionnaires following session one had a lower completion rate with 22 individuals completing the PSSUQ and 25 completing the AST questionnaire. Subsequent sessions had a very high attrition rate, where 8 individuals completed session two, 4 completed session three, and 1 individual completed sessions four to six. The stress module was completed by 16 individuals, 8 of which went on to complete subsequent sessions within the module. The PSSUQ and AST questionnaires were completed by 10 and 12 users, respectively. Subsequent sessions also had a high attrition rate, where 8 participants completed session two, 5 completed session three and four, and 3 users completed sessions five and six. The primary analysis included the nature of problems and stress triggers experienced by all participants. The secondary analysis focused on the usability and acceptability differences perceived by those who completed session one and did not return to the program with those who came back to complete more sessions to evaluate potential causes for treatment dropout rates.

Figure 1. Flow of participants through a self-guided behavior therapy for conflict, mood, and stress. AST: Acceptability of Self-Guided Treatment; PHQ-9: Patient Health Questionnaire-9; PSS: Perceived Stress Scale-14; PSSUQ: Post-Study System Usability Questionnaire.



Conflict Module

The Conflict Introduction Questionnaire was completed by 102 participants (Table 1) and was found to be interesting (mean 3.2). The Conflict Simulation Questionnaire indicated that participants (n=39) felt like the simulation was valuable (mean 3.2). From the Hypothesis Testing Questionnaire, participants

(n=26) indicated that the program was somewhat easy to understand (mean 2.5), and the feedback somewhat helped them understand why responses were correct or incorrect (mean 2.5). Participants (n=21) found the conflict-based negotiation activity enjoyable (mean 3.1) and highly valuable for learning about conflict management (mean 3.3).

Table 1. The evaluation of conflict module scored on a 0-4 scale ranging from “strongly disagree” (0) to “strongly agree” (4).^a

Item	Mean (SD)	P value	Median (range)
Conflict introduction (n=102)			
The conflict management briefing contained too little information.	1.2 (1.1)	<.001	1.0 (0-3)
The conflict management briefing contained too much information.	0.9 (0.9)	<.001	1.0 (0-4)
I learned a lot from the briefing.	2.8 (1.0)	<.001	3.0 (0-4)
The conflict management briefing was interesting.	3.2 (1.0)	<.001	3.0 (0-4)
I learned a lot from the conflict management briefing that I will probably use in future conflicts.	2.9 (1.0)	<.001	3.0 (0-4)
Conflict simulation (n=39)			
Dr Greenhalgh gave too much information in his spoken comments and advice during the simulation.	0.9 (0.9)	<.001	1.0 (0-3)
Dr Greenhalgh gave too little information in his spoken comments and advice during the simulation.	1.0 (1.0)	<.001	1.0 (0-3)
Overall, I found the simulation valuable for learning about conflict management.	3.2 (0.8)	<.001	3.0 (1-4)
Hypothesis testing (n=26)			
The process of hypothesis testing was easy to understand.	2.5 (1.1)	.048	2.5 (0-4)
The reasons for doing hypothesis testing was easy to understand.	2.9 (1.2)	<.001	3.0 (0-4)
The feedback on my choices helped me to understand why responses were correct or incorrect.	2.5 (1.3)	.06	3.0 (0-4)
I was confused about what I was supposed to do in the hypothesis testing activity.	1.5 (1.1)	.03	1.5 (0-4)
Conflict-based negotiation (n=21)			
Mr Weiss gave too much information in his spoken comments and advice during the interest-based negotiation activity.	1.2 (1.1)	0.2	1.0 (0-3)
Overall, I found doing the negotiation activity enjoyable.	3.1 (0.9)	<.001	3.0 (1-4)
Overall, I found the activity valuable for learning about conflict management.	3.3 (0.7)	<.001	3.0 (2-4)

^aSteve and John were two characters in the simulated conflict. A sign test was performed to evaluate whether the median acceptability scores were significantly different than a median of 2 (neutral).

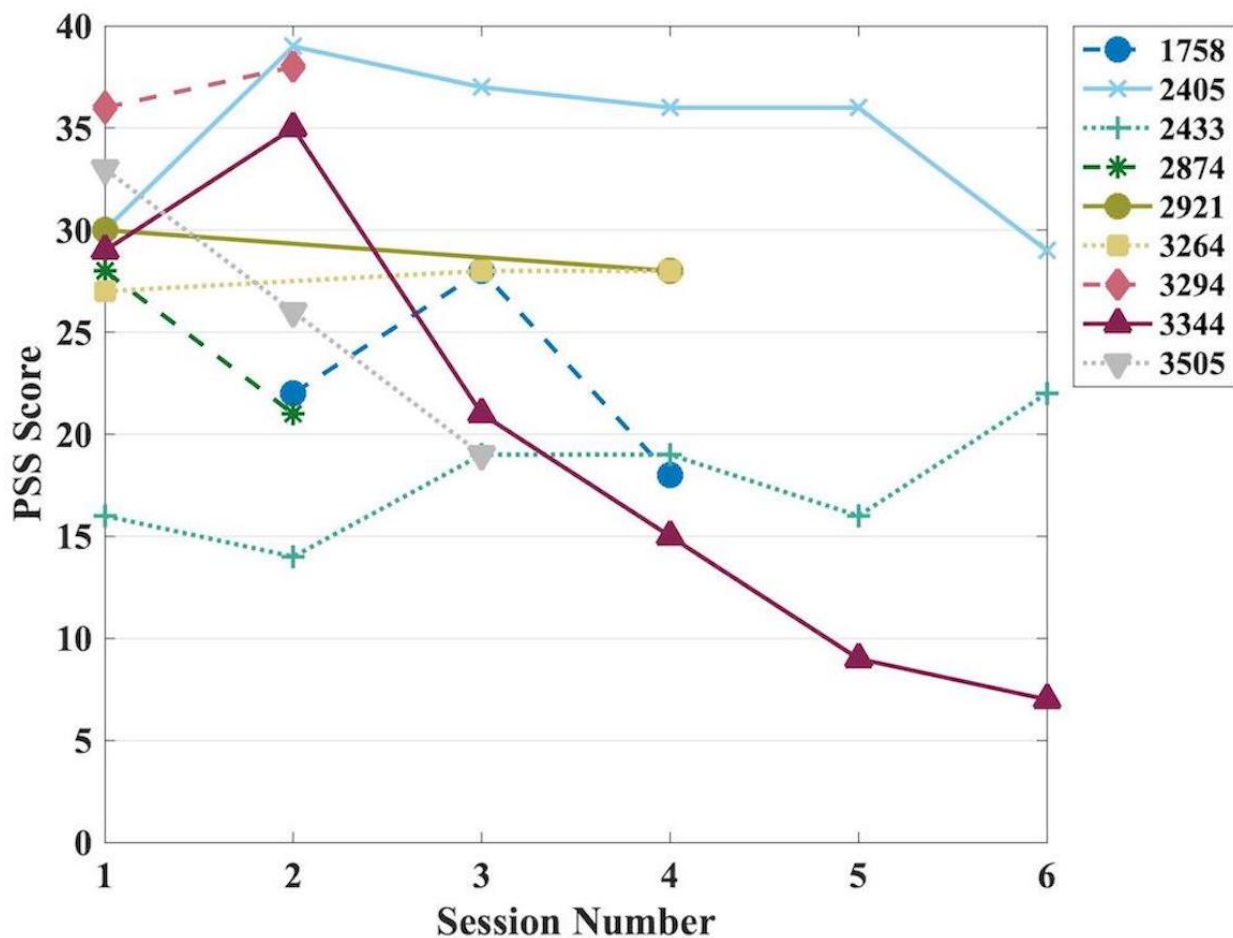
Stress Module

Effects on Stress

The mean PSS score baseline for session one, including all participants (n=61), was 30.0 (SD 6.6). The average PSS of session one for those who returned for more than one session

(n=8) was 28.6 (SD 5.9) while session two (n=8) was 27.3 (SD 9.1). The difference in scores for each participant between the first and last completed sessions was analyzed using the Wilcoxon signed rank test and was found to not be significant ($P=.19$). The individual analysis of the PSS scores across sessions one through six is shown in [Figure 2](#).

Figure 2. Change in stress levels across sessions one through six measured using PSS scores. The figure legend's four-digit values correspond to the identification numbers randomly assigned to users during program registration. PSS: Perceived Stress Scale-14.



To determine whether stress symptoms would improve significantly over time, we used a linear mixed-effects analysis of the relationship between time and PSS scores for each participant throughout all sessions completed. We entered time as the fixed effect and the intercept of subjects as random effects into the model. The decrease in symptoms in relation to time was significant ($\beta = -.14$, SE 0.062, 95% CI -0.27 to -0.013 ; $P = .03$).

Stress Troubleshooting Questionnaire

At the beginning of sessions two, four, and six, participants were asked a follow-up question to evaluate how successful participants were in achieving their goals. The question “How did it go solving this problem?” used a binary scale consisting of the responses “It went well” and “It went not so well.” A total of 16 individuals responded to this question in session two, of which 73% had a positive response. A total of 5 individuals responded to this question in session four, with an 80% positive response. Lastly, session six had a total of 3 participants answering with 100% positive responses.

Triggers and Emotional Response to Stress

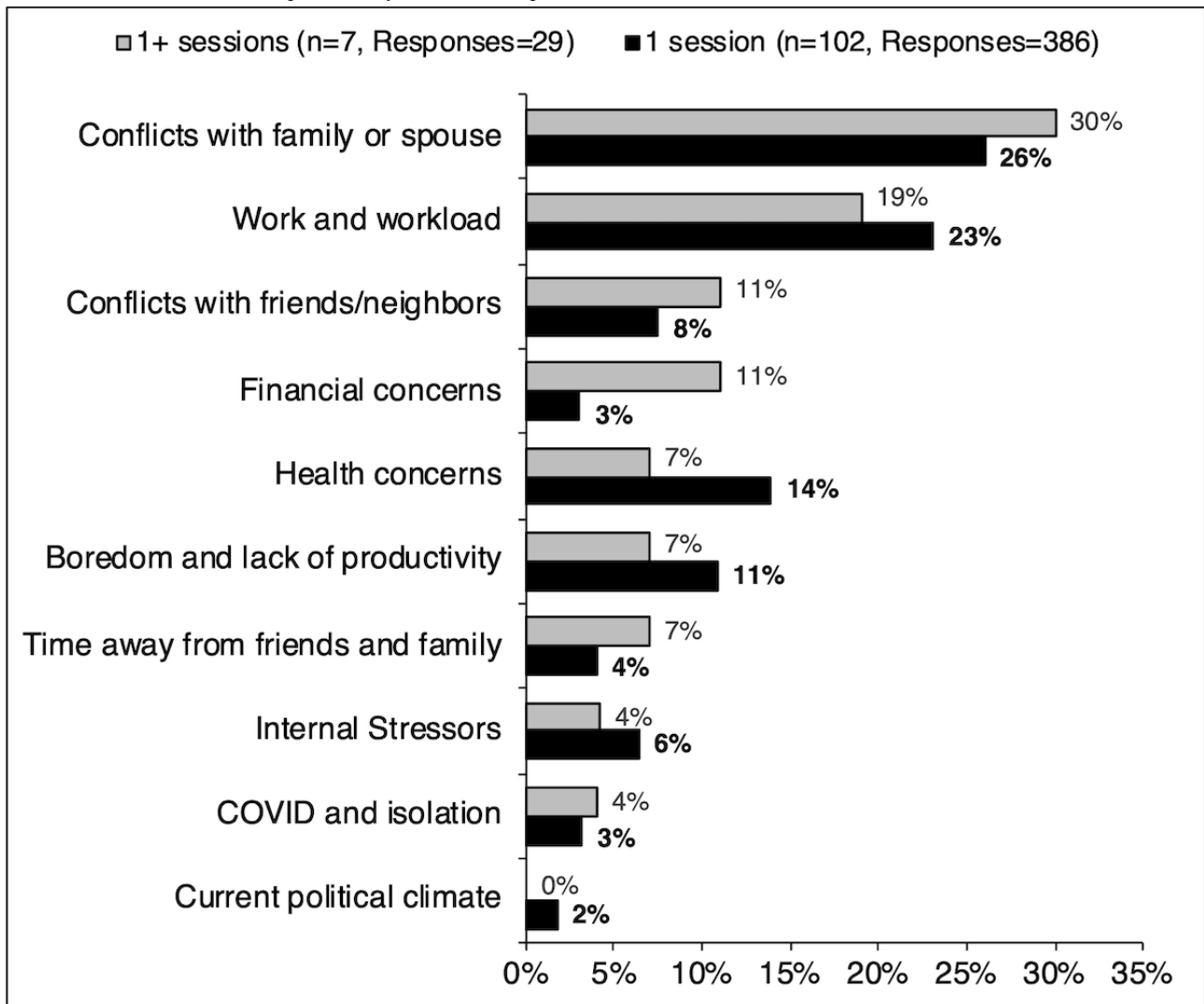
The stress triggers and the emotional response to perceived stressors were compared for those who completed one (n=102)

versus those who completed multiple sessions (n=7). Individuals were able to select multiple responses (ie, multiple triggers and emotional responses).

The stress triggers were grouped into 10 major categories: (1) conflicts with family or spouse; (2) work or workload; (3) conflicts with friends or neighbors; (4) financial concerns; (5) health concerns; (6) boredom and lack of productivity; (7) time away from friends and family; (8) internal stressors, which included negative views about oneself, uncertainty of future, and fear of failure; (9) COVID-19 and isolation; and (10) current political climate. The responses were categorized into each of the major groups, and only one category was counted per person (eg, if a participant noted multiple triggers of the same category, the analysis considered it as just one instance).

Conflict with family or spouse and work or workload were the two most common stressful triggers among participants. In contrast, conflicts with friends or neighbors and financial concerns were ranked third by those who completed multiple sessions, while health concerns and boredom or lack of productivity were ranked third by those who completed just one session. The response percentage rate for each category can be found in Figure 3.

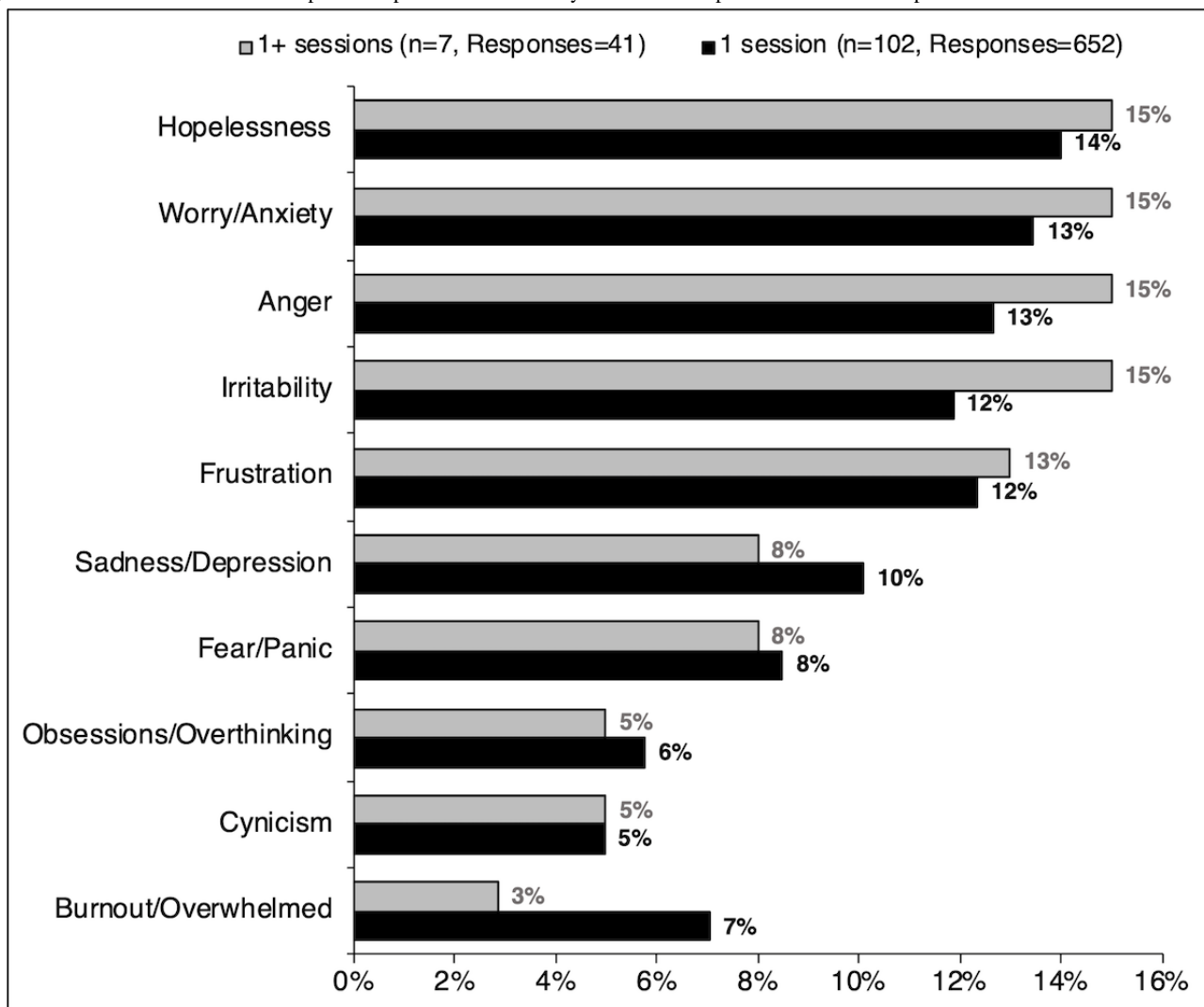
Figure 3. Most common stressors experienced by those who completed one versus more than one session.



The emotional response to perceived stressors were also grouped into 10 major categories: (1) hopelessness, (2) worry or anxiety, (3) anger, (4) irritability, (5) frustration, (6) sadness or depression, (7) fear or panic, (8) obsessions or overthinking, (9) cynicism, and (10) burnout or overwhelmed. The responses recorded allowed for multiple responses, and each response was

placed into the appropriate category (ie, repeated instances within a category were counted individually). The most common responses were hopelessness, worry or anxiety, anger, irritability, and frustration among participants. The response breakdown can be found in [Figure 4](#).

Figure 4. Most common emotional responses to perceived stressors by those who completed one versus multiple sessions.



Mood Module

Retention Rate and Session Duration

The time spent on each part of session one in the mood module was calculated in addition to the retention rate throughout the session. It had five major steps, including completing the PHQ-9, defining the problem, selecting a goal, brainstorming solutions, and developing an action plan. The time elapsed began when the participant started each section until the section was completed, including participant idle time and time away from the computer. The calculation included both individuals who

completed the section in one and multiple sittings. For this reason, participant completion times totaling over 100 hours were not included in the calculation, and the mode and range are included in Table 2. Each section had a retention rate calculated based on the number of individuals who completed the previous section but did not complete the subsequent one (eg, percentage of those who completed “Possible solutions” but did not continue on to the “Action Plan”). The mode for the first four sections ranged from 7 minutes to 15 minutes, the highest retention rate occurred between the possible solutions and action plan, and the highest attrition rate occurred between the PHQ-9 completion and problem definition.

Table 2. Mood session one duration and attrition rate.

Item	PHQ-9 ^a completion (n=289)	Problem definition (n=126)	Goal selection (n=87)	Possible solutions (n=56)	Action plan (n=42)
Time spent (h:min), mean (range)	1:22 (0:03-72:25)	0:34 (0:04-18:41)	0:37 (0:03-72:25)	2:28 (0:06-47:36)	1:07 (0:12-3:23)
Mode (min)	0:07	0:09	0:10	0:15	N/A ^b
Retention rate (%)	Baseline	43.6	69.0	64.4	75.0

^aPHQ-9: Patient Health Questionnaire-9.

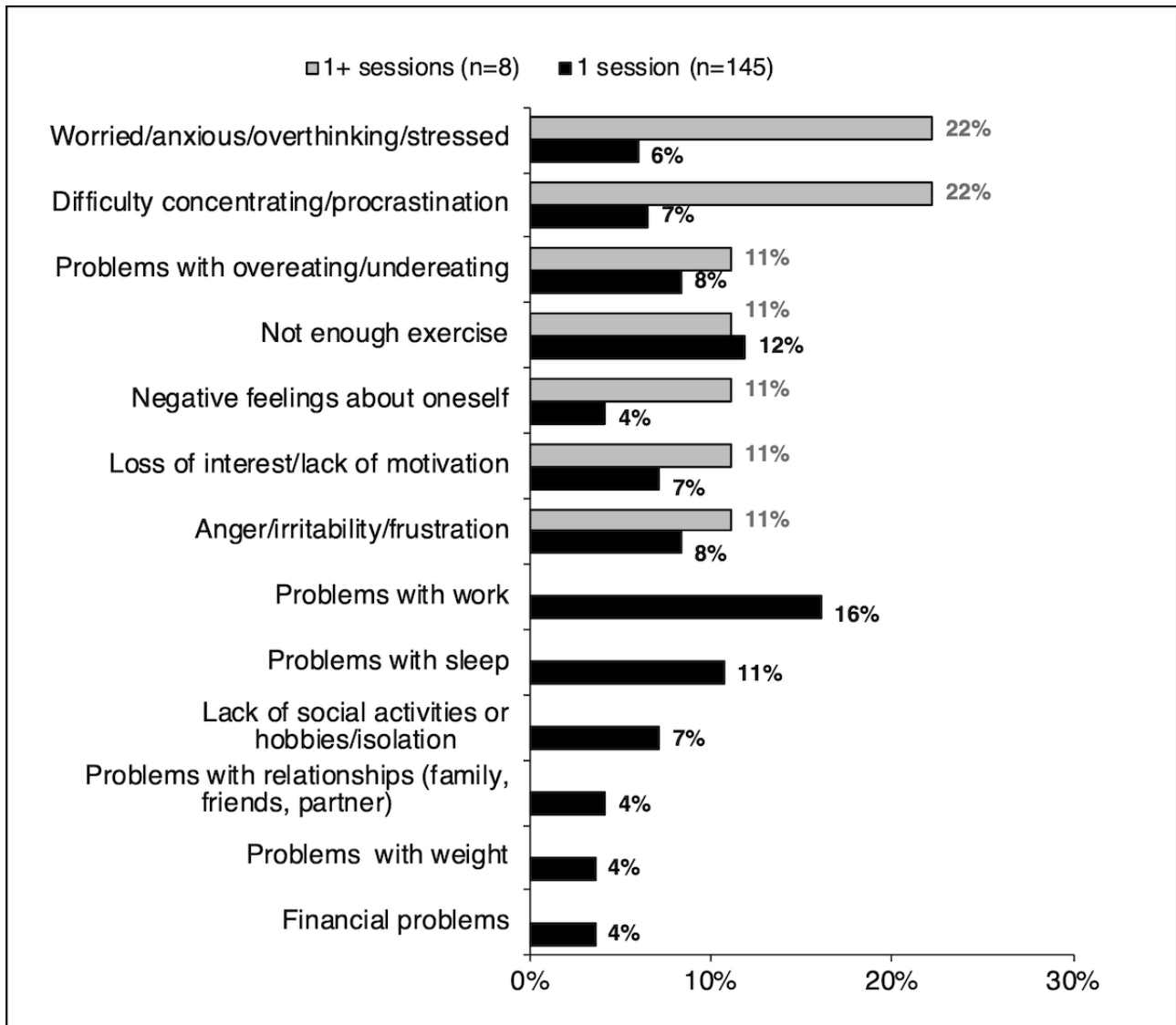
^bN/A: not applicable.

Problems Selected

The type of problems selected by participants in the program was evaluated to better understand the kind of issues experienced by participants and differences, if any, among individuals who only completed one session (n=145) versus those who completed multiple sessions (n=8). The problems were grouped into 13 categories: (1) worried, anxious, overthinking, or stressed; (2) difficulty concentrating or procrastination; (3) problems with overeating or undereating; (4) not enough exercise; (5) negative feelings about oneself; (6) loss of interest or lack of motivation;

(7) anger, irritability, or frustration; (8) problems with work; (9) problems with sleep; (10) lack of social activities or hobbies, or isolation; (11) problems with relationships (family, friends, partner); (12) problems with weight; and (13) financial problems. Those who completed multiple sessions ranked worried, anxious, overthinking, or stressed and difficulty concentrating or procrastination as the highest among the major problems. In contrast, those who completed just one session had problems with work and not enough exercise as the highest ranked. The complete analysis is shown in [Figure 5](#).

Figure 5. Most common problems selected by participants in the mood module.

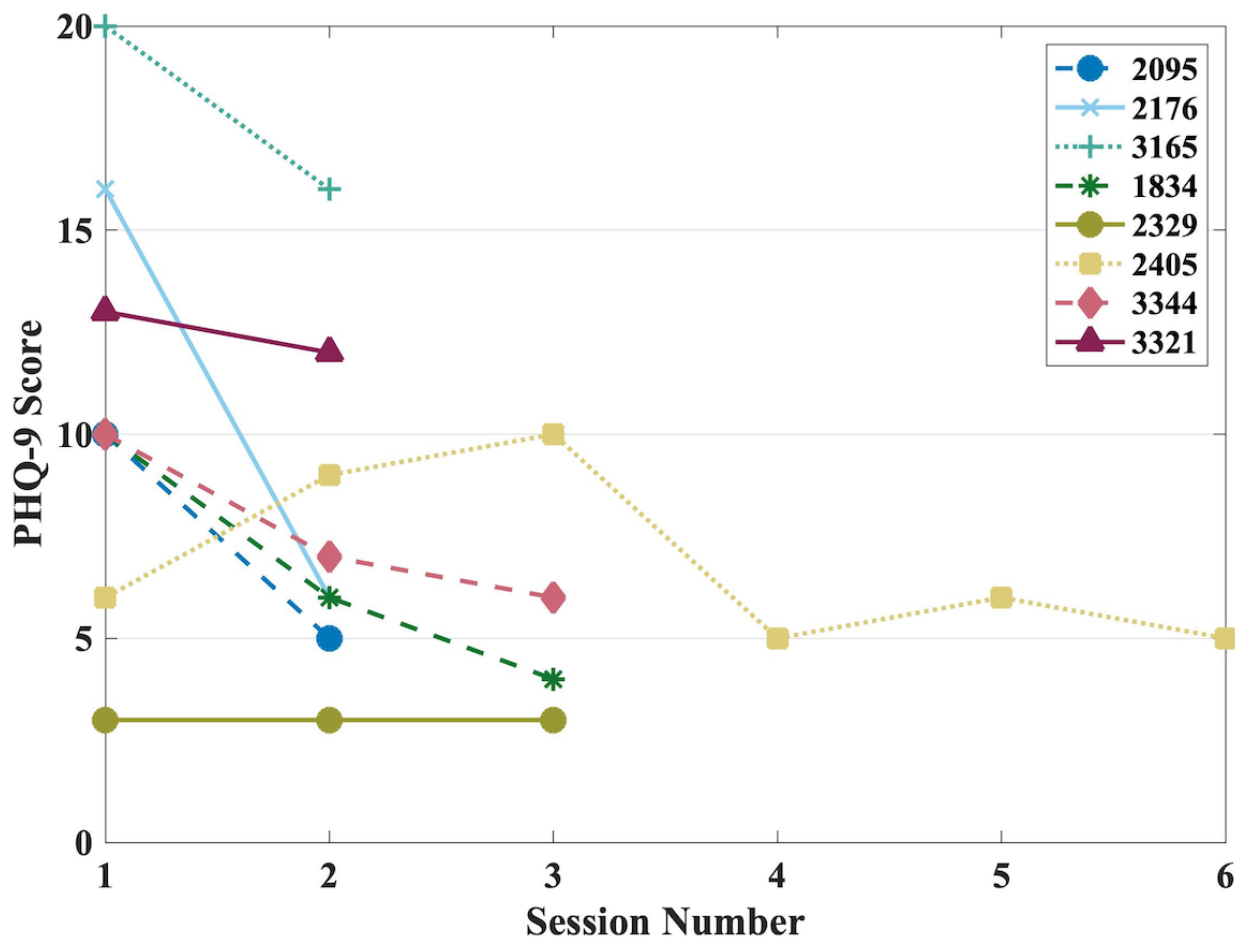


Effects on Depression

The change in depression severity levels was analyzed using the PHQ-9. The mean PHQ-9 score baseline for session one, including all participants (n=289), was 9.98 (SD 6.23). The average PHQ-9 of session one for those who returned for more than one session (n=8) was 11.0 (SD 5.4) while session two (n=8) was 8.0 (SD 4.2). The difference in scores between the

first two sessions, analyzed using the Wilcoxon signed rank test, was not significant ($P=.06$). The depression severity levels between the two sessions decreased by an average of 20.0% (SD 35.2%). The mean PHQ-9 score for session three (n=4) was 5.8 (SD 3.1), session 4 (n=1) was 5.0, session 5 (n=1) was 6.0, and session 6 (n=1) was 5.0. The individual analysis of the PHQ-9 scores across sections one through six is shown in [Figure 6](#).

Figure 6. Change in depression during sessions one through six measured with PHQ-9 scores. The figure legend's four-digit values correspond to the identification numbers randomly assigned to users during program registration. PHQ-9: Patient Health Questionnaire-9.



The linear mixed-effects model analysis of the relationship between time and PHQ-9 scores with time as the fixed effect and the intercept of subjects as random effects showed depressive symptoms decreased significantly over time ($\beta_1 = -.15$, SE 0.056, 95% CI -0.27 to -0.038 ; $P = .01$).

Enjoyable Activities Selected

At the end of each session, participants were asked to schedule enjoyable activities to be completed until the next session. The activities were classified based on the different types of joy-related feelings associated with different activity categories [35]. The categories included social (activities involving interaction with others), intellectual (school-related activities, going to a museum, or going to a concert), basic needs (activities that provide essential elements that their body requires, such as eating or bathing), physical (any activity that promotes physically active movements of the body), nurturance (activities

involving emotional or physical care of others), mastery (activities that involve learning or improving one’s skills), spirituality (religion-related activities or other forms of connection to the divine), and entertainment (miscellaneous activities such as watching TV or going to places not covered by the aforementioned categories) [35].

The average number of enjoyable activities selected by participants was analyzed based on the PHQ-9 depression level scores and the number of sessions completed. Those with higher depression levels who completed more than one session ($n=8$) selected more activities on average than those with lower PHQ-9 scores who completed only one session ($n=35$). A visual analysis is shown in Figure 7.

The most common type of activities selected among participants were entertainment, social, and physical activities. The response breakdown for each category is shown in Figure 8.

Figure 7. Mean number of activities selected by those who completed one versus multiple sessions based on Patient Health Questionnaire-9 depression levels.

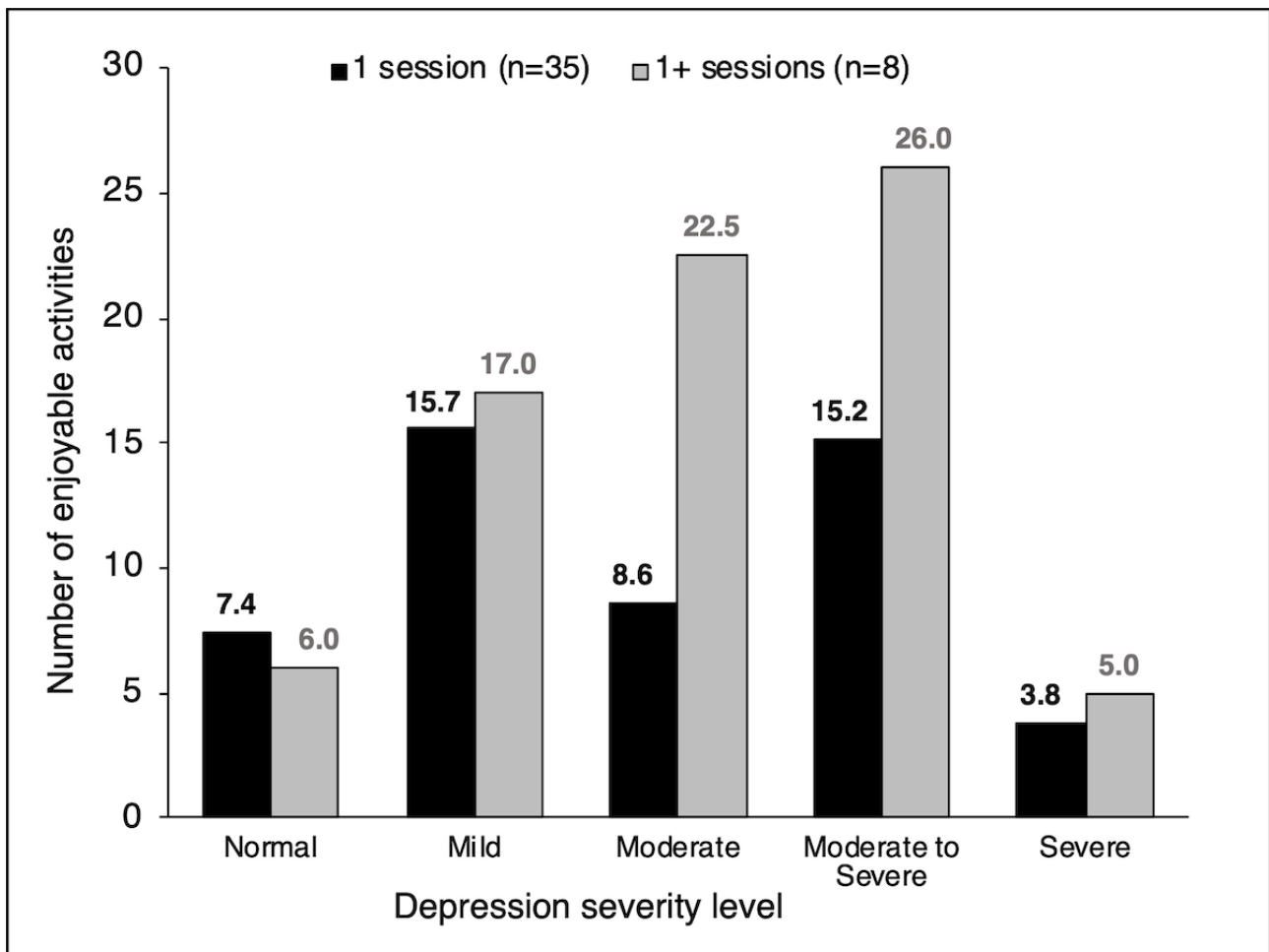
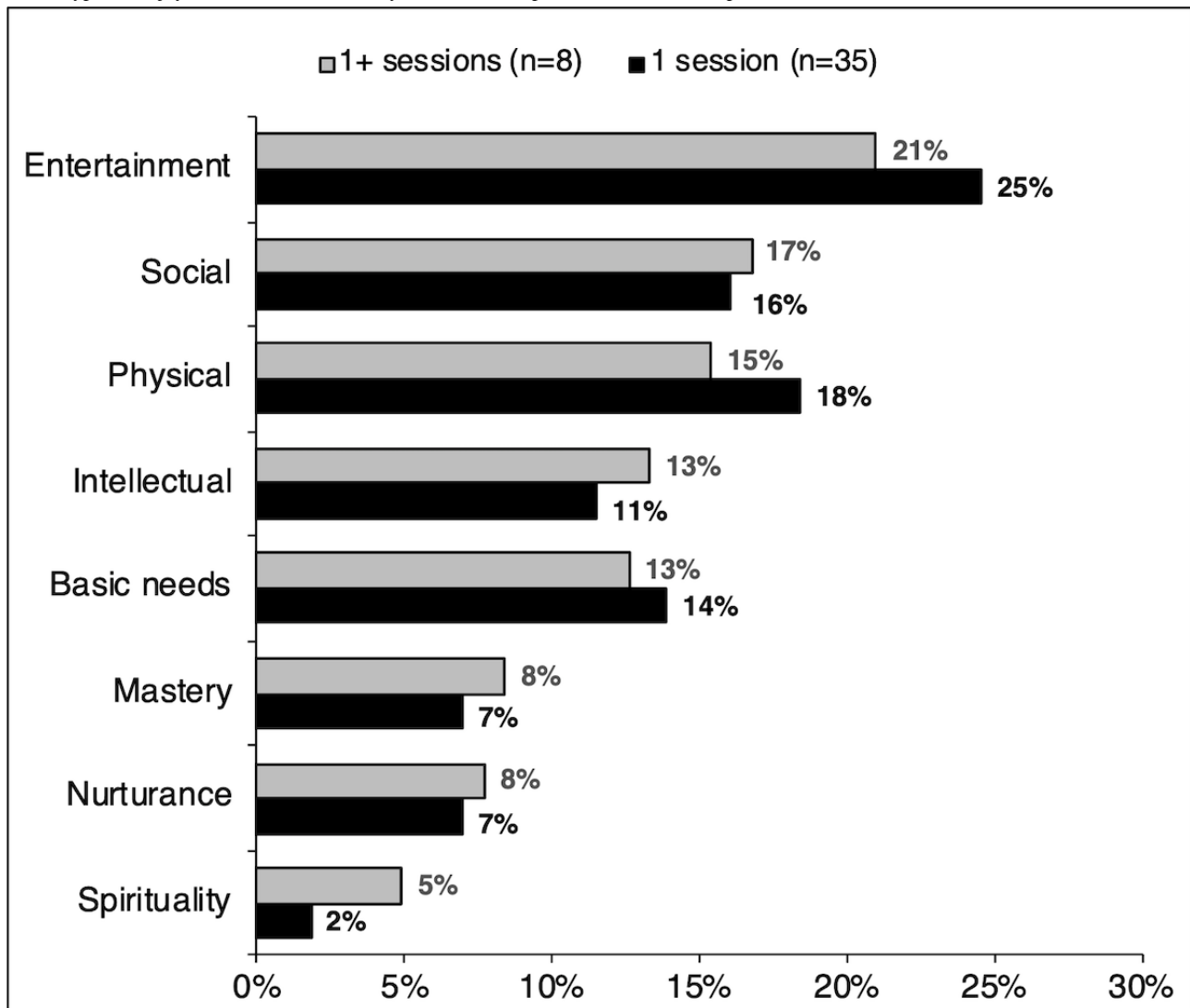


Figure 8. Types of enjoyable activities selected by those who completed one versus multiple sessions.

Usability and Acceptability

This secondary analysis consisted of evaluating the usability and acceptability of the program and differences perceived by those who completed session one and did not return to the program versus those who came back to complete more than one session.

Lower PSSUQ scores indicated a more agreeable usability experience. A total of 22 individuals completed this questionnaire at the end of session one in the mood module, of which 5 individuals went on to complete more than one session.

The mean overall usability score after session one given by those who completed one session (n=17) versus those who completed multiple sessions (n=5) was 2.7 (SD 1.8) and 3.2 (SD 2.2), respectively.

A total of 10 individuals completed the usability questionnaire at the end of session one in the stress module, of which 4 of those participants completed more than one session. The mean overall score after session one given by those who completed one session (n=6) and multiple sessions (n=4) was 3.4 (SD 2.5) and 4.1 (SD 2.4), respectively. Mean scores for each one of the PSSUQ subscales are presented in [Table 3](#).

Table 3. PSSUQ scores for self-guided treatment of the mood (n=22) and stress (n=10) modules of individuals who completed one session versus those who completed multiple sessions.^a

Items	Mood, mean (SD)				Stress, mean (SD)				PSSUQ ^b norms, mean (99% confidence limits)
	1 session (n=17)	<i>P</i> value	>1 session (n=5)	<i>P</i> value	1 session (n=6)	<i>P</i> value	>1 session (n=4)	<i>P</i> value	
System use (Q1-Q8)	2.6 (1.9)	.004	2.9 (2.4)	.19	3.3 (2.5)	.34	3.9 (2.5)	.50	2.8 (2.6, 3.0)
Information quality (Q9-Q15)	2.6 (1.9)	.004	3.4 (2.2)	.50	3.3 (2.5)	.34	4.2 (2.2)	.88	3.0 (2.8, 3.2)
Interface quality (Q16-Q19)	3.0 (1.7)	.03	3.6 (2.2)	.50	3.7 (2.4)	.34	4.3 (2.5)	.88	2.5 (2.3, 2.7)
Overall (Q1-Q19)	2.7 (1.8)	.004	3.2 (2.2)	.19	3.4 (2.5)	.34	4.1 (2.4)	.88	2.8 (2.6, 3.0)

^aItems are scored on a 7-point scale ranging from “Strongly agree” (1) to “Strongly disagree” (7), so lower scores are better. A sign test was performed to evaluate whether the median acceptability scores were significantly lower than a median of 4 (neutral).

^bPSSUQ: Post-Study System Usability Questionnaire.

Higher scores on the AST questionnaire indicate a higher degree of acceptability. A total of 25 individuals completed the questionnaire at the end of session one in the mood module. The mean acceptability score given by those who completed session one only (n=20) was 5.7 (SD 0.7), and the mean score given by individuals who completed multiple sessions (n=5) was 5.9 (SD 0.4). Both of these were significantly different from neutral.

A total of 12 individuals completed the questionnaire at the end of session one in the stress module, of which 5 individuals

completed multiple sessions. The mean acceptability score for the stress module after session one given by those who completed one session (n=7) was 5.5 (SD 0.7), and the mean score given by those who completed multiple sessions (n=5) was 5.3 (SD 0.4). These scores were significantly different from neutral.

The mean acceptability score for the conflict module (n=20) was 5.3 (SD 1.2) and statistically significant from neutral. Mean scores for each AST item are presented in [Table 4](#).

Table 4. Acceptability of Self-Guided Treatment Questionnaire scores for the mood, stress, and conflict modules.^a

Acceptability item	Mood, mean (SD)				Stress, mean (SD)				Conflict (n=20), mean (SD)	P value
	1 session (n=20)	P value	>1 session (n=5)	P value	1 session (n=7)	P value	>1 session (n=5)	P value		
I felt comfortable using the computer.	6.6 (0.8)	<.001	6.8 (0.4)	.03	6.1 (1.1)	.02	6.3 (1.0)	.06	5.7 (1.6)	<.001
Doing [problem-solving treatment/stress management/conflict training] using this program was acceptable to me.	6.2 (0.8)	<.001	6.2 (0.4)	.03	6.1 (1.1)	.02	6.2 (0.8)	.03	5.7 (1.4)	<.001
Using the program helped me to do [problem-solving treatment/stress management/conflict management]. ^b	6.1 (0.9)	<.001	6.4 (0.5)	.03	5.6 (1.1)	.03	6.2 (0.8)	.03	5.6 (1.5)	.01
I would rather do [problem-solving treatment/stress management/conflict management] with a therapist than with the computer. ^c	4.5 (1.6)	.09	4.0 (1.4)	.75	4.6 (1.6)	.31	4.6 (0.5)	.13	4.6 (1.6)	.03
I would rather use a computer to help myself privately than go to a therapist.	4.3 (1.5)	.21	5.0 (1.2)	.13	4.9 (1.1)	.13	4.2 (0.8)	.50	4.8 (1.5)	.046
Computer programs can help people with emotional problems such as depression. ^d	5.6 (1.1)	<.001	5.6 (1.0)	.06	5.4 (1.0)	.03	4.6 (0.5)	.13	5.4 (1.5)	<.001
I would feel comfortable using this program without a clinician's supervision.	5.6 (1.4)	.002	5.6 (1.5)	.19	6.0 (1.0)	.02	5.6 (1.1)	.06	N/A ^e	N/A
I felt safe using the program to do problem-solving treatment. ^f	6.2 (0.9)	<.001	6.6 (0.5)	.03	6.0 (1.0)	.02	5.8 (0.8)	.03	N/A	N/A
I would feel safe doing self-guided treatment for [depression/stress] on my own without a clinician's supervision.	5.6 (1.2)	<.001	5.0 (1.2)	.13	5.6 (1.1)	.03	5.2 (1.3)	.13	N/A	N/A
I would recommend this program to a friend [who was also in need of treatment for depression]. ^g	5.9 (1.0)	<.001	5.6 (0.9)	.06	4.7 (1.8)	.11	4.4 (0.5)	.25	5.7 (1.4)	<.001
Using a program like this could help someone to feel better. ^h	6.0 (0.8)	<.001	6.4 (0.5)	.03	5.4 (1.0)	.03	5.2 (0.8)	.06	5.8 (1.4)	<.001
I believe I would feel comfortable using the program at home on my own computer.	6.3 (0.9)	<.001	6.6 (0.5)	.03	6.1 (1.0)	.02	6.0 (1.0)	.03	5.8 (1.6)	<.001
I felt comfortable answering questions about [depression/stress] symptoms using this program.	6.2 (0.9)	<.001	6.6 (0.5)	.03	5.9 (0.9)	.02	5.6 (1.1)	.06	N/A	N/A
I would feel comfortable using this program without a clinician's assistance.	5.8 (1.4)	<.001	5.8 (1.6)	.19	6.0 (1.0)	.02	5.6 (1.1)	.06	N/A	N/A
Overall score	5.7 (0.7)	<.001	5.9 (0.4)	.03	5.5 (0.7)	.03	5.3 (0.4)	.03	5.3 (1.2)	<.001

^aThe conflict module contains 9 items. Items are scored on a 7-point scale ranging from "Strongly disagree" (1) to "Strongly agree" (7), so higher scores are better. A sign test was performed to evaluate whether the median acceptability scores were significantly higher than a median of 4 (neutral).

^bQuestion slightly modified in conflict module to “Using the program helped me improve my conflict management skills.”

^cQuestion slightly modified in conflict module to “I would rather do conflict-management training with a person rather than with a computer.” Score is inverted when calculating final overall score.

^dQuestion slightly modified in conflict module to “Computer programs can help people improve their conflict management skills.”

^eN/A: not applicable.

^fThe problem-solving treatment in the stress module focused on stress management and resilience training, while problem solving in the mood module focused on simulated live cognitive behavioral therapy treatment.

^gQuestion slightly modified in conflict module to “I would recommend this program to a friend.”

^hQuestion slightly modified in conflict module to “Using a program like this could help someone improve their conflict management skills.”

The difference in response patterns was further evaluated between individuals who completed the first session only and those who completed multiple sessions for the item “I would rather use a computer to help myself privately than go to a therapist.” In the stress module, 43% (3/7) of individuals who completed only one session chose “agree,” while 57% (4/7) chose “neither agree nor disagree.” A third of participants who completed multiple sessions chose “somewhat agree,” 50%

(3/6) chose “neither agree nor disagree,” while 17% (1/6) chose “somewhat disagree.” In the mood module, 45% (9/20) of participants who completed only the first session agreed with the statement and 30% (6/20) were neutral, while 25% (5/20) disagreed with the item. Individuals who completed multiple sessions in the mood module mostly agreed with the item at a rate of 60% (3/5), while 40% (2/5) remained neutral. The complete analysis can be found in [Table 5](#).

Table 5. Analysis of responses to the Acceptability of Self-Guided Treatment item “I would rather use a computer to help myself privately than go to a therapist.”

Item response	Stress 1 session (n=7), n (%)	Stress >1 session (n=6), n (%)	Mood 1 session (n=20), n (%)	Mood >1 session (n=5), n (%)	Conflict (n=20), n (%)
Strongly agree	0 (0)	0 (0)	1 (5)	1 (20)	2 (10)
Agree	3 (43)	0 (0)	3 (15)	0 (0)	4 (20)
Somewhat agree	0 (0)	2 (33)	5 (25)	2 (40)	4 (20)
Neither agree nor disagree	4 (57)	3 (50)	6 (30)	2 (40)	7 (35)
Somewhat disagree	0 (0)	1 (17)	1 (5)	0 (0)	3 (15)
Disagree	0 (0)	0 (0)	4 (20)	0 (0)	0 (0)
Strongly disagree	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)

Discussion

Principal Findings

This study evaluated the use and effectiveness patterns of a free and unrestricted web-based program designed to provide cognitive behavioral-based approaches to stress, depression, and conflict management during the COVID-19 pandemic. The program provided insight into the kinds of problems people were experiencing during the pandemic and seeking help for. In the stress module, the greatest source of stress was conflicts with family or spouse. Feelings of hopelessness, worry, anxiety, and anger were common. The problems selected for the mood program focused on issues of worry, anxiety, overthinking, and stress. Those who returned to complete more than one mood session tended to have problems with worry, stress, and anxiety.

Those who returned to complete more than one stress or mood module showed significant reductions in self-reported stress and depression, although they did not rate the program as high for usability as those who completed only one session. Individuals who completed multiple sessions of the mood module also rated the program higher for acceptability than those who completed only the first session and tended to agree that they would rather help themselves privately with a computer rather than go to a therapist. For the stress module, individuals

who completed only one session rated the program with slightly higher acceptability scores than those who returned to complete multiple subsequent sessions. On average, users thought the conflict training was useful and would recommend it to a friend. In the conflict module, the interest-based negotiation exercise was the highest rated.

The usability score describes the perceived ease of use of the technology applied in this program. It is based on a scale of 1 to 7 (ranging from “strongly agree” to “strongly disagree”), where a higher usability score infers a lower perceived ease of use by the participant. A psychometric evaluation of the PSSUQ established norms for this questionnaire [36]. The amount of usability data collected on the program was limited because the usability questionnaire was moved to after the first stress or mood session late in the study. The largest set of usability data was for session one of the mood program (n=17), which showed that the mood program was rated as significantly better than the norm (lower than the lower limit of the 99% confidence limit) for information quality and was within norms for system use and overall usability. The program, however, was rated worse than the norm on interface quality, although still significantly better than neutral (4) on the scale. The usability data on the stress module and for those who completed more than one visit was limited by low participant numbers (4-6 participants). There was a tendency, however, for the stress program to be rated

lower on usability than the mood program. This indicates that further usability assessments would be worthwhile to identify the areas needing improvement.

The PSSUQ score results provided one surprising finding. Those who completed more sessions (stress: $n=4$; mood $n=5$) perceived the program as less usable than those who completed just one session (stress $n=6$; mood $n=17$). This result was unforeseen, as one would expect participants rating the programs with a low usability score would be likely not to continue. The significance of this finding is limited by the small number of people in these groups who have usability data. Nevertheless, it is interesting that those who chose to continue were not necessarily those who rated the program highly for usability. For the mood module, one possible explanation for continuing was the participants' feelings about whether they wanted to use a computer rather than a therapist. Although 25% of those who completed only one mood session disagreed with the statement "I would rather use a computer to help myself privately than go to a therapist," none of those who did more than one session did.

Ratings on the AST showed that participants found all three modules to be acceptable. The mean score (on a scale of 1 to 7: strongly disagree to strongly agree) for the mood, stress, and conflict modules were 5.7 (SD 0.7), 5.5 (SD 0.7), and 5.3 (SD 1.2), respectively, which were all significantly different than average. These results are similar to mean scores for the same modules when they were used in isolated and confined environments (eg, the HI-SEAS and the Canadian Forces Arctic Station Alert) [18]. The HI-SEAS program was a Mars analog habitat where a crew of 6 people lived in a Mars-like habitat for 8 to 12 months. The Canadian Forces Alert station is the northernmost permanently inhabited place on Earth, and crews there are relatively isolated. The HISEAS and Alert research missions had similar results, with mean AST values of 5.7 (SD 0.9), 5.5 (SD 0.7), and 5.2 (SD 0.9), respectively. Noticeably, the scores given in the mood, stress, and conflict modules to the item "Doing problem-solving treatment/stress management/conflict training using this program was acceptable to me" were 6.2 (SD 0.8), 6.1 (SD 1.1), and 5.7 (SD 1.4), respectively, on a 1 to 7 scale, and scores to the item "I would recommend this program to a friend" were 5.9 (SD 1.0), 4.7 (SD 1.8), and 5.7 (SD 1.4), respectively, on a 1 to 7 scale. These results confirm the acceptability of the program with a general population and match results from previous studies [7,18,25,27,37].

The conflict module consisted of four sections, each concluding with a short questionnaire. The Conflict Introduction and Conflict Simulation Questionnaires were found to be interesting and valuable (mean 3.2, SD 1.0 and mean 3.2, SD 0.8, respectively; $P<.001$ on a 0-4 scale). From the Hypothesis Testing Questionnaire, participants indicated that the program was somewhat easy to understand (mean 2.5, SD 1.1; $P=.048$), and the feedback somewhat helped them understand why responses were correct or incorrect (mean 2.5, SD 1.3; $P=.06$). Lastly, the conflict-based negotiation activity was found to be both enjoyable (mean 3.1, SD 0.9; $P<.001$) and highly valuable for learning about conflict management (mean 3.3, SD 0.7; $P<.001$).

Symptoms of stress as measured by PSS scores between sessions one and two did not significantly decrease ($P=.20$); however, the scores indicated a significant improvement over time ($P=.03$). There was no difference in the response pattern of the selected stressful triggers and perceived emotional response to stress between those who completed one session versus multiple sessions. The major stressful triggers selected by participants included conflict with family or spouse and work or workload. Noticeably, COVID-19 and isolation, political climate, and time away from friends and family were also mentioned as stressful triggers. Lastly, the most common emotional response to perceived stressors were hopelessness, worry or anxiety, anger, and irritability.

There was a difference in the type of problems selected by participants who completed only one session versus those who went on to complete multiple sessions in the mood module. Specifically, those who completed only one session ranked problems with work and not enough exercise, and those who completed multiple sessions ranked worried, anxious, overthinking, or stressed and difficulty concentrating or procrastination as the major problems. Comparably, participants from a previous preliminary uncontrolled trial ranked relationships and financial problems as the highest [27]. This finding suggests more research is needed to identify if there are problems that are particularly suitable for being addressed by the program. The high dropout rate and limited number of people who completed more than one session restricted our interpretation of this finding.

There was a mean 20.0% (SD 35.2%) decrease in depression severity level, as measured by PHQ-9 scores between sessions one and two, which was not statistically significant ($P=.60$). However, there was a significant decrease in the depressive symptoms over time ($P=.01$). Finally, the average number of enjoyable activities selected by participants indicated that those with higher depression levels and who completed multiple sessions selected more enjoyable activities on average than those with lower PHQ-9 scores that completed only one session.

The results overall highlighted an ongoing problem with the use of online behavioral health tools. The fact that a large number of people came to view the programs based on reading or seeing a news article about the program indicates there is a strong interest in self-help tools for behavioral problems. In addition, 45% (9/20) of participants who completed only one session and 60% (3/5) of those who completed multiple sessions of the mood program indicated that they either agreed or strongly agreed with the statement that "I would rather use a computer to help myself privately than go to a therapist," indicating that a significant number of people want to use a computer-based or web-based approach for issues like depression. The acceptability and usability of the programs was good, and those who completed activities showed improvement in self-reported stress and depression. Despite all this, however, very few people returned to continue their work on problem solving or stress management after the first visit. One possible explanation for this is that, no matter how acceptable, usable, or effective a program may be, CBT requires time and effort to be successful. Just like other activities that provide benefits but require effort (eg, dieting or exercising), maintaining the

motivation to complete the programs is difficult. Programs like these likely need to be used within a supportive environment with a human touch that can provide encouragement and ongoing support [38].

Limitations

Data collected was uncontrolled since this was essentially a pragmatic trial during a time when there was heightened interest in the program. In contrast to previous studies of computer-based and online-based interventions evaluated in the setting of clinical trials, we analyzed anonymous data of a self-guided freely available program where users were self-referred.

This evaluation of the PATH program has many limitations, with the most significant likely being the high attrition rate, which is consistent with other research on self-help interventions [11,34,39-42].

Additionally, the data reported here are only from those who completed an activity in the program and then opted to complete a questionnaire, resulting in a selection bias. We do not have data from those who did not complete activities or who may have stopped because of difficulties with either usability or acceptability. The small sample size for those completing multiple sessions restricted our interpretations as to whether this program had a significant effect on reducing mental health symptoms. Given these limitations, any generalization from these results needs to be evaluated with caution. Nevertheless, the opportunity to evaluate a program like this during the COVID-19 pandemic was unique.

The programs were initially developed for use at NASA, and the conflict module included examples and situations based in the space program. The high acceptability and usability ratings for the conflict program indicate that these examples could be understood and applied to the public outside of NASA. Nevertheless, feedback from some participants who completed the conflict module included questions regarding the meaning of some of the words used, such as “EVA” (extravehicular activities). The vocabulary included in some of the scenarios could have created a barrier, hindering the full understanding of the module.

The sample was also not well characterized because, to maintain participant anonymity, only minimal data were collected about the users. This makes it difficult to gauge how confounding variables such as ethnicity, socioeconomic status, or education level would affect the results and attrition rate of this study.

Additionally, the Dartmouth PATH program had a high exposure within the Dartmouth hospital and college campus, which could have contributed to a sample skewed toward well-educated individuals. These factors could affect the results and dropout rates. Those with more experience with computers may be more likely to continue with the program compared to those with lower education backgrounds and less technological capability [43]. Future research will need to explore the generalizability of these findings by including a more in-depth demographic questionnaire at participant registration and an evaluation of computer and internet accessibility.

Conclusions

During the COVID-19 pandemic, those who came to the program seeking self-help for behavioral programs and engaged with the program rated conflicts with family or spouse as a major source of stress. Feelings of hopelessness, worry, anxiety, and anger were commonly reported. The problems selected for the mood program focused on worry, anxiety, overthinking, and stress. Despite the high attrition rate, this study shows that an open-access online behavioral program aiming to treat depression, stress, and conflict management can be effective and rated highly for usability and acceptability by users. This suggests the main issue is not content but instead finding the best implementation strategy. A significant proportion of users reported that they preferred to address behavioral health programs on their own using an online resource. Those who returned for additional visits tended to report more issues with worry, stress, and anxiety, and on average rated the programs less usable than those who completed just one session. For acceptability, there were differences in response patterns between the mood and stress module. Although those who returned for more visits in the mood module rated the program as more acceptable, the opposite trend was true in the stress module. The ultimate value of this program as a stand-alone resource will depend on understanding the reasons for the low completion rates and addressing them effectively. Strategies such as progress tracking, professional support, and weekly reminders may help increase users' adherence to online programs [44]. Attrition rate is also likely to decrease by integrating online-based interventions to a minimally guided approach through a therapist or coach. The minimum amount to which guidance will result in a significant improvement in retention rate and treatment efficacy will be an important focus of future studies.

Acknowledgments

We would like to thank Ryan Slaughaupt for his help with troubleshooting the programs. We greatly appreciate the efforts of David Hirsch who proposed publicizing the programs for use during COVID-19.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Video demonstration of the PATH program.

[MP4 File (MP4 Video), 141634 KB - [formative_v5i5e26989_app1.mp4](#)]

References

1. Kessler RC, Chiu WT, Demler O, Merikangas KR, Walters EE. Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. *Arch Gen Psychiatry* 2005 Jun;62(6):617-627 [FREE Full text] [doi: [10.1001/archpsyc.62.6.617](#)] [Medline: [15939839](#)]
2. Andrews G, Issakidis C, Sanderson K, Corry J, Lapsley H. Utilising survey data to inform public policy: comparison of the cost-effectiveness of treatment of ten mental disorders. *Br J Psychiatry* 2004 Jun;184:526-533. [doi: [10.1192/bjp.184.6.526](#)] [Medline: [15172947](#)]
3. Weissman MM, Verdelli H, Gameroff MJ, Bledsoe SE, Betts K, Mufson L, et al. National survey of psychotherapy training in psychiatry, psychology, and social work. *Arch Gen Psychiatry* 2006 Aug;63(8):925-934. [doi: [10.1001/archpsyc.63.8.925](#)] [Medline: [16894069](#)]
4. Hirschfeld RM, Keller MB, Panico S, Arons BS, Barlow D, Davidoff F, et al. The National Depressive and Manic-Depressive Association consensus statement on the undertreatment of depression. *JAMA* 1997;277(4):333-340. [Medline: [9002497](#)]
5. Andersson G, Carlbring P, Titov N, Lindefors N. Internet interventions for adults with anxiety and mood disorders: a narrative umbrella review of recent meta-analyses. *Can J Psychiatry* 2019 Jul;64(7):465-470 [FREE Full text] [doi: [10.1177/0706743719839381](#)] [Medline: [31096757](#)]
6. Andersson G, Cuijpers P. Internet-based and other computerized psychological treatments for adult depression: a meta-analysis. *Cogn Behav Ther* 2009;38(4):196-205. [doi: [10.1080/16506070903318960](#)] [Medline: [20183695](#)]
7. Andrews G, Cuijpers P, Craske MG, McEvoy P, Titov N. Computer therapy for the anxiety and depressive disorders is effective, acceptable and practical health care: a meta-analysis. *PLoS One* 2010 Oct 13;5(10):e13196 [FREE Full text] [doi: [10.1371/journal.pone.0013196](#)] [Medline: [20967242](#)]
8. Proudfoot J, Goldberg D, Mann A, Everitt B, Marks I, Gray JA. Computerized, interactive, multimedia cognitive-behavioural program for anxiety and depression in general practice. *Psychol Med* 2003 Feb;33(2):217-227. [doi: [10.1017/s0033291702007225](#)] [Medline: [12622301](#)]
9. Rost T, Stein J, Löbner M, Kersting A, Luck-Sikorski C, Riedel-Heller SG. User acceptance of computerized cognitive behavioral therapy for depression: systematic review. *J Med Internet Res* 2017 Sep 13;19(9):e309 [FREE Full text] [doi: [10.2196/jmir.7662](#)] [Medline: [28903893](#)]
10. Mackinnon A, Griffiths KM, Christensen H. Comparative randomised trial of online cognitive-behavioural therapy and an information website for depression: 12-month outcomes. *Br J Psychiatry* 2008 Feb;192(2):130-134. [doi: [10.1192/bjp.bp.106.032078](#)] [Medline: [18245031](#)]
11. 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](#)]
12. Karyotaki E, Kleiboer A, Smit F, Turner DT, Pastor AM, Andersson G, et al. Predictors of treatment dropout in self-guided web-based interventions for depression: an 'individual patient data' meta-analysis. *Psychol Med* 2015 Oct;45(13):2717-2726. [doi: [10.1017/S0033291715000665](#)] [Medline: [25881626](#)]
13. Gellatly J, Bower P, Hennessy S, Richards D, Gilbody S, Lovell K. What makes self-help interventions effective in the management of depressive symptoms? Meta-analysis and meta-regression. *Psychol Med* 2007 Sep;37(9):1217-1228. [doi: [10.1017/S0033291707000062](#)] [Medline: [17306044](#)]
14. Spek V, Cuijpers P, Nyklíček I, Riper H, Keyzer J, Pop V. Internet-based cognitive behaviour therapy for symptoms of depression and anxiety: a meta-analysis. *Psychol Med* 2007 Mar;37(3):319-328. [doi: [10.1017/S0033291706008944](#)] [Medline: [17112400](#)]
15. Cuijpers P, Donker T, Johansson R, Mohr DC, van Straten A, Andersson G. Self-guided psychological treatment for depressive symptoms: a meta-analysis. *PLoS One* 2011;6(6):e21274 [FREE Full text] [doi: [10.1371/journal.pone.0021274](#)] [Medline: [21712998](#)]
16. Richards D, Richardson T. Computer-based psychological treatments for depression: a systematic review and meta-analysis. *Clin Psychol Rev* 2012 Jun;32(4):329-342. [doi: [10.1016/j.cpr.2012.02.004](#)] [Medline: [22466510](#)]
17. Christensen H, Griffiths KM, Korten AE, Brittliffe K, Groves C. A comparison of changes in anxiety and depression symptoms of spontaneous users and trial participants of a cognitive behavior therapy website. *J Med Internet Res* 2004 Dec 22;6(4):e46 [FREE Full text] [doi: [10.2196/jmir.6.4.e46](#)] [Medline: [15631970](#)]
18. Anderson AP, Fellows AM, Binsted KA, Hegel MT, Buckley JC. Autonomous, computer-based behavioral health countermeasure evaluation at HI-SEAS Mars analog. *Aerosp Med Hum Perform* 2016;87(11):912-920. [doi: [10.3357/AMHP.4676.2016](#)] [Medline: [27779949](#)]
19. Buckley JC. *Space Physiology*. Oxford, England: Oxford University Press; 2006:978.
20. Kanas N, Sandal G, Boyd J, Gushin V, Manzey D, North R, et al. Psychology and culture during long-duration space missions. *Acta Astronautica* 2009 Apr;64(7-8):659-677. [doi: [10.1016/j.actaastro.2008.12.005](#)]
21. Lyons KD, Slaughenaupt RM, Mupparaju SH, Lim JS, Anderson AA, Stankovic AS, et al. Autonomous psychological support for isolation and confinement. *Aerosp Med Hum Perform* 2020 Nov 01;91(11):876-885. [doi: [10.3357/AMHP.5705.2020](#)] [Medline: [33334408](#)]

22. Czeisler M, Howard ME, Robbins R, Barger LK, Facer-Childs ER, Rajaratnam SMW, et al. COVID-19: public compliance with and public support for stay-at-home mitigation strategies. medRxiv. Preprint posted online on April 24, 2020. [doi: [10.1101/2020.04.22.20076141](https://doi.org/10.1101/2020.04.22.20076141)]
23. Early release of selected mental health estimates based on data from the January–June 2019 national health interview survey. Centers for Disease Control and Prevention. 2020 May. URL: <https://www.cdc.gov/nchs/data/nhis/earlyrelease/ERmentalhealth-508.pdf> [accessed 2020-10-26]
24. 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 2020-10-26]
25. Rose RD, Buckey JC, Zbozinek TD, Motivala SJ, Glenn DE, Cartreine JA, et al. A randomized controlled trial of a self-guided, multimedia, stress management and resilience training program. *Behav Res Ther* 2013 Feb;51(2):106-112. [doi: [10.1016/j.brat.2012.11.003](https://doi.org/10.1016/j.brat.2012.11.003)] [Medline: [23262118](https://pubmed.ncbi.nlm.nih.gov/23262118/)]
26. Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. *J Health Soc Behav* 1983 Dec;24(4):385-396. [Medline: [6668417](https://pubmed.ncbi.nlm.nih.gov/6668417/)]
27. Berman MI, Buckey JC, Hull JG, Linardatos E, Song SL, McLellan RK, et al. Feasibility study of an interactive multimedia electronic problem solving treatment program for depression: a preliminary uncontrolled trial. *Behav Ther* 2014 May;45(3):358-375 [FREE Full text] [doi: [10.1016/j.beth.2014.02.001](https://doi.org/10.1016/j.beth.2014.02.001)] [Medline: [24680231](https://pubmed.ncbi.nlm.nih.gov/24680231/)]
28. Sandoval LR, Buckey JC, Ainslie R, Tombari M, Stone W, Hegel MT. Randomized controlled trial of a computerized interactive media-based problem solving treatment for depression. *Behav Ther* 2017 May;48(3):413-425 [FREE Full text] [doi: [10.1016/j.beth.2016.04.001](https://doi.org/10.1016/j.beth.2016.04.001)] [Medline: [28390503](https://pubmed.ncbi.nlm.nih.gov/28390503/)]
29. 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](https://doi.org/10.1046/j.1525-1497.2001.016009606.x)] [Medline: [11556941](https://pubmed.ncbi.nlm.nih.gov/11556941/)]
30. Lewis JR. Psychometric evaluation of the Post-Study System Usability Questionnaire: the PSSUQ. *Proc Hum Factors Soc Annu Meeting* 2016 Aug 06;36(16):1259-1260. [doi: [10.1177/154193129203601617](https://doi.org/10.1177/154193129203601617)]
31. Lewis JR. IBM computer usability satisfaction questionnaires: psychometric evaluation and instructions for use. *Int J Human-Computer Interaction* 1995 Jan;7(1):57-78. [doi: [10.1080/10447319509526110](https://doi.org/10.1080/10447319509526110)]
32. Cartreine JA, Locke SE, Buckey JC, Sandoval L, Hegel MT. Electronic problem-solving treatment: description and pilot study of an interactive media treatment for depression. *JMIR Res Protoc* 2012 Sep 25;1(2):e11 [FREE Full text] [doi: [10.2196/resprot.1925](https://doi.org/10.2196/resprot.1925)] [Medline: [23611902](https://pubmed.ncbi.nlm.nih.gov/23611902/)]
33. Davis S, Hooke GR, Page AC. Identifying and targeting predictors of drop-out from group cognitive behaviour therapy. *Aust J Psychol* 2007 Feb 19;58(1):48-56. [doi: [10.1080/00049530500504096](https://doi.org/10.1080/00049530500504096)]
34. Melville KM, Casey LM, Kavanagh DJ. Dropout from Internet-based treatment for psychological disorders. *Br J Clin Psychol* 2010 Nov;49(Pt 4):455-471. [doi: [10.1348/014466509X472138](https://doi.org/10.1348/014466509X472138)] [Medline: [19799804](https://pubmed.ncbi.nlm.nih.gov/19799804/)]
35. Berenbaum H. Varieties of joy-related pleasurable activities and feelings. *Cogn Emotion* 2002 Jul;16(4):473-494. [doi: [10.1080/0269993014000383](https://doi.org/10.1080/0269993014000383)]
36. Lewis JR. Psychometric evaluation of the PSSUQ using data from five years of usability studies. *Int J Human-Computer Interaction* 2002 Sep;14(3-4):463-488. [doi: [10.1080/10447318.2002.9669130](https://doi.org/10.1080/10447318.2002.9669130)]
37. Fleischer J, Ayton J, Riley M, Binsted K, Cowan D, Fellows A, et al. Interactive interest-based negotiation training for managing conflict in isolated confined environments. 2019 Presented at: 70th International Astronautical Congress; October 21-25, 2019; Washington, DC.
38. Mohr DC, Cuijpers P, Lehman K. Supportive accountability: a model for providing human support to enhance adherence to eHealth interventions. *J Med Internet Res* 2011 Mar 10;13(1):e30 [FREE Full text] [doi: [10.2196/jmir.1602](https://doi.org/10.2196/jmir.1602)] [Medline: [21393123](https://pubmed.ncbi.nlm.nih.gov/21393123/)]
39. Farvolden P, Denisoff E, Selby P, Bagby RM, Rudy L. Usage and longitudinal effectiveness of a Web-based self-help cognitive behavioral therapy program for panic disorder. *J Med Internet Res* 2005 Mar 26;7(1):e7 [FREE Full text] [doi: [10.2196/jmir.7.1.e7](https://doi.org/10.2196/jmir.7.1.e7)] [Medline: [15829479](https://pubmed.ncbi.nlm.nih.gov/15829479/)]
40. Marrs R. A meta-analysis of bibliotherapy studies. *Am J Community Psychol* 1995 Dec;23(6):843-870. [doi: [10.1007/BF02507018](https://doi.org/10.1007/BF02507018)] [Medline: [8638553](https://pubmed.ncbi.nlm.nih.gov/8638553/)]
41. Gould RA, Clum GA. A meta-analysis of self-help treatment approaches. *Clin Psychol Rev* 1993 Jan;13(2):169-186. [doi: [10.1016/0272-7358\(93\)90039-O](https://doi.org/10.1016/0272-7358(93)90039-O)]
42. 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](https://doi.org/10.2196/jmir.1772)] [Medline: [21821503](https://pubmed.ncbi.nlm.nih.gov/21821503/)]
43. Andersson G, Titov N. Advantages and limitations of Internet-based interventions for common mental disorders. *World Psychiatry* 2014 Feb;13(1):4-11. [doi: [10.1002/wps.20083](https://doi.org/10.1002/wps.20083)] [Medline: [24497236](https://pubmed.ncbi.nlm.nih.gov/24497236/)]
44. Clarke G, Eubanks D, Reid E, Kelleher C, O'Connor E, DeBar LL, et al. Overcoming Depression on the Internet (ODIN) (2): a randomized trial of a self-help depression skills program with reminders. *J Med Internet Res* 2005 Jun 21;7(2):e16 [FREE Full text] [doi: [10.2196/jmir.7.2.e16](https://doi.org/10.2196/jmir.7.2.e16)] [Medline: [15998607](https://pubmed.ncbi.nlm.nih.gov/15998607/)]

Abbreviations

AST: Acceptability of Self-Guided Treatment
CBT: cognitive behavioral therapy
EVA: extravehicular activities
HI-SEAS: Hawaii Space Exploration Analog and Simulation
ICBT: internet-delivered cognitive behavioral therapy
NASA: National Aeronautics and Space Administration
PHQ-9: Patient Health Questionnaire-9
PSS: Perceived Stress Scale-14
PSSUQ: Post-Study System Usability Questionnaire

Edited by G Eysenbach; submitted 06.01.21; peer-reviewed by A Niles, P Gabs; comments to author 15.02.21; revised version received 08.03.21; accepted 07.05.21; published 31.05.21.

Please cite as:

Detweiler Guarino I, Cowan DR, Fellows AM, Buckey JC

Use of a Self-guided Computerized Cognitive Behavioral Tool During COVID-19: Evaluation Study

JMIR Form Res 2021;5(5):e26989

URL: <https://formative.jmir.org/2021/5/e26989>

doi: [10.2196/26989](https://doi.org/10.2196/26989)

PMID: [33973856](https://pubmed.ncbi.nlm.nih.gov/33973856/)

©Isadora Detweiler Guarino, Devin R Cowan, Abigail M Fellows, Jay C Buckey. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 31.05.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Using a Multisectoral Approach to Advance Health Equity in Rural Arizona: Community-Engaged Survey Development and Implementation Study

Mark Remiker^{1*}, MA; Samantha Sabo^{1*}, MPH, DrPH; Dulce Jiménez^{1*}, MPH; Alexandra Samarron Longorio^{1*}, RDN; Carmenlita Chief^{1*}, MPH; Heather Williamson^{1*}, MBA, DrPH; Nicolette Teufel-Shone^{1*}, DPhil

Center for Health Equity Research, Northern Arizona University, Flagstaff, AZ, United States

* all authors contributed equally

Corresponding Author:

Mark Remiker, MA

Center for Health Equity Research

Northern Arizona University

1395 Knoles Dr.

Flagstaff, AZ, 86011

United States

Phone: 1 2627487095

Email: mark.remiker@nau.edu

Abstract

Background: Over the past decade, public health research and practice sectors have shifted their focus away from identifying health disparities and toward addressing the social, environmental, and economic determinants of health equity. Given the complex and interrelated nature of these determinants, developing policies that will advance health equity requires collaboration across sectors outside of health. However, engaging various stakeholder groups, tapping into their unique knowledge systems, and identifying common objectives across sectors is difficult and time consuming and can impede collaborative efforts.

Objective: The Southwest Health Equity Research Collaborative at Northern Arizona University, in partnership with an 11-member community advisory council, is addressing this need with a joint community-campus effort to develop and implement a Regional Health Equity Survey (RHES) designed to generate an interdisciplinary body of knowledge, which will be used to guide future multisectoral action for improving community health and well-being.

Methods: Researchers and community partners used facilitated discussions and free listing techniques to generate survey items. The community partners pilot tested the survey instrument to evaluate its feasibility and duration before survey administration. Respondent-driven sampling was used to ensure that participants included leadership from across all sectors and regions of northern Arizona.

Results: Over the course of 6 months, 206 participants representing 13 sectors across the 5 counties of northern Arizona were recruited to participate in an RHES. Survey response rates, completion percentage, and sector representation were used to assess the effectiveness and feasibility of using a community-engaged approach for survey development and participant recruitment. The findings describe the current capacity to impact health equity by using a multisectoral approach in northern Arizona.

Conclusions: The Southwest Health Equity Research Collaborative effectively engaged community members to assist with the development and implementation of an RHES aimed at understanding and promoting multisectoral action on the root causes of health inequity. The results will help to build research and evaluation capacity to address the social, economic, and environmental conditions of health inequity in the region.

(*JMIR Form Res* 2021;5(5):e25577) doi:[10.2196/25577](https://doi.org/10.2196/25577)

KEYWORDS

health equity; community-engaged; multisector; survey development

Introduction

Background

Over recent decades, eliminating health disparities has been a major focus of public health efforts in the United States [1,2]. The social determinants of health (SDOH) framework is often used to guide health disparities research by defining the conditions in which people are born, grow, live, work, and age and demonstrating how these factors differentially shape health outcomes within and between populations. Although health disparities and SDOH approaches have offered valuable insights into the conditions and contexts that contribute to sickness and wellness among specific populations, these concepts are limited because they do not expose the important pathways by which social identity (eg, race and gender), the distribution of power and resources, and institutional policies shape opportunities for health. More recently, to address the underlying social inequalities that lead to differential health outcomes across population groups, public health research has shifted its focus toward a health equity framework [3-5].

In 2013, the Robert Wood Johnson Foundation launched a nationwide health equity effort called the *Culture of Health Initiative*, which aimed at making health a shared value; fostering cross-sector collaboration to achieve well-being; and creating healthier, more equitable communities [6]. Health equity initiatives have also been incorporated at the federal level in the United States through the creation of Offices of Minority Health and the goals of Healthy People 2020, which focus on achieving health equity by eliminating disparities and improving the health of all groups [7]. Despite these worthwhile efforts, health disparities and health inequities still loom large in the United States, particularly for people of color and rural communities [8,9].

The barriers to effective action on health equity may be due, in part, to a lack of intersectoral collaboration and consensus on how to identify and overcome the root causes of health inequity, which is defined as the underlying social, economic, and environmental inequalities that create different living conditions among and between populations. A multisectoral approach (MSA) to addressing health equity refers to “deliberate collaboration among various stakeholder groups and sectors (eg, public health, transportation, education, criminal justice) to jointly achieve a policy outcome” [10]. Using an MSA to improve health equity can have multiple benefits, including pooling resources, leveraging unique knowledge bases, expanding reach, and avoiding the duplication of work. This approach is highlighted in the *Health in All Policies* framework, which engages cross-sectoral partners in the promotion of health equity while simultaneously advancing other goals, such as promoting job creation and economic stability [11].

Objective

A major contributor to the lack of successful cross-sectoral collaboration is the problematic perception that addressing issues related to health equity is the sole responsibility of those working in health-related fields [12]. However, given that the root causes of health inequity are diverse, complex, evolving,

and interdependent in nature [13], making progress toward health equity will require collaboration across sectors [3,14].

To address this fundamental issue, we describe the community-engaged development and implementation of the Northern Arizona University (NAU), Southwest Health Equity Research Collaborative’s (SHERC) Regional Health Equity Survey (RHES) [15]. The RHES is designed to understand and strengthen research, practice, policy, and organizational capacity to address locally identified health equity issues using an MSA.

The SHERC is a National Institute of Health–funded Research Centers in Minority Institution (RCMI) initiative of the Center for Health Equity Research (CHER) at NAU. The overall goal of the SHERC is to increase basic biomedical, clinical, and behavioral research at NAU to address health disparities among diverse populations of the southwestern United States. The SHERC consists of 5 cores that interact synergistically: administration, investigator development, recruitment, research infrastructure, and community engagement. This paper focuses on the community engagement core’s (CEC) efforts to address collaborative engagement in health equity in northern Arizona.

The CEC endeavors to cultivate and sustain productive collaborations and partnerships with community-based organizations and leaders in meaningful ways that foster awareness and participation in health equity research. Broadly, the CEC is guided by the *Communities in Action–Pathways to Health Equity Model* [3] grounded in the Robert Wood Johnson *Culture of Health Action Framework* and the Prevention Institute’s *Framework of Emerging Systems to Achieve an Equitable Culture of Health* [6]. These asset-based frameworks recognize that health is impacted by multiple social determinants and that health inequity is produced by multilevel systems such as poverty, structural racism, and discrimination. Therefore, community-based solutions are necessary, but not sufficient, to achieve health equity.

The main objective of the CEC is to engage community-based organizations, community leaders, policy experts, and researchers from various sectors, including childhood development, education, criminal justice, public health, and policy, to (1) identify commonalities in health trends and social, structural, and environmental factors that contribute to health inequity and (2) understand and strengthen organizational capacity to address locally identified health equity issues using an MSA. A primary step in defining public health priorities and understanding the community’s current capacity to impact health inequities is through the systematic collection of information [16].

As a newly established research center, CHER took the first step to better understand public health priorities in 2017 through a collaboration with regional partners, which produced the groundbreaking Regional Health Equity Assessment (RHEA) [17]. Unique in its breadth and scope, the RHEA engaged more than 100 regional thought leaders in interviews, forums, and conferences; synthesized 57 community needs assessments; and produced a local analysis of public health data sets to prioritize health equity goals for northern Arizona. Building on the highly participatory work of the RHEA, and through new funding, the SHERC CEC launched the 2020 RHES designed to take the

next step of understanding and strengthening organizational capacity to address locally identified health equity issues in northern Arizona using an MSA.

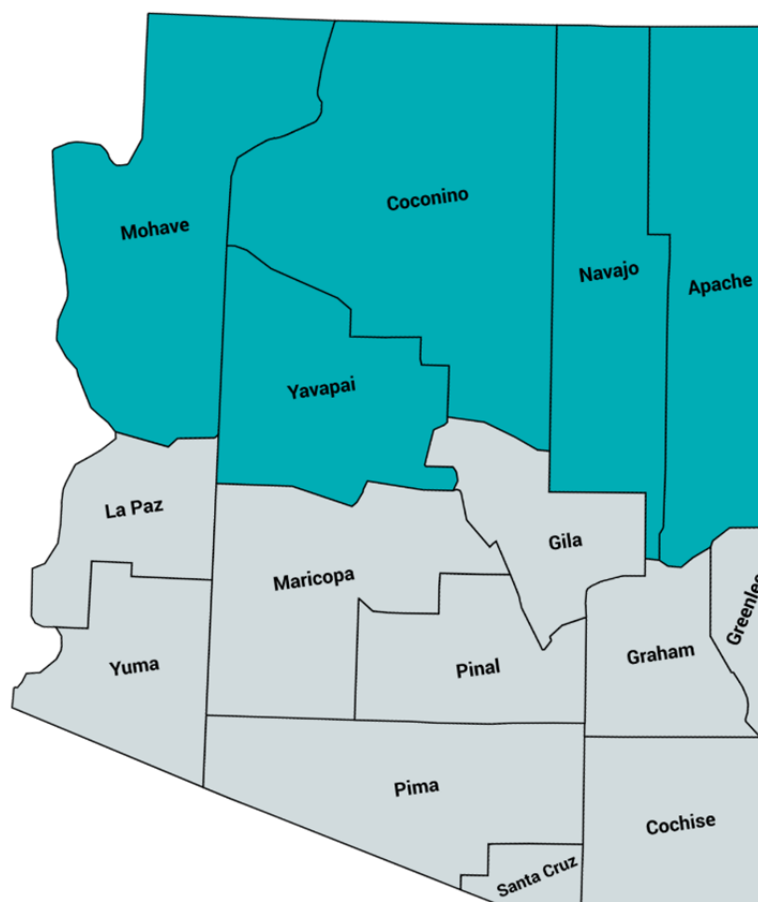
Methods

Regional Overview

Geographically positioned atop the Colorado Plateau, the northern Arizona region covers more than 6000 square miles of land; is home to 12 federally recognized American Indian tribes; and comprises the following 5 counties: Apache, Coconino, Mohave, Navajo, and Yavapai (Figure 1). This region is also characterized by great cultural and ethnic diversity. As of 2017, 62.5% of the people in northern Arizona identified as

White, 22.5% identified as American Indian, and 11% identified as Hispanic [17]. With 37% of the residents living in areas with a population of less than 2500 people, the northern Arizona region is largely rural [17]. In 2018, over 20% of the population was living in poverty, and the per capita annual income was US \$22,636—more than US \$30,000 less than the median household income of the entire state of Arizona [18]. The counties in northern Arizona vary greatly in demographics, such as ethnicity and age, and in degrees of access to all types of services, opportunities, and utilities necessary for healthy people and communities. This diversity makes it important to consider health issues as well as community assets and challenges in a locally specific context and makes this region a scientifically significant venue for the protocol identified in this paper.

Figure 1. Map of northern Arizona counties.



Community Advisory Council

Given the diversity of the northern Arizona region, it is crucial that any initiatives addressing health inequity be community driven. Community advisory councils (CACs) can benefit research institutions by ensuring that the research agenda aligns with priorities that are salient within the community. In addition to providing their unique perspectives and expertise to guide the development of research projects, CAC members can help to bridge gaps and build trust between the community and the research institution [19]. Early on, the CEC established an 11-member CAC composed of leaders from distinct sectors important to achieving health equity across northern Arizona, including early childhood development, education, criminal

justice, public health, and policy. Researchers and CAC members met face-to-face and remotely throughout the survey development and implementation process.

Survey Development and Implementation

The initial stage of the survey development occurred in April 2018 with a half-day, in-person survey workshop between personally, professionally, and geographically diverse members of the CAC and the CEC. The RHES workshop was guided by meaningful learning theory and popular education techniques, which acknowledge that adult learners integrate new knowledge into what is already known and create a cognitive structure to make sense of their surroundings and situations [20,21]. After an introduction to the overarching goals of the RHES,

participants in the CAC and CEC engaged in a facilitated discussion to collectively acknowledge operating systems of power and privilege and define the *root causes* of health inequity in our region. Working from this common understanding of perceived and experienced challenges to health equity in our region, members of the CAC and CEC participated in a free-listing activity aimed at generating specific items for the RHES. Free listing is a technique used to gather data about a specific domain or topic by asking people to list all the items they can think of related to the topic. By using sticky notes in the corresponding color and shape of components of a robust and healthy tree, we identified the specifics of the 5 primary

RHES constructs important to achieve, maintain, and scale health equity in our region: outcomes (leaves), innovations (flowers and fruits), measurement (tree bark), sustainability (rain and sunshine), and partnerships (bees and other cross-pollinators). [Figure 2](#) shows an image of the tree activity. The CEC research staff then transcribed the responses and sorted them into categories within the 5 major RHES constructs. The constructs identified by the CAC were combined with other existing health equity assessments (eg, the Bay Area Regional Health Inequity Initiative Organizational Self-Assessment) [22] to generate the initial set of RHES constructs and questions.

Figure 2. Community advisory council free-listing activity.



The survey questions underwent 2 rounds of edits by CAC members, SHERC project staff, and SHERC research leadership. The final RHES is composed of 48 open- and closed-ended questions covering topics including, but not limited to, the extent and focus of the current cross-sectoral partnerships, priority areas for future research, and the use of data in decision making. To aid participants in completing the survey, the CEC added a glossary of definitions and examples of major public health concepts such as *health disparity*, *SDOH*, and *root causes of health inequity*. The final RHES was generated using Qualtrics [23], a web-based survey platform, and pilot tested by the CEC staff and members of the CAC.

The population of interest for the RHES included community, organizational, and grassroots leaders from the 5 aforementioned northern Arizona counties. In line with the Vitalyst Health

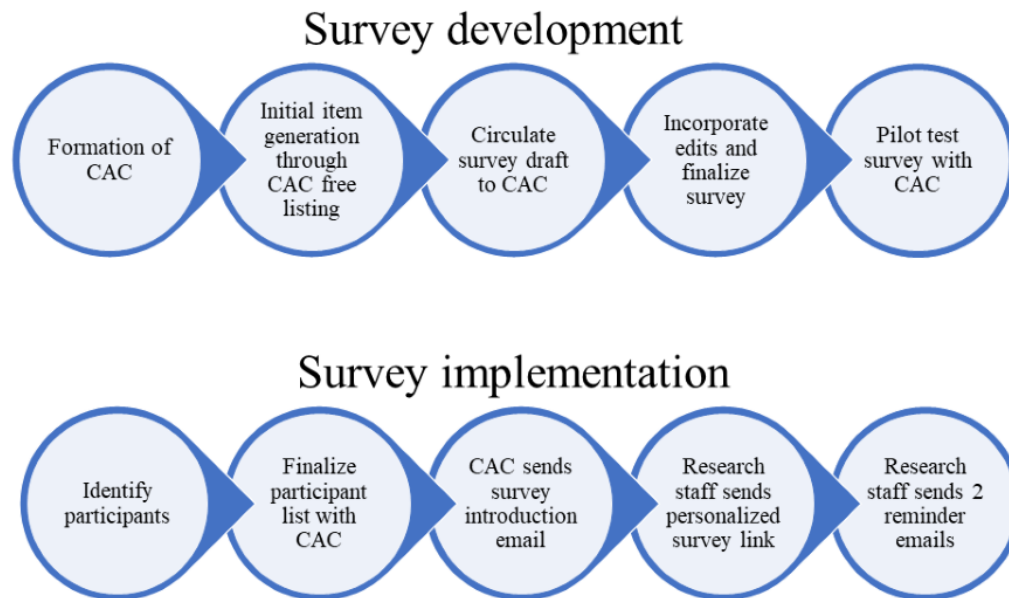
Foundation's elements of a healthy community [24], our sectors of interest included (1) community health and economic development; (2) health and human services; (3) law, justice, and public safety; (4) parks and recreation; (5) policy; (6) early childhood development; (7) transportation; (8) food systems; (9) housing; (10) education; (11) arts, music, and culture; (12) planning and zoning; and (13) cultural resources management.

A 3-pronged approach was used to identify potential participants for the RHES. First, extensive internet searches were conducted to identify individuals in positions of leadership across sectors and counties of northern Arizona. Second, the CAC members nominated leaders from their regions and sectors. Finally, the CEC staff presented the RHES and circulated RHES sign-up sheets at county-level leadership meetings attended by the target population. Attendees were encouraged to add the names of

sector leaders who were not present at the meeting. All potential participants' names were compiled, duplicate names were removed, and county-level participant lists were generated for each sector. Before administering the RHES, at least two county champions (eg, assistant county manager and local public health director) vetted each county's list, removing names of individuals who were no longer in their positions and filling in gaps in sectors where there was no representation.

Once participant lists were finalized, introductory emails were sent by the county champions, alerting all potential participants of the RHES, followed by a personalized email with links to the survey 1 day after the introductory emails were sent. Participants received 2 reminder emails 2 and 4 weeks after the initial invitation. All respondents were offered a US \$25 gift card as compensation for their participation. Figure 3 illustrates the survey development and implementation process.

Figure 3. Regional Health Equity Survey development and implementation flow chart. CAC: community advisory council.



Data Analysis

All descriptive statistics were analyzed using IBM SPSS (version 26). Depending on the responses, qualitative data from open-ended questions were analyzed using either a priori coding or emergent coding and a thematic analysis approach (ATLAS.ti 8, Scientific Software Development GmbH). The Vitalyst Health Foundation's elements of a healthy community [24] were applied to questions where the data were suited for a priori coding. Data were coded by one researcher, and consensus on codes and themes was achieved through intensive discussion with a second researcher throughout the analysis process.

Human Participant Compliance

The development and implementation of the RHES were reviewed and deemed nonhuman subjects research by the NAU's institutional review board (project number: 1198096-1).

Results

Participant Demographics

A total of 206 of the 560 invited multisectoral representatives (response rate 37%) from northern Arizona participated in the RHES. Of those who participated, 64.1% (132/206) completed

the entire survey, whereas 27.6% (57/206) answered 30% or less of the survey questions. Table 1 provides a summary of the survey respondents' demographics.

Although there was a relatively equal distribution across genders with all counties combined (female: 69/129, 53.4%; male 56/129, 43.4%), there was very little ethno-racial diversity, as most of the respondents identified as White (108/129, 83.7%). Survey results indicated that respondents were well established in their sectors (Table 1). On average, participants reported working in their respective sector for over 199.4 (SD 133.3) months, and had been at their current position for an average of 63.7 (SD 71.7) months.

Half of the participants (102/204, 50%) held government positions at the federal, state, county, and municipality levels, and approximately one-third of respondents (45/135, 33%) said that they had an active role or were the primary decision maker within their organization. The reported leadership positions of participants included, but were not limited to, county managers and department directors, chief of police, superintendents, presidents, chief executive officers, and executive directors. Most participants reported either working directly with community members (154/192, 80%) or supervising staff who worked directly with community members (140/192, 73%).

Table 1. Participant demographics by county.

Characteristics	County					
	Apache (n=8)	Coconino (n=94)	Mohave (n=34)	Navajo (n=28)	Yavapai (n=42)	Total (N=206)
Gender, n (%)						
Female	1 (13)	20 (21)	16 (47)	8 (29)	24 (57)	69 (53)
Male	4 (50)	26 (28)	7 (21)	11 (39)	8 (19)	56 (43)
Other	0 (0)	0 (0)	0 (0)	0 (0)	1 (2)	1 (1)
No answer	0 (0)	1 (1)	0 (0)	2 (7)	0 (0)	3 (2)
Race and ethnicity, n (%)						
American Indian or Alaskan Native	0 (0)	2 (2)	0 (0)	1 (4)	0 (0)	3 (2)
Asian or Pacific Islander	0 (0)	0 (0)	1 (3)	0 (0)	0 (0)	1 (1)
Black or African American	0 (0)	3 (3)	0 (0)	0 (0)	0 (0)	3 (2)
Hispanic or Latino	0 (0)	2 (2)	0 (0)	0 (0)	4 (10)	6 (5)
White	5 (62)	40 (43)	19 (56)	17 (61)	27 (64)	108 (84)
Other	0 (0)	0 (0)	1 (3)	0 (0)	2 (5)	3 (2)
No answer	0 (0)	0 (0)	2 (6)	3 (11)	0 (0)	5 (4)
Age (years), mean (SD)	52.6 (5.9)	45.8 (10.1)	52.7 (11.1)	50.9 (9.7)	49.4 (14.4)	49 (11.6)
Position time (months), mean (SD)	21.1 (15.6)	58.1 (71.9)	79.8 (78.5)	69.9 (59.2)	65.4 (77.2)	63.7 (71.7)
Sector time (months), mean (SD)	91.1 (87)	192.4 (116.2)	233.6 (147.3)	480 (199.6)	204.2 (150.8)	199.4 (133.3)

In some of the less populated counties, individuals may be responsible for leading multiple departments; thus, participants could identify with more than one sector. Of the respondents, approximately 71.8% (148/206) identified with only 1 sector. Although there was representation from all 13 sectors, most participants identified in part with either *health and human services* (95/206, 46.1%); *education* (48/206, 23.3%);

community and economic development (34/206, 16.5%); or *law, justice, and public safety* (34/206, 16.5%). A majority of sectors had representation from each county; however, there were 6 instances where 1 county had no sector representation. For instance, no leadership from Yavapai county identified with the *arts, music, and culture* sector. Table 2 shows the distribution of sector by county representation among survey respondents.

Table 2. Distribution of participants by county and self-identified sector.

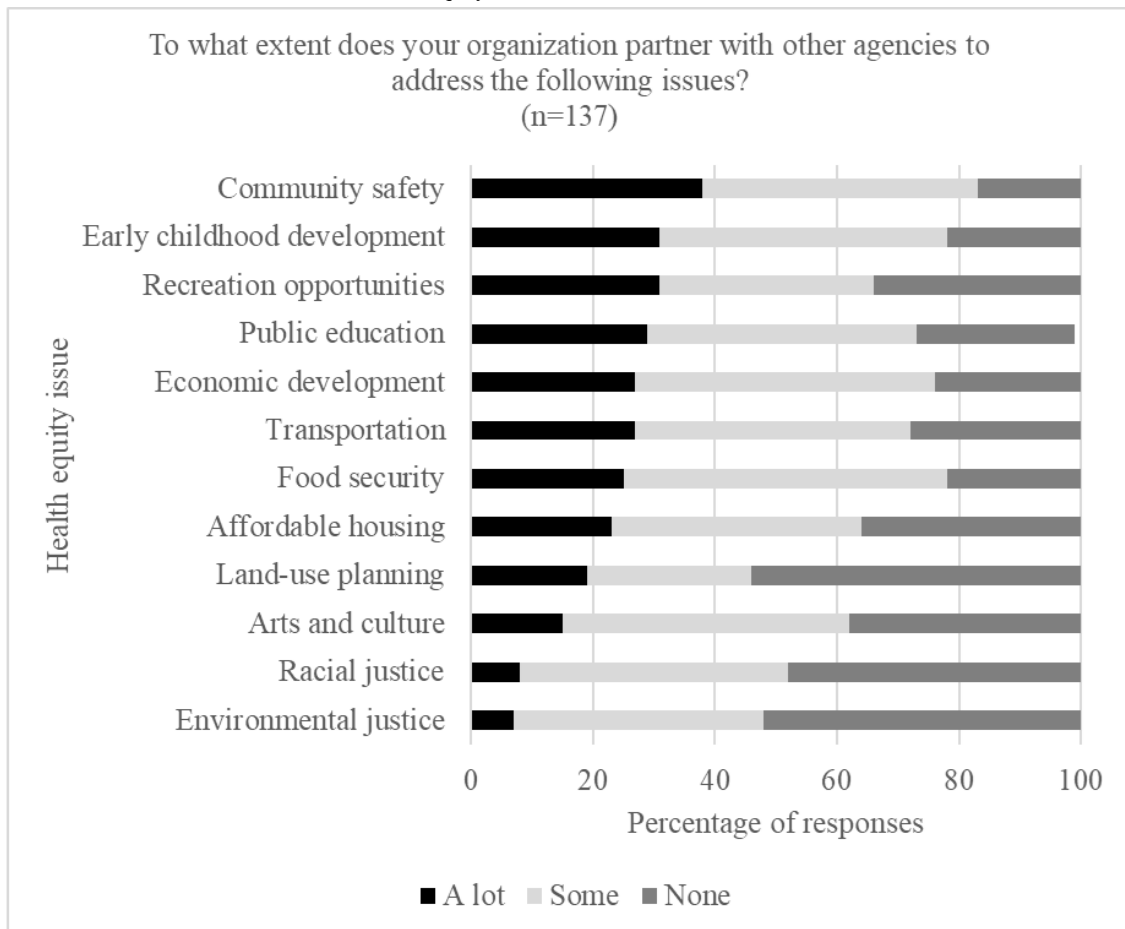
Participant self-identified sector	County					
	Apache (n=8), n (%)	Coconino (n=94), n (%)	Mohave (n=34), n (%)	Navajo (n=28), n (%)	Yavapai (n=42), n (%)	Total (N=206), n (%)
Health and human services	2 (25)	42 (45)	14 (41)	11 (4)	26 (62)	95 (46)
Education	1 (13)	18 (19)	11 (32)	7 (25)	11 (26)	48 (23)
Community and economic development	1 (13)	13 (14)	7 (21)	6 (21)	7 (17)	34 (17)
Law, justice, and public safety	2 (25)	21 (22)	0 (0)	5 (18)	6 (14)	34 (17)
Policy	1 (13)	13 (14)	3 (9)	1 (4)	7 (17)	25 (12)
Housing	0 (0)	11 (12)	1 (3)	2 (7)	4 (10)	18 (9)
Transportation	1 (13)	6 (6)	3 (9)	3 (11)	4 (10)	17 (8)
Food systems	1 (13)	1 (1)	6 (18)	1 (4)	4 (10)	13 (6)
Early childhood development	0 (0)	4 (4)	4 (12)	1 (4)	3 (7)	12 (6)
Parks and recreation	1 (13)	4 (4)	2 (6)	3 (11)	1 (2)	11 (5)
Planning and zoning	1 (13)	6 (6)	0 (0)	1 (4)	1 (2)	9 (4)
Arts, music, and culture	1 (13)	3 (3)	1 (3)	1 (4)	0 (0)	6 (3)
Cultural resource management	1 (13)	1 (1)	0 (0)	2 (7)	1 (2)	5 (2)

Cross-Sectoral Partnerships

The most frequently cited characteristics for developing a successful multisectoral partnership were communication, shared vision, and trust. Approximately 74.7% (130/174) of participants reported that they had previously partnered with a different sector to address environmental, social, and economic conditions that impact health, and 94.6% (123/130) of those individuals

indicated a past partnership with more than 1 sector (average number of partnerships 6.7, SD 3.5). As displayed in Figure 4, *community safety*, *early childhood development*, and *recreation opportunities* were the primary issues on which organizations most often collaborated with other sectors, whereas *land-use planning*, *racial justice*, and *environmental justice* were the issues least likely to garner cross-sectoral attention.

Figure 4. Extent of cross-sectoral collaborations on health equity issues.



Role of Research

To examine how research can effectively influence health equity in northern Arizona, participants were asked, “What role do you think research has in addressing the environmental, social, and economic conditions that impact health in the community you serve?” Most leaders asserted that research plays a significant role in addressing the root causes of health inequity, whereas very few participants felt the role of research was *little* or *none*. Often, participants described the limitations of research, expressing that although research plays an important role in identifying, understanding, and addressing needs or problems

in their communities, the right conditions must be met, including conducting research responsibly and ethically, using scientifically sound methods, and yielding actionable results to directly and positively impact the community. Specifically, leaders believed that research affords an opportunity to illuminate and understand the gaps and problems experienced by the community and serves to validate the community’s knowledge and lived experiences of their own needs; leaders further indicated that this information can be used for action. Textbox 1 outlines the areas of research that were identified by community leaders as a priority for future studies on health equity.

Textbox 1. Community-identified priority areas for health equity research.

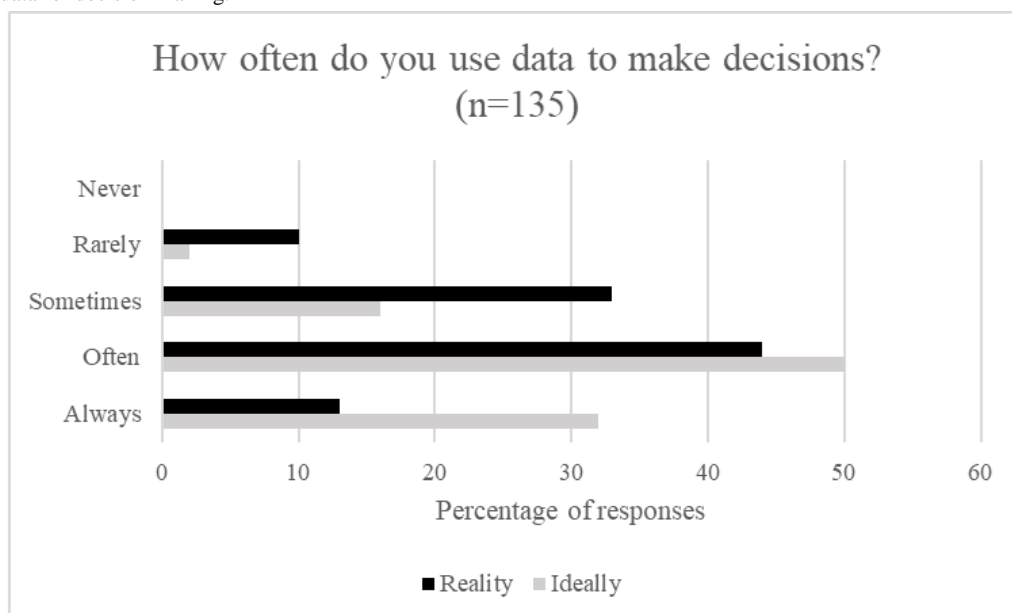
Area of research and specific research topics
<ul style="list-style-type: none"> • Economic opportunities: Poverty, disparities in income, job opportunity and lack of higher wage jobs, workforce development, economic development, and economic indicators • Health care: Access, affordability, and quality of health services and health plan coverage; long distances people must travel to seek care; and understaffing and difficulty attracting and retaining health care professionals, especially in rural areas • Behavioral health: Access to mental health and substance use services, including drug addiction, rehabilitation; addressing stigma related to mental health and substance use • Education: Educational opportunities from kindergarten through high school through higher education, affordability, and funding • Transportation: Access, affordability, and adequacy • Housing: Access, affordability, and homelessness • Food: Access, food security, and quality of healthy foods • Early childhood: Early childhood education and youth development • Social context: Social context around health inequities, understanding issues around culture, stigma related to health conditions, and social activities • Social justice: Effects of incarceration, historical trauma, and social justice in relation to other social determinants of health • Environment: Climate change • Tribal communities: Funding, focus and effectiveness of Indian Health Services, health care options on the reservation, and impact of native American culture on health maintenance • Rural communities: Access to services based on unique challenges experienced by rural communities • Aging and older people: Access to services

Use of Data in Decision Making

Data-informed health promotion is an emerging and ever-changing theme in public health research and practice. To examine the current data use across sectors, leaders were asked how often they used data to make decisions and what barriers they faced in the process. Across all leaders, 93.3% (126/135) reported having used data to make decisions; however, there exists a gap between how often data are currently used and how often leaders would ideally use data to guide their decision

making (Figure 5). Although 81.4% (110/135) of sector leaders would prefer to *always* or *often* use data to make decisions, only 56.2% (76/135) currently use data to make decisions. This pattern emerged for participants who identified with either a single sector or more than 1 sector. Furthermore, over 60.7% (82/135) of the respondents said that they did encounter barriers to using data to make decisions. When asked to identify the biggest barriers to using data, leaders most often cited a lack of useful available data, followed by an absence of the expertise needed to analyze the data.

Figure 5. Use of data for decision making.



Discussion

Principal Findings

In this paper, we describe how the SHERC at NAU effectively engaged community members to assist in the development and implementation of the RHES, with the goal of understanding and promoting multisectoral action on the root causes of health inequity in northern Arizona. Furthermore, we demonstrate how over 200 county-level leaders from various sectors, beyond public health and health care, were recruited to share their knowledge, attitudes, and actions to address the social, environmental, and economic conditions that impact health and well-being in the region.

Overall, our work revealed that using a community-engaged approach to survey local leadership can be an effective first step toward identifying and prioritizing areas of action on the root causes of health inequity. Specifically, using a community-engaged approach to develop and implement the RHES ensured that (1) survey questions resonated with community priorities and (2) respondents were recognized as representatives of their community. With participation across all sectors and counties of interest (Table 2) and an average reported time of 16.6 years working in their sector, we are confident that the outcomes of the RHES capture the perspectives of multisectoral leadership in northern Arizona and reveal the range of factors that contribute to health inequity in the region.

This study also exposes the benefits and challenges of developing and implementing cross-sectoral partnerships to address health inequities. Although collaboration with governmental and nongovernmental sectors outside of health is required to develop policies and programs to advance health equity, establishing and maintaining effective cross-sectoral partnerships is not an easy task [25]. For example, engaging relevant stakeholders, tapping into their unique knowledge systems, and identifying common objectives across sectors requires time and resources that are not often afforded to governmental and nonprofit agencies. As the local *anchor institution* and RCMi in northern Arizona, NAU is in a unique position to leverage their human resources and expertise to collect and disseminate cross-sectoral perspectives on regional issues of health inequity [3]. The approaches and methods used in this study serve as a model for other universities and RCMIs, in particular, to advance community-driven health equity agendas.

Implications for Research and Practice

Findings from this study suggest that county-level leaders in northern Arizona are currently working across sectors to address the root causes of health inequity; however, the extent to which they partner is limited, and the issues being addressed are bounded and unbalanced (Figure 3). These results indicate that there is capacity to impact health inequities using an MSA, but this work could benefit from more deliberate coordination across sectors. Recent research by Narain et al [26] shows that framing issues of health equity in ways that resonate with sectors outside of public health was valuable for promoting cross-sectoral work to improve health equity. Focusing on the cross-sectoral research

priorities that were identified in the RHES (Textbox 1) can, therefore, help to determine promising areas for collaborative action in northern Arizona.

Results from the RHES further indicated that data-driven decision making is highly valued among participating leaders, but most found data to be outdated or unavailable or they worked in an environment in which expertise to analyze data was lacking. Lack of access to health-related data is particularly salient in rural areas, such as northern Arizona [27]. Consequently, differential access to data has the potential to perpetuate and exacerbate health inequities in rural areas. As a local academic institution with expertise in methods of data collection and analysis, NAU is poised to collaborate with community partners to improve their ability to access and use high-quality data to inform decisions regarding health equity. The NAU SHERC has already taken steps toward this goal by inviting community partners to participate in monthly research methods workshops.

Finally, this study illustrated the potential utility of using a baseline assessment of organizational leaders to start a productive dialog on the various and unique ways in which each sector (eg, housing, transportation, justice, economic development, education, arts, and culture) can strengthen the health and well-being of their community. Following the Bay Area Regional Health Inequities Initiative framework [22], the RHES results will be used to guide research priority areas and practice and policy efforts of the SHERC, CHER, and NAU as a whole. Our next steps include (1) engaging our scientific and community advisory boards in the further interpretation of results and recommendations for research, practice, and policy; (2) hosting a series of regional community events to share survey results and dialog on the strategic next steps toward developing a regional health equity initiative that meaningfully engages stakeholders across sectors; and (3) promoting further community-campus collaboration through the development of community engagement studios [28].

Limitations

Although a community-engaged approach to survey development for health disparities research has clear benefits [16], it is important to recognize that it also comes with unique challenges. For example, it is difficult to develop a comprehensive yet concise survey on health equity that includes the diverse voices of the community and limits the burden on the participant. In this study, 27.6% (57/206) of leaders who started our survey answered only 30% or less of the survey questions, indicating that survey length may have been a deterrent for some individuals. Furthermore, respondent-driven sampling methods were used to recruit participants for the RHES, whereby county leaders were asked to nominate other colleagues to complete the survey. Although this approach helped us to compile a comprehensive list of local leaders, the findings are specific to the region of northern Arizona and should not be generalized to other areas of the United States. However, we believe that our survey development and implementation procedures can be used as a model for other institutions to conduct similar efforts in their communities. Importantly, we also acknowledge the lack of racial and ethnic

diversity in our respondents; however, we are uncertain if this is a limitation of our recruitment strategy or a true reflection of the lack of diversity in leadership in northern Arizona. Future studies can examine the health priorities among the diverse community members of northern Arizona to ascertain if their health priorities align with the priorities of those in leadership roles.

Conclusions

Without a clear consensus on the root causes of health equity and greater cross-sectoral collaboration, the development of

effective policy and practice objectives aimed at reducing health disparities and improving health equity will be limited [29]. As anchor institutions, local universities are well positioned to help lead multisector work aimed at eliminating health disparities and making advancements in the promotion of health equity [3]. In this study, we outlined the steps for engaging multisectoral leaders in survey development and distribution as a promising first step toward developing meaningful multisectoral collaborations across a diverse region of the southwestern United States.

Acknowledgments

Numerous people made significant contributions to the development and implementation of the work documented in this paper. In particular, the authors would like to thank the following members of our CAC: Amanda Guay, Emma Torres, Chelsey Donohoo, Emily Davalos, Candida Hunter, Stephen Julian, Mare Schumacher, Joyce Hamilton, Diana Gomez, Eric Wolverton, and Shepard Tsosie. This work was financially supported by the National Institute of Health (grant 1U54MD012388).

Conflicts of Interest

None declared.

References

1. Feachem R. Poverty and inequity: a proper focus for the new century. *Bull World Health Organ* 2000;78(1):1-2 [FREE Full text] [Medline: 10686728]
2. Centers for Disease Control and Prevention. Healthy people 2010. National Center for Health Statistics. 2015. URL: https://www.cdc.gov/nchs/healthy_people/hp2010.htm [accessed 2020-09-01]
3. National Academies of Sciences, Engineering, and Medicine, Health and Medicine Division, Board on Population Health and Public Health Practice, Committee on Community-Based Solutions to Promote Health Equity in the United States. *Communities in action: pathways to health equity*. In: Weinstein JN, Geller A, Negussie Y, Baciu A, editors. National Academies of Sciences, Engineering, and Medicine. Washington, DC: National Academies Press; 2017:1-582.
4. Centers for Disease Control and Prevention. Healthy people 2020. National Center for Health Statistics. 2010. URL: https://www.cdc.gov/nchs/healthy_people/hp2020.htm [accessed 2020-09-01]
5. Chandra A, Acosta J, Carman KG, Dubowitz T, Leviton L, Martin LT, et al. Building a national culture of health: background, action framework, measures, and next steps. *Rand Health Q* 2017 Jan;6(2):3 [FREE Full text] [Medline: 28845341]
6. Lavizzo-Mourey R. How to build a culture of health. *Q J Med* 2017 Feb 01;110(2):59-60. [doi: 10.1093/qjmed/hcw192] [Medline: 27803365]
7. Koh H, Piotrowski J, Kumanyika S, Fielding J. Healthy people: a 2020 vision for the social determinants approach. *Health Educ Behav* 2011 Dec;38(6):551-557. [doi: 10.1177/1090198111428646] [Medline: 22102542]
8. Crosby R, Wendel M, Vanderpool R, Casey B. *Rural Populations and Health: Determinants, Disparities, and Solutions*. San Francisco, California, United States: Jossey-Bass; 2012:1-416.
9. Singh GK, Siahpush M. Widening rural-urban disparities in life expectancy, U.S., 1969-2009. *Am J Prev Med* 2014 Feb;46(2):19-29. [doi: 10.1016/j.amepre.2013.10.017] [Medline: 24439358]
10. Salunke S, Lal D. Multisectoral approach for promoting public health. *Indian J Public Health* 2017;61(3):163. [doi: 10.4103/ijph.ijph_220_17]
11. Health in all policies training manual. World Health Organization. 2015. URL: <https://www.who.int/life-course/news/health-in-all-policies/en/> [accessed 2021-04-24]
12. Better health through equity: case studies in reframing public health work. American Public Health Association. 2015. URL: https://www.apha.org/~media/files/pdf/topics/equity/equity_stories.ashx [accessed 2021-04-24]
13. Williams DR, Collins C. Racial residential segregation: a fundamental cause of racial disparities in health. *Public Health Rep* 2001 Sep;116(5):404-416. [doi: 10.1016/s0033-3549(04)50068-7]
14. Sampson U, Kaplan R, Cooper R, Diez Roux AV, Marks JS, Engelgau MM, et al. Reducing health inequities in the U.S.: recommendations from the NHLB's health inequities think tank meeting. *J Am Coll Cardiol* 2016 Aug 02;68(5):517-524 [FREE Full text] [doi: 10.1016/j.jacc.2016.04.059] [Medline: 27470459]
15. Sabo S, Remiker M, Jimenez D, Samarron-Longorio A, Chief C, Williamson H, et al. A regional health equity survey report: building capacity to address health equity in Northern Arizona. Northern Arizona University. URL: https://nau.edu/wp-content/uploads/sites/160/RHES_county-report_FINAL.pdf [accessed 2021-04-24]

16. Kelley A, Piccione C, Fisher A, Matt K, Andreini M, Bingham D. Survey development: community involvement in the design and implementation process. *J Public Health Manag Pract* 2019;25 Suppl 5, Tribal Epidemiology Centers: Advancing Public Health in Indian Country for Over 20 Years:77-83 [FREE Full text] [doi: [10.1097/PHH.0000000000001016](https://doi.org/10.1097/PHH.0000000000001016)] [Medline: [31348193](https://pubmed.ncbi.nlm.nih.gov/31348193/)]
17. Advancing wellbeing in Northern Arizona: a regional health equity assessment. Center for Health Equity Research, Northern Arizona University. 2017. URL: https://nau.edu/wp-content/uploads/sites/160/2018/09/Wellbeing-in-Arizona_Final-9_22_17.pdf [accessed 2021-04-24]
18. Census. United States Census Bureau. URL: <https://www.census.gov/> [accessed 2020-10-22]
19. Newman S, Andrews J, Magwood G, Jenkins C, Cox M, Williamson D. Community advisory boards in community-based participatory research: a synthesis of best processes. *Prev Chronic Dis* 2011 May;8(3):A70 [FREE Full text] [Medline: [21477510](https://pubmed.ncbi.nlm.nih.gov/21477510/)]
20. Taylor P, Fransman J. Learning and teaching participation : exploring the role of Higher Learning Institutions as agents of development and social change. Institute of Development Studies. 2004. URL: <https://opendocs.ids.ac.uk/opendocs/handle/20.500.12413/4014> [accessed 2021-04-24]
21. Ausubel DP. *The Acquisition and Retention of Knowledge: A Cognitive View*. Netherlands: Springer; 2000:1-212.
22. Organizational self assessment toolkit. The Bay Area Regional Health Inequity Initiative. URL: <https://www.barhi.org/organizational-self-assessment-tool> [accessed 2021-04-24]
23. Qualtrics XM - Experience Management Software. 2015. URL: <https://www.qualtrics.com/> [accessed 2020-10-22]
24. Improving well-being in Arizona by addressing root causes and broader issues that affect health. Vitalyst Health Foundation. 2018. URL: <http://vitalysthealth.org/> [accessed 2020-10-22]
25. Ndumbe-Eyoh S, Moffatt H. Intersectoral action for health equity: a rapid systematic review. *BMC Public Health* 2013 Nov 09;13(1):1056 [FREE Full text] [doi: [10.1186/1471-2458-13-1056](https://doi.org/10.1186/1471-2458-13-1056)] [Medline: [24209299](https://pubmed.ncbi.nlm.nih.gov/24209299/)]
26. Narain K, Zimmerman F, Richards J, Fielding JE, Cole BL, Teutsch SM, et al. Making strides toward health equity: the experiences of public health departments. *J Public Health Manag Pract* 2019;25(4):342-347. [doi: [10.1097/PHH.0000000000000852](https://doi.org/10.1097/PHH.0000000000000852)] [Medline: [31136507](https://pubmed.ncbi.nlm.nih.gov/31136507/)]
27. Greenberg AJ, Haney D, Blake KD, Moser RP, Hesse BW. Differences in access to and use of electronic personal health information between rural and urban residents in the United States. *J Rural Health* 2018 Feb 11;34 Suppl 1:30-38 [FREE Full text] [doi: [10.1111/jrh.12228](https://doi.org/10.1111/jrh.12228)] [Medline: [28075508](https://pubmed.ncbi.nlm.nih.gov/28075508/)]
28. Joosten YA, Israel TL, Williams NA, Boone LR, Schlundt DG, Mouton CP, et al. Community engagement studios. *Acad Med* 2015;90(12):1646-1650. [doi: [10.1097/acm.0000000000000794](https://doi.org/10.1097/acm.0000000000000794)]
29. Mooney G. What does equity in health mean? *World Health Stat Q* 1987;40(4):296-303. [Medline: [3433797](https://pubmed.ncbi.nlm.nih.gov/3433797/)]

Abbreviations

CAC: community advisory council
CEC: community engagement core
CHER: Center for Health Equity Research
MSA: multisectoral approach
NAU: Northern Arizona University
RCMI: Research Centers in Minority Institution
RHEA: Regional Health Equity Assessment
RHES: Regional Health Equity Survey
SDOH: social determinants of health
SHERC: Southwest Health Equity Research Collaborative

Edited by G Eysenbach; submitted 06.11.20; peer-reviewed by D Yeung; comments to author 19.12.20; revised version received 25.02.21; accepted 16.04.21; published 12.05.21.

Please cite as:

Remiker M, Sabo S, Jiménez D, Samarron Longorio A, Chief C, Williamson H, Teufel-Shone N
Using a Multisectoral Approach to Advance Health Equity in Rural Arizona: Community-Engaged Survey Development and Implementation Study
JMIR Form Res 2021;5(5):e25577
URL: <https://formative.jmir.org/2021/5/e25577>
doi: [10.2196/25577](https://doi.org/10.2196/25577)
PMID: [33978596](https://pubmed.ncbi.nlm.nih.gov/33978596/)

©Mark Remiker, Samantha Sabo, Dulce Jiménez, Alexandra Samarron Longorio, Carmenlita Chief, Heather Williamson, Nicolette Teufel-Shone. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 12.05.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Reporting of Differences in Taste Between Branded and Unbranded Cigarettes by Smokers Blinded to Cigarette Branding: Within-Person, Randomized Crossover Study

Nasser F BinDhim^{1,2,3}, PhD; Nora A Althumiri¹, MSc; Mada H Basyouni^{1,4}, MPhil; Rasha A Almubark¹, MSc; Zaid Alkhamaali^{1,3}, PhD; Weam Banjar⁴, MSc; Mohammed Zamakhshary², MD; Khaled M AlKattan², MD

¹Sharik Association for Health Research, Riyadh, Saudi Arabia

²College of Medicine, Alfaisal University, Riyadh, Saudi Arabia

³Saudi Food and Drug Authority, Riyadh, Saudi Arabia

⁴Ministry of Health, Saudi Arabia, Riyadh, Saudi Arabia

Corresponding Author:

Nora A Althumiri, MSc

Sharik Association for Health Research

4 Anas ibn Malik

Riyadh, 11778

Saudi Arabia

Phone: 966 505435544

Email: nora@althumiri.net

Abstract

Background: Saudi Arabia implemented a plain tobacco packaging regulation, one of the World Health Organization's recommended initiatives to help reduce smoking rates, in August 2019. A few weeks after implementation, a large number of smokers complained via various media channels, especially social media (eg, Twitter), that an extreme change in cigarette taste had occurred, frequency of coughing had increased, and for some, shortness of breath had led to hospitalization.

Objective: The main objective is to determine whether smokers blinded to cigarette branding report differences in taste between branded and unbranded cigarettes. The secondary objective is to observe the frequency of immediate cough or shortness of breath.

Methods: This study employed a within-person, randomized crossover design that recruited current smokers 18 years and older who were cleared upon physical assessment before the experiment. Participants received 6 sequences of different random exposures (3 puffs) to 3 plain-packaged cigarettes (2 from their favorite brand and 1 from another brand as a control) and 3 branded cigarettes (2 from the favorite brand and 1 from another brand as a control). Participants wore virtual reality goggles accompanied by special software to alter visual reality and gloves to alter the touch sensation.

Results: This study recruited 18 participants, measured at 6 time points, to produce 108 experiments. Participants were not able to identify the correct type of cigarettes (plain or branded, estimate of fixed effect=-0.01, $P=.79$). Moreover, there were no differences in the ability of the participants to identify their favorite brand ($t_{107}=-0.63$, mean 0.47, $P=.53$). In terms of immediate coughing, out of the 108 experiments, 1 episode of short coughing was observed, which was attributed to the branded cigarette, not the plain-packaged cigarette.

Conclusions: After controlling the visual and touch sensations, participants were not able to differentiate between branded and plain-packaged cigarettes in terms of taste or inducing immediate shortness of breath or cough. Interestingly, participants were not able to identify their favorite brand.

(*JMIR Form Res* 2021;5(5):e24446) doi:[10.2196/24446](https://doi.org/10.2196/24446)

KEYWORDS

smoking; plain packaging; sensory; Saudi Arabia; tobacco; virtual reality; cigarettes

Introduction

Overview

The prevalence of cigarette smoking among adults in Saudi Arabia is between 16.1% and 21.1% [1-3]. Saudi Arabia implemented a plain tobacco packaging regulation in late August 2019. A few weeks after implementation, a huge number of smokers complained via various media channels (especially social media) that an extreme change in cigarette taste had occurred, frequency of coughing had increased, and for some, shortness of breath had led to hospitalization. The taste claims have arisen in other countries that have implemented plain packaging, such as Australia and the United Kingdom [4-6].

These complaints persisted for more than 90 days, starting in mid-November 2019 and continuing until the writing of the first draft of this manuscript, on March 2020. Rumors that the cigarettes currently sold in Saudi Arabia in plain packaging have toxic chemicals or other non-tobacco substances intended to harm smokers are circulating widely in the media and by word of mouth [7,8].

Saudi authorities requested that tobacco companies declare any changes of their cigarette content beyond the new plain-packaging requirements. British American Tobacco Middle East and Philip Morris International released public statements declaring they had made no changes to their products' contents (Multimedia Appendix 1) [6].

Nevertheless, citizens and visitors in Saudi Arabia have been anxious and concerned about the health consequences of consuming the current plain-packaged cigarettes [7]. Consequently, prices of branded cigarettes tripled, and smuggling increased dramatically [7,8].

The results of chemical analysis and examination of conformity with Saudi tobacco product and safety standards have shown the new plain-packaged cigarettes are within the standards, and no unusual level of toxicity was found [7]. However, the claims made in the media about the taste, immediate coughing, and shortness of breath have not been investigated.

A few studies have investigated this issue in other countries that implemented plain packaging [9-11]. These studies did not find significant differences in taste, but they highlighted the difficulties of measuring this variable, which may affect the results. The main difficulty is in the method of measuring the difference between the branded and the plain-packaged cigarettes without exposing participants to the brand they are trying during the study. No previous study was fully able to blind the participants to the cigarette branding, although the senses are known to affect the taste.

How Does Changing Vision and Touch Senses Affect Taste?

The stimulation of one sense organ influences, to some extent, the sensitivity of organs of another sense [12]. Human beings have 5 senses that are each interlinked. Changing or altering the vision or touch senses affects taste. In fact, taste involves a combination of gustatory and olfactory stimuli [13]. Although vision is not directly related to taste, changing vision alters the

perception that a person might have about something, which leads to a change in its taste [14]. When an individual tastes something that they see visually, there is a chance the taste will change based on what the individual has registered in the brain. Booth et al emphasize the role of sight: "A mouthful usually stimulates sight first and then touch, taste and smell." [15] However, there is a possibility that one's vision may lead to an incorrect interpretation of taste. For example, changing only the color of a drink has a direct effect on its taste [16]. Therefore, changing appearance affects taste based on a person's visual perception [17]. For example, in one study on changing the color of fruit drinks, color was found to influence sweetness and intensity of a typical flavor [18]. Moir [19] published more than 150 studies examining the influence of vision on taste and flavor. The majority of this research showed that changing the hue and/or intensity of the color added to a food or a beverage can affect the perceived identity and/or intensity of the flavor [19].

Apart from vision, altering touch also affects taste. The touch system contributes to this constancy because taste sensations appear to be localized by touch [20]. Touch, whether by mouth or by hand, has a far greater influence on perceptions of taste, quality, and satiety than we realize [13]. The tip of the tongue is an area of high receptor density, so taste sensation is strongest at the tip but begins to weaken farther back on the tongue [20]. Thus, for years, researchers have investigated how the act of touching products affects consumer response. Phenomena such as the firmness of a cup in which water is served affecting consumers' judgments of the water's taste have been shown [17]. The evidence of the effect of vision and touch on perception and taste may affect smokers as well when the external or internal packaging of the cigarette is changed.

Effect of Prominent Pictorial Warnings

A known sensory factor that influences smokers' product acceptance and satisfaction is pictorial warnings on tobacco products [21]. The recently implemented plain tobacco packaging regulation in Saudi Arabia was accompanied by the implementation of a new set of pictorial warnings that was more prominent compared with the previously implemented ones. This change was the first in terms of changing pictorial warnings since their introduction in Saudi Arabia in August 2011 [21].

A previous study on a Saudi sample in Saudi Arabia investigated the effect of prominent pictorial warnings compared to the old pictorial warning used before the plain packaging change and showed that the two prominent sets of pictorial warnings scored, on average, 13.1 and 10.2 rating points higher than the old pictorial warnings in increasing participants' worries on the Brief Worry Scale about Smoking (BWS). The BWS is a 4-item scale that measures worry about physical health as a consequence of smoking [21]. Furthermore, on average, the two prominent sets of pictorial warnings scored 12.5 and 10.1 rating points higher than the old pictorial warnings in increasing participants' negative reactions on the Self-Assessment Manikin (SAM) [21]. The SAM is a 3-item, nonverbal pictorial assessment method that measures the pleasure, arousal, and dominance associated with a person's reaction to a wide variety of stimuli [21]. These results indicate the expected stronger

emotional evocation in response to the change in pictorial warnings in Saudi Arabia.

As mentioned above, a great deal of scientific research on sensory influences on taste and flavor exists in the food-related domain, but to our knowledge, no research addresses the subject in relation to smoking. Thus, this trial strives to address the following concerns in the most scientific and ethically possible ways: (1) Do smokers who are blinded to cigarette branding report differences in taste between branded and unbranded cigarettes? (2) Do smokers who are blinded to cigarette branding experience differences in immediate coughing between branded and unbranded cigarettes? (3) Do smokers who are blinded to cigarette branding experience differences in shortness of breath between branded and unbranded cigarettes?

In addition, this paper discusses the lesson learned from the Saudi experience with regard to this trial's findings and other available evidence.

Methods

Study Design

A prospective within-person, randomized crossover design was used to address the objectives of this study.

Exposure and Procedure

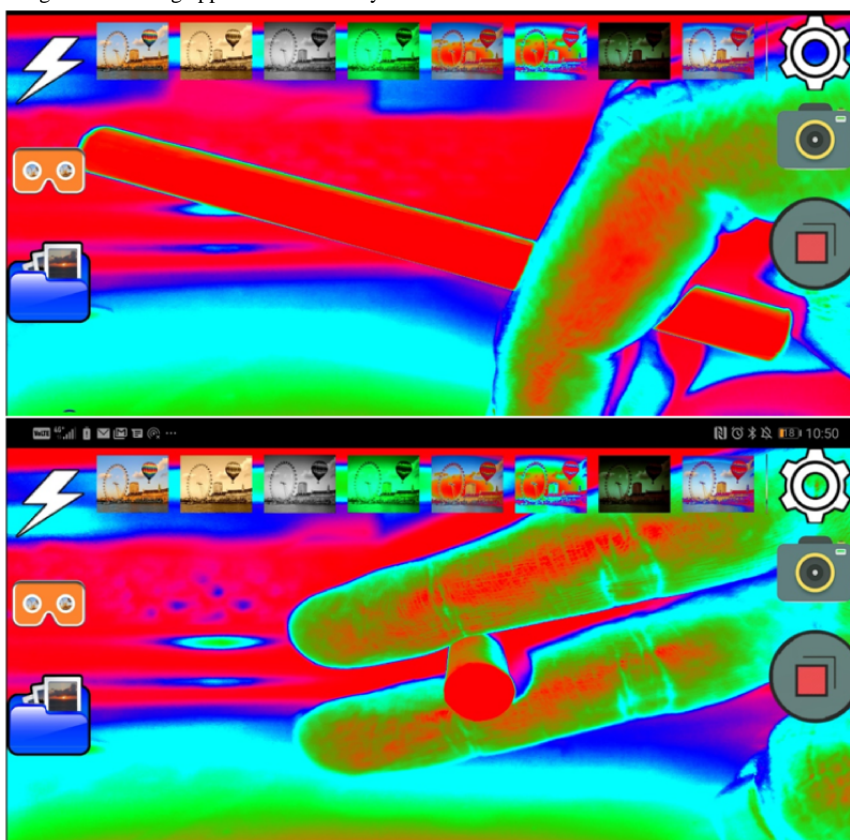
Participants received 6 sequences of different random exposures to 3 plain-packaged cigarettes (2 from the favorite brand and 1 from another brand as a control) and 3 branded cigarettes (2 from the favorite brand and 1 from another brand as a control), with a washout period of 5-10 minutes between each cigarette. To ensure reliability and reproducibility of the results, and due to the small sample size, all participants were invited to repeat the trial on another day.

The standard cigarette in Saudi Arabia is approximately 5.5 cm long, excluding the filter. To reduce harm to participants, they were exposed to approximately 3 puffs, which represents around 2 cm of cigarette. To ensure the participants did not exceed this amount, aluminum foil was wrapped around the rest of the cigarette. The use of 3 puffs or 2 cm was determined by the researchers to provide participants with a sufficient opportunity to judge the taste.

To ensure concealment, the cigarettes were provided in random order by a blinding handler who was involved in neither data collection nor data analysis.

To blind participants from recognizing the type of cigarette, each participant wore virtual reality goggles, accompanied by special software to alter the visual reality of the smoker (Figure 1). In addition, the participants wore medical gloves to alter the feeling of touching the cigarettes with their hands.

Figure 1. Screenshots of the cigarette blinding app for virtual reality.



Inclusion and Exclusion Criteria

Participants were current cigarette smokers 18 years or older who were cleared upon physical assessment before the experiment.

Candidates who were planning to quit smoking or in a quitting stage were not eligible to participate in this study, thereby reducing harm and ethically avoiding an alteration to their quitting process or plan. Candidates who had respiratory or cardiac disease, a taste-related disorder (eg, hypogeusia or ageusia), or an acute disease that affected taste or smell (eg, flu) were not eligible. Candidates who showed abnormal vital signs, which were checked before the study (fever, shortness of breath, or elevated blood pressure), were not eligible.

Recruitment

Candidates were invited from the Sharik research participants' database [22], which includes around 6000 smokers. Candidates received a phone call, and study information and consent were presented to them. If they agreed to participate, then their eligibility conditions were checked. If eligible, then an appointment was booked for them at the study site. Once the participants arrived, the researchers explained the details of the study to them, and written participant information sheets were provided. Participants who wanted to start the study were asked to sign the consent forms. Then, their vital signs were checked. Participant recruitment started in early January 2020 via phone interviews.

Data Collection and Outcome Measure

Demographics and Baseline

Data collection started with a quick interview survey that gathered their age, gender, Fagerstrom Test for Nicotine Dependence score, age they started smoking, and frequency of coughing that lasted 2 weeks.

Taste Change

The main outcome measure for the taste was taken by asking the participants to identify the type of cigarette they had smoked (plain or branded) based on taste. In addition, each smoker was asked to rate the taste of the cigarettes on a scale from 1 (*very bad taste*) to 7 (*very good taste*) and the heat perception of the cigarette smoke (burning sensation) from 1 (*acceptable*) to 7 (*unacceptable*). Finally, each smoker was asked whether the cigarette was his or her favorite brand.

Cough and Shortness of Breath

The outcome measure for immediate coughing was an observation of any coughing event during the smoking or washout period for each cigarette.

The outcome measure for shortness of breath was measured via self-reporting and peak flow test. We used an approved medical-grade electronic flow test device, following the recommended standard for performing the flow test [23]. Each participant received instructions and performed the flow test accurately before starting the experiment. The participants repeated the test before and after each cigarette for safety reasons. However, the main measure here was the comparison

between plain and branded cigarettes because it was delivered in random order for each participant.

Shortness of breath was defined on the peak flow if the reading was 40% less than the baseline before starting the first cigarette because readings of 50% less than the baseline are defined as the signal for medical alert [24].

After completing the experiment, the participants were asked about their perceptions of the taste and health concern claims circulated in the media about plain-packaged tobacco for comparison to their initial opinion before the experiment. The participants were asked, "Do you believe that the taste and content of the new plain-packaged cigarettes have changed compared to the old, branded cigarettes?"

Sample Size

Based on the smokers' complaints across media channels, the difference in taste between the plain and branded cigarettes seemed large to medium. Thus, a single-factor, repeated-measures design with a sample of 18 subjects, measured at 6 time points to produce 108 experiments, achieves 80% power to detect a contrast using a multivariate T^2 test at a .05 significance level at 0.45-0.35 effect size [25,26]. Effect size was selected based on Cohen d medium effect size due to the lack of sufficient information to calculate a more accurate effect size [27]. Sample size was calculated via PASS 2019, version 19.0.3 (NCSS), using the abovementioned inputs.

Data Analysis

Descriptive statistics were used to describe the sample demographics, and mixed-model analysis was used to analyze repeated measures of taste-related outcomes to account for within-person differences. A t test was used to analyze the ability of participants to identify their favorite brands, and t tests were used to analyze the follow-up data due to the small sample size, which prevented the use of repeated-measures mixed-model analysis.

Ethical Considerations

The study was performed in agreement with the Declaration of Helsinki. The Alfaisal University Institutional Review Board granted ethical approval with approval number IRB-20013. All participants signed the consent form approved by the Alfaisal University Institutional Review Board to participate in this study.

Results

Demographics and Baseline

Twenty-five participants were approached; out of these, 1 participant was excluded because he was in a quitting stage, and 2 were excluded for having a cardiovascular disease. Of the eligible participants, 5 did not show up. The 18 participants included 1 woman (6%) and 17 men (94%). Their mean age was 28.9 years (range 19-63), mean nicotine dependency score was 3.3 (range 2-5), and mean number of cigarettes smoked per day was 18.3 (range 12-24). In the prior 2 weeks, 16 out of 18 participants (89%) did not have a coughing attack, and 2 participants had 1 or more coughing attacks.

In terms of participants' pre-experiment opinions about changes in the new plain-packaged cigarettes' taste and content, 16 out of 18 participants (89%) thought they detected a change compared to the old branded cigarettes. However, after the experiment, all participants reported that they had changed their opinion and did not believe any differences existed between plain-packaged and branded cigarettes.

Taste Change

Mixed-model analysis showed no significant differences in participants' ability (between and within participants) to identify the correct type of cigarettes (plain versus branded cigarettes): the estimate of fixed effect was -0.01 ($P=.79$) (Table 1).

Table 1. Summary of the results for the main questions of this study.

Question	Measure	Results	P value
Do smokers who are blinded to cigarette branding report differences in taste between branded and unbranded cigarettes?	Participants' ability to identify plain or branded cigarettes correctly	Estimate of fixed effect= -0.01	.79
Do smokers who are blinded to cigarette branding report differences in taste between branded and unbranded cigarettes?	Taste and burning sensation ratings	Estimate of fixed effect for taste= -0.31 ; estimate of fixed effect for burning sensation= -0.25	.30; .42
Do smokers who are blinded to cigarette branding report differences in taste between branded and unbranded cigarettes?	Participants' ability to identify their favorite brand compared to other brands used in the experiment	$t_{107}=-0.63$, mean 0.47	.53
Do smokers who are blinded to cigarette branding experience differences in immediate coughing between branded and unbranded cigarettes?	Coughing incidents	Out of the 108 experiments, 1 incident of short coughing was observed	N/A ^a
Do smokers who are blinded to cigarette branding have differences in shortness of breath between branded and unbranded cigarettes?	Shortness of breath incidents	Out of the 108 experiments, 0 incidents of short coughing were observed	N/A

^aN/A: not applicable.

In terms of taste and burning sensation ratings, no significant differences were observed in the rating scores between the plain and the branded cigarettes, between and within participants: the estimate of fixed effect for taste was -0.31 ($P=.30$), and the estimate of fixed effect for burning sensation was -0.25 ($P=.42$).

Finally, no differences were seen in the same participants' ability to identify their favorite brand versus another nonfavorite brand ($t_{107}=-0.63$, mean 0.47, $P=.53$).

Cough and Shortness of Breath

In terms of immediate coughing, out of the 108 experiments, 1 short coughing episode was observed, which was attributed to a branded cigarette, not a plain-packaged cigarette.

None of the participants in the 108 experiments reported shortness of breath. In addition, comparing the changes in peak flow reading between the first cigarettes, none of the participants had shortness of breath. Finally, no cases of shortness of breath were recorded overall after the full experiment for all participants.

Discussion

Summary of Findings

This study investigated the claims of taste change, immediate continuous coughing, and immediate shortness of breath allegedly caused by plain-packaged cigarettes. After controlling participants' visual and touch perceptions, no significant differences were observed in their ability to identify plain versus branded cigarettes, and more surprisingly, no significant differences were seen in their ability to identify their favorite brand versus nonfavorite brands. No alarming findings emerged

related to immediate cough or shortness of breath. Most of the study participants were men, because the prevalence of women smokers is very low (1.5%) compared to men smokers (26.2%) in Saudi Arabia [1,2].

Summary of the Issue Escalation

Smokers in Saudi Arabia have complained about the look, feel, and taste of the newly introduced plain-packaged cigarettes. This was also made worse by introducing the new and emotionally evocative pictorial warnings at the same time, which proved to signal large health concerns in the same population in a previous study [21].

Moreover, another issue in the Saudi implementation process was the lack of a public awareness campaign before implementation. Without prior notice, the arrival of new cigarette packaging and emotionally evocative pictorial warnings sent smokers a shock, which could have directed them to seek answers and create theories.

In addition, after the issue had escalated, the authorities lightly explained that the plain packaging was a new regulation, but this explanation only addressed the outer packaging, not the internal (paper and filter) changes or the new pictorial warning [7,8,28]. This response left the majority with unanswered questions and concerns.

Although tobacco companies then released statements (Multimedia Appendix 1) declaring no changes in the cigarette content except what was required by the new regulation, they did not provide consumers with details of the changes called for in the new regulation; thus, the dilemma was dragged out. Later, the authorities revised the plain-packaging standards to

allow tobacco companies to reintroduce cigarettes with their original branded (paper and filter) look, while keeping the external plain packaging, which was a victory for tobacco companies. This change is a reduction of the World Health Organization plain-packaging standard and may reduce the effectiveness and benefits of the plain-packaging strategy.

Finally, it is worth noting that sensory perception and sensory research are priorities within the tobacco industry because they have direct effects on commercial concerns [29]. Sensory aspects contribute to smoker satisfaction and tobacco product acceptance, and they play an important role in controlling cigarette-puffing behavior. Tobacco companies have capitalized on distinct sensory preferences across gender, age, and ethnic groups by tailoring products for specific populations [29]. This study provided evidence that with the use of virtual reality and gloves to blind participants, they were not able to differentiate between their favorite brand and other nonfavorite brands. This highlights the fact that the tobacco industry understands such a topic in more depth and detail than regulators generally do.

Overall, this issue has some important lessons to be considered in any future tobacco policy changes.

Reorganizing the Chain of Events and Lessons Learned

The implementation of plain packaging in Saudi Arabia has represented a major chain of decisions that led to the current situation, which almost ruined the implementation of an effective public health policy. The major factors are as follows: (1) lack of a preimplementation awareness campaign, which played a major role in convincing the consumers that the plain-packaged cigarettes were counterfeit; (2) introducing the new pictorial warnings at the same time, having previous knowledge that it could cause consumer anxiety, with no preimplementation awareness campaign of this change either; (3) lack of prior risk assessment of the potential negative effects of such implementation on stakeholders; (4) lack of awareness about the plain-packaging standards, including explanation of why the externals and internals of the products had changed and a delay in the response to consumers' concerns, especially health-related ones; and (5) lack of scientifically sound evidence to explain the claims of taste changes and health concerns.

The most prominent factor here was the lack of proper risk assessment before implementation, which could have highlighted the potential risks before the implementation and improved the decision-making process. The impact of the new changes was

a life change for a smoker unused to such strong public health actions. As one of the study participants described it, "I woke up one day and found that all cigarettes brands have the same look and feel. [I thought] they must be counterfeit, there was no other explanation."

The mode of implementation also underestimated the role of social media in spreading fear and rumors to counteract the implementation of the new regulation, especially with the known history of how tobacco industries strive to undermine tobacco control regulations [30,31]. The claims about the plain packaging, including taste and quality, have been raised in other countries that implemented plain-packaging regulations, such as Australia and the United Kingdom [6,9,10], and the assumption that it was predictable for such claims to be reused in counter-policy campaigns was also overlooked.

Instead of being proactive with the consumers and engaging with them to explain the new regulation, the authorities maintained their silence for a long time, with a few contradicting announcements that started by denying any changes in physical and chemical content, then confirming the changes in quality and taste and blaming tobacco companies [7,28]. Tobacco companies, in return, stated they had made no changes. As a result, consumers did not know what had changed, how, or why. None of the statements by authorities or tobacco companies explained to consumers why all cigarettes looked the same (paper and filter) or why packets had new and emotionally provocative pictorial warnings.

Unfortunately, the same repellent effects (evoking smokers' health concerns and removing the appeal and charm of brands) that reflected the desire to encourage smokers to quit and to prevent nonsmokers from smoking, when introduced in this order and rushed into implementation, caused unexpected negative effects that neutralized the policy.

Conclusion

This study afforded a new dimension in understanding why some smokers find the taste of plain-packaged cigarettes differs from the branded cigarettes. After controlling the visual and touch sensations, participants were not able to differentiate between branded and plain-packaged cigarettes in terms of taste or inducing immediate shortness of breath or cough. Interestingly, participants were not able to identify their favorite brand.

Acknowledgments

This research project did not receive funding from a public or private entity during the project execution. The researchers provided all the research efforts in this study voluntarily. Open access publication-related cost was covered by Al-Dawaa Medical Services Co (DMSCO).

Authors' Contributions

All authors provided major contributions to the design and execution of this study. NFD designed the study and drafted the manuscript. MHB, NAA, and RAA administered and managed the study and data collection. MZ and KMA contributed to the study design and reviewed the manuscript. WB finalized and edited the manuscript for publication. All authors approved the final version for publication. NFD and ZA revised the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Tobacco companies' public statements declaring no changes to their products' content.

[[PDF File \(Adobe PDF File\), 97 KB - formative_v5i5e24446_app1.pdf](#)]

References

1. Algabbani AM, Alzubair R, Althumiri N, Alqahtani A, BinDhim N. The Prevalence of Cigarette Smoking in Saudi Arabia in 2018. *FDRSJ* 2018 Dec 12;1(1):1. [doi: [10.32868/rsj.v1i1.22](https://doi.org/10.32868/rsj.v1i1.22)]
2. Algabbani AM, Althumiri NA, Almarshad AM, BinDhim NF. National prevalence, perceptions, and determinants of tobacco consumption in Saudi Arabia. *FDRSJ* 2019 Nov 06;2(2):1. [doi: [10.32868/rsj.v2i2.33](https://doi.org/10.32868/rsj.v2i2.33)]
3. BinDhim NF, Althumairi NA, Basyouni MH. Sharik Health Indicators Surveillance System (HISS). 2020. URL: <https://sharikhealth.com/phm/results.php> [accessed 2020-10-01]
4. Miller CL, Brownbill AL, Dono J, Ettridge K. Presenting a strong and united front to tobacco industry interference : a content analysis of Australian newspaper coverage of tobacco plain packaging 2008-2014. *BMJ Open* 2018 Sep 17;8(9):e023485 [FREE Full text] [doi: [10.1136/bmjopen-2018-023485](https://doi.org/10.1136/bmjopen-2018-023485)] [Medline: [30224400](https://pubmed.ncbi.nlm.nih.gov/30224400/)]
5. Moodie C, Angus K, Mitchell D, Critchlow N. How tobacco companies in the United Kingdom prepared for, and responded to, standardised packaging of cigarettes and rolling tobacco. *Tob Control* 2018 Jul;27(e1):e85-e92. [doi: [10.1136/tobaccocontrol-2017-054011](https://doi.org/10.1136/tobaccocontrol-2017-054011)] [Medline: [29321273](https://pubmed.ncbi.nlm.nih.gov/29321273/)]
6. Evans-Reeves KA, Hiscock R, Lauber K, Gilmore AB. Prospective longitudinal study of tobacco company adaptation to standardised packaging in the UK: identifying circumventions and closing loopholes. *BMJ Open* 2019 Sep 24;9(9):e028506 [FREE Full text] [doi: [10.1136/bmjopen-2018-028506](https://doi.org/10.1136/bmjopen-2018-028506)] [Medline: [31551373](https://pubmed.ncbi.nlm.nih.gov/31551373/)]
7. New cigarette packs free of sawdust, SFDA confirm.: *Saudi Gazette*; 2019. URL: <http://saudigazette.com.sa/article/583828> [accessed 2020-01-01]
8. Saudi Arabia warns firms over cigarette quality.: *Arab News*; 2020. URL: <https://www.arabnews.com/node/1608481/saudi-arabia> [accessed 2020-01-01]
9. Guillaumier A, Bonevski B, Paul C, Durkin S, D'Este C. Socioeconomically disadvantaged smokers' ratings of plain and branded cigarette packaging: an experimental study. *BMJ Open* 2014 Feb 06;4(2):e004078 [FREE Full text] [doi: [10.1136/bmjopen-2013-004078](https://doi.org/10.1136/bmjopen-2013-004078)] [Medline: [24503299](https://pubmed.ncbi.nlm.nih.gov/24503299/)]
10. Maynard OM, Leonards U, Attwood AS, Bauld L, Hogarth L, Munafò MR. Effects of first exposure to plain cigarette packaging on smoking behaviour and attitudes: a randomised controlled study. *BMC Public Health* 2015 Mar 13;15:240 [FREE Full text] [doi: [10.1186/s12889-015-1586-8](https://doi.org/10.1186/s12889-015-1586-8)] [Medline: [25880278](https://pubmed.ncbi.nlm.nih.gov/25880278/)]
11. Moodie C, Mackintosh AM, Hastings G, Ford A. Young adult smokers' perceptions of plain packaging: a pilot naturalistic study. *Tob Control* 2011 Sep;20(5):367-373. [doi: [10.1136/tc.2011.042911](https://doi.org/10.1136/tc.2011.042911)] [Medline: [21752795](https://pubmed.ncbi.nlm.nih.gov/21752795/)]
12. Allen F, Schwartz M. The Effect of Stimulation of the Senses of Vision, Hearing, Taste, and Smell Upon the Sensibility of the Organs of Vision. *J Gen Physiol* 1940 Sep 20;24(1):105-121 [FREE Full text] [doi: [10.1085/jgp.24.1.105](https://doi.org/10.1085/jgp.24.1.105)] [Medline: [19873196](https://pubmed.ncbi.nlm.nih.gov/19873196/)]
13. Spence C. Multisensory flavor perception. *Cell* 2015 Mar 26;161(1):24-35 [FREE Full text] [doi: [10.1016/j.cell.2015.03.007](https://doi.org/10.1016/j.cell.2015.03.007)] [Medline: [25815982](https://pubmed.ncbi.nlm.nih.gov/25815982/)]
14. Korsmeyer C. Making sense of taste: Food and philosophy. New York: Cornell University Press; 2002.
15. Booth DA, Sharpe O, Freeman RPJ, Conner MT. Insight into sight, touch, taste and smell by multiple discriminations from norm. *Seeing Perceiving* 2011;24(5):485-511. [doi: [10.1163/187847511X588773](https://doi.org/10.1163/187847511X588773)] [Medline: [21902879](https://pubmed.ncbi.nlm.nih.gov/21902879/)]
16. Spence C, Levitan CA, Shankar MU, Zampini M. Does Food Color Influence Taste and Flavor Perception in Humans? *Chem. Percept* 2010 Mar 9;3(1):68-84. [doi: [10.1007/s12078-010-9067-z](https://doi.org/10.1007/s12078-010-9067-z)]
17. Krishna A, Morrin M. Does Touch Affect Taste? The Perceptual Transfer of Product Container Haptic Cues. *J Consum Res* 2008 Apr 01;34(6):807-818. [doi: [10.1086/523286](https://doi.org/10.1086/523286)]
18. Bayarri S, Calvo C, Costell E, Durán L. Influence of Color on Perception of Sweetness and Fruit Flavor of Fruit Drinks. *Food Sci Technol Int* 2016 Aug 26;7(5):399-404. [doi: [10.1106/jjwn-ffrq-jbmc-lq5r](https://doi.org/10.1106/jjwn-ffrq-jbmc-lq5r)]
19. Moir HC. Some observations on the appreciation of flavour in foodstuffs. *J Chem Technol Biotechnol* 1936 Feb 21;55(8):145-148. [doi: [10.1002/jctb.5000550803](https://doi.org/10.1002/jctb.5000550803)]
20. Todrank J, Bartoshuk LM. A taste illusion: taste sensation localized by touch. *Physiol Behav* 1991 Nov;50(5):1027-1031. [doi: [10.1016/0031-9384\(91\)90432-n](https://doi.org/10.1016/0031-9384(91)90432-n)] [Medline: [1805264](https://pubmed.ncbi.nlm.nih.gov/1805264/)]
21. BinDhim NF, McGeechan K, Alanazi AKT, Alanazi HMS, Alanazi SAJ, Al-Hadlaq SM, et al. Evaluating the pictorial warnings on tobacco products in Arabian Gulf countries against other international pictorial warnings. *Tob Control* 2017 May 12:261-265. [doi: [10.1136/tobaccocontrol-2016-053323](https://doi.org/10.1136/tobaccocontrol-2016-053323)] [Medline: [28500119](https://pubmed.ncbi.nlm.nih.gov/28500119/)]
22. Sharik Association for Health Research (SharikHealth). 2015. URL: <https://sharikhealth.com/> [accessed 2020-01-01]
23. NHS. Peak flow test. 2018. URL: <https://www.nhs.uk/conditions/peak-flow-test/> [accessed 2021-05-04]

24. American Lung Association. Measuring Your Peak Flow Rate. URL: <https://www.lung.org/lung-health-and-diseases/lung-disease-lookup/asthma/living-with-asthma/managing-asthma/measuring-your-peak-flow-rate.html> [accessed 2021-05-04]
25. Maxwell S, Delaney H, Kelley K. Designing Experiments and Analyzing Data: A Model Comparison Perspective. New York: Routledge; 2017.
26. Davis CS. Statistical methods for the analysis of repeated measurements. New York: Springer Science & Business Media; 2002.
27. Cohen J. Statistical power analysis for the behavioral sciences. Cambridge: Academic press; 2013.
28. Tobacco companies ordered to deal with difference in flavor.: Saudi Gazette; 2020. URL: <https://saudigazette.com.sa/article/586265> [accessed 2021-05-04]
29. Carpenter CM, Wayne GF, Connolly GN. The role of sensory perception in the development and targeting of tobacco products. *Addiction* 2007 Jan;102(1):136-147. [doi: [10.1111/j.1360-0443.2006.01649.x](https://doi.org/10.1111/j.1360-0443.2006.01649.x)] [Medline: [17207131](https://pubmed.ncbi.nlm.nih.gov/17207131/)]
30. Sebríe EM, Barnoya J, Pérez-Stable EJ, Glantz SA. Tobacco industry successfully prevented tobacco control legislation in Argentina. *Tob Control* 2005 Oct;14(5):e2 [FREE Full text] [doi: [10.1136/tc.2005.011130](https://doi.org/10.1136/tc.2005.011130)] [Medline: [16183967](https://pubmed.ncbi.nlm.nih.gov/16183967/)]
31. Hiilamo H. Tobacco industry strategy to undermine tobacco control in Finland. *Tob Control* 2003 Dec;12(4):414-423 [FREE Full text] [doi: [10.1136/tc.12.4.414](https://doi.org/10.1136/tc.12.4.414)] [Medline: [14660780](https://pubmed.ncbi.nlm.nih.gov/14660780/)]

Abbreviations

BWS: Brief Worry Scale about Smoking

SAM: Self-Assessment Manikin

Edited by G Eysenbach; submitted 21.09.20; peer-reviewed by R Krukowski, M Chaiton, I Shubina; comments to author 13.11.20; revised version received 19.11.20; accepted 12.04.21; published 14.05.21.

Please cite as:

BinDhim NF, Althumiri NA, Basyouni MH, Almubark RA, Alkhamaali Z, Banjar W, Zamakhshary M, AlKattan KM

Reporting of Differences in Taste Between Branded and Unbranded Cigarettes by Smokers Blinded to Cigarette Branding: Within-Person, Randomized Crossover Study

JMIR Form Res 2021;5(5):e24446

URL: <https://formative.jmir.org/2021/5/e24446>

doi: [10.2196/24446](https://doi.org/10.2196/24446)

PMID: [33988511](https://pubmed.ncbi.nlm.nih.gov/33988511/)

©Nasser F BinDhim, Nora A Althumiri, Mada H Basyouni, Rasha A Almubark, Zaied Alkhamaali, Weam Banjar, Mohammed Zamakhshary, Khaled M AlKattan. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 14.05.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Physical Activity Patterns and Neighborhood Characteristics of First-Generation Latina Immigrants Living in Arizona: Cross-sectional Study

Rodney P Joseph¹, PhD; Sonia Vega-López^{2,3}, PhD; SeungYong Han⁴, PhD

¹Center for Health Promotion and Disease Prevention, Edson College of Nursing and Health Innovation, Arizona State University, Phoenix, AZ, United States

²College of Health Solutions, Arizona State University, Phoenix, AZ, United States

³Southwest Interdisciplinary Research Center, Arizona State University, Phoenix, AZ, United States

⁴Edson College of Nursing and Health Innovation, Arizona State University, Phoenix, AZ, United States

Corresponding Author:

Rodney P Joseph, PhD

Center for Health Promotion and Disease Prevention

Edson College of Nursing and Health Innovation

Arizona State University

500 N 3rd St

Phoenix, AZ, 85004

United States

Phone: 1 6024960772

Fax: 1 6024961128

Email: rodney.joseph@asu.edu

Abstract

Background: Metabolic diseases, including obesity and type 2 diabetes, are a major health concern for Latina immigrants. Performing regular aerobic physical activity (PA) is a lifestyle behavior associated with the prevention and control of these conditions. However, PA levels of most Latina immigrants are below national guidelines. Neighborhood environmental factors may influence the PA levels of adults, but limited research has explored associations between the neighborhood environment and PA levels among Latina immigrants.

Objective: The objective of this study was to explore the PA patterns of first-generation US Latina immigrants and how neighborhood environmental factors are related to those PA patterns.

Methods: Using a cross-sectional study design, 50 first-generation Latina immigrants completed the International Physical Activity Questionnaire (IPAQ) and the Neighborhood Scales Questionnaire, which assessed 6 perceived neighborhood factors: (1) walking environment, (2) aesthetic quality, (3) safety, (4) violence, (5) social cohesion, and (6) activities with neighbors. Median self-reported metabolic equivalent (MET)-minutes/week of PA were used to summarize domain-specific (ie, work, domestic/household, leisure, and transportation) and intensity-specific (ie, walking, moderate, vigorous, moderate to vigorous) PA patterns. Logistic regression examined associations between neighborhood factors and engaging in leisure-time PA (ie, dichotomous outcome of some versus no leisure-time PA), transportation PA (ie, dichotomous outcome of some versus no transportation PA), and meeting national PA guidelines (ie, dichotomous outcome of meeting versus not meeting guidelines).

Results: Preliminary analyses showed that 10 participants reported excessively high PA levels and 1 participant had incomplete PA data; these women were excluded from analyses based on IPAQ scoring guidelines. The remaining 39 participants (mean age 40.5 years; mean length of US residency 4.6 years) reported a median of 4512 MET-minutes/week of total PA. The majority of PA was acquired through domestic activities (median 2160 MET-minutes/week), followed by leisure-time PA (median 396 MET-minutes/week), transportation PA (median 198 MET-minutes/week), and work PA (0 MET-minutes/week). Intensity-specific PA patterns showed a median of 594 MET-minutes/week of walking activity and 3500 MET-minutes/week of moderate-to-vigorous PA. Logistic regression models indicated that the neighborhood factors of walking environment, aesthetic quality, and safety were positively associated with engaging in leisure-time PA (odds ratios of 5.95, 95% CI 1.49-23.74; 2.45, 95% CI 1.01-5.93; and 3.30, 95% CI 1.26-8.67, respectively) and meeting national PA guidelines (odds ratios of 4.15, 95% CI 1.13-15.18; 6.43,

95% CI 1.45-28.39; and 2.53, 95% CI 1.00-6.36, respectively). The neighborhood factors of violence, social cohesion, and activities with neighbors were not significantly associated with PA outcomes.

Conclusions: Although most participants met national PA guidelines (ie, ≥ 500 MET-minutes/week of moderate-to-vigorous PA), the majority of their PA was achieved through domestic activities, with limited leisure, transportation, and work PA. Given that leisure-time PA in particular plays a significant role in improving health outcomes, findings suggest that many Latina immigrants could benefit from a leisure-time PA intervention. Such interventions should consider neighborhood environmental influences, as these factors may serve as determinants of PA.

(*JMIR Form Res* 2021;5(5):e25663) doi:[10.2196/25663](https://doi.org/10.2196/25663)

KEYWORDS

emigrants and immigrants; physical activity; exercise; residence characteristics; female; metabolic disease; Latina; immigrants; emigrants; health outcomes

Introduction

Metabolic disease conditions are a major health concern for Latina immigrants. Findings from the Hispanic Community Health Study/Study of Latinos (N=16,415) indicate that 45% of first-generation Latina immigrants (ie, Latinas born outside of the United States) are obese [1] and 17% have type 2 diabetes [2]. In comparison, national surveys estimate the prevalence of these conditions as 38% and 7%, respectively, among non-Latina White women and 40% and 9% among the US population as a whole [3,4]. Important to the discussion on the prevalence of metabolic diseases among Latina immigrants is the pattern in which they develop. Upon initial entry into the United States, Latina immigrants have an equal-to-lower risk prevalence of obesity and diabetes compared with US-born Latinas and non-Latina Whites [5-7]. However, as duration of US residency increases, so does risk for developing obesity and diabetes conditions [2,5-10].

The high metabolic disease burden among Latina immigrants represents a major public health concern. Latinx immigrants are the largest immigrant group in the United States, accounting for approximately 40% of the total US immigrant population, and are projected to remain the largest immigrant population through at least the year 2055 [11]. Thus, a better understanding of lifestyle behaviors that may contribute to the high metabolic disease prevalence among first-generation Latina immigrants with longer duration of US residency is essential to developing interventions to address this public health concern.

Regular aerobic physical activity (PA) is an independent lifestyle behavior associated with the prevention and control of both obesity [12-14] and diabetes [15]. However, the PA levels of most Latina immigrants are below the national guidelines of 150 minutes/week of moderate-intensity PA [16-18]; data on this population are limited due to the limited number of studies that differentiate by generational status. Existing data also show a decline in PA as duration of US residency increases [19]. This decline in PA parallels the increase in metabolic disease risk observed among this population of US Latinas. Interventions designed to promote sustained high levels of PA may be key in addressing the metabolic disease prevalence among Latina immigrants.

The development of effective PA interventions requires researchers to have extensive understanding of the PA patterns

of Latina immigrants as well as the social and environmental factors that may influence PA behaviors. The purpose of this study was to explore the PA patterns of first-generation Latina immigrants and the neighborhood environments in which they reside in the metropolitan area of Phoenix, Arizona. This research was conducted as part of a formative process to inform the development of culturally tailored PA interventions for first-generation Latina immigrants. First, we explored domain-specific PA engagement (ie, occupational, household, transportation, and leisure-time PA). Given the limited research on this topic among recently immigrated first-generation Latinas [16,18,20] and emerging evidence suggesting that domain-specific PA differentially influences health outcomes (ie, high levels of leisure-time PA have the most profound impact on improving health outcomes [21-23], while high levels of occupational PA may have a limited-to-negative effect on improving health outcomes [21,24-26]), such research is necessary to gain an understanding of the PA patterns of recently immigrated US Latinas. Second, we examined the social and physical characteristics of the neighborhoods in which recently immigrated Latinas reside. We were interested in neighborhood characteristics because most people spend the majority of their nonworking time in or around their residential neighborhood. Thus, the neighborhood characteristics of Latina immigrants are likely to influence the types of PA promoted through a PA intervention [16,18,20].

Methods

Study Design, Setting, and Participant Recruitment

A cross-sectional study design with self-report survey data was used. Data were collected as part of a broader study examining how immigration and integration experiences of recently immigrated first-generation Latinas influence perceptions of and opportunities for PA. Participants were a convenience sample of Latinas recruited from the metropolitan area of Phoenix. Recruitment strategies included flyer advertisements and in-person recruitment at predominately Hispanic-serving community outreach centers, health clinics, religious institutions, and local businesses. The Phoenix metropolitan area includes the cities of Phoenix, Mesa, and Chandler and is predominately comprised of urban and suburban neighborhoods. The region has an estimated 4.8 million residents, with 30% (ie, 1.5 million) identifying as Hispanic or Latinx [27]. Among residents identifying as Hispanic or Latinx, 400,000 are foreign-born and

approximately 60,000 immigrated to the United States after 2010 [27]. Women were eligible for participation in this study if they (1) self-identified as Latina, (2) immigrated into the United States within the past 10 years, (3) were aged ≥ 18 years, and (4) self-reported an ability to read in English or Spanish. Only women immigrating into the United States within the previous 10 years were included because the purpose of the larger study from which data were collected was to explore how contemporary social and contextual factors, including the impact of more restrictive immigration policies enacted during the US presidency of Donald Trump, were related to PA engagement among recently immigrated Latinas.

Women interested in study participation completed an eligibility screening interview either in-person or via telephone (according to the recruitment method). Eligible women were then given the option to complete the study questionnaire packet in English, Spanish, or a combined English/Spanish language using a completion method of their choice. Methods of survey completion included online via Qualtrics or by an oral interview. All participants were provided with an informed consent document prior to completing the survey. Participants completing the survey online provided informed consent by selecting the “continue” button on the webpage displaying the informed consent document that was accompanied by the following statement: “By clicking the Continue button, I acknowledge that I am at least 18 years old and that I consent to conducting the survey online.” Verbal informed consent was obtained from participants completing the survey via the oral interview or pencil-and-paper format. After completing the survey, participants were provided a US \$25 gift card for study participation. All study procedures were approved by the Institutional Review Board of Arizona State University.

Measures

Demographics

Participant demographic characteristics were collected using a form developed for this study. Age, country of origin, and years of residence in the United States were obtained using the following open-ended questions: “How old are you?”, “Where were you born?”, and “How long have you been living in the United States?” Characteristics of primary language at home, monthly household income, education level, marital status, and employment status were asked using closed-ended questions modeled after items used in the 2017 Behavioral Risk Factor Surveillance System survey [28]. All demographic questions are available in [Multimedia Appendix 1](#).

Duration and Frequency of PA

PA was assessed using the long version of the International Physical Activity Questionnaire (IPAQ) [29]. This 27-item questionnaire, previously validated in both English and Spanish [29,30], asks respondents the duration (minutes/week) and frequency (days/week) of PA that they engage in according to 5 domains: (1) work, (2) transportation, (3) domestic/household, (4) leisure time, and (5) sitting. Responses were used to estimate the minutes/week that participants engaged in each activity. Self-reported minutes/week of activity were then weighted based on estimated energy expenditure (ie, metabolic equivalents

[METs]) to provide an estimate of total weekly PA volume (ie, MET-minutes/week), as well as an estimate of energy expenditure for time spent in domain- and intensity-specific (ie, walking, moderate, and vigorous intensity) activities. When respondents indicated that they did not perform activity in a given domain, the value for that domain was set to zero (eg, if a participant indicated that she did not perform paid or unpaid work outside of the home, the value for work-related PA was set to 0). All data were scored and reported according to guidelines published in 2005 [31] and were calculated as a continuous measure of MET-minutes/week.

Neighborhood Environment

The neighborhood environmental factors of walking environment, aesthetic quality, safety, violence, social cohesion, and activities with neighbors were assessed using the Neighborhood Scales Questionnaire [32]. Scales measuring neighborhood walking environment (7 items), aesthetic quality (5 items), safety (3 items), and social cohesion (4 items) have respondents rate their agreement with various statements using a 5-point Likert scale (ie, 1=strongly agree, 2=agree, 3=neutral, 4=disagree, and 5=strongly disagree). Example statements include “I often see other people walking in my neighborhood” (from the walking environment scale) and “I feel safe walking in my neighborhood, day or night” (from the safety scale). Scales assessing neighborhood violence (4 items) and activities with neighbors (4 items) have participants rate the frequency of specific events on a 4-point Likert scale (1=often, 2=sometimes, 3=rarely, and 4=never). Example items from these scales include “During the past 6 months, how often was there a fight in your neighborhood in which a weapon was used?” (from the violence scale) and “How often do you and other people in your neighborhood visit in each other’s homes or speak with each other on the street?” (from the “activities with neighbors” scale). All scales were scored individually by calculating the mean score of individual scale items. For ease of data interpretation, scores were reverse ordered, when appropriate, to indicate that higher scores are associated with higher walkability, aesthetic quality, safety, social cohesion, violence, and activities with neighbors. The Neighborhood Scales Questionnaire was developed and validated in both English and Spanish [32] and has established test-retest reliability (ie, intraclass correlation coefficients ranging from .60 to .88) [32]. Internal consistency estimates (ie, Cronbach alpha coefficients) for the scales ranged from .78 to .92 in this study, which are comparable to previous research [32,33].

Statistical Analysis

Descriptive statistics (ie, means, 95% confidence intervals, medians, interquartile ranges, and frequencies) were used to summarize participant demographic, PA, and neighborhood variables. A series of regression models were used to examine associations between neighborhood environmental factors and the PA outcomes of leisure-time PA, transportation PA, and overall moderate-to-vigorous PA. Associations between neighborhood factors and work and household PA were not examined because we lacked a theoretical rationale for why neighborhood factors would influence these PA variables. Ordinal least squares (OLS) regression models, controlling for

age and level of education, were used to examine the independent effect of each neighborhood factor on continuous PA outcomes in MET-minutes/week. Logistic regression, controlling for age, was used to examine associations between neighborhood factors and dichotomous PA outcomes of engaging in leisure-time PA (ie, engaging in 0 versus >0 MET-minutes/week of leisure-time PA), transportation PA (ie, engaging in 0 versus >0 MET-minutes/week of transportation PA), and meeting national PA guidelines (ie, engaging in <500 versus ≥500 MET-minutes/week of moderate-to-vigorous PA). Level of education as a control variable was not included in logistic models because of perfect prediction. As a result of the collinearity of neighborhood variables, associations between each neighborhood variable (ie, walking environment, aesthetic quality, safety, violence, social cohesion, and activities with neighbors) and PA outcomes were examined separately. Stata/SE version 16.0 (StataCorp) was used for data analysis.

Results

Participants

A total of 50 first-generation Latinas participated in the study. However, preliminary data cleaning revealed that 10 participants reported unreasonably high PA data (ie, the sum of all walking,

moderate, and vigorous PA time was greater than 960 minutes or 16 hours/day) and 1 participant had incomplete IPAQ data. These women were excluded from data analysis according to IPAQ scoring guidelines [31], resulting in a final sample size of 39 participants. Sensitivity analyses (ie, chi-square test for categorical variables and *t* test for age and duration of US residency) were conducted to explore demographic differences between participants excluded from outcome analyses (*n*=11) and those included. Results showed that the women excluded from the study had a lower education level than those included (*P*=.003). No other demographic differences were observed between women included in the study and those excluded.

Among the 39 women included in outcome analyses, the mean age was 40.5 (SD 4.3) years and the mean duration of US residence was 4.6 (SD 1.0) years. The majority of the participants were from Mexico (27/39, 69%), with the remaining participants from various Central and South America countries. Approximately half (22/39, 56%) of the women were married and the majority spoke Spanish exclusively at home (32/39, 82%). Based on data provided by the IPAQ, only 7 (18%) of the 39 participants reported performing paid or unpaid work outside of the home. Complete demographic characteristics are presented in [Table 1](#).

Table 1. Demographic characteristics of study participants included in the outcome analysis (N=39).

Characteristic	Value
Age (years), mean (95% CI)	40.5 (36.2-44.8)
Age (years), median (minimum; maximum)	39.0 (2.4; 78.0)
Duration in the United States (months), mean (95% CI)	55.6 (43.5-67.7)
Duration in the United States (months), median (minimum; maximum)	52.0 (5.0; 128.0)
Country of origin, n (%)	
Mexico	27 (69)
Venezuela	5 (13)
Colombia	4 (10)
Cuba	1 (3)
Dominican Republic	1 (3)
Nicaragua	1 (3)
Primary language spoken at home, n (%)	
English	1 (3)
English and Spanish	6 (15)
Spanish	32 (82)
Monthly household income, n (%)	
≤\$1000	4 (10)
\$1001-\$2000	13 (33)
\$2001-\$3000	7 (18)
\$3001-\$4000	6 (15)
>\$4000	2 (5)
Don't know	4 (10)
Refused to respond	3 (8)
Education, n (%)	
Elementary	2 (5)
Middle school	2 (5)
High school	6 (15)
Some college or university	11 (28)
University graduate or postgraduate	18 (46)
Marital status, n (%)	
Single or no partner	6 (15)
Married	22 (56)
Cohabiting	3 (8)
Separated	1 (3)
Divorced	2 (5)
Widowed	5 (13)
Employment, n (%)	
Unemployed/looking for work	3 (8)
Unemployed/not looking for work	1 (3)
Homemaker	12 (31)
Student	1 (3)
Retired	2 (5)

Characteristic	Value
Other	2 (5)
Missing	18 (46)

PA Outcomes

Table 2 provides a detailed overview of intensity- and domain-specific PA patterns of our sample. Participants reported a median of 4512 MET-minutes/week of total PA. Intensity-specific PA outcomes showed that the majority of this PA was achieved through moderate-intensity activity (median

2820 MET-minutes/week). Median self-reported MET-minutes/week of walking and vigorous-intensity PA were 594 and 0, respectively. Overall, 85% (33/39) of participants reported PA levels that met the requirements of the PA guidelines (ie, ≥ 500 MET-minutes/week of moderate-to-vigorous PA).

Table 2. Summary statistics of intensity- and domain-specific metabolic equivalent (MET)-minutes/week of physical activity (PA).

Outcome	Participants, n (%)	MET-minutes/week of PA	
		Mean (95% CI)	Median (IQR)
Total PA ^a	39 (100)	6128 (4277-7979)	4512 (7271)
PA by intensity			
Walking	39 (100)	963 (532-1394)	594 (1386)
Moderate	39 (100)	3842 (2696-4988)	2820 (4410)
Vigorous	39 (100)	1323 (545-2101)	0 (1440)
Moderate to vigorous ^b	39 (100)	5165 (3456-6874)	3500 (6060)
Achieved PA guidelines^c			
Met PA guidelines	33 (85)	236 (18-454)	270 (396)
Did not meet PA guidelines	6 (15)	7199 (5225-9174)	5040 (7778)
Work PA			
Participants reporting working outside of the home	7 (18)	6009 (1171-10,846)	4590 (10,536)
Work PA—full sample ^d	39 (100)	1079 (65-2092)	0 (0)
Transportation PA			
Participants reporting transportation PA	23 (59)	584 (365-803)	360 (594)
Transportation PA—full sample ^d	39 (100)	345 (188-501)	198 (453)
Domestic/household PA			
Participations reporting domestic/household PA	34 (87)	3514 (2497-4531)	2850 (4024)
Domestic/household PA—full sample ^d	39 (100)	3063 (2102-4025)	2160 (4560)
Leisure-time PA			
Participants reporting engaging in leisure-time PA	26 (67)	24623 (1146-3779)	1042 (2997)
Leisure-time PA—full sample ^d	39 (100)	1642 (704-2580)	396 (1980)
Sitting time			
Weekday total	38 (97)	333 (267-399)	300 (240)
Weekend total	39 (100)	327 (250-405)	240 (240)

^aThe sum of walking and moderate and vigorous PA, which is equal to the sum of work, transportation, domestic/household, and leisure-time PA domains.

^bThe sum of moderate and vigorous PA.

^cAchieving PA guidelines was defined as engaging in ≥ 500 MET-minutes/week of moderate-to-vigorous PA; not meeting PA guidelines was defined as engaging in < 500 MET-minutes/week of moderate-to-vigorous PA.

^dParticipants reporting that they did not engage in a domain-specific activity had their value for that given domain set to zero; full-sample values for each domain include these participants with their zero values and thus provide an overall summary statistic for the domain that includes all study participants.

Examination of domain-specific PA patterns showed that domestic/household PA was the most frequent type of PA reported by participants (34/39, 87%) and, on average, how participants acquired the majority of their activity (ie, median 2160 MET-minutes/week). Leisure-time PA was the second most frequently reported domain of activity (26/39, 67%), followed by transportation (23/39, 59%) and work (7/39, 18%) PA. Additionally, although only 7 participants reported working outside of the home according to the IPAQ, the amount of PA acquired by these participants through work activity was higher than any other domain (4590 MET-minutes/week; see [Table 2](#)).

Associations Between Neighborhood Environmental Factors and PA

[Table 3](#) shows the descriptive values for scales assessing neighborhood environmental factors. OLS regression results,

presented in [Table 4](#), revealed no significant associations between neighborhood environmental factors and continuous PA measures. [Table 5](#) presents logistic regression models examining associations between neighborhood environmental factors and dichotomous PA outcomes of engaging in transportation PA versus not engaging in transportation PA, engaging in leisure-time PA versus not engaging in leisure-time PA, and meeting PA guidelines versus not meeting PA guidelines. Results showed that the neighborhood factors of walking environment, aesthetic quality, and safety were significantly associated with engaging in leisure-time PA (OR 5.95, 95% CI 1.49 to 23.74; OR 2.45, 95% CI 1.01 to 5.93; and OR 3.30, 95% CI 1.26 to 8.67, respectively) and meeting national PA guidelines (OR 4.15, 95% CI 1.13 to 15.18; OR 6.43, 95% CI 1.15 to 28.39; and OR 2.53, 95% CI 1.00 to 6.36, respectively). No other significant associations were found.

Table 3. Summary statistics of perceived neighborhood factors.

Perceived neighborhood factor	Neighborhood Scales Questionnaire [32] scale scores		
	Observations, n	Mean (95% CI)	Median (minimum; maximum)
Walking environment ^a	38	3.9 (3.6-4.1)	3.9 (1.9; 5.0)
Aesthetic quality ^a	39	3.7 (3.4-4.0)	3.8 (1.0; 5.0)
Safety ^a	38	3.4 (3.1-3.7)	3.7 (1.0; 5.0)
Social cohesion ^a	35	3.1 (2.8-3.4)	3.3 (1.0; 5.0)
Violence ^b	33	1.4 (1.2-1.7)	1.0 (1.0; 5.0)
Activities with neighbors ^b	38	1.7 (1.5-1.9)	1.6 (1.0; 3.4)

^aScore range of 1 to 5.

^bScore range of 1 to 4.

Table 4. Ordinal least squares (OLS) regression analyses examining associations between neighborhood factors and dichotomous physical activity (PA) outcomes.

PA outcome and neighborhood factor	OLS regression results for continuous outcomes ^a	
	Observations, n	Coefficient (95% CI)
PA domain: transportation		
Walking environment	38	-31.45 (-217.21 to 154.30)
Aesthetic quality	39	-5.99 (-161.78 to 149.79)
Safety	38	47.58 (-91.97 to 187.13)
Social cohesion	35	-32.95 (-194.04 to 128.15)
Violence	33	20.30 (-156.13 to 196.73)
Activities with neighbors	38	102.70 (-97.72 to 303.12)
PA domain: leisure		
Walking environment	38	734.54 (-595.48 to 2064.57)
Aesthetic quality	39	558.27 (-538.28 to 1654.81)
Safety	38	724.93 (-266.29 to 1716.14)
Social cohesion	35	189.49 (-819.58 to 1198.57)
Violence	33	-1201.72 (-2537.39 to 133.95)
Activities with neighbors	38	-348.37 (-1778.59 to 1081.85)
PA intensity level: moderate to vigorous		
Walking environment	38	1934.56 (-418.61 to 4287.73)
Aesthetic quality	39	1662.83 (-273.18 to 3598.83)
Safety	38	1168.56 (-634.70 to 2971.83)
Social cohesion	35	593.57 (-1386.55 to 2573.70)
Violence	33	-1679.53 (-4229.44 to 870.37)
Activities with neighbors	38	509.87 (-2052.37 to 3072.10)

^aControlled for age and level of education.

Table 5. Logistic regression analyses examining associations between neighborhood factors and dichotomous physical activity (PA) outcomes.^{a,b}

PA outcome and neighborhood factor	Frequency, n (%)	OR (95% CI)
PA domain: transportation		
Did not engage in transportation PA	16 (41)	1.00
Engaged in transportation PA	23 (59)	
Walking environment		1.60 (0.61-4.23)
Aesthetic quality		0.83 (0.38-1.78)
Safety		1.22 (0.60-2.47)
Social cohesion		1.45 (0.66-3.16)
Violence		1.17 (0.46-2.99)
Activities with neighbors		1.75 (0.59-5.16)
PA domain: leisure-time		
Did not engage in leisure-time PA	13 (33)	1.00
Engaged in leisure-time PA	26 (67)	
Walking environment		5.95* (1.49-23.74)
Aesthetic quality		2.45* (1.01-5.93)
Safety		3.30* (1.26-8.67)
Social cohesion		1.92 (0.81-4.55)
Violence		0.38 (0.13-1.10)
Activities with neighbors		0.67 (0.24-1.83)
Achieved national PA guidelines		
Did not meet PA guidelines	7 (18)	1.00
Met PA guidelines	32 (82)	
Walking environment		4.15* (1.13-15.18)
Aesthetic quality		6.43* (1.45-28.39)
Safety		2.53* (1.00-6.36)
Social cohesion		1.29 (0.53-3.15)
Violence		0.43 (0.16-1.20)
Activities with neighbors		0.99 (0.24-4.02)

^aControlled for age.

^bPA outcome variables for logistic regression analyses included engaging in transportation PA (ie, >0 metabolic equivalent [MET]-minutes/week of transportation PA) versus not engaging in transportation PA (0 MET-minutes/week of transportation PA), engaging in leisure-time PA (ie, >0 MET-minutes/week of leisure-time PA) versus not engaging in leisure-time PA (ie, 0 MET-minutes/week), and meeting national PA guidelines (ie, ≥500 MET-minutes/week of moderate-to-vigorous PA) versus not meeting national PA guidelines (ie, <500 MET-minutes/week of moderate-to-vigorous PA).

*Significant at a *P* value <.05.

Discussion

Principal Findings

This study explored PA patterns among first-generation Latina immigrants residing in the metropolitan area of Phoenix, Arizona, and the influence of perceived neighborhood environmental factors on these PA patterns. This work adds to the limited body of research on domain-specific PA patterns among first-generation Latina immigrants, with the majority of previous PA studies among Latinas having focused exclusively on leisure-time PA [34] and failing to differentiate by generational status. Given that first-generation immigrants likely

have different lived experiences and lifestyle activities than Latinas who were born in the United States, work is needed in order to inform the development of PA interventions for this population.

Results showed that the majority of our sample (85%) reported PA levels that met or exceeded national PA guidelines. Domain-specific outcomes indicated that participants accrued the majority of their PA through domestic and household activities, with limited PA performed for leisure or transportation purposes. The perceived neighborhood environmental factors of walking environment, aesthetic quality, and safety were positively associated with engaging in leisure-time PA and

meeting national PA guidelines. The neighborhood factors of violence, social cohesion, and activities with neighbors were not related to PA outcomes.

Comparisons With Prior Work

Our finding that domestic and household activities accounted for the majority of PA performed reflects the outcomes of several previous studies examining PA patterns among Latinas [8,20] and supports the notion that Latinas, regardless of generational status, perform extensive caretaking and household activities as part of their daily routine. Likewise, low levels of leisure-time PA and a limited number of participants engaging in paid or unpaid work outside the home have also been previously reported among studies comprised predominantly of first-generation Latinas [19,20,33]. Leisure-time PA, in particular, has profound benefits for reducing cardiometabolic disease conditions [23,35,36]. Moreover, when compared with other PA domains (ie, work and transportation), leisure-time PA may be more amendable to change in the context of a PA intervention because it is within the volitional control of many Latinas. In future work, researchers should explore intervention strategies to increase leisure-time PA among first-generation Latinas. Such work may be key to reducing the disproportionate metabolic disease burden in this high-risk population.

Our examination of overall and intensity-specific PA patterns showed that our sample reported engaging in rather high levels of overall PA (ie, 4512 MET-minutes/week), with the majority of the PA being at a moderate intensity. We speculate that these high PA levels may be a result of overreporting, as this is a commonly reported occurrence with self-report PA measures [37]. Nicaise and colleagues [20] illustrated this issue in a previous study with low-income Latinas using the IPAQ and accelerometers. When assessing PA levels using the IPAQ, the authors found that 73% of Latinas met PA guidelines, with the bulk of the PA being accrued through domestic or household activities [20]—similar to the outcomes of our study. However, when examining PA levels with accelerometers, this percentage was reduced to 20%, suggesting that participants likely overestimated the intensity and the amount of PA accrued through domestic and household activities [20]. We speculate that a similar phenomenon may have impacted our study outcomes. Another key finding of our study was that our sample of Latinas reported limited vigorous-intensity PA, which also mirrors the outcome of several previous studies [18,20]. We hypothesize that this outcome was related to the limited amount of leisure-time PA reported by our sample, as vigorous PA is predominately achieved through purposeful exercise as opposed to daily activities [38].

When examining associations between neighborhood environmental factors and PA outcomes, linear regression models failed to show any significant associations. This might be due to the nonlinear patterns of the associations and a low statistical power due to the small sample size. However, logistic regression models revealed that the neighborhood factors of walking environment, aesthetic quality, and safety were associated with reported engagement in leisure-time PA and meeting national PA guidelines. These findings confirm previous studies showing that these factors are positively associated with

PA engagement [19,39-42]. We also found it interesting that these 3 factors are ones that recent immigrants likely have limited control over (ie, these factors are primarily driven by public policies and allocation of local government funds). This outcome highlights the importance of considering all levels of the social ecological model, including public policy and urban design planning, when developing PA promotion interventions.

The lack of significant associations between the 6 neighborhood environmental factors and engagement in some versus no transportation PA in logistic models may be related to transportation PA being a necessity for daily activities rather an option (ie, participants may have had no other type of transportation). However, we did not include survey items asking participants about their primary mode of transportation, which is needed to draw a more precise conclusion on this outcome. Likewise, the lack of associations between social cohesion, violence, and activities with neighbors and PA was reported previously in a study among Latinx residing in Massachusetts [40], suggesting that these 3 factors may not be key determinants of PA engagement among Latinas. Future studies with larger sample sizes are needed before definitive conclusions can be drawn.

Limitations of the Study

Limitations of our study include the use of a relatively small sample from a single metropolitan geographic region. The sample size was determined based on available funds, and the geographic region chosen was based on the population that the subsequent intervention would be developed to target. The small sample limited our ability to conduct additional subgroup analyses to further examine associations among duration of US residency (ie, ≤ 5 years versus >5 years), neighborhood environmental factors, and PA. Further examination of neighborhood characteristics and PA patterns based on duration of residency status in the United States would be interesting to explore in future work. Likewise, given that our sample comprised women residing in urban and suburban areas of metropolitan Phoenix, the study findings should not be generalized to residents residing in rural areas. Additional research with a larger, more geographically diverse sample is warranted to confirm and expand on our findings. Another limitation was that we relied exclusively on self-report measures to assess PA and neighborhood characteristics. Including an objective PA measure would have provided further context regarding the PA patterns of our sample. However, this was not possible because of the limited resources of the study. Likewise, preliminary data analyses resulted in 11 women being excluded from the study based on IPAQ scoring guidelines for either reporting nonplausible PA data ($n=10$) or providing incomplete data ($n=1$). Sensitivity analyses showed that these participants reported lower education levels than participants who provided valid data. This suggests that the IPAQ may not be appropriate for Latinas of lower education levels, further limiting generalization of the study findings. Additionally, perceived neighborhood characteristics may differ from objectively measured characteristics (eg, neighborhood characteristics measured using geographic information systems, local crime/violence data). Future research among first-generation Latinas should consider the use of both subjective and objective

PA measures and neighborhood measures, as both types of measures provide important insights into the PA patterns and neighborhood characteristics of Latinas. A final limitation of this study is that it did not include items assessing BMI or health status. Such information would have provided additional insight on the generalizability of the study findings.

Strengths of the Study

Despite its limitations, the current study has several strengths. First, this is one of few studies to examine domain-specific PA patterns among first-generation Latina immigrants, as most PA studies among Latinas have only included a measure of leisure-time PA [34] and failed to differentiate by generational status. Another strength of the study is that our sample was comprised of women from diverse countries of origin (ie, 69% from Mexico and the remainder from various Central and South American countries). The composition of our sample likely reflects the changing Latinx immigrant population in the United

States (ie, immigration from Central and South American countries has increased in recent years, with immigration from Mexico slowly declining) [11].

Conclusions

Findings suggest that many first-generation Latinas could benefit from a leisure-time PA intervention. Given that leisure-time PA has pronounced benefits for promoting positive health outcomes, intervention efforts targeting leisure-time PA may be an effective method for researchers and public health professionals to reduce obesity and diabetes health disparities among Latina immigrants. Such interventions should consider the neighborhood environments in which first-generation Latinas reside, as these factors will likely influence the types of PA promoted in an intervention. Results of this study will be used to inform development of a culturally tailored PA intervention for the reduction of metabolic disease risk among first-generation Latina immigrants.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Demographic questionnaire.

[[DOCX File, 14 KB - formative_v5i5e25663_app1.docx](#)]

References

1. Kaplan RC, Avilés-Santa ML, Parrinello CM, Hanna DB, Jung M, Castañeda SF, et al. Body mass index, sex, and cardiovascular disease risk factors among Hispanic/Latino adults: Hispanic community health study/study of Latinos. *J Am Heart Assoc* 2014 Jul 09;3(4):e000923 [FREE Full text] [doi: [10.1161/JAHA.114.000923](https://doi.org/10.1161/JAHA.114.000923)] [Medline: [25008353](https://pubmed.ncbi.nlm.nih.gov/25008353/)]
2. Schneiderman N, Llabre M, Cowie CC, Barnhart J, Carnethon M, Gallo LC, et al. Prevalence of diabetes among Hispanics/Latinos from diverse backgrounds: the Hispanic Community Health Study/Study of Latinos (HCHS/SOL). *Diabetes Care* 2014 Aug;37(8):2233-2239 [FREE Full text] [doi: [10.2337/dc13-2939](https://doi.org/10.2337/dc13-2939)] [Medline: [25061138](https://pubmed.ncbi.nlm.nih.gov/25061138/)]
3. Hales C, Carroll M, Fryar CD, Ogden CL. Prevalence of Obesity Among Adults and Youth: United States, 2015-2016. *NCHS Data Brief* 2017 Oct(288):1-8 [FREE Full text] [Medline: [29155689](https://pubmed.ncbi.nlm.nih.gov/29155689/)]
4. Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Division of Diabetes Translation. National Diabetes Statistics Report, 2017: Estimates of Diabetes and Its Burden in the United States. 2017. URL: <https://dev.diabetes.org/sites/default/files/2019-06/cdc-statistics-report-2017.pdf> [accessed 2021-04-25]
5. Goel MS, McCarthy EP, Phillips RS, Wee CC. Obesity among US immigrant subgroups by duration of residence. *JAMA* 2004 Dec 15;292(23):2860-2867. [doi: [10.1001/jama.292.23.2860](https://doi.org/10.1001/jama.292.23.2860)] [Medline: [15598917](https://pubmed.ncbi.nlm.nih.gov/15598917/)]
6. Isasi CR, Ayala GX, Sotres-Alvarez D, Madanat H, Penedo F, Loria CM, et al. Is acculturation related to obesity in Hispanic/Latino adults? Results from the Hispanic community health study/study of Latinos. *J Obes* 2015;2015:186276 [FREE Full text] [doi: [10.1155/2015/186276](https://doi.org/10.1155/2015/186276)] [Medline: [25893114](https://pubmed.ncbi.nlm.nih.gov/25893114/)]
7. Koya DL, Egede LE. Association between length of residence and cardiovascular disease risk factors among an ethnically diverse group of United States immigrants. *J Gen Intern Med* 2007 Jun;22(6):841-846 [FREE Full text] [doi: [10.1007/s11606-007-0163-y](https://doi.org/10.1007/s11606-007-0163-y)] [Medline: [17503110](https://pubmed.ncbi.nlm.nih.gov/17503110/)]
8. Wolin KY, Colangelo LA, Chiu BC, Gapstur SM. Obesity and immigration among Latina women. *J Immigr Minor Health* 2009 Oct;11(5):428-431 [FREE Full text] [doi: [10.1007/s10903-007-9115-1](https://doi.org/10.1007/s10903-007-9115-1)] [Medline: [18183486](https://pubmed.ncbi.nlm.nih.gov/18183486/)]
9. Oza-Frank R, Stephenson R, Narayan KMV. Diabetes prevalence by length of residence among US immigrants. *J Immigr Minor Health* 2011 Feb;13(1):1-8. [doi: [10.1007/s10903-009-9283-2](https://doi.org/10.1007/s10903-009-9283-2)] [Medline: [19688263](https://pubmed.ncbi.nlm.nih.gov/19688263/)]
10. Oza-Frank R, Cunningham SA. The weight of US residence among immigrants: a systematic review. *Obes Rev* 2010 Apr;11(4):271-280. [doi: [10.1111/j.1467-789X.2009.00610.x](https://doi.org/10.1111/j.1467-789X.2009.00610.x)] [Medline: [19538440](https://pubmed.ncbi.nlm.nih.gov/19538440/)]
11. Budiman A. Key findings about U.S. immigrants. Pew Research Center. 2020 Aug 20. URL: <https://www.pewresearch.org/fact-tank/2019/06/17/key-findings-about-u-s-immigrants/> [accessed 2021-04-22]
12. Swift DL, Johannsen NM, Lavie CJ, Earnest CP, Church TS. The role of exercise and physical activity in weight loss and maintenance. *Prog Cardiovasc Dis* 2014;56(4):441-447 [FREE Full text] [doi: [10.1016/j.pcad.2013.09.012](https://doi.org/10.1016/j.pcad.2013.09.012)] [Medline: [24438736](https://pubmed.ncbi.nlm.nih.gov/24438736/)]

13. Jakicic JM, Davis KK. Obesity and physical activity. *Psychiatr Clin North Am* 2011 Dec;34(4):829-840. [doi: [10.1016/j.psc.2011.08.009](https://doi.org/10.1016/j.psc.2011.08.009)] [Medline: [22098807](https://pubmed.ncbi.nlm.nih.gov/22098807/)]
14. Jakicic J. Physical activity and weight loss. *Nestle Nutr Inst Workshop Ser* 2012;73:21-36. [doi: [10.1159/000341283](https://doi.org/10.1159/000341283)] [Medline: [23128763](https://pubmed.ncbi.nlm.nih.gov/23128763/)]
15. Colberg SR, Sigal RJ, Fernhall B, Regensteiner JG, Blissmer BJ, Rubin RR, American College of Sports Medicine, American Diabetes Association. Exercise and type 2 diabetes: the American College of Sports Medicine and the American Diabetes Association: joint position statement executive summary. *Diabetes Care* 2010 Dec;33(12):2692-2696 [FREE Full text] [doi: [10.2337/dc10-1548](https://doi.org/10.2337/dc10-1548)] [Medline: [21115771](https://pubmed.ncbi.nlm.nih.gov/21115771/)]
16. Sweatt SK, Willig AL, Agne AA, Powell JL, Cherrington AL. Physical Activity Patterns of Latina Immigrants Living in Alabama. *J Racial Ethn Health Disparities* 2015 Sep;2(3):365-372 [FREE Full text] [doi: [10.1007/s40615-015-0083-1](https://doi.org/10.1007/s40615-015-0083-1)] [Medline: [26413457](https://pubmed.ncbi.nlm.nih.gov/26413457/)]
17. Arredondo EM, Sotres-Alvarez D, Stoutenberg M, Davis SM, Crespo NC, Carnethon MR, et al. Physical Activity Levels in U.S. Latino/Hispanic Adults: Results From the Hispanic Community Health Study/Study of Latinos. *Am J Prev Med* 2016 Apr;50(4):500-508 [FREE Full text] [doi: [10.1016/j.amepre.2015.08.029](https://doi.org/10.1016/j.amepre.2015.08.029)] [Medline: [26597505](https://pubmed.ncbi.nlm.nih.gov/26597505/)]
18. Evenson KR, Sarmiento OL, Ayala GX. Acculturation and physical activity among North Carolina Latina immigrants. *Soc Sci Med* 2004 Dec;59(12):2509-2522. [doi: [10.1016/j.socscimed.2004.04.011](https://doi.org/10.1016/j.socscimed.2004.04.011)] [Medline: [15474205](https://pubmed.ncbi.nlm.nih.gov/15474205/)]
19. Perez LG, Chavez A, Marquez DX, Soto SC, Haughton J, Arredondo EM. Associations of Acculturation With Self-Report and Objective Physical Activity and Sedentary Behaviors Among Latinas. *Health Educ Behav* 2017 Jun;44(3):431-438 [FREE Full text] [doi: [10.1177/1090198116669802](https://doi.org/10.1177/1090198116669802)] [Medline: [27679665](https://pubmed.ncbi.nlm.nih.gov/27679665/)]
20. Nicaise V, Marshall S, Ainsworth BE. Domain-specific physical activity and self-report bias among low-income Latinas living in San Diego County. *J Phys Act Health* 2011 Sep;8(7):881-890. [doi: [10.1123/jpah.8.7.881](https://doi.org/10.1123/jpah.8.7.881)] [Medline: [21885878](https://pubmed.ncbi.nlm.nih.gov/21885878/)]
21. Holtermann A, Hansen JV, Burr H, Sjøgaard K, Sjøgaard G. The health paradox of occupational and leisure-time physical activity. *Br J Sports Med* 2012 Mar;46(4):291-295. [doi: [10.1136/bjism.2010.079582](https://doi.org/10.1136/bjism.2010.079582)] [Medline: [21459873](https://pubmed.ncbi.nlm.nih.gov/21459873/)]
22. Samitz G, Egger M, Zwahlen M. Domains of physical activity and all-cause mortality: systematic review and dose-response meta-analysis of cohort studies. *Int J Epidemiol* 2011 Oct;40(5):1382-1400. [doi: [10.1093/ije/dyr112](https://doi.org/10.1093/ije/dyr112)] [Medline: [22039197](https://pubmed.ncbi.nlm.nih.gov/22039197/)]
23. Autenrieth CS, Baumert J, Baumeister SE, Fischer B, Peters A, Döring A, et al. Association between domains of physical activity and all-cause, cardiovascular and cancer mortality. *Eur J Epidemiol* 2011 Feb;26(2):91-99. [doi: [10.1007/s10654-010-9517-6](https://doi.org/10.1007/s10654-010-9517-6)] [Medline: [21153912](https://pubmed.ncbi.nlm.nih.gov/21153912/)]
24. Harari G, Green MS, Zelber-Sagi S. Combined association of occupational and leisure-time physical activity with all-cause and coronary heart disease mortality among a cohort of men followed-up for 22 years. *Occup Environ Med* 2015 Sep;72(9):617-624. [doi: [10.1136/oemed-2014-102613](https://doi.org/10.1136/oemed-2014-102613)] [Medline: [25805756](https://pubmed.ncbi.nlm.nih.gov/25805756/)]
25. Holtermann A, Marott JL, Gyntelberg F, Sjøgaard K, Mortensen OS, Prescott E, et al. Self-reported occupational physical activity and cardiorespiratory fitness: Importance for cardiovascular disease and all-cause mortality. *Scand J Work Environ Health* 2016 Jul 01;42(4):291-298 [FREE Full text] [doi: [10.5271/sjweh.3563](https://doi.org/10.5271/sjweh.3563)] [Medline: [27100403](https://pubmed.ncbi.nlm.nih.gov/27100403/)]
26. Wanner M, Martin BW, Autenrieth CS, Schaffner E, Meier F, Brombach C, et al. Associations between domains of physical activity, sitting time, and different measures of overweight and obesity. *Prev Med Rep* 2016 Jun;3:177-184 [FREE Full text] [doi: [10.1016/j.pmedr.2016.01.007](https://doi.org/10.1016/j.pmedr.2016.01.007)] [Medline: [27419012](https://pubmed.ncbi.nlm.nih.gov/27419012/)]
27. American Community Survey: 2015—2019 ACS 5-Year Data Profile. United States Census Bureau.: United States Census Bureau; 2015. URL: <https://www.census.gov/acs/www/data/data-tables-and-tools/data-profiles/> [accessed 2021-01-28]
28. 2017 Behavioral Risk Factor Surveillance System Questionnaire. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/brfss/questionnaires/pdf-ques/2017_BRFSS_Pub_Ques_508_tagged.pdf [accessed 2021-04-26]
29. Craig CL, Marshall AL, Sjöström M, Bauman AE, Booth ML, Ainsworth BE, et al. International physical activity questionnaire: 12-country reliability and validity. *Med Sci Sports Exerc* 2003 Aug;35(8):1381-1395. [doi: [10.1249/01.MSS.0000078924.61453.FB](https://doi.org/10.1249/01.MSS.0000078924.61453.FB)] [Medline: [12900694](https://pubmed.ncbi.nlm.nih.gov/12900694/)]
30. Roman-Viñas B, Serra-Majem L, Hagströmer M, Ribas-Barba L, Sjöström M, Segura-Cardona R. International Physical Activity Questionnaire: Reliability and validity in a Spanish population. *Eur J Sport Sci* 2010 Sep;10(5):297-304. [doi: [10.1080/17461390903426667](https://doi.org/10.1080/17461390903426667)]
31. The IPAQ Group. Guidelines for data processing and analysis of the International Physical Activity Questionnaire (IPAQ) - Short and Long Forms. International Physical Activity Questionnaire. 2005 Nov. URL: <https://docs.google.com/viewer?a=v&pid=sites&srcid=ZGVmYXVsdGRvbWFpbm90aGVpcGFxfGd4OjE0NDgxMDk3NDU1YWZRZTM> [accessed 2021-04-26]
32. Mujahid MS, Diez Roux AV, Morenoff JD, Raghunathan T. Assessing the measurement properties of neighborhood scales: from psychometrics to ecometrics. *Am J Epidemiol* 2007 Apr 15;165(8):858-867. [doi: [10.1093/aje/kwm040](https://doi.org/10.1093/aje/kwm040)] [Medline: [17329713](https://pubmed.ncbi.nlm.nih.gov/17329713/)]
33. Keller C, Ainsworth B, Records K, Todd M, Belyea M, Vega-López S, et al. A comparison of a social support physical activity intervention in weight management among post-partum Latinas. *BMC Public Health* 2014 Sep 19;14:971 [FREE Full text] [doi: [10.1186/1471-2458-14-971](https://doi.org/10.1186/1471-2458-14-971)] [Medline: [25233867](https://pubmed.ncbi.nlm.nih.gov/25233867/)]
34. Tovar M, Walker JL, Rew L. Factors Associated With Physical Activity in Latina Women: A Systematic Review. *West J Nurs Res* 2018 Feb;40(2):270-297. [doi: [10.1177/0193945916681004](https://doi.org/10.1177/0193945916681004)] [Medline: [27920349](https://pubmed.ncbi.nlm.nih.gov/27920349/)]

35. Smith AD, Crippa A, Woodcock J, Brage S. Physical activity and incident type 2 diabetes mellitus: a systematic review and dose-response meta-analysis of prospective cohort studies. *Diabetologia* 2016 Dec;59(12):2527-2545 [FREE Full text] [doi: [10.1007/s00125-016-4079-0](https://doi.org/10.1007/s00125-016-4079-0)] [Medline: [27747395](https://pubmed.ncbi.nlm.nih.gov/27747395/)]
36. Aune D, Norat T, Leitzmann M, Tonstad S, Vatten LJ. Physical activity and the risk of type 2 diabetes: a systematic review and dose-response meta-analysis. *Eur J Epidemiol* 2015 Jul;30(7):529-542. [doi: [10.1007/s10654-015-0056-z](https://doi.org/10.1007/s10654-015-0056-z)] [Medline: [26092138](https://pubmed.ncbi.nlm.nih.gov/26092138/)]
37. Prince SA, Adamo KB, Hamel ME, Hardt J, Connor Gorber S, Tremblay M. A comparison of direct versus self-report measures for assessing physical activity in adults: a systematic review. *Int J Behav Nutr Phys Act* 2008 Nov 06;5:56 [FREE Full text] [doi: [10.1186/1479-5868-5-56](https://doi.org/10.1186/1479-5868-5-56)] [Medline: [18990237](https://pubmed.ncbi.nlm.nih.gov/18990237/)]
38. Tudor-Locke C, Johnson W, Katzmarzyk P. Frequently reported activities by intensity for U.S. adults: the American Time Use Survey. *Am J Prev Med* 2010 Oct;39(4):e13-e20. [doi: [10.1016/j.amepre.2010.05.017](https://doi.org/10.1016/j.amepre.2010.05.017)] [Medline: [20837277](https://pubmed.ncbi.nlm.nih.gov/20837277/)]
39. Barnett DW, Barnett A, Nathan A, Van Cauwenberg J, Cerin E, Council on EnvironmentPhysical Activity (CEPA) – Older Adults working group. Built environmental correlates of older adults' total physical activity and walking: a systematic review and meta-analysis. *Int J Behav Nutr Phys Act* 2017 Aug 07;14(1):103 [FREE Full text] [doi: [10.1186/s12966-017-0558-z](https://doi.org/10.1186/s12966-017-0558-z)] [Medline: [28784183](https://pubmed.ncbi.nlm.nih.gov/28784183/)]
40. Silfee VJ, Rosal MC, Sreedhara M, Lora V, Lemon SC. Neighborhood environment correlates of physical activity and sedentary behavior among Latino adults in Massachusetts. *BMC Public Health* 2016 Sep 13;16:966 [FREE Full text] [doi: [10.1186/s12889-016-3650-4](https://doi.org/10.1186/s12889-016-3650-4)] [Medline: [27619205](https://pubmed.ncbi.nlm.nih.gov/27619205/)]
41. Van Holle V, Van Cauwenberg J, Van Dyck D, Deforche B, Van de Weghe N, De Bourdeaudhuij I. Relationship between neighborhood walkability and older adults' physical activity: results from the Belgian Environmental Physical Activity Study in Seniors (BEPAS Seniors). *Int J Behav Nutr Phys Act* 2014 Aug 23;11:110 [FREE Full text] [doi: [10.1186/s12966-014-0110-3](https://doi.org/10.1186/s12966-014-0110-3)] [Medline: [25148845](https://pubmed.ncbi.nlm.nih.gov/25148845/)]
42. Pratt M, Yin S, Soler R, Njai R, Siegel PZ, Liao Y. Does perceived neighborhood walkability and safety mediate the association between education and meeting physical activity guidelines? *Prev Chronic Dis* 2015 Apr 09;12:E46 [FREE Full text] [doi: [10.5888/pcd12.140570](https://doi.org/10.5888/pcd12.140570)] [Medline: [25855989](https://pubmed.ncbi.nlm.nih.gov/25855989/)]

Abbreviations

IPAQ: International Physical Activity Questionnaire

MET: metabolic equivalent

OLS: ordinal least squares

PA: physical activity

Edited by G Eysenbach; submitted 10.11.20; peer-reviewed by S Nyenhuis, M Sreedhara, G Signorelli; comments to author 29.12.20; revised version received 30.03.21; accepted 04.04.21; published 17.05.21.

Please cite as:

Joseph RP, Vega-López S, Han S

Physical Activity Patterns and Neighborhood Characteristics of First-Generation Latina Immigrants Living in Arizona: Cross-sectional Study

JMIR Form Res 2021;5(5):e25663

URL: <https://formative.jmir.org/2021/5/e25663>

doi: [10.2196/25663](https://doi.org/10.2196/25663)

PMID: [33999003](https://pubmed.ncbi.nlm.nih.gov/33999003/)

©Rodney P Joseph, Sonia Vega-López, SeungYong Han. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 17.05.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Survival Analysis of Patients With COVID-19 in India by Demographic Factors: Quantitative Study

Sampurna Kundu^{1*}, MPhil; Kirti Chauhan^{1*}, MPhil; Debarghya Mandal^{1*}, MSc

Department of Mathematical Demography and Statistics, International Institute for Population Sciences, Mumbai, India

*all authors contributed equally

Corresponding Author:

Sampurna Kundu, MPhil

Department of Mathematical Demography and Statistics

International Institute for Population Sciences

Govandi Station Road

Mumbai, 400088

India

Phone: 91 9073111858

Email: sampurna34@gmail.com

Abstract

Background: Studies of the transmission dynamics of COVID-19 have depicted the rate, patterns, and predictions of cases of this pandemic disease. To combat transmission of the disease in India, the government declared a lockdown on March 25, 2020. Even after this strict lockdown was enacted nationwide, the number of COVID-19 cases increased and surpassed 450,000. A positive point to note is that the number of recovered cases began to slowly exceed that of active cases. The survival of patients, taking death as the event that varies by age group and sex, is noteworthy.

Objective: The aim of this study was to conduct a survival analysis to establish the variability in survivorship of patients with COVID-19 in India by age group and sex at different levels, that is, the national, state, and district levels.

Methods: The study period was taken from the date of the first reported case of COVID-19 in India, which was January 30, 2020, up to June 30, 2020. Due to the amount of underreported data and removal of missing columns, a total sample of 26,815 patients was considered. Kaplan-Meier survival estimation, the Cox proportional hazard model, and the multilevel survival model were used to perform the survival analysis.

Results: The Kaplan-Meier survival function showed that the probability of survival of patients with COVID-19 declined during the study period of 5 months, which was supplemented by the log rank test ($P < .001$) and Wilcoxon test ($P < .001$) to compare the survival functions. Significant variability was observed in the age groups, as evident from all the survival estimates; with increasing age, the risk of dying of COVID-19 increased. The Cox proportional hazard model reiterated that male patients with COVID-19 had a 1.14 times higher risk of dying than female patients (hazard ratio 1.14; SE 0.11; 95% CI 0.93-1.38). Western and Central India showed decreasing survival rates in the framed time period, while Eastern, North Eastern, and Southern India showed slightly better results in terms of survival.

Conclusions: This study depicts a grave scenario of decreasing survival rates in various regions of India and shows variability in these rates by age and sex. In essence, we can safely conclude that the critical appraisal of the survival rate and thorough analysis of patient data in this study equipped us to identify risk groups and perform comparative studies of various segments in India.

International Registered Report Identifier (IRRID): RR2-10.1101/2020.08.01.20162115

(*JMIR Form Res* 2021;5(5):e23251) doi:[10.2196/23251](https://doi.org/10.2196/23251)

KEYWORDS

survival analysis; COVID-19; patient data; Kaplan-Meier; hazard model; modeling; survival; mortality; demographic; India; transmission

Introduction

The entire world has been greatly challenged by the sudden outbreak of COVID-19, as the human race has no remedial measures to combat the lethal impact of the disease. According to the World Health Organization (WHO), the global pandemic of COVID-19 is derived from SARS-CoV-2, a member of a large family of viruses, named coronaviruses; these viruses cause respiratory infections ranging from the common cold to high fever, leading to disease. This blue planet has witnessed many epidemics, such as that caused by severe acute respiratory syndrome coronavirus (SARS-CoV) from 2002 to 2003 and H1N1 influenza in 2009 [1], due to various pernicious viruses in the last two decades; however, COVID-19 is incomparable with previous epidemics because of the indomitable growth rate of the disease and its high fatality rate. China was the first country to experience high numbers of cases at the beginning of the pandemic; presently, Chinese authorities have “flattened the curve” with continuous testing and aggressive quarantine measures [2]. Outside China, South Korea was the country that had the largest initial outbreak; they managed to slow the spread of COVID-19 and flatten the curve without imposing lockdown in their country [3]. The only method used to slow and contain the outbreak in Korea was mass diagnostic testing and quarantining. The WHO declared that incorporating self-isolation, sanitizing, washing hands repeatedly and abstaining from touching the mouth, face, or nose to stop the spread of COVID-19 [4]. To combat transmission of the disease in India, the government declared a lockdown on March 25, 2020. However, the disease has spread rapidly across the entire country, and as of June 30, 2020, there were 1,385,494 cases, with 32,096 deaths and 886,235 recoveries [5]. For a developing country such as India, the COVID-19 pandemic is a serious problem facing the nation, and the main sufferers are marginalized sections of society. Even after a strict lockdown was established nationwide, the number of cases increased and surpassed 450,000. However, the fatality rates later decreased, and several studies have shown that the lockdown did slow the rate of increase in a number of cases [6]. A positive point to be noted is that the number of recovered cases is slowly exceeding that of active cases.

In the study of dynamics of infectious diseases, compartmental models and the basic reproductive number (r_0) have been observed to be the mostly commonly used over the past year [7]. Basic mathematical models such as the Gompertz, exponential, and logistic growth models have shown to be quite effective in understanding the growth patterns of the disease [8]. One of the main demerits of the Indian database for COVID-19 is the underreporting of cases due to misreporting and the lower number of tests [9]. Amid the growing number of deaths due to COVID-19, researchers worldwide have associated these deaths with additional important cofactors, namely the effect on the older population and the impact of pollution and smoking as well as the development of acute respiratory distress syndrome [10,11]. In one study [12], a district-level analysis showed that 92 districts in India are in red zones of the disease. These red zones are mostly found in

the states of Maharashtra and Gujarat; in another study [13], it was predicted by the autoregressive integrated moving average (ARIMA) model that the number of cases will increase alarmingly.

Studies have described the impact of lockdown, the transmission dynamics of the disease, and forecasts of the pandemic. The survival of patients, taking death as the event that varies by age group and sex, is noteworthy. The aim of this study was to conduct a survival analysis to establish the variability in survivorship by age group and sex at different levels, that is, the national, state, district, and patient levels. This quantitative analysis (with analysis of data from patients with COVID-19) is exceptional not only for its gravity and pertinence but also for its subtle nuances and penetrating approach.

Methods

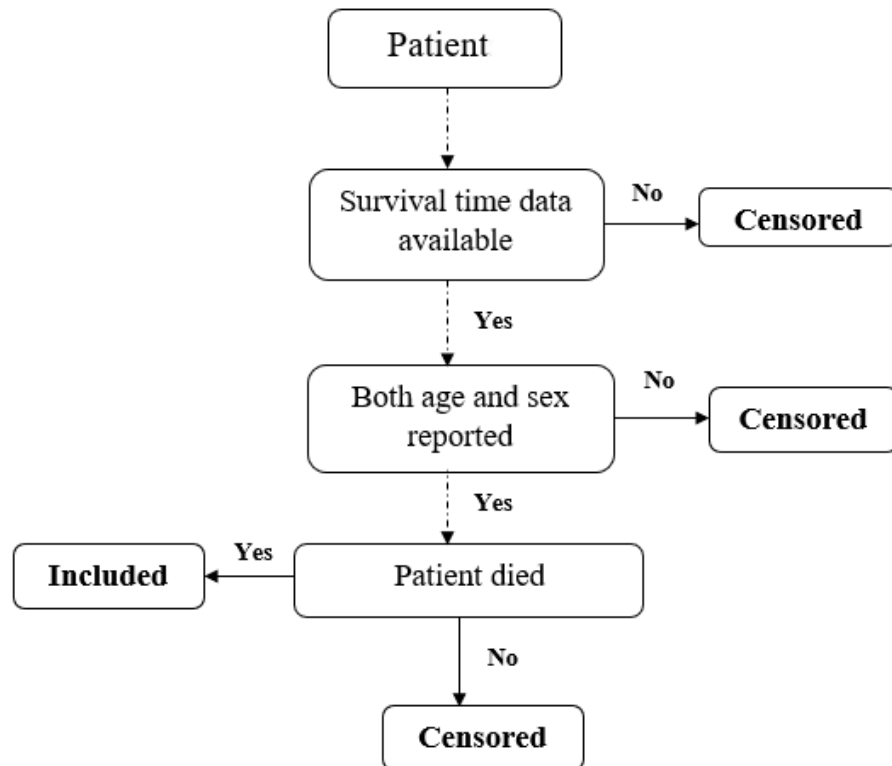
Data and Analysis

The data for this study were retrieved from the data sharing portal of India [5]. Patient-level data, consisting of time-to-event data, were used for the study. Here, the study period is from the date of the first case report in India, which was January 30, 2020, to June 30 of that year (ie, 5 months or 150 days). The entry point of each patient was different, and the event of interest in this study was death. If this event had not occurred, the survival time was taken to be censored. Due to the amount of underreported data and dropping of missing columns, a total of 26,815 sample patients were considered. The inclusion criteria for each patient were the date on which they tested positive for COVID-19, the date of the change of status, and reported age and sex. Survival time was computed by taking the difference between the date on which each sample patient tested positive for the infection and the date of the change of status. A flowchart of the selection of patient data for the study is shown in Figure 1.

The Kaplan-Meier survival estimator method, Cox proportional hazard model, and multilevel survival model were used to perform the survival analysis. Firstly, the Kaplan-Meier survival estimator method was used to estimate the survival function from the survival data. To compare the survival functions for different groups (ie, by sex, age group, and region), the log-rank test and Gehan-Breslow-Wilcoxon test were used. To estimate the survival functions in the presence of various covariates, the Cox proportional hazard model was used, with sex, age, and region as the covariates, assuming that the hazard is independent of time. The Cox proportional hazard model can be expressed as below:

$$h(t) = h_0(t)\exp(\beta_1\text{sex} + \beta_2\text{age} + \beta_3\text{region})$$

where t represents the survival time and $h(t)$ is the hazard function determined by a set of 3 covariates (sex, age, region). The coefficients (β_1 , β_2 , β_3) measure the impact (ie, the effect size) of the covariates. The term h_0 is called the baseline hazard, and it corresponds to the value of the hazard if all values of x_i are equal to zero.

Figure 1. Flowchart of the inclusion of patients in the study.

Finally, multilevel mixed effects survival analysis was performed, as clustering of lower level units at higher level units is a common scenario in such studies. Here, patients were clustered at the district level; then, the districts were clustered at the state level, and all the states were clustered at the national level. We considered $i=1,2,\dots,N$ clusters (eg, states and districts), with each cluster having $j=1,2,\dots,n_i$ patients. Let S_{ij} be the true survival time of the j^{th} patient in the i^{th} cluster, $T_{ij} = \min(S_{ij}, C_{ij})$ be the observed survival time, and C_{ij} be the censoring time. The proportional-hazards mixed-effects survival model can be written as below:



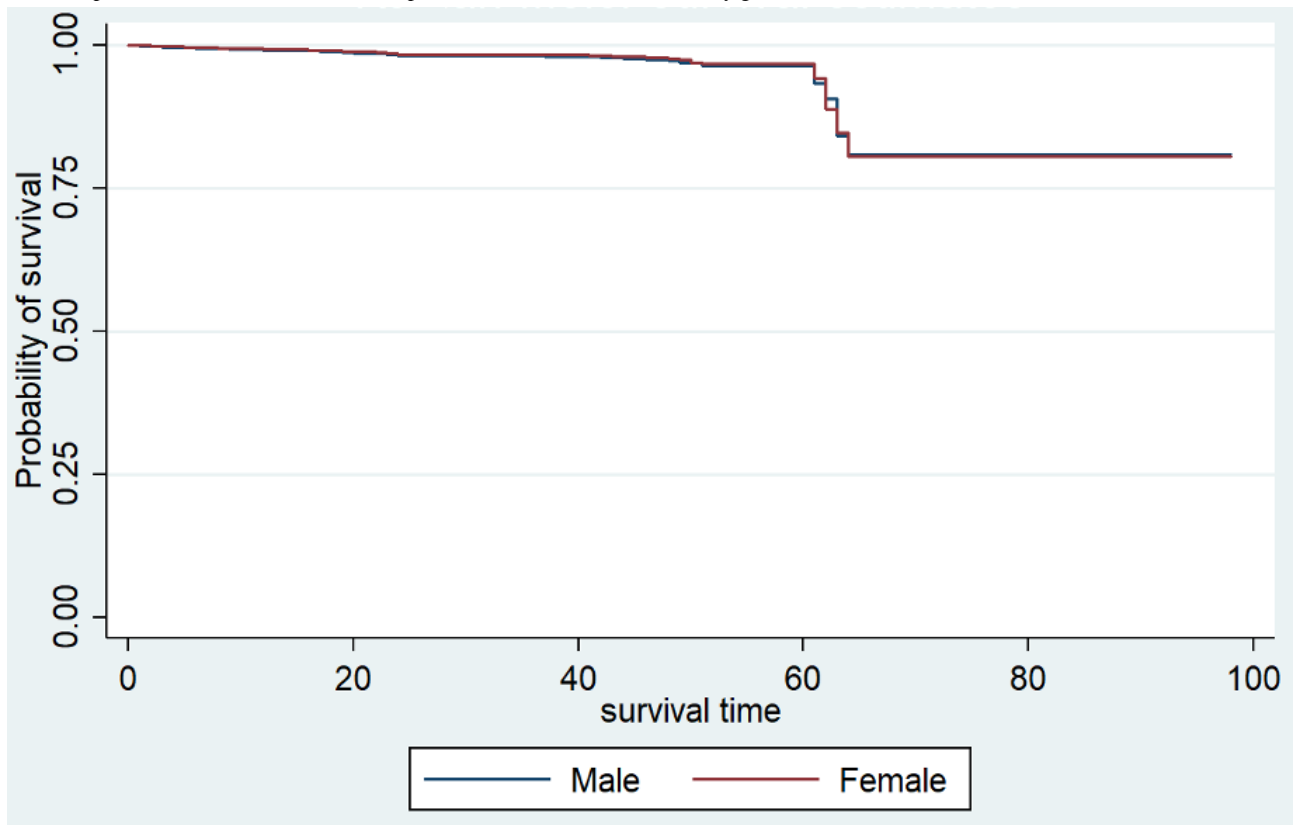
where $h_0(t)$ is the baseline hazard function of a standard parametric model (eg, here, we use Weibull at each level because according to the Akaike information criterion, it is the most appropriate model to use). Therefore, a 3-level cluster analysis will help eliminate the variability at each level due to intercorrelation between the units and can provide better estimates of the survival function.

Ethical Approval

The research was conducted using a publicly available database. The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees in human experimentation and with the Helsinki Declaration of 1978 as revised in 2008.

Results

Kaplan-Meier estimates were obtained initially to estimate the survival functions of patients in India with COVID-19 by sex, age group, and region. It can be seen from [Figure 2](#) that the survival curves from the Kaplan-Meier estimator for male and female sex are almost the same. In [Table 1](#), it can be observed that according to both the log rank test ($P>.001$) and Wilcoxon test ($P>.001$) for comparing the survival functions, the difference is not significant, indicating that there is no significant difference between the survival curves of male and female patients with COVID-19.

Figure 2. Kaplan-Meier estimate of survival of patients with COVID-19 in India by gender.**Table 1.** Kaplan-Meier estimator: comparison of survival functions for patients with COVID-19 in India by sex.

Test	$\chi^2 (1)$	<i>P</i> value
Log rank	1.6	.20
Wilcoxon	2.6	.11

As shown in Figure 3, the survival curves from the Kaplan-Meier estimator by 5-year age group are significantly different. This result was further supplemented by the log rank test ($P < .001$) and Wilcoxon test ($P < .001$) for comparing the

survival functions (Table 2); both tests gave highly significant results, indicating that there are significant differences among the survival curves of various age groups.

Figure 3. Kaplan-Meier estimate of survival of patients with COVID-19 in India by age group (years).

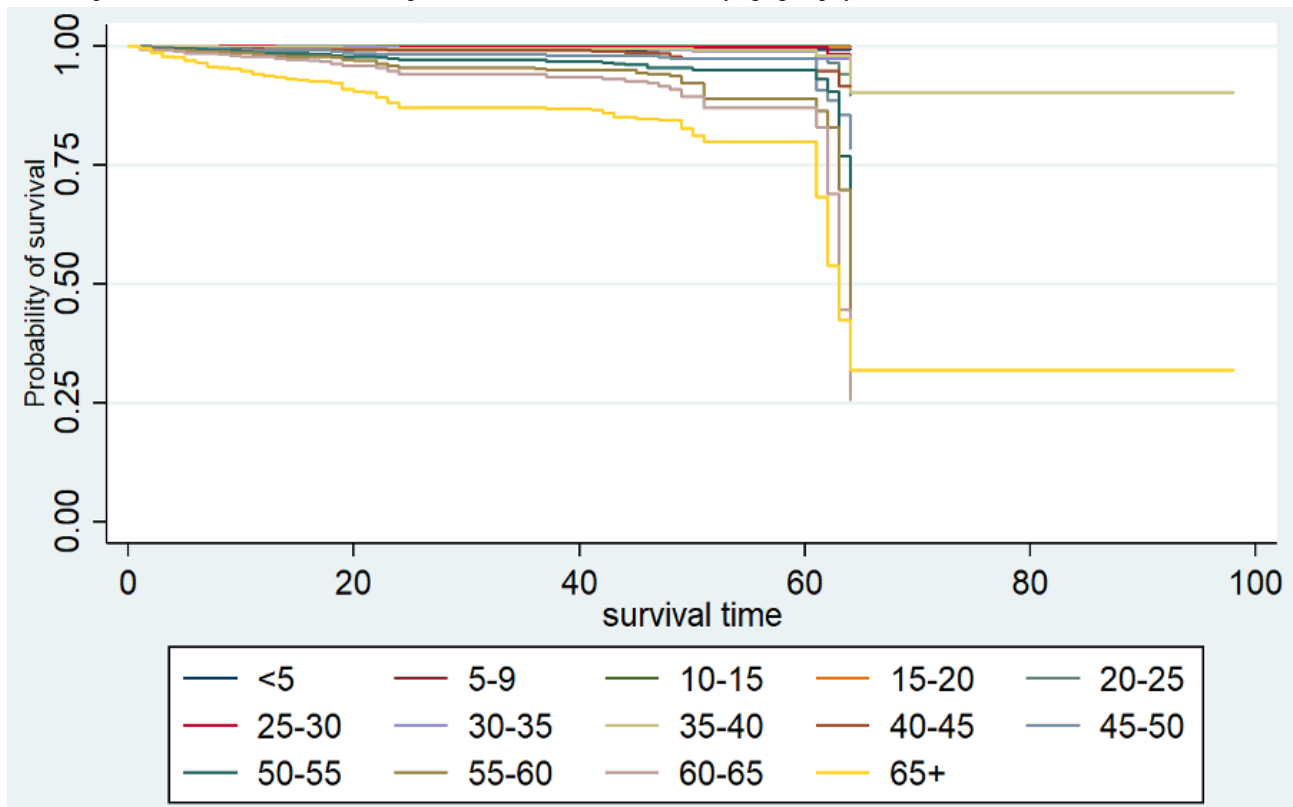


Table 2. Kaplan-Meier estimator: comparison of survival functions for patients with COVID-19 in India by age group.

Test	$\chi^2 (13)$	P value
Log rank	1302.7	<.001
Wilcoxon	1072.26	<.001

Figure 4 depicts that the survival curves from the Kaplan-Meier estimator for different regions of India are significantly different. From Table 3, it can be inferred that the log rank test ($P<.001$) and Wilcoxon test ($P<.001$) for comparing the survival functions are both highly significant, indicating that there are significant differences among the survival curves by region due to regional

variations. Therefore, we can find that both age and region are significantly associated with the survival rate of COVID-19 without adjusting for other covariates. Figure 5 depicts a comparison of the survival curves among the states most affected by COVID-19, which shows the low survival rates in Maharashtra, Delhi, Gujarat, West Bengal, and Rajasthan.

Figure 4. Kaplan-Meier estimate of survival of patients with COVID-19 in India by region.

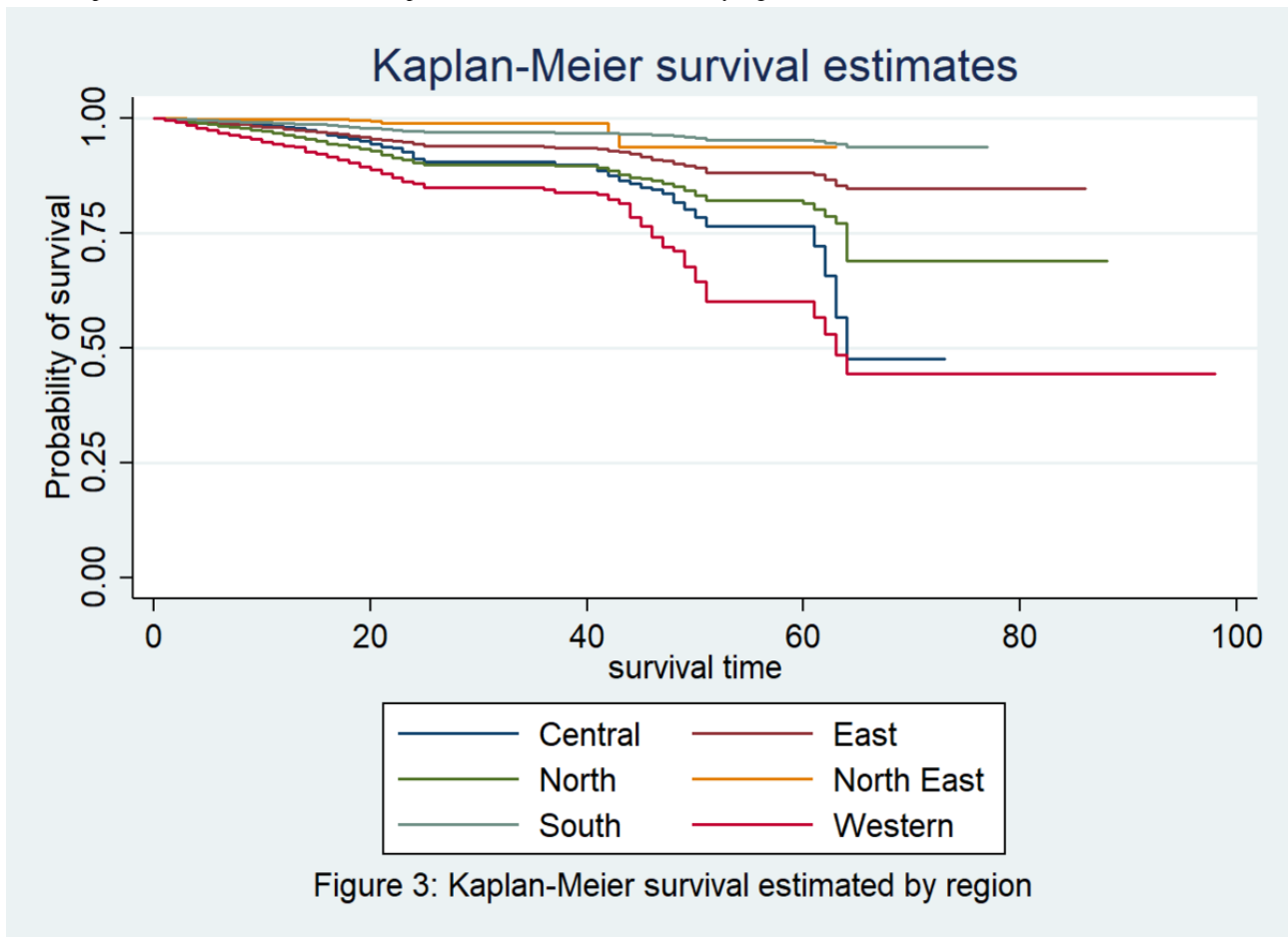


Table 3. Kaplan-Meier estimator: comparison of survival functions for patients with COVID-19 in India by region.

Test	$\chi^2 (5)$	P value
Log rank	1997.43	<.001
Wilcoxon	1175.27	<.001

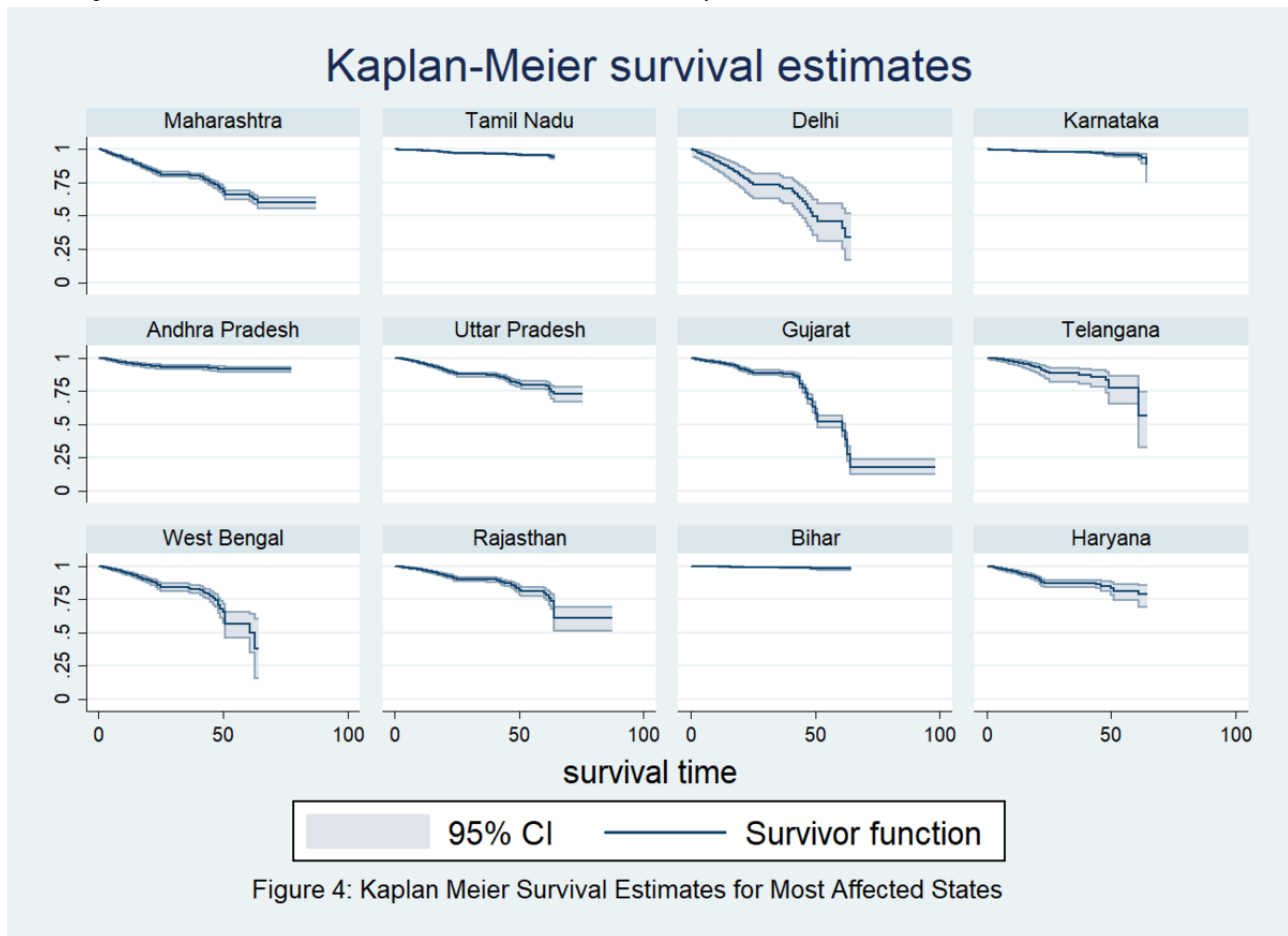
Figure 5. Kaplan-Meier survival estimates for the states in India most affected by COVID-19.

Table 4 presents the results of the survival analysis using the Cox proportional hazard model and reiterates that male patients with COVID-19 have a 1.14 times higher risk of dying than female patients (hazard ratio [HR] 1.14; SE 0.11; 95% CI 0.93-1.38). Coming to the agewise comparison, we observed that patients aged 45-49 years to ≥ 65 years had a 5.83, 10.08, 15.31, 22.03, and 39.21 times higher risk of dying due to COVID-19 than those in the age group of 0-5 years, respectively. The highest risk of death from the disease was among people in the ≥ 65 years age group, with 39.2 times higher risk ($P < .001$) but a larger confidence interval (95% CI 9.73-157.97). When analyzing the survival curves by region, it can be seen that patients in the East, North East, and South Indian regions were at 59%, 14%, and 26% lower risk of dying from COVID-19 infection than those in Central India, respectively, whereas patients in the West Indian region had a 1.9 times higher risk of dying than those in Central India.

A multilevel survival model was applied at the overall national level (Weibull regression) and the state and district level (mixed

effects Weibull regression), taking age group and sex as the covariates. The results of the multilevel survival analysis, that is, after controlling the variability due to clustering of the lower level at the higher level, can be seen in Table 5; the hazard ratio at the India level shows that male patients are at 1.27 times higher risk than female patients of dying from COVID-19 (HR 1.27; SE 0.13), which is almost the same at the state level (HR 1.32; SE 0.13) as well as the district level (HR 1.21; SE 0.13). Significant variability in survival was observed for the age groups older than 45 years. At all levels, we found that the hazard ratio increased with increasing age but decreased across each level. For instance, in the ≥ 65 years age group, the patient is at 39.3 times higher risk of dying overall than a patient in a younger age group (HR 39.3; SE 27.94); meanwhile, at the state level, the hazard ratio is 32.28, and that at the district level is 23.55. Now, from the variance of the errors of model, we can infer that the heterogeneity is greater at the district level (σ_e^2 6.85; SE 1.35) than at the state level (σ_e^2 2.28; SE 0.83).

Table 4. Cox proportional hazard model showing unadjusted hazard ratios and 95% confidence intervals for deaths occurring due to COVID-19 in India.

Independent variable	Unadjusted hazard ratio (SE)	95% CI
Sex		
Female	Reference	N/A ^a
Male	1.14 (0.11)	0.93-1.38
Age group (years)		
0-4	Reference	N/A
5-9	0.00 (N/A)	N/A
10-14	0.00 (N/A)	N/A
15-19	0.74 (0.64)	0.14-4.03
20-24	0.85 (0.67)	0.18-4.01
25-29	0.87 (0.68)	0.19-4.01
30-34	1.44 (1.09)	0.32-6.36
35-39	1.52 (1.15)	0.34-6.73
40-44	3.66 (2.69)	0.87-15.44
45-49	5.83 (4.24) ^b	1.40-24.22
50-54	10.08 (7.27) ^c	2.45-41.42
55-59	15.31 (11.00) ^c	3.74-62.59
60-64	22.03 (15.81) ^c	5.40-89.90
≥65	39.21 (27.88) ^c	9.73-157.97
Region		
Central		
East	0.59 (0.05) ^c	0.50-0.70
North	0.98 (0.09)	0.82-1.16
North East	0.14 (0.04) ^c	0.08-0.24
South	0.26 (0.02) ^c	0.22-0.31
West	1.90 (0.16) ^c	1.61-2.23

^aN/A: not applicable.

^bSignificant at a 1% level of significance.

^cSignificant at a 0.1% level of significance.

Table 5. Multilevel survival analysis by hazard ratio at the national, state, and district levels for deaths among patients with COVID-19 in India.

Independent variable	Level, hazard ratio (SE)		
	National	State ^a	District ^b
Sex			
Female	Reference	Reference	Reference
Male	1.27 (0.13) ^c	1.32 (0.13) ^c	1.21 (0.13)
Age group (years)			
0-4	Reference	Reference	Reference
5-9 years	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
10-14	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
15-19	0.71 (0.62)	0.86 (0.75)	0.49 (0.45)
20-24	0.80 (0.63)	0.84 (0.66)	0.68 (0.54)
25-29	0.81 (0.63)	0.91 (0.71)	0.71 (0.56)
30-34	1.26 (0.96)	1.53 (1.16)	1.22 (0.93)
35-39	1.22 (0.94)	1.37 (1.05)	1.13 (0.87)
40-44	3.36 (2.47)	4.00 (2.94)	3.37 (2.49)
45-49	5.20 (3.78) ^c	6.22 (4.52) ^c	4.89 (3.56) ^c
50-54	8.53 (6.16) ^d	10.70 (7.73) ^d	7.77 (5.63) ^c
55-59	13.99 (10.06) ^d	14.76 (10.62) ^d	11.49 (8.29) ^d
60-64	20.52 (14.74) ^c	18.40 (13.23) ^c	14.19 (10.24) ^c
≥65	39.30 (27.94) ^c	32.28 (23.00) ^c	23.55 (16.84) ^c

^aError variance: 2.28 (0.83).

^bError variance: 6.85 (1.35).

^cSignificant at a 1% level of significance.

^dSignificant at a 0.1% level of significance.

Discussion

Epitomizing the whole study, it is notable that in the stipulated time period, our observation clearly revealed that the survival rate was continually declining, and to date, that trend has not abated. It is worth mentioning that age, sex, and regional variability were important determinants at each step. Also, from this study, it is very clear that the male population in India is more vulnerable to COVID-19, likely due to prevalent comorbidities and the dominant presence of men outside the home (also, our data support the fact that the survival rate of the female population is higher). This study also traced a different pattern for India than for other countries, as the younger population is greater in our country than in most countries where the number of affected people is numerous.

In this study, we strived to identify reliable features associated with survival patterns, and we inevitably scanned the roles of sex, age, and regional variability as controlling factors of the survival rate. For the survival analysis, the study period was 5 months, with death being the event of interest in our analysis. As we evaluated the Kaplan-Meier survival function, we observed that the probability of survival continually declined during the study period of 5 months. During the study period,

no stabilization could be observed. Female patients were found to have better survival rates compared to their male counterparts, as is evident from the Cox proportional hazard results, which may be due to sex differentials in cellular compositions and the immunological microenvironment of the lungs [14,15]. Although we only observed a miniscule difference in the survival curves of male and female patients, it was stated in earlier studies that men with COVID-19 are at higher risk of death and health outcomes, independent of age [16], as men have greater disease burden (diabetes, hypertension, or cardiovascular diseases); therefore, men have shown markedly increased risk of developing severe COVID-19 in comparison to women. Also, a greater proportion of the confirmed cases are male rather than female; this finding is expected in a country with a gender hegemony in which work participation, mobility, and migration are predominately higher for men than for women [17], which makes men more vulnerable to the infection.

Significant variability was observed among age groups, as is evident from all the survival estimates, which show that with increasing age, the risk of dying from COVID-19 increases [18]. It was reported in a study that among comorbid patients with COVID-19, nearly 21% had hypertension, 11% had diabetes, and 7% had cardiovascular disease, which increased their risk

of mortality [19]. In contrast to data from other countries, in India, only 15% of confirmed cases are aged >60 years, and the majority of these patients are in the age bracket of 25-59 years; this is most probably because the older population is the most affected by this pandemic and India has a fairly young population, which may contribute to a lower case fatality rate [9]. Approximately 84% of the patients with COVID-19 were men, and 82% patients overall were above 40 years of age, as reported in an Indian Council of Medical Research study [20]. India is one of the largest countries in the world, and it is highly diverse in every respect. Every province possesses its own demographic features, typical climatic character, and above all, its own lifestyle. Needless to say, these factors play a pivotal

role. Therefore, variation in survival rate is easily traced. Although Western and Central India show continually decreasing survival rates in the framed time period, Eastern, Northeastern and Southern India show slightly better results in terms of survival. Maharashtra, Gujarat, Delhi, Rajasthan, and West Bengal showed alarmingly low survival rates as well.

Finally, this study has depicted a grave scenario of continual degradation of the COVID-19 survival rate in various regions. In essence, we can safely conclude that critical appraisal of the survival rate and thorough analysis of patient data in this study equipped us to identify risk groups and perform comparative studies of various segments of the population in India.

Acknowledgments

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sector.

Conflicts of Interest

None declared.

References

1. Cascella M, Rajnik M, Cuomo A, Dulebohn SC, Di Napoli R. Features, evaluation, and treatment of coronavirus (COVID-19). StatPearls. URL: <https://www.statpearls.com/ArticleLibrary/viewarticle/52171> [accessed 2021-05-03]
2. Kraemer MUG, Yang CH, Gutierrez B, Wu CH, Klein B, Pigott DM, Open COVID-19 Data Working Group, et al. The effect of human mobility and control measures on the COVID-19 epidemic in China. *Science* 2020 May 01;368(6490):493-497 [FREE Full text] [doi: [10.1126/science.abb4218](https://doi.org/10.1126/science.abb4218)] [Medline: [32213647](https://pubmed.ncbi.nlm.nih.gov/32213647/)]
3. How the Republic of Korea flattened the curve without nationwide lockdowns or border closures: report. International Organization for Migration. 2020 May 19. URL: <https://www.iom.int/news/how-republic-korea-flattened-curve-without-nationwide-lockdowns-or-border-closures-report> [accessed 2021-05-04]
4. Coronavirus disease (COVID-19) pandemic. World Health Organization. URL: <https://www.who.int/emergencies/diseases/novel-coronavirus-2019> [accessed 2020-05-07]
5. COVID-19 India. URL: <https://www.covid19india.org/> [accessed 2020-07-05]
6. Dwivedi LK, Rai B, Shukla A, Dey T, Ram U, Shekhar C, et al. Assessing the Impact of Complete Lockdown on COVID-19 Infections in India and its Burden on Public Health Facilities. International Institute for Population Sciences. 2020 Apr 25. URL: https://iipsindia.ac.in/sites/default/files/iips_covid19_aicl.pdf [accessed 2020-04-25]
7. Delamater P, Street EJ, Leslie TF, Yang YT, Jacobsen KH. Complexity of the basic reproduction number (R). *Emerg Infect Dis* 2019 Jan;25(1):1-4. [doi: [10.3201/eid2501.171901](https://doi.org/10.3201/eid2501.171901)] [Medline: [30560777](https://pubmed.ncbi.nlm.nih.gov/30560777/)]
8. Ram U, Kumar P. Pattern and Future Magnitude of COVID 19 Cases in Selected Red Zone Areas of India. International Institute for Population Sciences. 2020 Apr 22. URL: https://www.iipsindia.ac.in/sites/default/files/iips_coivd19_pfm.pdf [accessed 2020-04-22]
9. Dhillon P, Kundu S, Shekhar C, Ram U, Dwivedi LK, Yadav S, et al. Case-Fatality Ratio and Recovery Rate of COVID-19: Scenario of Most Affected Countries and Indian States. International Institute for Population Sciences. URL: https://iipsindia.ac.in/sites/default/files/iips_covid19_cfr.pdf [accessed 2020-04-25]
10. Conticini E, Frediani B, Caro D. Can atmospheric pollution be considered a co-factor in extremely high level of SARS-CoV-2 lethality in Northern Italy? *Environ Pollut* 2020 Jun;261:114465 [FREE Full text] [doi: [10.1016/j.envpol.2020.114465](https://doi.org/10.1016/j.envpol.2020.114465)] [Medline: [32268945](https://pubmed.ncbi.nlm.nih.gov/32268945/)]
11. Xu X, Chen P, Wang J, Feng J, Zhou H, Li X, et al. Evolution of the novel coronavirus from the ongoing Wuhan outbreak and modeling of its spike protein for risk of human transmission. *Sci China Life Sci* 2020 Mar;63(3):457-460 [FREE Full text] [doi: [10.1007/s11427-020-1637-5](https://doi.org/10.1007/s11427-020-1637-5)] [Medline: [32009228](https://pubmed.ncbi.nlm.nih.gov/32009228/)]
12. Unisa S, Rai B, Rai S, Saraswat A, Mondal S, Dhillon P, et al. Potential Factors of Transmission and Spread of COVID 19 in Indian States and Districts. International Institute for Population Sciences. 2020 May 27. URL: https://iipsindia.ac.in/sites/default/files/iips_coivd19_pfts.pdf [accessed 2020-05-27]
13. Tyagi R, Bramhankar M, Pandey M, Kishore M. COVID 19: real-time Forecasts of confirmed cases, active cases, and health infrastructure requirements for India and its states using the ARIMA model. medRxiv Preprint posted online on June 11, 2020. [doi: [10.1101/2020.05.17.20104588](https://doi.org/10.1101/2020.05.17.20104588)]

14. Wei X, Xiao YT, Wang J, Chen R, Zhang W, Yang Y, et al. Sex differences in severity and mortality among patients with COVID-19: evidence from pooled literature analysis and insights from integrated bioinformatic analysis. ArXiv Preprint posted online on March 30, 2020 [[FREE Full text](#)]
15. Heo J, Park JA, Han D, Kim HJ, Ahn D, Ha B, et al. COVID-19 outcome prediction and monitoring solution for military hospitals in South Korea: development and evaluation of an application. J Med Internet Res 2020 Nov 04;22(11):e22131 [[FREE Full text](#)] [doi: [10.2196/22131](https://doi.org/10.2196/22131)] [Medline: [33048824](https://pubmed.ncbi.nlm.nih.gov/33048824/)]
16. Jin J, Bai P, He W, Wu F, Liu XF, Han DM, et al. Gender differences in patients with COVID-19: focus on severity and mortality. Front Public Health 2020;8:152 [[FREE Full text](#)] [doi: [10.3389/fpubh.2020.00152](https://doi.org/10.3389/fpubh.2020.00152)] [Medline: [32411652](https://pubmed.ncbi.nlm.nih.gov/32411652/)]
17. Banerjee A, Raju S. Gendered mobility: women migrants and work in urban India. Econ Polit Wkly 2009 Jul 11;44(28):115-123.
18. Velavan T, Meyer CG. The COVID-19 epidemic. Trop Med Int Health 2020 Mar;25(3):278-280 [[FREE Full text](#)] [doi: [10.1111/tmi.13383](https://doi.org/10.1111/tmi.13383)] [Medline: [32052514](https://pubmed.ncbi.nlm.nih.gov/32052514/)]
19. Singh A, Gupta R, Misra A. Comorbidities in COVID-19: outcomes in hypertensive cohort and controversies with renin angiotensin system blockers. Diabetes Metab Syndr 2020;14(4):283-287 [[FREE Full text](#)] [doi: [10.1016/j.dsx.2020.03.016](https://doi.org/10.1016/j.dsx.2020.03.016)] [Medline: [32283499](https://pubmed.ncbi.nlm.nih.gov/32283499/)]
20. Gupta N, Praharaj I, Bhatnagar T, Vivian Thangaraj JW, Giri S, Chauhan H, ICMR COVID Team. Severe acute respiratory illness surveillance for coronavirus disease 2019, India, 2020. Indian J Med Res 2020;151(2 & 3):236-240 [[FREE Full text](#)] [doi: [10.4103/ijmr.IJMR_1035_20](https://doi.org/10.4103/ijmr.IJMR_1035_20)] [Medline: [32362647](https://pubmed.ncbi.nlm.nih.gov/32362647/)]

Abbreviations

ARIMA: autoregressive integrated moving average

HR: hazard ratio

SARS-CoV: severe acute respiratory syndrome coronavirus

WHO: World Health Organization

Edited by G Eysenbach; submitted 06.08.20; peer-reviewed by M Zheng, GE Iyawa, A Azzam, F Solano Zapata, A Dawood, B Jugdutt; comments to author 22.12.20; revised version received 10.02.21; accepted 13.04.21; published 06.05.21.

Please cite as:

Kundu S, Chauhan K, Mandal D

Survival Analysis of Patients With COVID-19 in India by Demographic Factors: Quantitative Study

JMIR Form Res 2021;5(5):e23251

URL: <https://formative.jmir.org/2021/5/e23251>

doi: [10.2196/23251](https://doi.org/10.2196/23251)

PMID: [33882017](https://pubmed.ncbi.nlm.nih.gov/33882017/)

©Sampurna Kundu, Kirti Chauhan, Debarghya Mandal. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 06.05.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Learning From the Experiences of COVID-19 Survivors: Web-Based Survey Study

Temiloluwa Prioleau¹, PhD

Dartmouth College, Hanover, NH, United States

Corresponding Author:

Temiloluwa Prioleau, PhD

Dartmouth College

9 Maynard Street

Hanover, NH, 03755

United States

Phone: 1 6036468730

Email: tprioleau@dartmouth.edu

Abstract

Background: There are still many unanswered questions about the novel coronavirus; however, a largely underutilized source of knowledge is the millions of people who have recovered after contracting the virus. This includes a majority of undocumented cases of COVID-19, which were classified as mild or moderate and received little to no clinical care during the course of illness.

Objective: This study aims to document and glean insights from the experiences of individuals with a first-hand experience in dealing with COVID-19, especially the so-called mild-to-moderate cases that self-resolved while in isolation.

Methods: This web-based survey study called C19 Insider Scoop recruited adult participants aged 18 years or older who reside in the United States and had tested positive for COVID-19 or antibodies. Participants were recruited through various methods, including online support groups for COVID-19 survivors, advertisement in local news outlets, as well as through professional and other networks. The main outcomes measured in this study included knowledge of contraction or transmission of the virus, symptoms, and personal experiences on the road to recovery.

Results: A total of 72 participants (female, n=53; male, n=19; age range: 18-73 years; mean age: 41 [SD 14] years) from 22 US states were enrolled in this study. The top known source of how people contracted SARS-CoV-2, the virus known to cause COVID-19, was through a family or household member (26/72, 35%). This was followed by essential workers contracting the virus through the workplace (13/72, 18%). Participants reported up to 27 less-documented symptoms that they experienced during their illness, such as brain or memory fog, palpitations, ear pain or discomfort, and neurological problems. In addition, 47 of 72 (65%) participants reported that their symptoms lasted longer than the commonly cited 2-week period even for mild cases of COVID-19. The mean recovery time of the study participants was 4.5 weeks, and exactly one-half of participants (50%) still experienced lingering symptoms of COVID-19 after an average of 65 days following illness onset. Additionally, 37 (51%) participants reported that they experienced stigma associated with contracting COVID-19.

Conclusions: This study presents preliminary findings suggesting that emphasis on family or household spread of COVID-19 may be lacking and that there is a general underestimation of the recovery time even for mild cases of illness with the virus. Although a larger study is needed to validate these results, it is important to note that as more people experience COVID-19, insights from COVID-19 survivors can enable a more informed public, pave the way for others who may be affected by the virus, and guide further research.

(*JMIR Form Res* 2021;5(5):e23009) doi:[10.2196/23009](https://doi.org/10.2196/23009)

KEYWORDS

patient-reported outcomes; coronavirus; COVID-19; outcome; crowdsourcing; social media; internet; survivor; experience

Introduction

The COVID-19 pandemic has significantly impacted the majority of countries in the world, with a total of more than

41.1 million confirmed cases and over 1.1 million deaths reported worldwide as of October 22, 2020 [1]. Months into the pandemic, there are still many unknowns about the novel coronavirus, its transmission, and the experiences of people who have been infected by the virus [2,3]. However, it is known

that the disease severity of COVID-19 can range from mild to critical. Studies from the Centers for Disease Control and Prevention (CDC) in the United States and China show that around 80% of the COVID-19–infected population experience “mild-to-moderate” symptoms [4,5]. Due to the sheer volume of infected persons and the limited capacity of the health care systems, it was and still is recommended that patients with mild-to-moderate cases of COVID-19 manage their illness in isolation [2]. Many of these patients, especially in the early months of the outbreak (ie, February and March 2020 in the United States [6]), received little to no clinical care [7]; hence, there is a critical gap in knowledge about their experiences with COVID-19. According to the CDC, “recognition of factors associated with [the] amplified spread during the early acceleration period will help inform future decisions” and “strengthen systems to detect potential transmission resurgence” [6].

Unlike hospitalized patients, the population of people with mild-to-moderate symptoms of COVID-19 remains largely understudied and are among the “undocumented masses”; yet this population contributes significantly to the rapid transmission of the virus [3,8,9]. To fill in the gap, this study aims to enable the research and larger community to glean insights from the experiences of COVID-19 survivors, especially those whose journey with the virus is not captured in clinical or medical records. Through personal stories, this study aims to bring awareness to how people in the United States contracted the coronavirus, the range of symptoms experienced during illness, duration of illness, and their experiences on the road to recovery. Such knowledge can inform guidelines for the vast majority of cases of COVID-19 that are considered to be of mild-to-moderate severity, especially given the current resurgence in the United States [10]. In this work, we developed and deployed a web-based platform for collating experiential data from persons who recovered and/or are on the road to recovery from COVID-19. As previously mentioned, the primary aim of this study is to garner insights from persons with first-hand experience with COVID-19, especially cases of mild-to-moderate severity that are less documented in literature. A total of 72 subjects from 22 US states with laboratory-confirmed positive tests for COVID-19 (68/72, 94%) and presumptively positive as identified by clinical personnel (4/72, 6%) were recruited to share their experience with COVID-19. Details on the recruitment and participant characteristics are described in the Methods section.

Methods

Study Description

This research study was approved by Committee for Protection of Human Subjects (CPHS) at Dartmouth College, New Hampshire, United States, and is reported in accordance with the CHERRIES (Checklist for Reporting Results of Internet E-Surveys) statement for web-based surveys [11]. We recruited COVID-19 survivors defined as persons who tested positive for coronavirus and/or who were later confirmed to have had the virus by testing positive for antibodies. Participants were recruited through multiple sources, including online support

groups such as Survivor Corps [12], advertisement in local newspapers, and by sharing about the research study through professional and other networks. In total, 72 participants were recruited from May 10 through October 1, 2020, although the majority (ie, 54/72, 73%) of the participants were recruited between May 10 and June 18, 2020. About half of the participants in this study (35/72, 49%) were recruited from the Survivor Corps online support group following the participant’s self-identification of having had COVID-19. For this population, a recruitment message was sent directly to prospective participants informing them about the study and requesting their participation if they were eligible. All interested participants who learned about the study were directed to the project website and invited to first complete a presurvey on the *C19 Insider Scoop* project website to screen for the following eligibility criteria: 18 years or older, living in the United States, and tested positive for COVID-19 and/or COVID-19 antibodies. Participants who met these eligibility requirements were then provided a password by email and a link to the full survey, through which they shared their story of COVID-19. In total, 105 prospective participants completed the presurvey and 72 participants completed the full survey, thereby providing data to be presented in this study. Eligible participants who completed the study requirements (ie, presurvey and full survey) were provided the option of receiving a monetary incentive for participation.

The full survey was developed using the well-established Qualtrics survey software under Dartmouth College’s license to ensure secure data storage behind institutional firewalls. The survey questions in this study were informed by various sources, including the CDC’s coronavirus case report form [13], the National Institutes of Health (NIH) repository of COVID-19 research tools [14], and knowledge gaps identified in the literature [7]. The full survey included questions covering the themes of descriptive characteristics, contraction or transmission of the virus, symptoms and coping strategies, and the road to recovery. In addition, adaptive questioning was implemented such that some questions were only displayed based on responses to previous items. On two occasions, additions were made to the full survey during data collection. More specifically, on May 21, 2020, the symptoms of headaches, muscle or body aches, and dizziness were added to the default checklist provided in this study, after 20 participants had completed the full survey. On May 28, 2020, questions on antibody testing were added, after 32 participants had completed the full survey. Participants who completed the survey before each of these revisions were made have no responses to the new questions added after their participation in the survey. An exact copy of the full survey used in this study is presented in [Multimedia Appendix 1](#).

A complete demographic summary of the participants is presented in [Table 1](#). Of the 72 (female, n=53; male, n=19) participants, 68 (94%) received a positive laboratory test and/or positive antibody test for COVID-19, and the remaining 4 (6%) were confirmed presumptively positive by clinical personnel. Participants’ age ranged from 18 to 73 years. They were residents of a total of 22 US states, with the highest representation from New York. The race demographics included 43 (60%) White, 14 (19%) Black or African American, 6 (8%)

Asian, and 2 (3%) American Indian or Alaska Native participants, and the remaining 7 (10%) reported “other” or mixed race. In addition, a total of 23 (32%) participants had at least one pre-existing medical condition, and 6 (8%) participants had more than one pre-existing medical condition. The most common pre-existing conditions reported included asthma (15/72, 21%), high blood pressure (4/72, 6%), and immunosuppressive conditions (3/72, 4%). Other descriptive

factors such as the participants’ highest level of education, household income, and occupation are detailed in [Table 1](#).

The following subsections describe other categories of the full survey, the results of which will be presented, including (1) onset, testing, and contraction of COVID-19; (2) a deeper look at COVID-19 symptoms; (3) the road to recovery from COVID-19; and (4) insights from COVID-19 survivors.

Table 1. Demographic summary of participants (ie, COVID-19 survivors) in the C19 Insider Scoop Study (N=72).

Characteristics	Value
Age (years), mean (SD)	41 (14)
Age range (years)	18-73
Sex, n (%)	
Female	53 (74)
Male	19 (26)
Race, n (%)	
White	43 (60)
Black or African American	14 (19)
Asian	6 (8)
American Indian or Alaska Native	2 (3)
Other or Mixed race	7 (10)
US state of residence, n (%)	
New York	25 (35)
California	5 (7)
Georgia	5 (7)
Massachusetts	4 (6)
Virginia	4 (6)
Texas	3 (4)
New Hampshire	3 (4)
8 other states ^a	2 (3)
7 other states ^b	1 (1)
Pre-existing medical condition, n (%)	
No, with pre-existing condition	48 (68)
Yes, with pre-existing condition	23 (32)
Not reported	1 (1)
COVID-19 testing, n (%)	
Positive laboratory or antibody test	68 (94)
Presumptively positive	4 (6)
Highest level of education, n (%)	
4-year degree	25 (35)
Professional degree	18 (25)
Some college	14 (19)
High school or GED ^c	6 (8)
Doctorate	5 (7)
2-year degree	4 (6)
Household income level (US \$), n (%)	
<20,000	13 (18)
40,000-79,999	9 (12)
80,000-139,999	25 (35)
140,000-199,999	15 (21)
≥200,000	9 (12)

Characteristics	Value
Not reported	1 (1)
Occupation, n (%)	
Nonessential worker	20 (28)
Essential worker (health care)	14 (19)
Essential worker (non-health care)	14 (19)
Unemployed	12 (17)
Student	7 (10)
Retired	5 (7)

^aAlabama, Ohio, Illinois, Maryland, Louisiana, North Carolina, Washington, Vermont

^bIndiana, Missouri, Arkansas, Wisconsin, New Jersey, Connecticut, Pennsylvania

^cGED: Tests of General Educational Development.

Onset, Testing, and Contraction of COVID-19

Within the full survey, there were a total of 8 questions that addressed topics related to the onset, testing, and contraction of COVID-19, including the following:

1. Did you test positive for COVID-19?
2. When did you start to feel ill or experience symptoms?
3. When did you take the COVID-19 test?
4. Were you working from home before you showed symptoms and/or tested positive for COVID-19?
5. Were you using any precautionary measures before you contracted COVID-19? If so, please share what measures you had in place (eg, strict adherence to social distancing, frequent use of masks in public spaces).
6. Do you know how you contracted COVID-19? If yes, please share as much detail as possible regarding how you contracted COVID-19.
7. Do you know of any others who may have contracted COVID-19 from you?

Responses to the above questions were summarized and are reported in a single subsection in the Results section.

A Deeper Look at COVID-19 Symptoms

Within the full survey, there were a total of 2 questions that addressed topics related to COVID-19 symptoms, including the following:

1. Did you visit a hospital or clinical care for treatment during the course of your illness?
 - [If you were hospitalized], how many days were you hospitalized for?
2. What symptoms did you experience [while ill with COVID-19]? Check all that apply.
 - [If “other” symptoms is selected from the above checklist] What other symptoms did you experience?

Responses to the above questions were summarized and are reported in a single subsection in the Results section.

The Road to Recovery From COVID-19

Within the full survey, there were a total of 4 questions that addressed topics related to the road to recovery from COVID-19, including the following:

1. Are you fully recovered from COVID-19?
 - [If you still have lingering symptoms], please list the lingering symptoms that you currently have.
2. When do you believe you fully recovered from COVID-19 and associated symptoms?
3. How many weeks did it take you to recover from majority of the symptoms associated with COVID-19?
4. Have you experienced any stigma associated with having COVID-19?
 - [If yes], please share more about your experience with stigmas associated with having COVID-19.

Responses to the above questions were summarized and are reported in a single subsection in the Results section.

Insights From COVID-19 Survivors

Within the full survey, there was 1 open-ended question that addressed topics related to insights from COVID-19 survivors, namely:

Are there any comments or insights that you want to share with regards to your experience with COVID-19?

Responses to the above question were summarized and are reported in a single subsection in the Results section.

Results

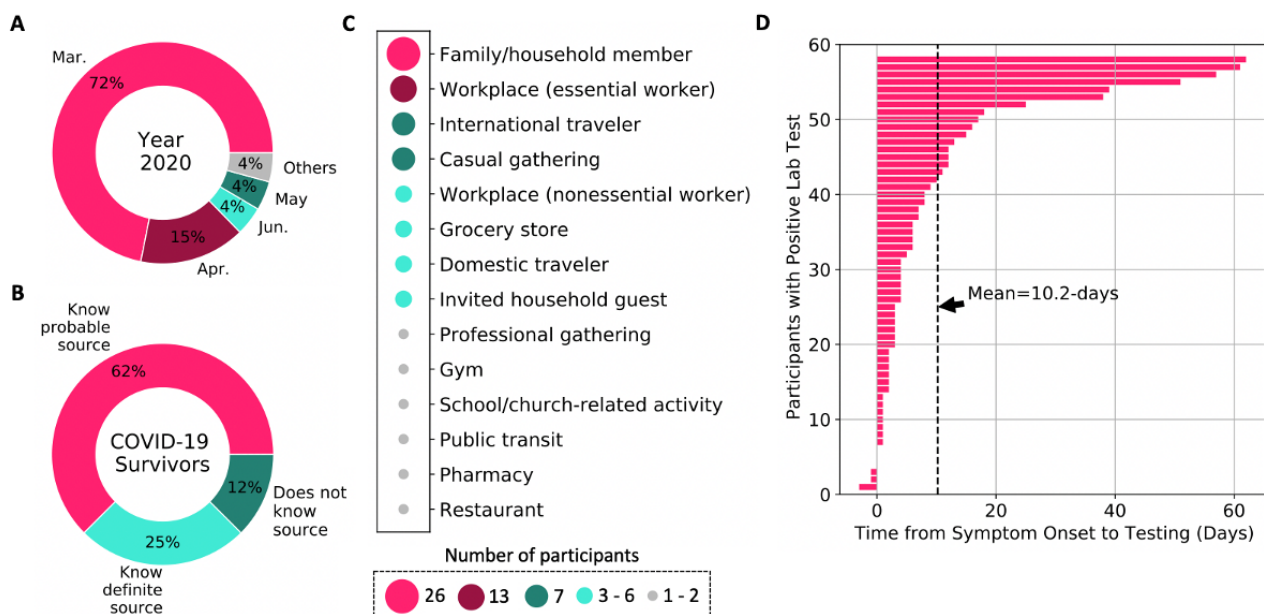
Onset, Testing, and Contraction of COVID-19

Figure 1A shows that 51 of 72 (72%) participants experienced onset of COVID-19 symptoms in March 2020, about 40 days after the first reported case in the United States [15]. Moreover, 30 of 72 (42%) participants were working from home at the time of their symptom onset, whereas 42 (58%) were not working from home. In addition, 46 of 72 (64%) participants reported to have been practicing the recommended precautionary measures prior to the onset of their symptoms, such as use of masks (n=25), social distancing (n=22), frequent handwashing

($n=14$), and use of gloves ($n=12$). However, it is important to note that the majority of participants experienced symptom onset during the first accelerated spread of COVID-19 in the United States, in March 2020, during which recommendations for personal protective practices were only beginning to unfold [6]. In retrospect, 63 of 72 (87%) participants could narrow down the source of how they might have contracted SARS-CoV-2; 62% of these participants shared probable sources, whereas 25% shared what they considered to be definite sources of their infection. From the reported data shown in Figure 1B and 1C, the most prevalent source of virus contraction was from a family or household member. More specifically, 26 of 72 (36%) participants attributed the source of their virus to a family or household member. Additionally, 27 (38%) participants reported that at least one other person in their family

or household had contracted COVID-19 from them. This finding suggests that there is a moderate probability of the coronavirus spreading among persons sharing a living space even when the appropriate precautions are taken to minimize the risk. However, emphasis on family or household transmission of COVID-19 is limited. The second most prevalent source of virus transmission was the essential worker who contracted the virus while on the job. This was reported by 13 of 72 (18%) participants. These two sources of virus transmission were also the top sources identified from persons who shared more definite sources of their infection. The results obtained show other sources of virus transmission reported by the participants, such as casual gatherings (7/72, 10%), workplace for nonessential workers (6/72, 8%), and even the grocery store (4/72, 6%).

Figure 1. Overview of COVID-19 onset, testing, and contraction for participants in this study. (A) Month of symptom onset; (B) knowledge about the source of infection; (C) prevalent sources of virus transmission; and (D) time lag from symptom onset to testing for COVID-19.



During the early months of the outbreak in the United States, limitations in testing was identified as one of the multiple factors that contributed to the rapid spread of the disease [6,8]. Figure 1D shows that following infection, participants experienced large variability in the time it took to gain access to COVID-19 testing; this ranged from 3 days before symptom onset to 62 days after symptom onset among this study's participants. The average time from symptom onset to testing was 10.2 days, with about 10% of the participants getting tested more than 35 days after their symptoms began. Although this study required a COVID-19-positive test to be eligible for participation, we learned during the recruitment phase that a large number of people were continuously denied access to testing during their illness. For example, a patient-led study with over 600 participants who experienced COVID-19 symptoms show that "47.8% were either denied testing or not tested for another reason" [16]. The delay in testing and uncertainty that persons with COVID-19 symptoms experienced is a possible contributor to the further spread of the virus. In addition, a few participants reported that they received treatment for sinus infections, bronchitis, and other conditions without testing, as a result they

may not have isolated early in their journey with the virus due to lack of knowledge.

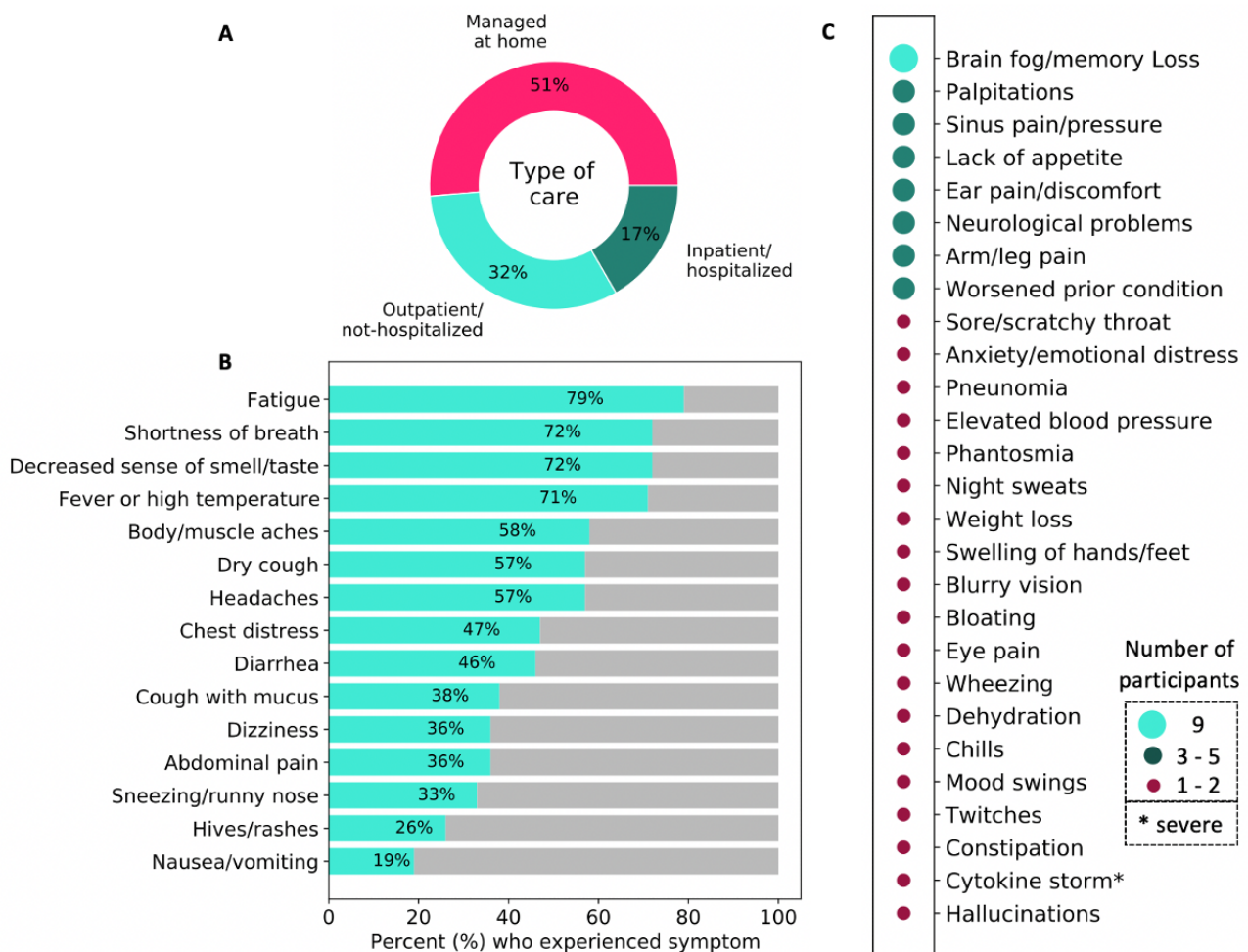
A Deeper Look at COVID-19 Symptoms

There is no dearth on the reported characteristics of COVID-19 for severe cases among hospitalized patients [5,7,17,18]. However, severe cases account for only about 20% of the total population of confirmed COVID-19 cases, and several differences have been identified between mild and severe cases [2,19]. Therefore, it is critical to also understand the journey of nonhospitalized persons with COVID-19 who account for 80% or more of the confirmed cases. The majority of participants in this study (60/72, 83%) managed their symptoms at home and/or in isolation (see Figure 2A); these participants represent people with mild-to-moderate severity of COVID-19. This is in alignment with the estimates in the current literature that about 80% of the COVID-19 cases are mild to moderate as opposed to severe [2,18]. Participants were asked to identify all symptoms they experienced during their illness with COVID-19. Figure 2B shows the prevalence of the well-recognized COVID-19 symptoms reported in this study population ($N=72$),

with the top-4 reported symptoms being fatigue (57/72, 79%), shortness of breath (52/72, 72%), decreased sense of smell or taste (52/72, 72%), and fever (51/72, 71%). These more common symptoms have also been identified in earlier studies, albeit with different percentages for frequency of occurrence in various populations [18,20,21]. More interestingly, in addition to the well-recognized symptoms of COVID-19, 30 of 72 (41%) participants reported other less-documented symptoms that they experienced during their illness. Figure 2C shows a list of other patient-reported symptoms associated with COVID-19 and the prevalence of each symptom among the study population. The most prevalent patient-reported symptoms include brain fog or memory loss, which was reported by 9 (13%) participants; palpitations (ie, elevated heart rate) and sinus pain or pressure, both reported by 5 (7%) participants; and lack of appetite, ear pain or discomfort, and neurological problems, all of which were reported by 4 (6%) participants each. Many of these other symptoms have not been emphasized in prior work. From the

list of 27 additional patient-reported symptoms, cytokine storm is considered the most severe and can lead to rapid deterioration, multiple organ failure, and mortality [22-24]. Cytokine storm is characterized by a lethal overreaction of the immune system in response to an infection or disease. The single participant who experienced cytokine storm (male, 62 years) was tested for COVID-19 on April 4, 2020, and then examined by a registered nurse; however, signs of the ensuing havoc were not identified at that point. Following clinical evaluation, the participant was advised to self-quarantine; however, within 72 hours, he was admitted to the hospital as “COVID-19 had triggered a massive cytokine storm.” He was hospitalized for 36 days, on the ventilator for 9 days, and needed surgery to repair the collapse of his left lung due to COVID-19. Fortunately, he lives to tell his story and strongly urges that COVID-19 patients should be evaluated, early in their illness and possibly during testing, for signs of cytokine storm to prevent rapid deterioration and save lives.

Figure 2. Overview of the type of care received and COVID-19 symptoms reported by participants in this study. (A) The type of care received by COVID-19 survivors; (B) list and prevalence of more common COVID-19 symptoms; and (C) other less-documented symptoms of COVID-19.



The Road to Recovery From COVID-19

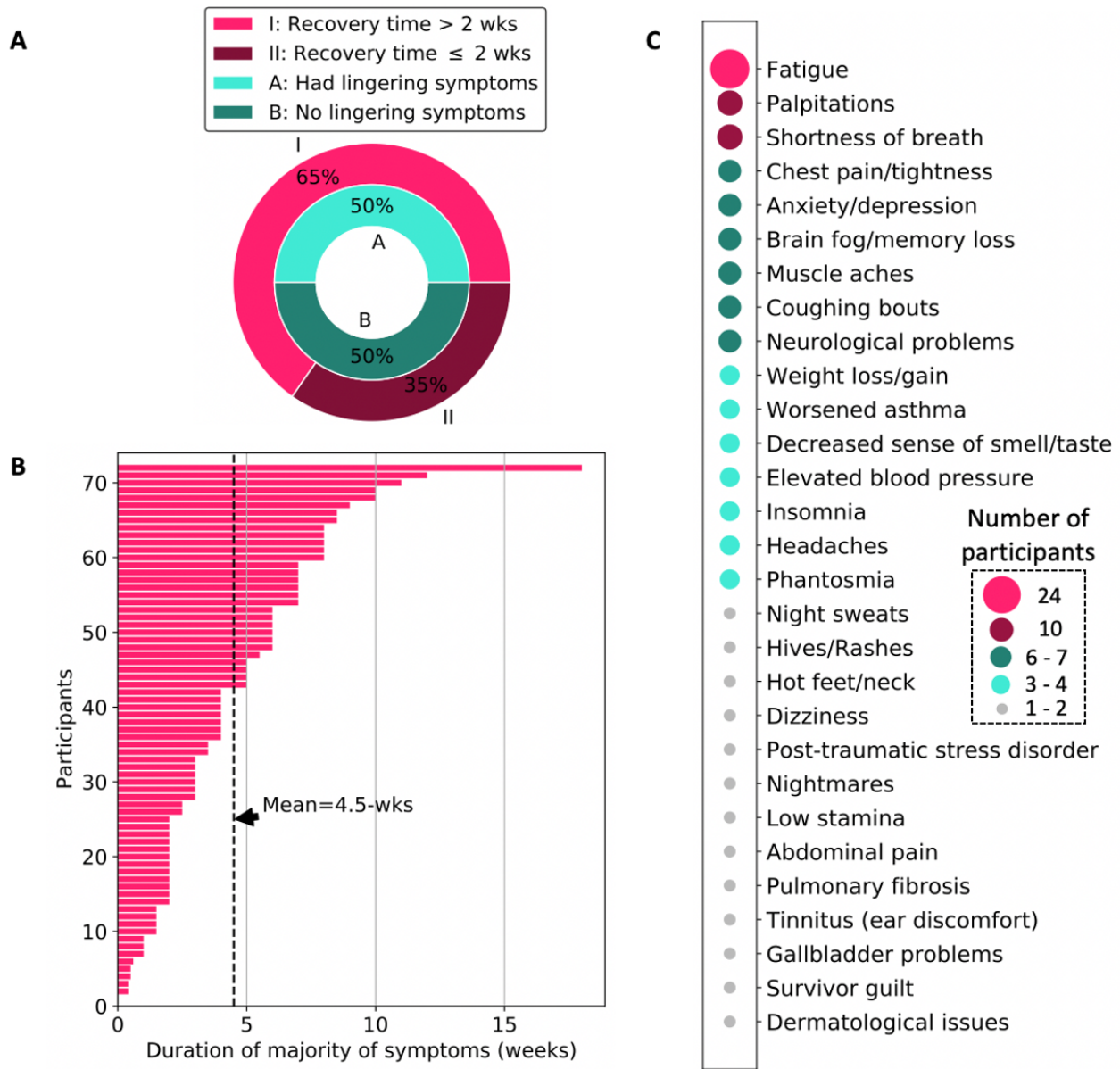
There are many unknowns about the road to recovery for persons who contract COVID-19. According to a February 2020 remark from the World Health Organization (WHO) Director-General, the recovery time is 2 weeks for persons with mild disease and 3-6 weeks for persons with severe or critical disease [25].

However, our results show that 47 of 72 (65%) participants experienced a recovery time longer than 2 weeks (see Figure 3A). More specifically, we found that the mean time to recover from the “majority of symptoms” across all participants was 4.5 weeks; the average recovery time for nonhospitalized patients (ie, patients with mild-to-moderate disease severity) was 4.2 weeks, whereas the average recovery time for

hospitalized patients (ie, patients with a more severe experience of COVID-19) was 6.2 weeks or more. A few hospitalized patients reported that they had not recovered from majority of their symptoms at the time of participating in the study; hence, the length of their illness at the time of participation was used as a proxy (ie, date of participation minus date of symptom onset). Our study also shows that COVID-19 survivors

experienced high variability in their recovery time from the illness, ranging from 0 days as reported by asymptomatic patients to 18 weeks as reported by a 58-year-old female participant with one pre-existing condition and 3 hospital visits for COVID-19-related symptoms, including dehydration and vestibular migraine.

Figure 3. Overview of the road to recovery from COVID-19 for participants in this study. (A) Percentage of COVID-19 survivors who experienced lingering symptoms and a recovery time longer than the commonly cited 2-week period; (B) duration of the majority of symptoms reported (mean duration: 4.5 weeks); and (C) list and prevalence of lingering symptoms associated with COVID-19.



In addition to the longer recovery times identified in this study, we found that 36 of 72 (50%) participants still experienced lingering symptoms of COVID-19 even after an average of 65 days following their illness onset. A list of such lingering symptoms and changes to general health identified by participants of this study is illustrated in Figure 3C. The most common issues identified include fatigue reported by 24 of 72 (33%) participants; followed by shortness of breath and palpitations reported by 10 (14%) participants; and chest pain or tightness, anxiety or depression, brain fog or memory loss, muscle aches, and coughing bouts—all reported by 6-7

(8%-10%) participants each. As shown in Table 2, 37 of 72 (51%) participants reported having experienced stigma following their contraction of COVID-19, whereas 35 (49%) reported otherwise. There were 9 themes that stood out from further descriptions of the stigmas experienced by those who contracted COVID-19. The most prevalent stigma was avoidance by others after recovery (12/72, 17%), followed by people undermining the illness or experience with COVID-19 (5/72, 7%). Other examples of stigma reported include hostility and dismissal from clinical or medical staff and being blamed for contracting and spreading the virus.

Table 2. Experience with stigma reported by participants (ie, COVID-19 survivors) in the C19 Insider Scoop Study (N=72).

Observation	Value, n (%)
Experienced stigma associated with having COVID-19	
Yes	37 (51)
No	35 (49)
Examples of stigmas reported	
Avoidance by others (eg, friends or neighbors) after recovery	12 (17)
People undermining the illness or experience with COVID-19	5 (7)
Hostility and dismissal from clinical or medical staff	4 (6)
Blame for contracting and spreading the virus	4 (6)
Shame and name-calling	3 (4)
Discrimination at places of work and living	3 (4)
Blame for the stay-at-home orders and the economic impact	1 (1)
Reprimand from household members at risk of exposure	1 (1)
Organizations expressing unjust entitlement to medical information	1 (1)

Insights From COVID-19 Survivors

COVID-19 survivors are an invaluable resource not being fully utilized as a source of knowledge. Through the surveys conducted in this study, these survivors shared insights that they believe will help the society at large better cope with the realities of the ongoing pandemic. Many of these insights are targeted toward the larger population who may not yet have been infected by the coronavirus; however, given the recent rise in COVID-19 cases in the United States and other countries [10], these insights are critical for the general public.

Salient insights obtained from these COVID-19 survivors are listed below (with a few direct quotes from participants):

1. “Any symptom can be the virus so self-isolate even if you have a tickle in your throat...It is easy to mistake a mild case of COVID-19 for a cold or allergies” in the early stages.
2. Understand the “fact that doctors and nurses do not have all the answers” so set your expectations accordingly.
3. The journey with COVID-19 can be scary and lonely, as well as physically, mentally, and emotionally taxing. “You have to advocate for yourself every step of the way.”
4. Young and healthy people without any underlying medical conditions can also have a rough journey with COVID-19. Take it seriously. Protect yourselves and others.
5. “The disease presents differently in all patients. It is not a one size fits all.” For some it is short-lived, whereas others experience long-term health issues due to COVID-19.
6. If you contract the virus, take it day by day. Be patient with the recovery process as it can be long and volatile. Symptoms and lingering symptoms can come in waves.
7. Do not wait too long to go to the hospital (if needed), as a delay has led to worse outcomes for some. Additionally, some patients have gone to the hospital and been dismissed or discharged prematurely. Trust your instinct as the sole person with first-hand knowledge of your own illness.

8. Do not allow distrust of the health care system to be your excuse for having a worse outcome. Ask questions during clinical care to ensure that you are being prescribed the best treatment for your specific situation.
9. Some medications used in severe cases can cause hallucination and uncanny dreams.
10. Stigma during or after COVID-19 is a reality. Others may have various reactions to people who contracted the virus, especially during the initial phase of returning back to society. Some people may be cruel and unkind with their actions or words.
11. COVID-19 tests and antibody tests can be inaccurate. There are persons who test positive for COVID-19 and negative for antibodies, and vice versa.
12. It is not sufficient for organizations to rely on body temperature as a primary means for determining COVID-19 symptoms. Some survivors never experienced fever, whereas others only experienced fever for a short time (eg, 24 hours).

Discussion

Principal Findings

This paper presents a descriptive study that summarizes the experiences of first-hand COVID-19 survivors with a particular focus on how they contracted the virus, the range of symptoms observed, the duration of illness, and their experiences on the road to recovery. Our findings show that the top known source of how people contracted COVID-19 was through a family or household member. Participants also reported a range of symptoms that they experienced during their illness, including up to 27 less-documented symptoms such as brain or memory fog, palpitations, ear pain or discomfort, and neurological problems. Another key finding is that the majority of participants (47/72, 65%) experienced a recovery time longer than the commonly cited 2-week period. The mean recovery time in this study was 4.5 weeks, which was more specifically 4.2 weeks for nonhospitalized patients (ie, mild-to-moderate cases) and

6.2 weeks or more for severe cases (ie, patients who were hospitalized for COVID-19 management). In addition, 51% (37/72) of the participants reported that they experienced stigma associated with having COVID-19. Examples of stigmas shared by our study participants include avoidance by others after recovery and people undermining their illness or experience with COVID-19. Finally, participants shared insights from their personal journey with COVID-19 in the hope to inform and encourage the general public, especially others who may not yet have been infected by the coronavirus.

Comparison with Prior Work

Majority of the early literature that presented the characteristics of COVID-19 primarily focused on persons with severe cases of the disease or hospitalized patients [5,18,20,21]. This created a critical gap in the literature with regard to the majority of COVID-19 cases (~80%) that are categorized as mild or moderate (ie, nonhospitalized cases). One of the first questions related to the COVID-19 pandemic is how is it spreading among people? Early sources of virus transmission in the United States were primarily from international travelers returning to the country, such as the “First [Reported] Case of 2019 Novel Coronavirus in the United States” [15]. However, community spread became prevalent not too long after. Later reports by the CDC identified transmission through a family member or work colleague as a prevalent source of virus contraction as well as transmission through social gatherings such as dining at a restaurant [26,27]. Results from this study show agreement that family or household- and work-related transmission of COVID-19 is prominent.

Other studies have now begun to uncover characteristics of mild-to-moderate cases of COVID-19. For example, Liu et al [19] reported viral shedding patterns observed in patients with mild and severe COVID-19. They identified some key differences between patients with mild and severe cases of COVID-19, thereby supporting the need for studying both populations comprehensively. Xu et al [28] presented a telemedicine system to continuously monitor the progression of home-quarantined patients with COVID-19, 74 of whom had confirmed infection. They identified symptoms and trajectories that were indicative of the need for hospitalization. More closely related to this work, Jeon et al [29] used a combination of biomedical literature and social media data to characterize the symptoms of COVID-19. They identified 25 novel symptoms by analyzing social media posts, specifically Twitter posts. Some of the less-common symptoms identified in their work align with the less-documented symptoms reported by the participants of this study, such as ear and eye problems, weight loss, and memory disorder. However, this study’s findings also include symptoms that were not identified by Jeon et al [29], such as worsened pre-existing condition (eg, asthma) and elevated blood pressure.

Even fewer studies in the literature have focused on understanding the time to recovery after COVID-19 infection. One of the earliest sources of information on recovery time was a February 2020 remark from the WHO Director-General, which stated an average recovery time of 2 weeks for persons with

mild disease and 3-6 weeks for those with severe or critical disease [25]. Subsequently, a CDC study found that approximately 35% of symptomatic outpatients with COVID-19 had not returned to their baseline health 14-21 days after the test date [9]. Carfi et al [30] assessed persistent symptoms in patients after acute COVID-19 and found that 87.4% of patients with severe disease reported at least 1 symptom, particularly fatigue and dyspnea, after an average of 60 days from symptom onset. Results from this study show the average recovery time for mild cases is 4.2 weeks and that for severe cases is 6.2 weeks or more. This finding is consistent with the findings from the CDC report [9] and those reported by Carfi et al [30]; however, our results provide concrete numbers based on participants in this study.

Limitations

There are several limitations to this study. First, the sample size of the study is small compared to the total number of people in the United States that have contracted COVID-19. However, this study benefits from diverse recruiting strategies and thus has representation from 22 of the 50 states in the country. Second, this study relies on participants’ ability to recall and share their experience with COVID-19. Self-report-based studies have well-known limitations such as recall bias. However, we believe this bias is minimized because the majority of the study participants experienced symptom onset in March 2020 and were recruited between the months of May and June 2020, suggesting that these participants were recruited about 2 months after they contracted the virus. In addition, the average duration of illness was 4.5 weeks across all participants, and participation in this study required recovery from COVID-19-related illness. Hence, we expect that the majority of the participants enrolled into the study within their first month after recovery, thus maximizing their chance to accurately remember their experience.

Conclusions

This study presents preliminary findings from a small cohort of COVID-19 survivors. It is well known that many people do recover from COVID-19; however, the results from this study show that the journey for some may be long and uncomfortable. To support an accurate depiction of the journey with COVID-19, it is important to recognize that the recovery time can be more than 2 weeks for patients with mild symptoms. In addition, a notable proportion of COVID-19 survivors are considered “long haulers” [31,32]. This refers to people who experience lingering symptoms weeks and even months after initial contraction of the virus. More research is needed to understand the reasons for such diverse experiences with the novel coronavirus. Moreover, thousands of people have died due to infection with the same virus with disproportionate percentages reported among Black and Hispanic populations in the United States [33]. More research is therefore needed to better understand the trajectory of cases that ended in fatality and help prevent a similar outcome for others. One of such efforts should be toward the early detection of the cytokine storm triggered by SARS-CoV-2, as this has been shown to cause fatality for some of the affected population.

Acknowledgments

The author acknowledges the many participants and advocates of the *C19 Insider Scoop* project with special thanks to S. Prioleau for designing and building the project website for data collection. This work is supported by Dartmouth COVID-19 Spark funding and a “RAPID” grant from the National Science Foundation (NSF) (award number: 2031546). The author TP conceived the study, developed the survey, curated the data, performed analyses, and wrote the manuscript. This author has full access to all the data of this study and takes responsibility for the integrity of the data and accuracy of the data analysis.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Survey questions administered to study participants.

[[PDF File \(Adobe PDF File\), 158 KB - formative_v5i5e23009_app1.pdf](#)]

References

1. WHO Coronavirus Disease (COVID-19) Dashboard. World Health Organization. URL: <https://covid19.who.int/> [accessed 2020-10-22]
2. Interim Clinical Guidance for Management of Patients with Confirmed Coronavirus Disease (COVID-19). Centers for Disease Control and Prevention. URL: <https://www.cdc.gov/coronavirus/2019-ncov/hcp/clinical-guidance-management-patients.html> [accessed 2020-07-02]
3. Li R, Pei S, Chen B, Song Y, Zhang T, Yang W, et al. Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2). *Science* 2020 May 01;368(6490):489-493 [FREE Full text] [doi: [10.1126/science.abb3221](https://doi.org/10.1126/science.abb3221)] [Medline: [32179701](https://pubmed.ncbi.nlm.nih.gov/32179701/)]
4. Wu Z, McGoogan JM. Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China: Summary of a report of 72 314 cases from the Chinese Center for Disease Control and Prevention. *JAMA* 2020 Apr 07;323(13):1239-1242. [doi: [10.1001/jama.2020.2648](https://doi.org/10.1001/jama.2020.2648)] [Medline: [32091533](https://pubmed.ncbi.nlm.nih.gov/32091533/)]
5. CDC COVID-19 Response Team. Severe outcomes among patients with coronavirus disease 2019 (COVID-19) - United States, February 12-March 16, 2020. *MMWR Morb Mortal Wkly Rep* 2020 Mar 27;69(12):343-346 [FREE Full text] [doi: [10.15585/mmwr.mm6912e2](https://doi.org/10.15585/mmwr.mm6912e2)] [Medline: [32214079](https://pubmed.ncbi.nlm.nih.gov/32214079/)]
6. Schuchat A, CDC COVID-19 Response Team. Public health response to the initiation and spread of pandemic COVID-19 in the United States, February 24-April 21, 2020. *MMWR Morb Mortal Wkly Rep* 2020 May 08;69(18):551-556 [FREE Full text] [doi: [10.15585/mmwr.mm6918e2](https://doi.org/10.15585/mmwr.mm6918e2)] [Medline: [32379733](https://pubmed.ncbi.nlm.nih.gov/32379733/)]
7. 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](https://doi.org/10.15585/mmwr.mm6915e3)] [Medline: [32298251](https://pubmed.ncbi.nlm.nih.gov/32298251/)]
8. Cohen J, Kupferschmidt K. Countries test tactics in 'war' against COVID-19. *Science* 2020 Mar 20;367(6484):1287-1288. [doi: [10.1126/science.367.6484.1287](https://doi.org/10.1126/science.367.6484.1287)] [Medline: [32193299](https://pubmed.ncbi.nlm.nih.gov/32193299/)]
9. Tenforde M, Kim S, Lindsell C, Billig Rose E, Shapiro N, Files DC, CDC COVID-19 Response Team, et al. Symptom duration and risk factors for delayed return to usual health among outpatients with COVID-19 in a Multistate Health Care Systems Network - United States, March-June 2020. *MMWR Morb Mortal Wkly Rep* 2020 Jul 31;69(30):993-998 [FREE Full text] [doi: [10.15585/mmwr.mm6930e1](https://doi.org/10.15585/mmwr.mm6930e1)] [Medline: [32730238](https://pubmed.ncbi.nlm.nih.gov/32730238/)]
10. New COVID-19 Cases Worldwide. John Hopkins Coronavirus Resource Center. URL: <https://coronavirus.jhu.edu/data/new-cases> [accessed 2020-10-22]
11. Eysenbach G. Improving the quality of web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). *J Med Internet Res* 2004 Sep 29;6(3):e34 [FREE Full text] [doi: [10.2196/jmir.6.3.e34](https://doi.org/10.2196/jmir.6.3.e34)] [Medline: [15471760](https://pubmed.ncbi.nlm.nih.gov/15471760/)]
12. Berrent D. Survivor Corps. URL: <https://www.survivorcorps.com> [accessed 2020-04-19]
13. Human Infection with 2019 Novel Coronavirus Case Report Form. Centers for Disease Control and Prevention. URL: <https://www.cdc.gov/coronavirus/2019-ncov/downloads/pui-form.pdf> [accessed 2020-04-16]
14. NIH Public Health Emergence and Disaster Research Response (DR2) - NIH Repository of COVID-19 Research Tools. National Institute of Environmental Health Sciences. URL: <https://dr2.nlm.nih.gov/> [accessed 2020-04-19]
15. Holshue ML, DeBolt C, Lindquist S, Lofy KH, Wiesman J, Bruce H, et al. First case of 2019 novel coronavirus in the United States. *N Engl J Med* 2020 Mar 05;382(10):929-936. [doi: [10.1056/nejmoa2001191](https://doi.org/10.1056/nejmoa2001191)]
16. Assaf G, Davis H, Mccorkell L, Wei H, Brooke O. An Analysis of the Prolonged COVID-19 Symptoms Survey by Patient-Led Research Team. Patient-Led Research Collaborative. 2020 May 11. URL: <https://patientresearchcovid19.com/research/report-1/> [accessed 2021-04-27]

17. Guan W, Ni Z, Hu Y, Liang W, Ou C, He J, China Medical Treatment Expert Group for Covid-19. Clinical characteristics of coronavirus disease 2019 in China. *N Engl J Med* 2020 Apr 30;382(18):1708-1720 [[FREE Full text](#)] [doi: [10.1056/NEJMoa2002032](https://doi.org/10.1056/NEJMoa2002032)] [Medline: [32109013](#)]
18. Chen N, Zhou M, Dong X, Qu J, Gong F, Han Y, et al. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. *Lancet* 2020 Feb 15;395(10223):507-513 [[FREE Full text](#)] [doi: [10.1016/S0140-6736\(20\)30211-7](https://doi.org/10.1016/S0140-6736(20)30211-7)] [Medline: [32007143](#)]
19. Liu Y, Yan L, Wan L, Xiang T, Le A, Liu J, et al. Viral dynamics in mild and severe cases of COVID-19. *Lancet Infect Dis* 2020 Jun;20(6):656-657 [[FREE Full text](#)] [doi: [10.1016/S1473-3099\(20\)30232-2](https://doi.org/10.1016/S1473-3099(20)30232-2)] [Medline: [32199493](#)]
20. Wang D, Hu B, Hu C, Zhu F, Liu X, Zhang J, et al. Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus-infected pneumonia in Wuhan, China. *JAMA* 2020 Mar 17;323(11):1061-1069 [[FREE Full text](#)] [doi: [10.1001/jama.2020.1585](https://doi.org/10.1001/jama.2020.1585)] [Medline: [32031570](#)]
21. Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *The Lancet* 2020 Feb;395(10223):497-506. [doi: [10.1016/s0140-6736\(20\)30183-5](https://doi.org/10.1016/s0140-6736(20)30183-5)]
22. Moore JB, June CH. Cytokine release syndrome in severe COVID-19. *Science* 2020 May 01;368(6490):473-474. [doi: [10.1126/science.abb8925](https://doi.org/10.1126/science.abb8925)] [Medline: [32303591](#)]
23. Coperchini F, Chiovato L, Croce L, Magri F, Rotondi M. The cytokine storm in COVID-19: An overview of the involvement of the chemokine/chemokine-receptor system. *Cytokine Growth Factor Rev* 2020 Jun;53:25-32 [[FREE Full text](#)] [doi: [10.1016/j.cytogfr.2020.05.003](https://doi.org/10.1016/j.cytogfr.2020.05.003)] [Medline: [32446778](#)]
24. Sinha P, Matthay MA, Calfee CS. Is a "cytokine storm" relevant to COVID-19? *JAMA Intern Med* 2020 Sep 01;180(9):1152-1154. [doi: [10.1001/jamainternmed.2020.3313](https://doi.org/10.1001/jamainternmed.2020.3313)] [Medline: [32602883](#)]
25. WHO Director-General's opening remarks at the media briefing on COVID-19 - 24 February 2020. World Health Organization. 2020 Feb 23. URL: <https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---24-february-2020> [accessed 2020-10-22]
26. Tenforde M, Billig Rose E, Lindsell C, Shapiro N, Files D, Gibbs KW, CDC COVID-19 Response Team. Characteristics of adult outpatients and inpatients with COVID-19 - 11 Academic Medical Centers, United States, March-May 2020. *MMWR Morb Mortal Wkly Rep* 2020 Jul 03;69(26):841-846 [[FREE Full text](#)] [doi: [10.15585/mmwr.mm6926e3](https://doi.org/10.15585/mmwr.mm6926e3)] [Medline: [32614810](#)]
27. Fisher KA, Tenforde MW, Feldstein LR, Lindsell CJ, Shapiro NI, Files DC, CDC COVID-19 Response Team. Community and close contact exposures associated with COVID-19 among symptomatic adults ≥18 Years in 11 Outpatient Health Care Facilities - United States, July 2020. *MMWR Morb Mortal Wkly Rep* 2020 Sep 11;69(36):1258-1264 [[FREE Full text](#)] [doi: [10.15585/mmwr.mm6936a5](https://doi.org/10.15585/mmwr.mm6936a5)] [Medline: [32915165](#)]
28. Xu H, Huang S, Qiu C, Liu S, Deng J, Jiao B, et al. Monitoring and Management of Home-Quarantined Patients With COVID-19 Using a WeChat-Based Telemedicine System: Retrospective Cohort Study. *J Med Internet Res* 2020 Jul 02;22(7):e19514 [[FREE Full text](#)] [doi: [10.2196/19514](https://doi.org/10.2196/19514)] [Medline: [32568727](#)]
29. Jeon J, Baruah G, Sarabadani S, Palanica A. Identification of risk factors and symptoms of COVID-19: Analysis of biomedical literature and social media data. *J Med Internet Res* 2020 Oct 02;22(10):e20509 [[FREE Full text](#)] [doi: [10.2196/20509](https://doi.org/10.2196/20509)] [Medline: [32936770](#)]
30. Carfi A, Bernabei R, Landi F, Gemelli Against COVID-19 Post-Acute Care Study Group. Persistent symptoms in patients after acute COVID-19. *JAMA* 2020 Aug 11;324(6):603-605 [[FREE Full text](#)] [doi: [10.1001/jama.2020.12603](https://doi.org/10.1001/jama.2020.12603)] [Medline: [32644129](#)]
31. Rubin R. As their numbers grow, COVID-19 "Long Haulers" stump experts. *JAMA* 2020 Sep 23;324(14):1381-1383. [doi: [10.1001/jama.2020.17709](https://doi.org/10.1001/jama.2020.17709)] [Medline: [32965460](#)]
32. Mahase E. Covid-19: What do we know about "long covid"? *BMJ* 2020 Jul 14;370:m2815. [doi: [10.1136/bmj.m2815](https://doi.org/10.1136/bmj.m2815)] [Medline: [32665317](#)]
33. Price-Haywood EG, Burton J, Fort D, Seoane L. Hospitalization and mortality among Black patients and White patients with Covid-19. *N Engl J Med* 2020 Jun 25;382(26):2534-2543. [doi: [10.1056/nejmsa2011686](https://doi.org/10.1056/nejmsa2011686)]

Abbreviations

- CDC:** Centers for Disease Control and Prevention
- CHERRIES:** Checklist for Reporting Results of Internet E-Surveys
- CPHS:** Committee for Protection of Human Subjects
- NIH:** National Institutes of Health
- WHO:** World Health Organization

Edited by G Eysenbach; submitted 29.07.20; peer-reviewed by B Chen, R Zhong, M Wallwiener; comments to author 10.09.20; revised version received 22.10.20; accepted 13.04.21; published 11.05.21.

Please cite as:

Prioleau T

Learning From the Experiences of COVID-19 Survivors: Web-Based Survey Study

JMIR Form Res 2021;5(5):e23009

URL: <https://formative.jmir.org/2021/5/e23009>

doi: [10.2196/23009](https://doi.org/10.2196/23009)

PMID: [33878012](https://pubmed.ncbi.nlm.nih.gov/33878012/)

©Temiloluwa Prioleau. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 11.05.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Changes in Perceptions and Use of Mobile Technology and Health Communication in South Africa During the COVID-19 Lockdown: Cross-sectional Survey Study

Alex Emilio Fischer¹, BSc, MPH; Tanya Van Tonder², BA; Siphamandla B Gumede¹, MSc; Samanta T Lalla-Edward¹, PhD

¹Ezintsha, Faculty of Health Sciences, University of Witwatersrand, Johannesburg, South Africa

²Opinion Solutions, Johannesburg, South Africa

Corresponding Author:

Alex Emilio Fischer, BSc, MPH

Ezintsha

Faculty of Health Sciences

University of Witwatersrand

Sunnyside Office Park

31 Princess of Wales Terrace Parktown

Johannesburg, 2193

South Africa

Phone: 27 073 776 2705

Email: fischermHealth@gmail.com

Abstract

Background: In late March 2020, South Africa implemented a 5-stage COVID-19 Risk Adjusted Strategy, which included a lockdown that required all residents to remain home to prevent the spread of COVID-19. Due to this lockdown, individuals have been forced to find and use alternatives for accomplishing tasks including shopping, socializing, working, and finding information, and many have turned to the internet and their mobile devices.

Objective: This study aimed to describe how South Africans consume and internalize information surrounding the COVID-19 outbreak in order to determine whether the COVID-19 lockdown and social isolation have influenced technology behavior, particularly in terms of health communication and information.

Methods: From June 24 to August 24, 2020, people in South Africa were invited to complete a survey through the Upinion mobile app, an online data collection resource. The survey collected information on demographics, and technology use during the lockdown, and COVID-19 knowledge.

Results: There were 405 participants, of which 296 (73.06%) were female. A total of 320 (79.01%) participants had a tertiary school education, 242 (59.75%) were single, and 173 (42.72%) had full-time employment. The lockdown forced 363 (89.63%) participants to use more technology, especially for work (n=140, 24.05%) and social media/communication (n=133, 22.85%). Security or privacy issues (n=46, 38.98%) and unfamiliarity with technology (n=32, 27.12%) were identified as the most common issues faced by the 127 (31.36%) participants who were unsure about using technology prior to the lockdown. Almost all participants (n=392, 96.79%) stated that they would continue using technology after the lockdown. Multimedia (n=215, 53.09%), mobile phone content (n=99, 24.44%), and health organizations and professionals (n=91, 22.47%) were the main sources of COVID-19 information. Most participants (n=282, 69.63%) felt that they had enough information. Two-thirds (n=275, 67.90%) of participants stated that they had used their mobile phones for health information before the lockdown, with web searches (n=109, 26.91%), social media (n=58, 14.32%), and government and institutional websites (n=52; 12.84%) serving as their main sources of information. Overall, the mean COVID-19 knowledge score was 8.8 (out of 10), and 335 (82.72%) had adequate knowledge (scored ≥ 8). Males were less likely to identify the correct transmission routes, and single participants were less likely to identify the signs and symptoms of the coronavirus. Tertiary school graduates were 4 times more likely to correctly identify the routes and 2 times more likely to identify how to stop the spread of the virus. People aged 43-56 years were 4 times more likely to identify how the coronavirus can be prevented, and participants ≥ 57 years were 2.6 times more likely to obtain a knowledge score of 10 when compared to those under 29 years of age.

Conclusions: This study has shown that the COVID-19 lockdown has forced people to increase technology use, and people plan to continue using technology after the lockdown is lifted. Increased technology use was seen across a variety of fields; however, barriers including privacy, unfamiliarity, and data costs were identified. This population showed high COVID-19 knowledge, although the use of web searches and social media, instead of government and institutional websites, increases the potential for health misinformation to be spread.

(*JMIR Form Res* 2021;5(5):e25273) doi:[10.2196/25273](https://doi.org/10.2196/25273)

KEYWORDS

coronavirus; SARS-CoV-2; COVID-19; technology; mHealth; app

Introduction

On March 11, 2020, the World Health Organization declared the COVID-19 outbreak a pandemic [1]; thereafter, many countries followed China's "lockdown" approach to reduce new cases. In March 2020, South Africa began a 21-day, level 5 lockdown as part of a 5-stage COVID-19 Risk Adjusted Strategy. During this period, only hospitals, clinics, grocery stores, and pharmacies remained open, and only essential personnel (eg, doctors, nurses, police) were permitted to leave their homes, although there were some controlled exceptions for medical care or essential supplies like food and medicine [2].

During the lockdown, cases and preventative measures have been well documented and investigated, both globally and in South Africa [3,4]. However, the behavioral effects of the lockdown are not as well known. With social distancing, individuals have been forced to find and use alternatives to accomplish tasks such as shopping, working, attending school, and staying informed, and many have turned to their mobile devices. In China there was a 30% increase in app use during their lockdown [5-7], while a global analysis of five popular social media platforms (Twitter, Instagram, YouTube, Reddit, and Gab) identified 8 million COVID-19-related posts and comments over the first 45 days of 2020 [8].

In a sense, this is all forced use of technology since people have limited alternatives to meet their needs, and to engage with this captive audience, many governments and institutions have introduced mobile health (mHealth) interventions to disseminate information during the pandemic [9-12]. Specifically, the South African government has implemented a website that provides information on COVID-19, the Risk Adjusted Strategy, preventative measures, news and updates, and links to other resources [13]. These additional resources include a WhatsApp support line, an emergency telehealth hotline, social media message campaigns, and updates from the South African Government and National Department of Health websites [3,13,14].

With all of this electronic communication resulting from COVID-19, researchers have taken the opportunity to investigate how it has influenced digital health, and a variety of studies have already been conducted. Some studies have harnessed big data to predict outbreak hotspots with algorithm-based web mining [8,15-17], while others have looked at how individuals share and consume COVID-19-related content [18]. A study from India showed that more than half of participants (n=58,

56.3%) had adequate information regarding COVID-19; however, their primary source of information was from multimedia (radio, TV, newspaper) (n=57, 55.3%), and only 22 (21.4%) relied on the internet as their main source for information [19].

Despite high mobile penetration in low- and middle-income countries [20,21], there are still individuals who have not embraced technology for various reasons, including security and privacy concerns, data costs, and an inability to understand modern electronics [22]. With the limitations set by the lockdown, increased exposure to technology may have altered some people's perceptions and use of technology. This study aimed to describe how South Africans consume and internalize information surrounding the COVID-19 outbreak to assess whether the COVID-19 lockdown and social isolation has influenced technology behavior, particularly for health communication and information.

Methods

Study Design

This South African cross-sectional study was conducted electronically, administered through the Upinion mobile app, an online data collection resource. Participants were included if they were an existing or new Upinion user with current access to surveys on the app, ≥18 years of age, and able to provide online consent. Individuals were excluded if they were not able to access the Upinion app, were younger than 18 years, or refused to participate.

Data Collection

From June 24 to August 24, 2020, existing and new Upinion users were invited to complete a survey through Upinion notifications and advertisements on social media platforms, respectively. Once an individual agreed to participate in the current study, they were able to provide informed consent through the app and then register for the survey group [23]. The participant was then given access to the survey, which was completed through their mobile phone. During the survey, all answers were recorded electronically in the backend of the app.

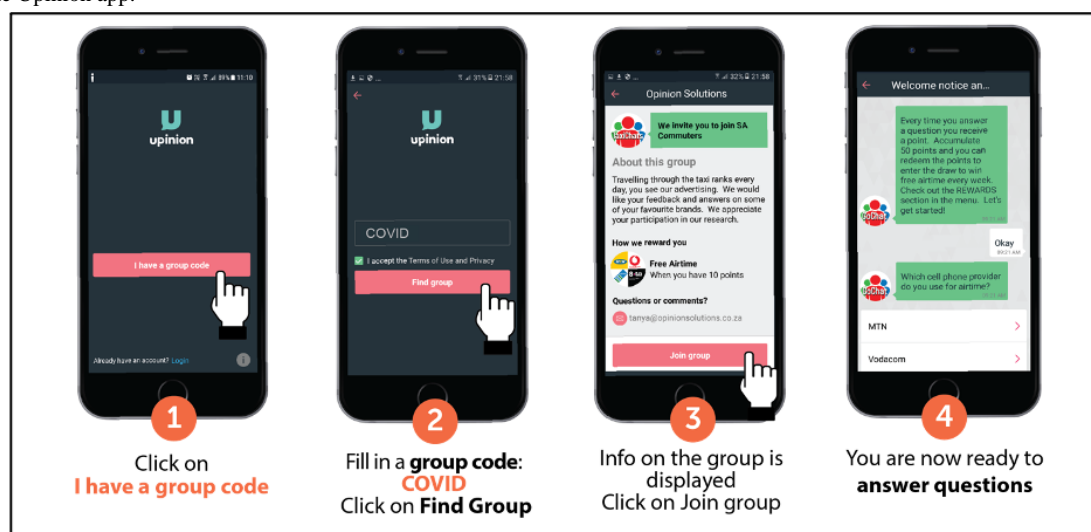
A mobile app was used to collect data as this was deemed the easiest way to gather responses, while obeying the lockdown restrictions and ensuring the safety of both participants and data collectors. This method of online distribution of a survey and accompanying electronic consent has been used with increasing frequency, particularly during the COVID-19 pandemic for studies with similar methodologies [19,24-27].

The Upinion App

The Upinion messaging and data collection app was developed in 2014 by Upinion, a people-centric research technology company based in the Netherlands, and its use in Southern African Development Community countries is licensed to Opinion Solutions. The app was developed as a way to collect feedback from affected communities in any response effort in order to provide better and more efficient support. It serves as an outlet for those affected by crisis to share their unique problems, needs and solutions, so that nongovernmental organizations have a grass-roots understanding of the situation on the ground, allowing for tailored interventions. This has been

used by nonprofit organizations like Oxfam to identify the needs of refugee communities [23], and research institutes like the Wits Reproductive Health and HIV Institute to administer health-related surveys directly via participants' mobile phones [28]. Upinion does not collect personal data, but rather personal data is collected through survey questions and the participant shares this voluntarily. Upinion encrypts all mobile phone numbers and IP addresses in compliance with General Data Protection Regulation and is also ISO/IEC (International Organization for Standardization/International Electrotechnical Commission) 27001 certified. Screenshots of the Upinion app are presented in [Figure 1](#).

Figure 1. The Upinion app.



Survey Development

This survey was adapted from the survey *Demographic Data and Structure Knowledge Questionnaire Regarding Prevention of COVID-19*, used in a similar study from India [19]. The original questionnaire consists of two sections—the first comprising 8 questions to explore demographic information and the second 10 questions that focused on COVID-19 knowledge. In our survey ([Multimedia Appendix 1](#)), we have modified the sections on demographic information and COVID-19 knowledge to reflect the South African context, and a third section was added to explore participants' technology use during the COVID-19 outbreak.

Data Analysis

Upinion has a built-in dashboard to monitor responses in real time; however, the final data set was exported to Excel (Microsoft Corp) for cleaning and coding, then exported to Stata V.15 (StataCorp) for analysis. Demographic information, technology use, and COVID-19 knowledge questions were all described as frequency and percentages. A mean knowledge score (with standard deviation) was also calculated across all 10 knowledge questions, with a score below 6 considered inadequate knowledge, 6-8 considered moderately adequate knowledge, and a score above 8 considered adequate knowledge [19].

The Pearson chi-square test was used to assess trends of association between outcome variables (COVID-19 knowledge

and technology use) and demographic characteristics. Logistic regression models (bivariate [not included in this paper] and multivariable models) were constructed for the outcome variables to control for confounders and identify independent predictors. These predictors were reported as crude (not included in this paper) and adjusted odds ratios (aOR), with 95% CI and *P* values (<.05 was considered significant).

Ethical Consideration and Approval

Ethics approval was obtained from the University of the Witwatersrand Human Research Ethics Committee (nonmedical) (reference number 200512). Survey respondents did not receive any compensation for participation.

Results

Demographics

Participants' demographic data are presented in [Table 1](#). Of the 405 participants, 84 (20.74%) were 28 years or younger, 165 (40.74%) were between the ages of 29 and 42 years, 110 (27.16%) were between the ages of 43 and 56 years, and 46 (11.36%) were 57 years or older. There were 296 (73.06%) females, 320 (79.01%) participants had completed tertiary school education, and 242 (59.75%) were single. A total of 173 (42.72%) participants had full-time employment, 74 (18.27%) were casually employed, 29 (7.16%) were students, and 129 (32.85%) were unemployed.

Table 1. Demographic characteristics.

Characteristic	Participants (N=405), n (%) ^a
Age (years)	
18-28	84 (20.74)
29-42	165 (40.74)
43-56	110 (27.16)
≥57	46 (11.36)
Sex	
Female	296 (73.09)
Male	109 (26.91)
Education	
Primary school or less	1 (0.25)
Secondary school	84 (20.74)
Tertiary school (any)	320 (79.01)
Marital status	
Married	163 (40.25)
Single	242 (59.75)
Employment status	
Casually employed	74 (18.27)
Full-time employment	173 (42.72)
Student	29 (7.16)
Unemployed	129 (31.85)

^aPercentages may not add up to 100.00% due to rounding.

Technology Use

A total of 363 (89.63%) participants stated that the lockdown had forced them to use more technology, and the greatest increases in use were for work (n=140, 24.05%), social media/communication (n=133, 22.85%), shopping (n=78, 13.4%), and news and information (n=70, 12.03%). Nearly one-third (n=127, 31.36%) of participants stated that they were unsure about using technology before the lockdown, with security and privacy issues (n=46, 38.98%) and unfamiliarity with technology (n=32, 27.12%) identified as the most common concerns. More than half (n=209, 51.60%) the participants had positive feelings about the increased forced technology use,

while almost all (n=392, 96.79%) participants stated that they would continue using technology after the lockdown. When asked about information regarding COVID-19, 282 (69.63%) felt that they had enough information and knowledge, with multimedia (n=215, 53.09%), mobile phone content (n=99, 24.44%), and health organizations and professionals (n=91, 22.47%) as their main source of COVID-19 information. Two-thirds (n=275, 67.90%) of participants stated that they had used their mobile phones for health information before the COVID-19 outbreak, with web searches (n=109, 26.91%), social media posts (n=58, 14.32%), government and institutional websites (n=52, 12.84%), and mobile apps (n=58, 14.32%) serving as their main sources of health information (Table 2).

Table 2. Technology use.

Technology questions	Participants (N=405), n (%) ^a
Has the lockdown forced you to use more technology?	
Yes	363 (89.63)
No	42 (10.37)
If yes, what do you use technology for?	
Job searching	33 (5.67)
Social media/communication	133 (22.85)
Education	58 (9.97)
Shopping	78 (13.40)
Entertainment	48 (8.25)
Work	140 (24.05)
News and information	70 (12.03)
Banking	16 (2.75)
Religion	6 (1.03)
Were you unsure about using technology/online methods before?	
Yes	127 (31.36)
No	278 (68.64)
If yes, what made you feel uncomfortable?	
Security/privacy issues	46 (38.98)
Unfamiliar with technology	32 (27.12)
Lack of personal connection/accountability	16 (13.56)
Cost of data and devices	10 (8.47)
Reliability issues	14 (11.86)
How do you feel about the increased forced use of technology?	
Positive feelings	209 (51.60)
Neutral/mixed feelings	129 (31.85)
Negative feelings	67 (16.54)
Will you continue to use technology after the lockdown?	
Yes	392 (96.79)
No	13 (3.21)
Do you have enough information/knowledge regarding COVID-19?	
Yes	282 (69.63)
No	123 (30.37)
What is your main source of information on COVID-19?	
Health organizations and professionals	91 (22.47)
Mobile phone content	99 (24.44)
Multimedia (radio, television, newspaper)	215 (53.09)
Have you used your mobile phone for health information before the COVID-19 outbreak?	
Yes	275 (67.90)
No	130 (32.10)
If yes, what was your main source of health information?	
Email	1 (0.25)
Government/institutional websites	52 (12.84)

Technology questions	Participants (N=405), n (%) ^a
Messaging platforms (WhatsApp, SMS)	17 (4.20)
Mobile apps	38 (9.38)
Social media posts	58 (14.32)
Web searches (Google)	109 (26.91)

^aPercentages may not add up to 100.00% due to rounding.

Logistic regression analysis identified relationships between demographics and 4 technology use variables ([Multimedia Appendix 2](#)). When asked if the lockdown had forced participants to use more technology, participants with a tertiary school education were 2.5 times more likely to increase their technology use than those with a primary or secondary school education (aOR 2.580; 95% CI 1.212-5.489, $P=.01$), and full-time employees were also less likely to increase their technology use compared to those casually employed (aOR 0.275; 95% CI 0.078-0.966, $P=.04$).

Regarding the main source of COVID-19 information, multimedia, health organizations and professionals, and mobile phone content all had demographic associations. Tertiary school graduates were less likely to use multimedia as their main source of COVID-19 information compared to those with primary or secondary school education (aOR 0.536; 95% CI 0.319-0.900, $P=.02$). Multimedia was almost 2 times more likely to be the main source of information in respondents aged 29-42 years, when compared to those younger than 29 years (aOR 1.862; 95% CI 1.062-3.378, $P=.04$). Single participants were less likely to use health organizations and professionals (aOR 0.537; 95% CI 0.318-0.906, $P=.02$) as their main source of COVID-19 information. Mobile phone content was also associated with age, with the 57-70-year-old group being least likely (aOR 0.339; 95% CI 0.128-0.896, $P=.03$) to use their mobile as the main source of health information compared to those younger than 29 years of age.

The associations seen among participants who responded that they had enough information or knowledge about COVID-19 included age, being male, being single, and having a tertiary education.

The 57-70-year-old group were approximately 6 times (aOR 5.661; 95% CI 1.894-16.925, $P=.002$) more likely to have adequate information compared to those younger than 29 years of age. Males were almost twice as likely (aOR 1.892; 95% CI 1.094-3.272, $P=.02$) than females to have enough COVID-19 information as were those having a tertiary school education (aOR 1.885; 95% CI 1.111-3.198, $P=.02$) over those with a

secondary education or lower, while single participants were less likely (aOR 0.509; 95% CI 0.297-0.873, $P=.01$) to have adequate information.

The oldest age group and students were the least likely to use their phone for health information prior to the pandemic (older adults: aOR 0.184; 95% CI 0.075-0.449, $P<.001$; students: aOR 0.277; 95% CI 0.103-0.740, $P=.01$).

COVID-19 Knowledge

When asked about COVID-19, 358 (88.40%) participants correctly identified it as a contagious respiratory virus, and 392 (96.79%) correctly stated that it was transmitted through respiratory droplets. Over three-quarters ($n=319$, 78.77%) of participants correctly chose all the ways that the virus could be spread; the rest thought it was only spread by coughing or sneezing ($n=52$, 18.84%), by touching objects that have COVID-19 droplets on them ($n=17$, 4.2%), or through close contact with an infected individual ($n=16$, 3.95%). All of the common COVID-19 symptoms (cough, sore throat, fever, and shortness of breath) were correctly identified by 379 (93.58%) participants; the same percentage correctly identified all encouraged prevention techniques (avoid touching one's face, avoid contact with sick people, and wash hands thoroughly). When asked about handwashing duration, 20 seconds was correctly selected by the majority ($n=340$, 83.95%). For the question on how to stop the spread of COVID-19, 368 (90.86%) correctly chose social distancing, self-isolation, and regular handwashing as their response, and when asked how to stop the chance of spreading the virus, 383 (94.57%) correctly chose coughing and sneezing into their elbow, social distancing and self-isolation, and regular handwashing as their response. Most participants ($n=308$, 76.05%) correctly stated that they would call the emergency hotline or WhatsApp support line if they thought they had COVID-19 symptoms, although 79 (19.51%) incorrectly stated that they would rush to the nearest hospital for testing. Lastly, practicing social distancing, self-isolation, and washing one's hands thoroughly were all correctly identified by 369 (91.11%) participants as the key to prevent the spread of COVID-19 ([Table 3](#)).

Table 3. Structured COVID-19 questionnaire.

COVID-19 questions	Participants (N=405), n (%) ^a
What is the novel coronavirus (COVID-19)?	
It is a bioweapon	11 (2.72)
It is a sexually transmitted infection	4 (0.99)
It is a very contagious respiratory virus	358 (88.40)
It is just another term for the common cold	22 (5.43)
It is transmitted through respiratory droplets	10 (2.47)
What are the transmission routes of COVID-19?	
It is transmitted by eating Chinese food	4 (0.99)
It is transmitted through direct blood contact	6 (1.48)
It is transmitted through respiratory droplets	392 (96.79)
It is transmitted through sexual intercourse	3 (0.74)
How can COVID-19 be spread?	
By touching objects that have COVID-19 respiratory droplets	17 (4.20)
Through close contact with an infected individual	16 (3.95)
Through coughing or sneezing	52 (12.84)
All of the above	319 (78.77)
(Blank)	1 (0.25)
What are the signs and symptoms of COVID-19?	
Cough and sore throat	11 (2.72)
Fever	15 (3.70)
Shortness of breath	12 (2.96)
All of the above	367 (90.62)
The coronavirus can be prevented by	
Avoid touching your face	8 (1.98)
Avoiding contact with sick people	7 (1.73)
Wash your hands thoroughly	11 (2.72)
All of the above	379 (93.58)
Wash your hands with soap or sanitizer for at least	
5 seconds	5 (1.23)
10 seconds	17 (4.20)
20 seconds	340 (83.95)
1 minute	43 (10.62)
To stop the spread of the coronavirus, you should	
Practice social distancing	17 (4.20)
Practice social distancing and wash your hands thoroughly	1 (0.25)
Self-isolate	16 (3.95)
Self-isolate and practice social distancing	1 (0.25)
Wash your hands thoroughly	2 (0.49)
All of the above	368 (90.86)
How can you stop the chances of spreading the coronavirus?	
Cough or sneeze into a tissue or your elbow	4 (0.99)

COVID-19 questions	Participants (N=405), n (%) ^a
Self-isolate and practice social distancing	13 (3.21)
Wash your hands thoroughly	5 (1.23)
All of the above	383 (94.57)
What will you do if you suspect that you have symptoms of COVID-19?	
Call the emergency hotline or WhatsApp support line	308 (76.05)
Go to the pharmacy to get medication	9 (2.22)
Rush to the nearest hospital for testing	79 (19.51)
Stay in close physical contact with friends/family for support	8 (1.98)
(Blank)	1 (0.25)
What is the key to prevent the spread of COVID-19?	
Practice social distancing	10 (2.47)
Self-isolate	18 (4.44)
Wash your hands thoroughly	7 (1.73)
All of the above	369 (91.11)
Total knowledge score	
Inadequate (score ≤5)	19 (4.69)
Moderately adequate (score=6,7)	51 (12.59)
Adequate (score ≥8)	335 (82.72)

^aPercentages may not add up to 100.00% due to rounding.

Overall, the mean knowledge score was 8.8 (SD 1.53). There were only 19 (4.69%) participants with inadequate knowledge, 51 (12.59%) with moderately adequate knowledge, and 335 (82.72%) with adequate knowledge (Table 3).

Logistic regression analysis identified relationships between demographics and 4 COVID-19 knowledge variables. Males were less likely to identify the correct transmission routes of COVID-19 (aOR 0.216; 95% CI 0.063-0.744, $P=.02$) than females, while those with a tertiary education were 4 times more likely to correctly identify the routes (aOR 4.414; 95% CI 1.308-14.900, $P=.02$) than those with only primary or secondary education. Tertiary school graduates were also 2 times more likely to identify how to stop the spread of the virus (aOR 2.215; 95% CI 1.041-4.714, $P=.04$), compared to participants with only primary or secondary education. Single participants were less likely to identify the signs and symptoms of COVID-19 (aOR 0.182; 95% CI 0.052-0.631, $P=.01$) than married participants. The 43-56 years age category was 4 times more likely to identify how COVID-19 can be prevented (aOR 3.987; 95% CI 1.011-15.718, $P=.048$) compared to those under 29 years of age (Multimedia Appendix 3).

Lastly, association analyses conducted separately between demographics and the outcome variables (COVID-19 knowledge scores and technology use) only identified a significant relationship in participants ≥57 years being 2.6 times more likely to obtain a knowledge score of 10 (aOR 2.60; 95% CI 1.1-6.0, $P=.03$) when compared to participants 28 years and under.

Discussion

Principal Findings

This study is the first to describe how South Africans interact with technology and consume health information during the current COVID-19 outbreak. Our findings were in line with a similar study from India [19]. Multimedia was the main source of COVID-19 information for both countries (India: $n=57$, 55.4% vs South Africa: $n=215$, 53.09%), followed by the internet in India ($n=22$, 21.4%) and mobile phone content in South Africa ($n=99$, 24.44%). Despite more people in India stating that they had adequate COVID-19 information (India: $n=98$, 95.1% vs South Africa: $n=282$, 69.63%), the South African mean knowledge score of 8.8 was slightly higher than that of India (8.01). The South African study also showed that the lockdown has forced the majority of participants to increase their technology use and these findings are in line with similar increases in technology use from around the world [5-7,29-31]. Participants with a tertiary school education were more likely to increase their technology use than those with less education, who were less likely to use multimedia as their main source of COVID-19 information. This is in line with a study from sub-Saharan Africa, which showed that the positive effects of mobile phone use is diminished by poor primary education [32]. However, in addition, these findings may be explained by socioeconomic factors associated with more education, as college graduates earn higher wages and are better equipped to cope with economic shocks [33]. Full-time employees were less likely to increase their technology use compared to those who were causally employed, although this may be due to a higher

baseline of technology use for full-time employees due to the growing demands of the knowledge economy [34].

The rise in South African technology use has also been validated by the nation's data usage, which increased by more than one-third over the first few days of the lockdown [35]. This increase in technology use led the government to quickly digitize education through a combination of free electronic readers and zero-rated educational apps and websites. This has allowed schools to move to an online curriculum, which has facilitated the return to studies via home-based schooling for many students, by mid-March 2020 [29]. Similarly, apps and websites are also being used by the National Department of Health to relay COVID-19 information to the public [3,13,14]; however, there are many other online sources for COVID-19 information.

Government or institutional websites [3,4,13,14] publish evidence-based information and fact check their findings; however, more participants stated that their main mobile source of health information was web searches or social media posts. Unfortunately, web searches and social media posts are not regulated, and the sharing of misinformation has created an infodemic surrounding COVID-19 [8,25]. This misinformation includes false news articles, conspiracy theories surrounding the virus creation, ineffective home remedies for treatment, and downplaying of the need for prevention control, such as social distancing and mask use. The propagation of this misinformation can actually present a health risk and may undermine the countermeasures implemented by governments and credible institutions [8,36]. Despite a high overall knowledge score, misinformation may have played a role in this study, as two questions (*How COVID-19 can be spread?* and *What will you do if you suspect that you have symptoms of COVID-19?*) scored below adequate. These questions may identify knowledge gaps where increased outreach is needed to educate the population, especially for the second question, where 79 (19.51%) participants stated that they would rush to the nearest hospital for testing instead of calling the emergency hotline or WhatsApp support line for further instructions. There are a number of documented ways to engage users on mobile platforms, and the government can use them to dispel misinformation by guiding people to accurate information sources. Social media outreach, with dialogue loops, is a particularly effective way to engage with individuals, and this type of social media outreach can be tailored with specific messages that target specific subpopulations [37,38].

Misinformation may have disproportionately affected participants under the age of 29 years, especially when compared to those above 57 years. The older group was less likely to use their mobile as the main source of health information, yet they were 6 times more likely to have enough COVID-19 information, and 2.6 times more likely to obtain a knowledge score of 10. In South Africa, youth under 30 years are almost 20% more likely to use their phone to access the internet than their parents, which would expose the younger age group to more online misinformation than the oldest age group [39]. Single participants were less likely to use health organizations

and professionals as their source of COVID-19 information, and not using a trusted source may have also led to misinformation, as they were less likely to have enough COVID-19 information and correctly identify COVID-19 signs and symptoms. Having enough COVID-19 information may not be a true indicator of knowledge though since males were twice as likely to say they had enough COVID-19 information but were less likely to identify the correct routes of COVID-19 transmission.

This study has also reiterated some known barriers to mobile use in South Africa, such as security and privacy issues, unfamiliarity with technology, and data costs. Due to an increase in data usage, some local networks have temporarily lowered data costs [35], but long-term affordable data plans are required to ensure equitable mobile usage for the duration of this lockdown and in the future [40]. Security and privacy issues have been well documented in South Africa, especially for mHealth platforms [22,41,42]. However, previous studies have shown that personal identification number (PIN)-protected mobile platforms for delivering sensitive health information are feasible and acceptable in South Africa [42,43]. Furthermore, a Japanese study that investigated online consumption suggests that the process of making online purchases for the first time during the lockdown has facilitated people becoming familiar with technology, thus alleviating some perceived barriers [44]. This information provides context to the 392 (96.8%) participants who stated they will continue to use technology after the pandemic. However, follow-up studies must be conducted to quantify this.

Limitations

A selection bias may be present due to the device and data requirements needed to access this survey, which was conducted online via a convenience sample. As this survey was adapted from a pre-existing survey, it was not validated or pilot tested in South Africa before this study. Furthermore, participants were asked to self-report their technology use, and no measurements were taken to validate these statements.

Conclusion

This study has shown that the COVID-19 lockdown has forced many people to increase technology use, and almost all participants will continue to use technology post lockdown. Increased technology use was seen across a variety of fields; however, well-known barriers were cited, including privacy and security concerns, unfamiliarity with technology, and data costs. This population showed high COVID-19 knowledge, but the use of web searches and social media posts, instead of government and institutional websites, provides the potential for health misinformation about COVID-19 to be spread. This was particularly evident in some subdemographic groups, including participants under 29 years, single participants, participants without tertiary education, and males. These groups should be targeted with further education and preventative measures.

Acknowledgments

The authors would like to thank all of the survey participants.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Technology use during the COVID-19 lockdown survey.

[[DOCX File , 22 KB - formative_v5i5e25273_app1.docx](#)]

Multimedia Appendix 2

Logistic regressions of technology use.

[[DOCX File , 20 KB - formative_v5i5e25273_app2.docx](#)]

Multimedia Appendix 3

Logistic regressions of COVID-19 knowledge.

[[DOCX File , 19 KB - formative_v5i5e25273_app3.docx](#)]

References

1. WHO Director-General's opening remarks at the media briefing on COVID-19 - 11 March 2020. World Health Organization. 2020 Mar 11. URL: <https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020> [accessed 2021-05-12]
2. President Cyril Ramaphosa: Escalation of measures to combat Coronavirus COVID-19 pandemic. South African Government. 2020 Mar 23. URL: <https://www.gov.za/speeches/president-cyril-ramaphosa-escalation-measures-combat-coronavirus-covid-19-pandemic-23-mar> [accessed 2021-05-12]
3. COVID-19. National Department of Health, South Africa. 2020. URL: <http://health.gov.za/covid19/index.html>
4. Coronavirus disease (COVID-19) Situation Report – 209. World Health Organization. 2020 Aug 17. URL: https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200816-covid-19-sitrep-209.pdf?sfvrsn=5dde1ca2_2 [accessed 2021-05-12]
5. Depoux A, Martin S, Karafillakis E, Preet R, Wilder-Smith A, Larson H. The pandemic of social media panic travels faster than the COVID-19 outbreak. *J Travel Med* 2020 May 18;27(3) [FREE Full text] [doi: [10.1093/jtm/taaa031](https://doi.org/10.1093/jtm/taaa031)] [Medline: [32125413](https://pubmed.ncbi.nlm.nih.gov/32125413/)]
6. Gao J, Dai J, Zheng P, Chen H. Mental health problems and social media exposure during COVID-19 outbreak. *SSRN Journal*. Preprint posted online Feb 20, 2020 [FREE Full text] [doi: [10.2139/ssrn.3541120](https://doi.org/10.2139/ssrn.3541120)]
7. Koetsier J. Coronavirus boosted mobile app spend in Q1 to the highest in history. *Forbes*. 2020 Apr 2. URL: <https://www.forbes.com/sites/johnkoetsier/2020/04/02/coronavirus-impact-houseparty-installs-up-735x-app-use-up-20-app-store-spend-highest-in-history/?sh=31ab858cae58> [accessed 2020-04-16]
8. Cinelli M, Quattrocioni W, Galeazzi A, Valensise CM, Brugnoli E, Schmidt AL, et al. The COVID-19 social media infodemic. *Sci Rep* 2020 Oct 06;10(1):16598 [FREE Full text] [doi: [10.1038/s41598-020-73510-5](https://doi.org/10.1038/s41598-020-73510-5)] [Medline: [33024152](https://pubmed.ncbi.nlm.nih.gov/33024152/)]
9. Cho H, Ippolito D, Yu Y. Contact Tracing Mobile Apps for COVID-19: Privacy Considerations and Related Trade-offs. *arXiv*. Preprint posted online Mar 25, 2020 [FREE Full text]
10. Hollander JE, Carr BG. Virtually Perfect? Telemedicine for Covid-19. *N Engl J Med* 2020 Apr 30;382(18):1679-1681. [doi: [10.1056/NEJMp2003539](https://doi.org/10.1056/NEJMp2003539)] [Medline: [32160451](https://pubmed.ncbi.nlm.nih.gov/32160451/)]
11. The World Health Organization launches WHO Health Alert on WhatsApp. *WhatsApp*. URL: <https://www.whatsapp.com/coronavirus/who> [accessed 2020-04-05]
12. Wong JEL, Leo YS, Tan CC. COVID-19 in Singapore-Current Experience: Critical Global Issues That Require Attention and Action. *JAMA* 2020 Apr 07;323(13):1243-1244. [doi: [10.1001/jama.2020.2467](https://doi.org/10.1001/jama.2020.2467)] [Medline: [32077901](https://pubmed.ncbi.nlm.nih.gov/32077901/)]
13. National Department of Health, South Africa. 2020. URL: <https://sacoronavirus.co.za> [accessed 2020-04-07]
14. COVID-19 / Novel Coronavirus. South African Government. 2020. URL: <https://www.gov.za/Coronavirus> [accessed 2020-04-07]
15. Jahanbin K, Rahmanian V. Using Twitter and web news mining to predict COVID-19 outbreak. *Asian Pac J Trop Med* 2020;13(8):378-380 [FREE Full text] [doi: [10.4103/1995-7645.279651](https://doi.org/10.4103/1995-7645.279651)]
16. Li C, Chen L, Chen X, Zhang M, Pang CP, Chen H. Retrospective analysis of the possibility of predicting the COVID-19 outbreak from Internet searches and social media data, China, 2020. *Euro Surveill* 2020 Mar;25(10):2000199 [FREE Full text] [doi: [10.2807/1560-7917.ES.2020.25.10.2000199](https://doi.org/10.2807/1560-7917.ES.2020.25.10.2000199)] [Medline: [32183935](https://pubmed.ncbi.nlm.nih.gov/32183935/)]
17. Niler E. An AI Epidemiologist Sent the First Warnings of the Wuhan Virus. *Wired*. 2020 Jan 25. URL: <https://www.wired.com/story/ai-epidemiologist-wuhan-public-health-warnings/> [accessed 2020-04-06]

18. Nayar KR, Sadasivan L, Shaffi M, Vijayan B, P Rao A. Social Media Messages Related to COVID-19: A Content Analysis. SSRN Journal. Preprint posted online Mar 25, 2020 [FREE Full text] [doi: [10.2139/ssrn.3560666](https://doi.org/10.2139/ssrn.3560666)]
19. Ranjan R, Ranjan G. Knowledge Regarding Prevention of Novel Coronavirus (COVID-19): An Electronic Cross-Sectional Survey among Selected Rural Community. International Journal of Trend in Scientific Research and Development 2020;4(3):422-426 [FREE Full text]
20. Bloomfield GS, Vedanthan R, Vasudevan L, Kithai A, Were M, Velazquez EJ. Mobile health for non-communicable diseases in Sub-Saharan Africa: a systematic review of the literature and strategic framework for research. Global Health 2014 Jun 13;10(1):49 [FREE Full text] [doi: [10.1186/1744-8603-10-49](https://doi.org/10.1186/1744-8603-10-49)] [Medline: [24927745](https://pubmed.ncbi.nlm.nih.gov/24927745/)]
21. The Mobile Economy: Sub-Saharan Africa 2017. GSMA Intelligence. 2017. URL: www.gsmaintelligence.com/research/?file=7bf3592e6d750144e58d9dcfac6adfab&download [accessed 2021-05-12]
22. Vaportzis E, Clausen MG, Gow AJ. Older Adults Perceptions of Technology and Barriers to Interacting with Tablet Computers: A Focus Group Study. Front Psychol 2017 Oct 04;8:1687 [FREE Full text] [doi: [10.3389/fpsyg.2017.01687](https://doi.org/10.3389/fpsyg.2017.01687)] [Medline: [29071004](https://pubmed.ncbi.nlm.nih.gov/29071004/)]
23. Amplify voices in crisis situations. Upinion. URL: <https://upinion.com> [accessed 2021-04-08]
24. Geldsetzer P. Knowledge and Perceptions of COVID-19 Among the General Public in the United States and the United Kingdom: A Cross-sectional Online Survey. Ann Intern Med 2020 Jul 21;173(2):157-160 [FREE Full text] [doi: [10.7326/M20-0912](https://doi.org/10.7326/M20-0912)] [Medline: [32196071](https://pubmed.ncbi.nlm.nih.gov/32196071/)]
25. Pennycook G, McPhetres J, Zhang Y, Rand D. Fighting COVID-19 Misinformation on Social Media: Experimental Evidence for a Scalable Accuracy-Nudge Intervention. PsyArXiv Preprints. Preprint posted online March 17, 2020. [doi: [10.31234/osf.io/uhbk9](https://doi.org/10.31234/osf.io/uhbk9)]
26. Lunt H, Connor S, Skinner H, Brogden G. Electronic informed consent: the need to redesign the consent process for the digital age. Intern Med J 2019 Jul 11;49(7):923-929. [doi: [10.1111/imj.14339](https://doi.org/10.1111/imj.14339)] [Medline: [31295775](https://pubmed.ncbi.nlm.nih.gov/31295775/)]
27. Rothwell E, Wong B, Rose NC, Anderson R, Fedor B, Stark LA, et al. A randomized controlled trial of an electronic informed consent process. J Empir Res Hum Res Ethics 2014 Dec 02;9(5):1-7 [FREE Full text] [doi: [10.1177/1556264614552627](https://doi.org/10.1177/1556264614552627)] [Medline: [25747685](https://pubmed.ncbi.nlm.nih.gov/25747685/)]
28. Majam M, Venter WDF. Taxi rank baseline Sexual Health HIV Survey (HSTAR006) - Protocol. Johannesburg, South Africa: Wits RHI, University of Witwatersrand; Jan 8, 2018.
29. Mhlanga D, Moloi T. COVID-19 and the Digital Transformation of Education: What Are We Learning on 4IR in South Africa? Education Sciences 2020 Jul 09;10(7):180. [doi: [10.3390/educsci10070180](https://doi.org/10.3390/educsci10070180)]
30. Semuels A. Many Companies Won't Survive the Pandemic. Amazon Will Emerge Stronger Than Ever. The Times. 2020 Jul 28. URL: <https://time.com/5870826/amazon-coronavirus-jeff-bezos-congress/> [accessed 2020-10-07]
31. Koeze EN. The virus changed the way we internet. New York Times. 2020 Apr 7. URL: <https://www.nytimes.com/interactive/2020/04/07/technology/coronavirus-internet-use.html> [accessed 2020-10-07]
32. Asongu SA, Odhiambo NM. Basic formal education quality, information technology, and inclusive human development in sub-Saharan Africa. Sustainable Development 2018 Nov 28;27(3):419-428. [doi: [10.1002/sd.1914](https://doi.org/10.1002/sd.1914)]
33. World Bank Education Overview : Higher Education (English). The World Bank. 2018. URL: <http://documents.worldbank.org/curated/en/610121541079963484/World-Bank-Education-Overview-Higher-Education> [accessed 2021-05-12]
34. Baum S, Ma J, Payea K. Education Pays, 2013: The Benefits of Higher Education for Individuals and Society. Trends in Higher Education Series. College Board 2013:1-48 [FREE Full text]
35. Chen D. Coronavirus: Overcoming the Educational Digital Divide in South Africa. Tech Financials. 2020 Apr 2. URL: <https://www.techfinancials.co.za/2020/04/02/coronavirus-overcoming-the-educational-digital-divide-in-south-africa/> [accessed 2021-10-08]
36. Allington D, Duffy B, Wessely S, Dhavan N, Rubin J. Health-protective behaviour, social media usage and conspiracy belief during the COVID-19 public health emergency. Psychol Med 2020 Jun 09;1-7. [doi: [10.1017/s003329172000224x](https://doi.org/10.1017/s003329172000224x)]
37. Poushter J. Smartphone Ownership and Internet Usage Continues to Climb in Emerging Economies. Pew Research Center. 2016 Feb 22. URL: <https://www.pewresearch.org/global/2016/02/22/smartphone-ownership-and-internet-usage-continues-to-climb-in-emerging-economies/> [accessed 2018-10-22]
38. Chen Q, Min C, Zhang W, Wang G, Ma X, Evans R. Unpacking the black box: How to promote citizen engagement through government social media during the COVID-19 crisis. Comput Human Behav 2020 Sep;110:106380 [FREE Full text] [doi: [10.1016/j.chb.2020.106380](https://doi.org/10.1016/j.chb.2020.106380)] [Medline: [32292239](https://pubmed.ncbi.nlm.nih.gov/32292239/)]
39. Gough A, Hunter RF, Ajao O, Jurek A, McKeown G, Hong J, et al. Tweet for Behavior Change: Using Social Media for the Dissemination of Public Health Messages. JMIR Public Health Surveill 2017 Mar 23;3(1):e14 [FREE Full text] [doi: [10.2196/publichealth.6313](https://doi.org/10.2196/publichealth.6313)] [Medline: [28336503](https://pubmed.ncbi.nlm.nih.gov/28336503/)]
40. Fischer AE, Sebidi J, Barron P, Lalla-Edward ST. The MomConnect Nurses and Midwives Support Platform (NurseConnect): A Qualitative Process Evaluation. JMIR Mhealth Uhealth 2019 Feb 13;7(2):e11644 [FREE Full text] [doi: [10.2196/11644](https://doi.org/10.2196/11644)] [Medline: [30758298](https://pubmed.ncbi.nlm.nih.gov/30758298/)]
41. Venter WDF, Fischer A, Lalla-Edward ST, Coleman J, Lau Chan V, Shubber Z, et al. Improving Linkage to and Retention in Care in Newly Diagnosed HIV-Positive Patients Using Smartphones in South Africa: Randomized Controlled Trial. JMIR Mhealth Uhealth 2019 Apr 02;7(4):e12652 [FREE Full text] [doi: [10.2196/12652](https://doi.org/10.2196/12652)] [Medline: [30938681](https://pubmed.ncbi.nlm.nih.gov/30938681/)]

42. DiAndreth L, Krishnan N, Elf JL, Cox S, Tilchin C, Nthulana M, et al. Formative research for an mHealth program to improve the HIV care continuum in South Africa. *AIDS Care* 2020 Jun 12;32(6):744-748. [doi: [10.1080/09540121.2019.1640850](https://doi.org/10.1080/09540121.2019.1640850)] [Medline: [31298566](https://pubmed.ncbi.nlm.nih.gov/31298566/)]
43. Maraba N, Hoffmann CJ, Chihota VN, Chang LW, Ismail N, Candy S, et al. Using mHealth to improve tuberculosis case identification and treatment initiation in South Africa: Results from a pilot study. *PLoS One* 2018 Jul 3;13(7):e0199687 [FREE Full text] [doi: [10.1371/journal.pone.0199687](https://doi.org/10.1371/journal.pone.0199687)] [Medline: [29969486](https://pubmed.ncbi.nlm.nih.gov/29969486/)]
44. Watanabe T, Yuki O. Online Consumption During the COVID-19 Crisis: Evidence from Japan. Working Paper Series No.023. Tokyo, Japan: University of Tokyo; 2020. URL: <https://www.centralbank.e.u-tokyo.ac.jp/wp-content/uploads/2020/06/cb-wp023.pdf> [accessed 2021-05-12]

Abbreviations

aOR: adjusted odds ratio

IEC: International Electrotechnical Commission

ISO: International Organization for Standardization

mHealth: mobile health

Edited by G Eysenbach; submitted 26.10.20; peer-reviewed by B Maharaj, XF Zhong; comments to author 23.11.20; revised version received 04.12.20; accepted 16.04.21; published 17.05.21.

Please cite as:

Fischer AE, Van Tonder T, Gumede SB, Lalla-Edward ST

Changes in Perceptions and Use of Mobile Technology and Health Communication in South Africa During the COVID-19 Lockdown: Cross-sectional Survey Study

JMIR Form Res 2021;5(5):e25273

URL: <https://formative.jmir.org/2021/5/e25273>

doi: [10.2196/25273](https://doi.org/10.2196/25273)

PMID: [33956640](https://pubmed.ncbi.nlm.nih.gov/33956640/)

©Alex Emilio Fischer, Tanya Van Tonder, Siphamandla B Gumede, Samanta T Lalla-Edward. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 17.05.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

The Impact of the COVID-19 Pandemic on Physical and Mental Health in China and Spain: Cross-sectional Study

Cuiyan Wang^{1*}, PhD; María Inmaculada López-Núñez^{2*}, PhD; Riyu Pan¹, PhD; Xiaoyang Wan¹, PhD; Yilin Tan¹, PhD; Linkang Xu¹, PhD; Faith Choo³, MD; Roger Ho⁴, MD; Cyrus Ho³, MD; Marta E Aparicio García², PhD

¹Institute of Cognitive Neuroscience, Faculty of Education, Huaibei Normal University, Huaibei, China

²Department of Social Work and Differential Psychology, Faculty of Psychology, Complutense University of Madrid, Madrid, Spain

³Department of Psychological Medicine, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore

⁴Department of Psychological Medicine, Institute of Health Innovation and Technology, National University of Singapore, Singapore, Singapore

*these authors contributed equally

Corresponding Author:

Roger Ho, MD

Department of Psychological Medicine

Institute of Health Innovation and Technology

National University of Singapore

Level 9, NUHS Tower Block

1E Kent Ridge Road

Singapore, 119228

Singapore

Phone: 65 97321097

Email: pcmrhcm@nus.edu.sg

Abstract

Background: Differences in physical and mental health impact across continents during the COVID-19 pandemic are unknown.

Objective: This study compared the levels of impact of COVID-19 on mental health among people from Spain and China and correlated mental health parameters with variables relating to symptoms similar to COVID-19, COVID-19 knowledge, and precautionary measures.

Methods: We collected information on demographic data, physical symptoms, contact history with persons with a confirmed COVID-19 diagnosis, COVID-19 knowledge, and precautionary measures. Participants completed the Impact of Event Scale-Revised (IES-R) and the Depression, Anxiety and Stress Scale-21 Items (DASS-21). To analyze the differences in the mental health parameters, the mean scores between Chinese and Spanish respondents were compared using the independent samples *t* test. The differences in categorical variables between the two samples were analyzed by the chi-square test. Linear regression was used to calculate the univariate associations between the independent variables and mental health parameters for both groups separately, with adjustments made for age, gender, and education.

Results: A total of 1528 participants (Spain: *n*=687; China: *n*=841) were recruited. The mean age of the Chinese respondents was 24.73 years (SD 7.60; range 18-65 years), and the mean age of the Spanish respondents was 43.06 years (SD 11.95; range 18-76 years). Spanish participants reported significantly more symptoms similar to COVID-19 infection (eg, fever, sore throat, and breathing difficulties), contact history with COVID-19, higher perceived risk of contracting COVID-19, frequent use of medical services, and less confidence in medical services compared with their Chinese counterparts (*P*<.001). Spanish participants reported significantly higher DASS-21 stress and depression scores, while Chinese participants reported significantly higher IES-R scores (*P*<.001). Chinese participants encountered more discrimination from other countries (*P*<.001). Significantly more Chinese participants reported using face masks than Spanish ones (*P*<.001). More exposure to health information was associated with adverse mental health in Spain (depression: *P*=.02; anxiety: *P*=.02; stress: *P*=.001).

Conclusions: Our study found that Spanish respondents reported higher levels of stress and depression as well as more symptoms and use of medical services. In preparation for the next pandemic, Spain needs to establish a prompt policy to implement rapid response and enhance medical services to safeguard physical and mental health.

(JMIR Form Res 2021;5(5):e27818) doi:[10.2196/27818](https://doi.org/10.2196/27818)

KEYWORDS

anxiety; China; coronavirus; COVID-19; depression; developing countries; knowledge; masks; pandemic; physical; precaution; psychological impact; Spain; stress

Introduction

Background

The city of Wuhan was ground zero of the COVID-19 outbreak, which spread to all 23 provinces of China. As a result, China became the first epicenter in Asia during the COVID-19 pandemic [1]. China's first confirmed COVID-19 case was reported on November 17, 2019. As of May 5, 2020, there were 82,880 confirmed cases, 4633 deaths, and 77,766 recovered cases in China [2]. The Chinese government imposed lockdown measures to restrict travel to prevent transmission. Since the onset of the COVID-19 outbreak, the Chinese government's response efforts have been swift, and 3 weeks into the epidemic, in an unprecedented move to slow the spread of the virus, a lockdown was imposed in Wuhan on January 23, 2020 [3]. In the Lunar New Year of 2020, the quarantine was extended to additional provinces and cities during the peak travel period. Many Chinese people stayed at home and socially isolated themselves to prevent infection. Due to many new cases in February 2020, the Central Government of China deployed thousands of medical personnel to a rapidly completed hospital specially designed to treat patients with COVID-19 [4]. The Chinese health authority imparted unbiased and clear guidelines on the use of face masks during the early stage of the pandemic [5].

Spain is one of Europe's COVID-19 epicenters [2]. Spain's first patient was detected on January 31, 2020, in La Gomera (Canary Islands). However, it was not until February 24 when the virus spread to the peninsula, with cases in Madrid, Catalonia, and the Valencian Community. Since then, the number of COVID-19 infections has progressively increased. There has been a rapid increase in the number of new cases of COVID-19 in Spain since March 2020 [6]. On March 30, 2020, Spain overtook China in terms of the number of infectees [7]. Although there were major differences in government response between Spain and China, no research has been conducted to compare Chinese and Spanish physical and mental health at the beginning of the pandemic. In the early stage of the COVID-19 outbreak, Spanish authorities allowed travel to other European countries without any restrictions as well as mass celebrations [6]. On March 12, 2020, the Spanish government agreed to implement educational and social distancing measures throughout Spain. Two days later, a state of alarm was declared for 14 days to establish measures to protect citizens' health and safety, contain the progression of the disease, and strengthen the public health system [8]. Among these measures was the mandatory confinement of all citizens. On March 28, the state of alarm was extended until April 12 [9]. On April 9, a third extension was approved until April 25 [10]. On April 18, the government announced a new extension until May 10. On that day, there were a total of 191,726 infections and 20,043 deaths in Spain [11]. More than 10,000 elderly people died in nursing homes, comprising half of Spain's official death toll [12].

This Study

The lockdown measures and economic recession associated with the COVID-19 pandemic has put the general public in China and Spain under more pressure than ever before. There has been an atmosphere of anxiety and depression in Spain and China due to the pandemic [13]. Recent studies have reported on depression, anxiety, and stress levels among Chinese [14-17], Thai [18], and Vietnamese [19] populations during the pandemic in Asia. A recent study found that the impact of anxiety on adopting social distancing did not vary between people in the United Kingdom and Hong Kong [20]. The psychological impact of other preventive measures between people in Asia and Europe require further study. As a result, we proposed a study to conduct an in-depth analysis of physical and mental health in China and Spain during the pandemic. The theoretical framework of this study was based on the chain mediation model, which shows the need for health information, and the perceived impacts of the pandemic were sequential mediators between physical symptoms resembling COVID-19 infection and consequent mental health status [21].

We aimed (1) to compare levels of physical health (eg, COVID-19-related symptoms) and mental health status (eg, depression, anxiety, stress, and psychological impact) between Spanish and Chinese survey respondents during the pandemic; and (2) to correlate psychological impact, depression, anxiety, and stress scores with variables relating to physical symptoms, COVID-19 knowledge and views (eg, knowledge of COVID-19 transmission, confidence in medical services, etc), prevention measures (eg, face mask use, hand hygiene), and information needs (eg, symptoms, prevention methods, the effectiveness of drugs and vaccines) among Spanish and Chinese respondents. The null hypothesis was that there would be no difference in survey scores between Spanish and Chinese participants. The other null hypothesis was that physical symptoms resembling COVID-19, higher confidence in medical services, perceived higher chances of survival, satisfaction with health information, and adoption of precautionary measures had no significant association with lower survey scores in both countries.

Methods

Study Participants

This cross-cultural study compared the mental and physical health status and the psychological impact between Spanish and Chinese people. The recruitment period was from February 28 to March 1, 2020, in China, a month after the Chinese government declared a lockdown of Wuhan. For Spain, data collection was from April 14 to 18, 2020, a month after the Spanish government declared the state of alarm. A snowball sampling strategy was utilized to recruit participants from the general public in Spain and China. Recruitment started with a set of initial respondents who were associated with the Huaibei Normal University of China and the Complutense University of Madrid, who referred other participants by email and social

networks including colleagues, relatives, classmates, and friends. These individuals, in turn, referred other participants across different cities in China and Spain. Inclusion criteria restricted participation to respondents who could provide online written consent, had access to the internet, and were above 18 years of age. Exclusion criteria included respondents who were illiterate or had no access to the internet.

Study Procedure

Potential participants were electronically invited by respondents who had completed the questionnaires. The respondents completed the online survey via SurveyStar in China and Google Forms (Google LLC) in Spain.

Study Outcomes

The variables used in this study have been used previously in research on the severe acute respiratory syndrome (SARS); thus, the content of variables have been validated [22]. This study used a questionnaire with psychometric properties established at the beginning and peak of the COVID-19 epidemic [3,15]. The COVID-19 questionnaire had 43 items and covered the following areas: (1) demographic information, (2) symptoms similar to COVID-19 infection in the past 2 weeks, (3) contact history with COVID-19 in the past 14 days, (4) knowledge about COVID-19, and (5) preventive measures against COVID-19 in the past 2 weeks. Responses depended on the questions (eg, yes or no, Likert scale based on frequency). To ensure high data quality, the structured questionnaire was developed in English in such a way that it included all relevant variables to meet the objective; to ensure consistency in the questionnaire, it was first prepared in English and then translated into Chinese and Spanish by fluent speakers of both languages and then translated back to English. Data were checked for completeness and consistency before processing and analysis.

Demographic data on age, gender, education, household size, marital status, parental status, and residential city in the past 2 weeks were collected. Physical symptoms similar to COVID-19 infection included cough, fever, gastrointestinal symptoms, and other symptoms. Respondents also stated their past medical history, health service use, quarantine by health authority, and recent testing for COVID-19. Knowledge and views related to COVID-19 included confidence in medical services, health information satisfaction, the practice of hand hygiene and face mask use, time spent viewing health information, transmission route, and the likelihood of contracting and surviving COVID-19.

The Chinese and Spanish versions of the Impact of Event Scale-Revised (IES-R) and the Depression, Anxiety and Stress Scale-21 Items (DASS-21) were used to assess mental health. The IES-R has 22 items, and it was previously validated in the European and Asian populations to assess impact when exposed to the COVID-19 pandemic [23-25]. The IES-R measures avoidance, intrusion, and hyperarousal [26]. The IES-R items used the following Likert-scales responses: 0=not at all, 1=a little bit, 2=moderately, 3=quite a bit, and 4=extremely. A total IES-R score of 0 to 23 is considered normal, 24 to 32 is mild, 33 to 36 is moderate, and ≥ 37 is severe [27]. A total IES-R score of >24 is the cut-off score for posttraumatic stress disorder

(PTSD) symptoms [28]. In this study, the Cronbach alphas of IES-R for the Chinese and Spanish surveys were .949 and .948, respectively.

DASS-21 scores were calculated based on a previous Asian study [29]. DASS-21 has 21 items, and it has been used to assess mental health in Chinese [30,31] and Spanish [32] populations. The responses were indicated using the following Likert scale: 0=did not apply to me at all, 1=applied to me to some degree or some of the time, 2=applied to me to a considerable degree or a good part of the time, and 3=applied to me very much or most of the time. In this study, the Cronbach alpha for the Chinese version of DASS-21 was as follows: stress: .888, anxiety: .845, and depression: .878; for the Spanish version, the Cronbach alpha was as follows: stress: .895, anxiety: .876, and depression: .89. DASS-21 and IES-R have been previously used in COVID-19 research [3,17,24,33].

Statistical Analysis

Statistical analysis was performed on SPSS Statistics 21.0 (IBM Corp). Descriptive statistics were reported for demographics, COVID-19 symptoms, health service utilization, knowledge about COVID-19, and preventive measures. To analyze the differences in the mental health parameters, the differences in mean scores between the Chinese and Spanish respondents were compared by the independent samples *t* test. The differences in categorical variables between the two samples were analyzed by the chi-square test. Linear regression was used to calculate the univariate associations between independent variables and mental health parameters for the Chinese and Spanish respondents separately with adjustment for age, gender, and education. The significance level was set at $P<.05$.

Ethics Statement

The institutional review board of the Complutense University of Madrid (Pr_2019_20_027) and the Huaibei Normal University (HBU-IRB-2020-002) approved the studies for Spain and China, respectively. All respondents signed the informed consent electronically. The collected information was anonymous and treated as confidential.

Results

Comparison of Mental Health Status Between Spanish and Chinese Participants

In Spain, 687 respondents completed the questionnaires (completion rate: 94.0%). We excluded 4 incomplete questionnaires in the Chinese sample, which left a total of 841 out of 845 (99.5%) valid questionnaires. Respondents from China came from 159 cities, and respondents from Spain came from all 19 autonomous regions. As a result, the total number of respondents from both countries was 1528. In terms of age, 338 respondents from China were aged 18-21 years, 400 respondents were 22-30 years, 37 respondents were 31-40 years, 51 respondents were 41-49 years, and 15 respondents were ≥ 50 years. In Spain, 27 respondents were aged 18-21 years, 91 respondents were 22-30 years, 158 respondents were 31-40 years, 200 respondents were 41-49 years, and 211 respondents were ≥ 50 years.

Table 1 compares the mental health status of the Spanish and Chinese samples. For the DASS-21 stress subscale (China: mean 7.79, SD 7.90; Spain: mean 14.23, SD 10.04), Spanish respondents reported significantly higher scores than Chinese respondents ($t_{1287.27}=-13.70$, $P<.001$, 95% CI -7.36 to -5.52). For the DASS-21 depression subscale (China: mean 6.28, SD 7.30; Spain: mean 8.61, SD 8.80), Spanish respondents had significantly higher depression scores ($t_{1330.70}=-5.55$, $P<.001$, 95% CI -3.15 to -1.51). For the DASS-21 anxiety subscale (China: mean 6.07, SD 6.87; Spain: mean 6.78, SD 8.23), there was no significant difference between the two countries ($t_{1335.75}=-1.79$, $P=.07$, 95% CI -1.48 to 0.07). For the IES-R scale (China: mean 30.69, SD 16.21; Spain: mean 27.62, SD 18.67), Spanish participants had significantly lower scores than

their Chinese counterparts ($t_{1367.38}=3.39$, $P=.001$, 95% CI $1.30-4.85$). Nevertheless, the mean IES-R scores of respondents from both countries were greater than 24 points (ie, the cut-off score for PTSD symptoms). Table S1 in [Multimedia Appendix 1](#) compares the mental health status of both samples across different age groups. Both Chinese and Spanish respondents showed significant differences in IES-R scores based on age category, with the 18-21 years age group having significantly higher IES-R scores than the 41-49 years and ≥ 50 years age groups (Spain: 18-21 years, $P=.001$; 41-49 years, $P=.02$; China: 18-21 years, $P=.02$; 41-49 years, $P=.49$). For the DASS-21 scores, Spanish respondents showed significant differences, with the 18-21 years age group reporting significantly higher depression ($P=.002$), anxiety ($P=.02$), and stress ($P<.001$) scores than those ≥ 50 years (Table S1, [Multimedia Appendix 1](#)).

Table 1. Scores of the Depression, Anxiety and Stress Scale–21 Items (DASS-21) and Impact of Event Scale-Revised (IES-R) for Spanish and Chinese respondents.

Country	DASS-21			IES-R, score (SD)
	Depression, score (SD)	Anxiety, score (SD)	Stress, score (SD)	
China	6.28 (7.30)	6.07 (6.87)	7.79 (7.90)	30.69 (16.21)
Spain	8.61 (8.80)	6.78 (8.23)	14.23 (10.04)	27.62 (18.67)

Association Between Variables

Demographic Characteristics and Mental Health Status

More than half the Spanish respondents were women ($n=541$, 78.7%), had a family size of 3 to 5 people ($n=368$, 53.6%), were well educated ($n=516$, 75.8% with a bachelor or higher degree), and were married ($n=74$, 10.8%). Similarly, more than half the Chinese respondents were women ($n=631$, 75%), had a family size of 3 to 5 people ($n=676$, 80.4%), were well educated ($n=754$, 89.2% with a bachelor or higher degree), and were married ($n=699$, 83.5%). A significantly higher proportion of Chinese respondents were younger, lived with more than 6 people in the same household, and had a child over 16 years ($P<.001$) (Table S2, [Multimedia Appendix 1](#)).

Among the Spanish respondents, the male gender was a protective factor associated with a lower score for the DASS-21 stress ($P=.004$) and depression ($P=.02$) subscales and IES-R ($P=.001$) while age <50 years and having a child over 16 years of age were risk factors associated with a higher score for IES-R ($P=.02$) and some of the DASS-21 subscales (stress: $P<.001$; anxiety: $P=.03$) (Table S1, [Multimedia Appendix 2](#)). Among the Chinese respondents, the male gender was associated with a lower IES-R score ($P=.01$) but a higher DASS-21 depression score ($P=.002$). Chinese people who lived in a household with 3 to 5 people ($P=.04$) and more than 6 people ($P=.03$) were associated with a higher score of IES-R compared to those who lived alone. This association was not observed in the Spanish sample.

Physical and Mental Health Parameters

Previous studies have established the association between physical symptoms and psychological outcomes during the COVID-19 outbreak [24]. For symptoms resembling COVID-19, significantly more Spanish individuals reported fever ($P<.001$),

chills ($P<.001$), headache ($P<.001$), myalgia ($P<.001$), cough ($P<.001$), breathing difficulty ($P<.001$), dizziness ($P<.001$), coryza ($P<.001$), sore throat ($P<.001$), nausea and vomiting ($P<.001$), recent consultation with a doctor ($P<.001$), recent hospitalization and quarantine (IES-R: $P=.03$; anxiety: $P=.01$), recent COVID-19 testing ($P<.001$), chronic illness ($P<.001$), direct ($P<.001$) and indirect ($P<.001$) contact with someone who was COVID-19 positive, and direct contact with contaminated materials ($P<.001$) compared to the Chinese sample (Table S3, [Multimedia Appendix 1](#)). Nevertheless, significantly more Spanish respondents rated good physical health compared with the Chinese sample ($P<.001$).

Linear regression showed that headache, myalgia, and dizziness were associated with higher DASS-21 stress, anxiety, and depression and IES-R scores in both countries (Spain—headache: IES-R, $P=.002$; stress, $P=.001$; anxiety, $P<.001$; depression, $P=.001$; myalgia: IES-R, $P<.001$; stress, $P<.001$; anxiety, $P=.002$; depression, $P=.02$; dizziness: IES-R, $P=.03$; stress, $P=.05$; anxiety, $P=.003$; depression, $P=.02$; China—headache: IES-R, $P=.02$; stress, $P<.001$; anxiety, $P<.001$; depression, $P=.001$; myalgia: IES-R, $P=.007$; stress, $P<.001$; anxiety, $P<.001$; depression, $P<.001$; dizziness: IES-R, $P<.001$; stress, $P<.001$; anxiety, $P<.001$; depression, $P=.007$) (Table S2, [Multimedia Appendix 2](#)). In contrast, myalgia, sore throat, and gastrointestinal symptoms were associated with higher DASS-21 stress, anxiety, and depression scores in both countries (Spain—myalgia: stress, $P<.001$; anxiety, $P=.002$; depression, $P=.02$; sore throat: stress, $P=.01$; anxiety, $P<.001$; depression, $P=.03$; gastrointestinal symptoms: stress, $P=.004$; anxiety, $P\leq.001$; depression, $P=.007$; China—myalgia: stress, $P<.001$; anxiety, $P<.001$; depression, $P<.001$; sore throat: stress, $P<.001$; anxiety, $P<.001$; depression, $P<.001$; gastrointestinal symptoms: stress, $P<.001$; anxiety, $P=.002$; depression, $P<.001$). Recent consultation with a doctor and quarantine were associated with greater psychological impact and anxiety

(consultation with a doctor; IES-R, $P=.02$; stress, $P=.002$; quarantine: IES-R, $P=.03$; anxiety, $P=.01$) in Spanish respondents only.

Knowledge and Views of COVID-19 and Mental Health Parameters

Spanish and Chinese respondents held significantly different views regarding knowledge and views related to COVID-19 (Table S4, [Multimedia Appendix 1](#)). For the routes of transmission, there were significantly more Spanish respondents who agreed that droplets transmitted the coronavirus ($P=.01$), as well as contact via contaminated objects ($P<.001$), but significantly more Chinese individuals agreed with airborne transmission ($P<.001$). In terms of detection and risk of contracting COVID-19, there were significantly more Spanish people who were not confident about the competency of doctors to diagnose COVID-19 ($P<.001$) and perceived a greater risk of contracting COVID-19 ($P<.001$) as well as greater chances of survival after COVID-19 infection ($P<.001$). More Spanish respondents were concerned about their relatives contracting COVID-19 ($P<.001$). There were significantly more Spanish participants who were unsatisfied with the amount of health information they received ($P<.001$). There were significantly more Chinese respondents who felt discriminated against by other countries ($P<.001$).

Spanish and Chinese respondents showed different results in the linear regression analysis (Table S3, [Multimedia Appendix 2](#)). In terms of the level of confidence in doctors to diagnose COVID-19, Chinese individuals who were somewhat confident and very confident were associated with lower scores on the DASS-21 depression subscale (very confident: $P=.002$; somewhat confident: $P=.01$). Chinese individuals who perceived a lower chance of contracting COVID-19 were associated with a lower DASS-21 depression score ($P=.04$), and those who perceived a very high likelihood of survival were associated with lower DASS-21 depression and IES-R scores (depression: $P=.02$; IES-R: $P=.04$). Chinese individuals who were satisfied with the health information they received were significantly associated with lower DASS-21 depression, anxiety, and stress scores (depression: very confident, $P<.001$; somewhat confident, $P=.002$; anxiety: very confident, $P=.003$; somewhat confident, $P=.01$; stress: very confident, $P<.001$; somewhat confident, $P=.001$). The above significant association was not found among Spanish respondents (confidence in doctors—depression: very confident, $P=.21$, somewhat confident, $P=.44$; chance of contracting COVID-19: $P=.14$; likelihood of survival after contracting COVID-19: depression, $P=.32$; IES-R, $P=.96$; satisfaction with health information—depression: very satisfied, $P=.67$; somewhat satisfied, $P=.42$; anxiety: very satisfied, $P=.92$; somewhat satisfied, $P=.66$; stress: very satisfied, $P=.65$; somewhat satisfied, $P=.45$). In contrast, Spanish respondents who spent less than 1 hour per day monitoring COVID-19 information were significantly associated with lower DASS-21 depression, anxiety, and stress scores and IES-R scores (depression: $P=.01$; anxiety: $P=.001$; stress: $P=.02$; IES-R: $P=.02$). Spanish individuals who felt discriminated were associated with higher DASS-21 depression, anxiety, and stress scores and IES-R scores (depression: $P=.01$; anxiety: $P=.003$; stress: $P<.001$; IES-R: $P<.001$).

Precautionary Measures and Mental Health Parameters

Spanish and Chinese respondents demonstrated significantly different precautionary measures (Table S5, [Multimedia Appendix 1](#)). There were significantly more Spanish individuals who reported covering their mouths when coughing and sneezing ($P<.001$), avoiding the sharing of utensils ($P=.01$), and practicing hand hygiene ($P<.001$). More Spanish respondents felt that the COVID-19 pandemic caused unnecessary worry ($P<.001$). In contrast, there were significantly more Chinese respondents who washed their hands immediately after coughing or sneezing ($P<.001$), agreed to wear a face mask ($P<.001$), and practiced hand hygiene after touching contaminated objects ($P<.001$).

The linear regression analysis found that the avoidance of sharing utensils, covering one's mouth when coughing and sneezing, practicing hand hygiene, and using a face mask were significantly associated with lower DASS-21 and IES-R scores (avoidance of sharing utensils: depression, $P=.01$; stress, $P=.02$; IES-R, $P=.01$; covering mouth when sneezing and coughing: depression, $P<.001$; anxiety, $P=.01$; stress, $P=.01$; hand hygiene—wash hands with soap: depression, $P<.001$; anxiety, $P=.002$; stress, $P=.001$; wash hands immediately after coughing and sneezing: depression, $P<.001$; anxiety, $P<.001$; stress, $P<.001$; IES-R, $P=.01$; wash hands immediately after touching contaminated objects: depression, $P=.001$; anxiety, $P=.01$; stress, $P=.002$; wearing a face mask: depression, $P<.001$; anxiety, $P=.02$; stress, $P<.001$) among Chinese respondents (Table S4, [Multimedia Appendix 2](#)). Among Spanish respondents, covering one's mouth when coughing and sneezing and hand hygiene measures were associated with lower DASS-21 anxiety and stress subscale scores, respectively (covering mouth when coughing and sneezing: anxiety, $P=.02$; stress, $P=.02$; hand hygiene—wash hand with soap: anxiety, $P=.04$; stress, $P=.01$). Wearing a face mask was also associated with lower IES-R and DASS-21 stress scores in Spanish respondents (IES-R: $P=.01$; stress: $P=.047$). Spanish respondents who felt the COVID-19 pandemic caused unnecessary worry were associated with higher DASS-21 depression, anxiety, and stress scores (depression: $P=.01$; anxiety: $P=.03$; stress: $P=.04$). In contrast, Chinese respondents who felt that the COVID-19 pandemic had caused too much unnecessary worry were associated with lower IES-R scores ($P=.01$).

Health Information About COVID-19 and Mental Health Parameters

Spanish and Chinese participants were significantly different in terms of information needs pertaining to COVID-19. There were significantly more Chinese participants who needed regular updates; more personalized information; and data on symptoms related to COVID-19, prevention methods, management and treatment methods, the effectiveness of drugs and vaccines, the number of infected by geographical locations, travel advice, and transmission methods compared to Spanish participants ($P<.001$) (Table S6, [Multimedia Appendix 1](#)). In contrast, there were significantly more Spanish respondents who needed information on other countries' strategies and responses than Chinese respondents ($P<.001$).

Health information on prevention methods, as well as more personalized information on transmission methods, information related to COVID-19 in general, and the effectiveness of drugs and vaccines, were associated with higher IES-R scores or higher scores for one of the DASS-21 subscales among Spanish respondents only (prevention methods: IES-R, $P < .001$; stress, $P < .001$; anxiety, $P < .001$; mode of transmission: IES-R, $P < .001$; stress, $P < .001$; anxiety, $P < .001$; depression, $P = .01$; personalized information: IES-R, $P < .001$; stress, $P = .003$; anxiety, $P = .001$; depression, $P = .004$; effectiveness of medication: IES-R, $P = .003$) (Table S5, [Multimedia Appendix 2](#)). Information on management methods and transmission methods was associated with higher IES-R scores in both Spanish and Chinese respondents (Spain: management methods, $P < .001$; transmission methods, $P < .001$; China: management methods, $P = .04$; transmission methods, $P = .01$). Travel advice and information on other countries' responses were associated with lower depression scores only in China (travel advice: $P = .02$; information on other countries' responses: $P = .03$).

Discussion

Principal Findings

This is the first study to compare the physical and mental health of citizens from two COVID-19 epicenters in Europe and Asia. Spain and China faced different problems. In Spain, the number of COVID-19 cases and deaths was 5494 (98 times that of China) and 558 (186 times that of China) per 1 million population, respectively, in May 2020 [2]. The research findings rejected the null hypothesis. A significantly higher proportion of Spanish respondents reported symptoms similar to COVID-19 infection, contact history with COVID-19, and higher perceived risk of contracting COVID-19. More Spanish respondents also utilized medical services recently, but they reported less confidence in their medical services. Spanish respondents were more dissatisfied with health information; the time and types of health information were associated with adverse mental health. Spanish respondents reported significantly higher levels of depression and stress during the pandemic than Chinese participants, especially younger Spanish people (18-40 years).

Protective factors against adverse mental health in Spain included the male gender, living with an older child, and spending less time monitoring COVID-19 information. In China, the number of COVID-19 cases and deaths were 58 and 3 per 1 million population, respectively, in May 2020 [2]. Although China has a relatively lower number of COVID-19 cases and death per capita, significantly more Chinese respondents reported a higher psychological impact and discrimination by other countries. Protective factors against adverse mental health in China include smaller household sizes, higher confidence levels in their doctors, wearing face masks, and obtaining information on travel advice and other countries' responses during the pandemic. For both countries, risk factors for adverse mental health include the presence of symptoms similar to COVID-19 infection (eg, headache, myalgia, sore throat, gastrointestinal symptoms). Protective factors for both countries included certain precautionary behaviors (eg, covering one's

mouth when coughing or sneezing, certain hand hygiene measures).

This study has provided an understanding of the differences between China and Spain, which will allow us to better prepare for the next pandemic. The higher levels of adverse mental health and lower confidence in Spanish medical services were due to the rising number of infected health care workers in Spain and the lack of effective coordination between different hospitals [34]. In contrast, the Chinese government rapidly deployed medical personnel and treated patients with COVID-19 at rapidly built hospitals [4]. This prompt action had restored public confidence in the health care system. Spain and other European countries shared four common characteristics in the early stage of the pandemic: (1) the lack of personal protection equipment (PPE) for health care workers, (2) the delay in response strategy, (3) an overstretched health care system with a shortage of hospital beds, and (4) the failure to protect vulnerable nursing home residents from COVID-19 [35]. Following the COVID-19 pandemic, European countries need to rectify the insufficiencies in the number of available health care professionals, hospital beds, and medical equipment (eg, PPE, mechanical ventilators) caused by long-term underinvestment in health services following the 2008 financial crisis [36].

Many research studies supported the benefits of face masks in blocking virus transmission in aerosols but were opposed by erroneous judgment [37]. Wearing a face mask can impede the spread of the virus from asymptomatic patients with COVID-19 [38]. Nevertheless, this study found that Spanish respondents were significantly less likely to wear face masks than Chinese ones. The Chinese government provided health information on COVID-19 transmission via contact, droplets, and through the air; personal preventive measures (eg, wearing a face mask, hand hygiene, other personal precaution); and public health measures (eg, good ventilation, social distancing, COVID-19 testing for the general population) [16]. Furthermore, the Chinese government emphasized the benefits of mask wearing due to airborne transmission and made such practice mandatory in February 2020 to reduce the coronavirus spread [39]. Besides potential benefits on physical health, wearing face masks could offer psychological benefits associated with lower IES-R and DASS-21 stress scores among Chinese and Spanish respondents. A recent epidemiological study compared the incidence of COVID-19 per 1 million in Hong Kong with a high prevalence of community-wide masking with non-mask-wearing countries. This study found that the incidence of COVID-19 in Hong Kong was 129.0 per 1 million population, which was much lower than the incidence rate of Spain (2983.2 per million population, 23 times the incidence rate of Hong Kong) [40]. During the beginning stage of the COVID-19 pandemic, Europeans held ambivalent views toward face masks due to cultural reasons. It was generally difficult for Europeans to accept the need for mask use by healthy people since mask wearing suggests vulnerability to sickness and concealment of identity. Due to the collective culture of China, Chinese individuals are more aware of the importance of wearing masks and the responsibility associated with the need to protect their health and the health of others—this sentiment

is more pronounced in China than it is in Europe. Our findings suggest that there is a need for health education with scientific information from Spanish health authorities on the use and benefits of face masks during the pandemic and implement measures to reduce social stigma.

Chinese participants reported a significantly higher level of the psychological impact associated with the COVID-19 pandemic. China, the epicenter in Asia in the early stage of the pandemic, faced other unique problems not shared by its European counterparts. A significantly higher number of Chinese people felt they were discriminated against by other countries due to COVID-19 through negative media comments. The editor-in-chief of *The Lancet*, Richard Horton, expressed concern for the discrimination faced by China, saying that while it is important to understand the origin and interspecies transmission of the coronavirus, it is both unhelpful and unscientific to attribute China as the origin of the COVID-19 pandemic and seek a patient zero, as such efforts could be highly stigmatizing and discriminatory [41]. Global cooperation involving an exchange of expertise, adopting effective prevention strategies, and sharing resources and technologies across different countries to form a united front to tackle the COVID-19 pandemic remains a work in progress. As physical symptoms resembling COVID-19 infection (eg, headache, myalgia, sore throat, gastrointestinal symptoms) were associated with adverse mental health in both epicenters, the lack of testing for the coronavirus could worsen anxiety and stress. There is an urgent need to develop accurate, rapid diagnostic tests for use in general practitioners' clinics and community settings. Nevertheless, there is a need for a culturally sensitive approach to improve mental health during the pandemic. Further research is required to understand the perception of Spanish individuals toward health information related to COVID-19 and explore the underlying reasons between adverse mental health and health information (eg, symptoms, prevention methods, transmission methods, and personalized information). Furthermore, online psychological intervention such as internet-delivered cognitive behavioral therapy is a cost-effective treatment [42] and can reduce psychological symptoms [43] and challenge negative thoughts. Online psychological treatment can be conducted without face-to-face contact, thereby adhering to lockdown and social distancing measures during the pandemic.

Strengths and Limitations

The main strength of this study lies in the fact that we performed in-depth analyses and studied the relationship between psychological outcomes and other variables related to COVID-19 in the two largest epicenters in Europe and Asia during the early stage of the pandemic. However, there are several limitations to be considered when interpreting the results. One major limitation was the potential risk of sampling bias,

although respondents came from 159 cities from China and all 19 autonomous regions in Spain. This bias could be due to the online administration of the questionnaires; further, the majority of respondents from both countries had a good educational background and internet access. We could not reach out to potential respondents without internet access. The research findings should be interpreted with caution as the respondents might not be representative of the population, and the demographic data were not normally distributed. Nevertheless, it is important to compare mental health between two epicenters across two continents during the early stage of the pandemic as such a study cannot be done after the pandemic. Another major limitation is that Spanish and Chinese respondents were not matched by age as the Spanish respondents were significantly older and more likely to be unmarried compared to their Chinese counterparts. The differences observed between the Spanish and Chinese participants could be due to intrinsic differences in age and other demographic factors. Another limitation is that we were unable to calculate the response rate. For potential respondents who were not keen to participate in the online survey, no response was recorded, and we could not collect any information from them due to ethical requirements. Furthermore, we could only include one country from Asia and Europe because not all universities had agreements for collaboration and research data exchange. Our findings cannot be generalized to other epicenters in Asia and Europe (eg, Italy and the Lombardy region) [44]. Finally, this is a cross-sectional study, and we demonstrated an association but not a causal relationship between the independent variables and psychological outcomes.

Conclusions

COVID-19 epicenters in Europe and Asia faced different challenges. Due to a higher number of COVID-19 cases per capita in Spain and lack of previous health care system experience in dealing with the SARS outbreak, Spanish respondents reported more physical symptoms, contact history with COVID-19, higher perceived risk of contracting COVID-19, and frequent use of but less confidence in medical services. Spanish participants reported higher levels of stress and depression, while Chinese participants reported higher levels of psychological impact. Chinese respondents encountered more discrimination from other countries. There were cultural differences regarding the use of face masks and satisfaction with health information related to COVID-19. Spanish respondents were less likely to wear face masks. Spanish respondents were more dissatisfied with health information; the timing and types of health information were associated with adverse mental health parameters. Our findings also confirmed that physical symptoms resembling COVID-19 infection (eg, headache, myalgia, sore throat, gastrointestinal symptoms) were associated with adverse mental health in China and Spain.

Acknowledgments

We would like to express our appreciation to all the participants in our study. This study was funded by NUS iHealthtech Other Operating Expenses (R-722-000-004-731) and NUS Department of Psychological Medicine Other Operating Expenses (R-177-000-003-001).

Authors' Contributions

CW, MIL-N, RH, and MEAG were responsible for conceptualization. CW, MIL-N, RP, XW, YT, LX, CH, and MEAG were involved in data curation. Formal analysis was performed by CW, MIL-N, RP, XW, YT, and LX. RH was responsible for funding acquisition. XW and YT carried out the investigation, and YT was responsible for methodology. CW and MIL-N were responsible for project administration, XW for software; MIL-N, RP, RH, and MEAG for supervision, and FC for visualization. MIL-N, FC, and RH wrote the original draft; RH, CH, and MEAG reviewed and edited the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Supplementary tables (part 1).

[[DOCX File , 1859 KB - formative_v5i5e27818_app1.docx](#)]

Multimedia Appendix 2

Supplementary tables (part 2).

[[DOCX File , 103 KB - formative_v5i5e27818_app2.docx](#)]

References

1. Mahase E. Covid-19: WHO declares pandemic because of "alarming levels" of spread, severity, and inaction. *BMJ* 2020 Mar 12;368:m1036. [doi: [10.1136/bmj.m1036](https://doi.org/10.1136/bmj.m1036)] [Medline: [32165426](https://pubmed.ncbi.nlm.nih.gov/32165426/)]
2. COVID-19 Coronavirus Pandemic. Worldmeter. 2020. URL: <https://www.worldometers.info/coronavirus/> [accessed 2020-12-04]
3. Wang C, Pan R, Wan X, Tan Y, Xu L, Ho CS, et al. Immediate Psychological Responses and Associated Factors during the Initial Stage of the 2019 Coronavirus Disease (COVID-19) Epidemic among the General Population in China. *Int J Environ Res Public Health* 2020 Mar 06;17(5):1729 [FREE Full text] [doi: [10.3390/ijerph17051729](https://doi.org/10.3390/ijerph17051729)] [Medline: [32155789](https://pubmed.ncbi.nlm.nih.gov/32155789/)]
4. Salo J. China orders 1,400 military doctors, nurses to treat coronavirus. *NY Post*. 2020 Feb 2. URL: <https://nypost.com/2020/02/02/china-orders-14000-military-doctors-nurses-to-treat-coronavirus/> [accessed 2020-03-22]
5. Wang C, Chudzicka-Czupala 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 Sep 9;11:569981-569913 [FREE Full text] [doi: [10.3389/fpsy.2020.569981](https://doi.org/10.3389/fpsy.2020.569981)] [Medline: [33033485](https://pubmed.ncbi.nlm.nih.gov/33033485/)]
6. Perez-Bermejo M, Murillo-Llorente MT. The Fast Territorial Expansion of COVID-19 in Spain. *J Epidemiol* 2020 May 05;30(5):236 [FREE Full text] [doi: [10.2188/jea.JE20200123](https://doi.org/10.2188/jea.JE20200123)] [Medline: [32281555](https://pubmed.ncbi.nlm.nih.gov/32281555/)]
7. Spain overtakes China in COVID-19 infections. *Channel News Asia*. 2020 Mar 30. URL: <https://www.channelnewsasia.com/news/world/spain-covid-19-overtake-china-coronavirus-infections-12590454> [accessed 2021-03-25]
8. Saez M, Tobias A, Varga D, Barceló MA. Effectiveness of the measures to flatten the epidemic curve of COVID-19. The case of Spain. *Sci Total Environ* 2020 Jul 20;727:138761 [FREE Full text] [doi: [10.1016/j.scitotenv.2020.138761](https://doi.org/10.1016/j.scitotenv.2020.138761)] [Medline: [32330703](https://pubmed.ncbi.nlm.nih.gov/32330703/)]
9. Resolution of March 25, 2020, of the Spanish Parliament ordering the publication of the agreement authorizing the extension of the state of alarm declared by the Royal Decree 463 of March 14, 2020, on March 28, 2020. Number 2609. *Latham & Watkins*. 2020 Mar 16. URL: <https://www.lw.com/thoughtLeadership/spain-imposes-royal-decree-463-2020-t-manage-covid19-health-crisis> [accessed 2020-09-24]
10. Resolution of April 9, 2020, of the Spanish Parliament ordering the publication of the agreement authorizing the extension of the state of alarm declared by the Royal Decree 463 of March 14, 2020, on April 11, 2020. Number 101. *Boletín Oficial del Estado*. 2020 Apr 11. URL: <https://www.boe.es/boe/dias/2020/04/11/pdfs/BOE-A-2020-4406.pdf> [accessed 2020-09-24]
11. Covid-19 situation in Spain. *ISCIII CNE - National Epidemiology Center*. 2020. URL: <https://covid19.isciii.es/> [accessed 2020-09-24]
12. Rada AG. Covid-19: the precarious position of Spain's nursing homes. *BMJ* 2020 Apr 20;369:m1554. [doi: [10.1136/bmj.m1554](https://doi.org/10.1136/bmj.m1554)] [Medline: [32312710](https://pubmed.ncbi.nlm.nih.gov/32312710/)]
13. Ho CS, Chee CY, Ho RC. Mental Health Strategies to Combat the Psychological Impact of Coronavirus Disease 2019 (COVID-19) Beyond Paranoia and Panic. *Ann Acad Med Singap* 2020 Mar 31;49(3):155-160. [doi: [10.47102/annals-acadmedsg.202043](https://doi.org/10.47102/annals-acadmedsg.202043)]
14. Wang Y, Di Y, Ye J, Wei W. Study on the public psychological states and its related factors during the outbreak of coronavirus disease 2019 (COVID-19) in some regions of China. *Psychol Health Med* 2021 Jan;26(1):13-22. [doi: [10.1080/13548506.2020.1746817](https://doi.org/10.1080/13548506.2020.1746817)] [Medline: [32223317](https://pubmed.ncbi.nlm.nih.gov/32223317/)]

15. Wang C, Pan R, Wan X, Tan Y, Xu L, McIntyre RS, et al. A longitudinal study on the mental health of general population during the COVID-19 epidemic in China. *Brain Behav Immun* 2020 Jul;87:40-48 [FREE Full text] [doi: [10.1016/j.bbi.2020.04.028](https://doi.org/10.1016/j.bbi.2020.04.028)] [Medline: [32298802](https://pubmed.ncbi.nlm.nih.gov/32298802/)]
16. Tan W, Hao F, McIntyre RS, Jiang L, Jiang X, Zhang L, et al. Is returning to work during the COVID-19 pandemic stressful? A study on immediate mental health status and psychoneuroimmunity prevention measures of Chinese workforce. *Brain Behav Immun* 2020 Jul;87:84-92 [FREE Full text] [doi: [10.1016/j.bbi.2020.04.055](https://doi.org/10.1016/j.bbi.2020.04.055)] [Medline: [32335200](https://pubmed.ncbi.nlm.nih.gov/32335200/)]
17. Hao F, Tan W, Jiang L, Zhang L, Zhao X, Zou Y, et al. Do psychiatric patients experience more psychiatric symptoms during COVID-19 pandemic and lockdown? A case-control study with service and research implications for immunopsychiatry. *Brain Behav Immun* 2020 Jul;87:100-106. [doi: [10.1016/j.bbi.2020.04.069](https://doi.org/10.1016/j.bbi.2020.04.069)] [Medline: [32353518](https://pubmed.ncbi.nlm.nih.gov/32353518/)]
18. Mongkhon, Ruengorn C, Awiphan R, Thavorn K, Hutton B, Wongpakaran N, et al. Exposure to COVID-19-Related Information and its Association With Mental Health Problems in Thailand: Nationwide, Cross-sectional Survey Study. *J Med Internet Res* 2021 Feb 12;23(2):e25363 [FREE Full text] [doi: [10.2196/25363](https://doi.org/10.2196/25363)] [Medline: [33523828](https://pubmed.ncbi.nlm.nih.gov/33523828/)]
19. Ngoc Cong Duong K, Nguyen Le Bao T, Thi Lan Nguyen P, Vo Van T, Phung Lam T, Pham Gia A, et al. Psychological Impacts of COVID-19 During the First Nationwide Lockdown in Vietnam: Web-Based, Cross-Sectional Survey Study. *JMIR Form Res* 2020 Dec 15;4(12):e24776 [FREE Full text] [doi: [10.2196/24776](https://doi.org/10.2196/24776)] [Medline: [33284778](https://pubmed.ncbi.nlm.nih.gov/33284778/)]
20. Bowman L, Kwok KO, Redd R, Yi Y, Ward H, Wei WI, et al. Comparing Public Perceptions and Preventive Behaviors During the Early Phase of the COVID-19 Pandemic in Hong Kong and the United Kingdom: Cross-sectional Survey Study. *J Med Internet Res* 2021 Mar 08;23(3):e23231 [FREE Full text] [doi: [10.2196/23231](https://doi.org/10.2196/23231)] [Medline: [33539309](https://pubmed.ncbi.nlm.nih.gov/33539309/)]
21. Wang C, Chudzicka-Czupala A, Tee ML, Núñez MIL, Tripp C, Fardin MA, et al. A chain mediation model on COVID-19 symptoms and mental health outcomes in Americans, Asians and Europeans. *Sci Rep* 2021 Mar 19;11(1):6481-6491 [FREE Full text] [doi: [10.1038/s41598-021-85943-7](https://doi.org/10.1038/s41598-021-85943-7)] [Medline: [33742072](https://pubmed.ncbi.nlm.nih.gov/33742072/)]
22. Leung GM, Lam TH, Ho LM, Ho SY, Chan BHY, Wong IOL, et al. The impact of community psychological responses on outbreak control for severe acute respiratory syndrome in Hong Kong. *J Epidemiol Community Health* 2003 Nov 01;57(11):857-863 [FREE Full text] [doi: [10.1136/jech.57.11.857](https://doi.org/10.1136/jech.57.11.857)] [Medline: [14600110](https://pubmed.ncbi.nlm.nih.gov/14600110/)]
23. Zhang MW, Ho CS, Fang P, Lu Y, Ho RC. Usage of social media and smartphone application in assessment of physical and psychological well-being of individuals in times of a major air pollution crisis. *JMIR Mhealth Uhealth* 2014 Mar 25;2(1):e16. [doi: [10.2196/mhealth.2827](https://doi.org/10.2196/mhealth.2827)] [Medline: [25098255](https://pubmed.ncbi.nlm.nih.gov/25098255/)]
24. Chew NW, Lee GK, Tan BY, Jing M, Goh Y, Ngiam NJ, et al. A multinational, multicentre study on the psychological outcomes and associated physical symptoms amongst healthcare workers during COVID-19 outbreak. *Brain Behav Immun* 2020 Aug;88:559-565 [FREE Full text] [doi: [10.1016/j.bbi.2020.04.049](https://doi.org/10.1016/j.bbi.2020.04.049)] [Medline: [32330593](https://pubmed.ncbi.nlm.nih.gov/32330593/)]
25. Papadakaki M, Tsalkanis A, Sarris M, Pierrakos G, Eleonora Ferraro O, Stamouli M, et al. Physical, psychological and economic burden of two-wheel users after a road traffic injury: Evidence from intensive care units of three EU countries. *J Safety Res* 2018 Dec;67:155-163. [doi: [10.1016/j.jsr.2018.10.005](https://doi.org/10.1016/j.jsr.2018.10.005)] [Medline: [30553418](https://pubmed.ncbi.nlm.nih.gov/30553418/)]
26. Zhang MW, Ho CS, Fang P, Lu Y, Ho RC. Methodology of developing a smartphone application for crisis research and its clinical application. *THC* 2014 Aug 01;22(4):547-559. [doi: [10.3233/thc-140819](https://doi.org/10.3233/thc-140819)]
27. Creamer M, Bell R, Failla S. Psychometric properties of the Impact of Event Scale - Revised. *Behav Res Ther* 2003 Dec;41(12):1489-1496. [doi: [10.1016/j.brat.2003.07.010](https://doi.org/10.1016/j.brat.2003.07.010)] [Medline: [14705607](https://pubmed.ncbi.nlm.nih.gov/14705607/)]
28. Lee SM, Kang WS, Cho A, Kim T, Park JK. Psychological impact of the 2015 MERS outbreak on hospital workers and quarantined hemodialysis patients. *Compr Psychiatry* 2018 Nov;87:123-127 [FREE Full text] [doi: [10.1016/j.comppsy.2018.10.003](https://doi.org/10.1016/j.comppsy.2018.10.003)] [Medline: [30343247](https://pubmed.ncbi.nlm.nih.gov/30343247/)]
29. Le TA, Le MQT, Dang AD, Dang AK, Nguyen CT, Pham HQ, et al. Multi-level predictors of psychological problems among methadone maintenance treatment patients in difference types of settings in Vietnam. *Subst Abuse Treat Prev Policy* 2019 Sep 18;14(1):39-49 [FREE Full text] [doi: [10.1186/s13011-019-0223-4](https://doi.org/10.1186/s13011-019-0223-4)] [Medline: [31533764](https://pubmed.ncbi.nlm.nih.gov/31533764/)]
30. Ho CS, Tan EL, Ho RC, Chiu MY. Relationship of Anxiety and Depression with Respiratory Symptoms: Comparison between Depressed and Non-Depressed Smokers in Singapore. *Int J Environ Res Public Health* 2019 Jan 08;16(1):163 [FREE Full text] [doi: [10.3390/ijerph16010163](https://doi.org/10.3390/ijerph16010163)] [Medline: [30626156](https://pubmed.ncbi.nlm.nih.gov/30626156/)]
31. Quek T, Ho C, Choo C, Nguyen L, Tran B, Ho R. Misophonia in Singaporean Psychiatric Patients: A Cross-Sectional Study. *Int J Environ Res Public Health* 2018 Jul 04;15(7):1410 [FREE Full text] [doi: [10.3390/ijerph15071410](https://doi.org/10.3390/ijerph15071410)] [Medline: [29973546](https://pubmed.ncbi.nlm.nih.gov/29973546/)]
32. González-Cabrera J, Tourón J, Machimbarrena, Gutiérrez-Ortega M, Álvarez-Bardón A, Garaigordobil. Cyberbullying in Gifted Students: Prevalence and Psychological Well-Being in a Spanish Sample. *Int J Environ Res Public Health* 2019 Jun 19;16(12):2173 [FREE Full text] [doi: [10.3390/ijerph16122173](https://doi.org/10.3390/ijerph16122173)] [Medline: [31248205](https://pubmed.ncbi.nlm.nih.gov/31248205/)]
33. Tan B, Chew NWS, Lee GKH, Jing M, Goh Y, Yeo LLL, et al. Psychological Impact of the COVID-19 Pandemic on Health Care Workers in Singapore. *Ann Intern Med* 2020 Aug 18;173(4):317-320 [FREE Full text] [doi: [10.7326/M20-1083](https://doi.org/10.7326/M20-1083)] [Medline: [32251513](https://pubmed.ncbi.nlm.nih.gov/32251513/)]
34. Jones S. Spain: doctors struggle to cope as 514 die from coronavirus in a day. *The Guardian*. 2020 Mar 24. URL: <https://www.theguardian.com/world/2020/mar/24/spain-doctors-lack-protection-coronavirus-covid-19> [accessed 2020-08-05]
35. "Slow response": how Britain became worst-hit in Europe by virus. *MedicalXpress*. 2020 May 5. URL: <https://medicalxpress.com/news/2020-05-response-britain-worst-hit-europe-virus.html> [accessed 2020-05-08]

36. Legido-Quigley H, Mateos-García JT, Campos VR, Gea-Sánchez M, Muntaner C, McKee M. The resilience of the Spanish health system against the COVID-19 pandemic. *The Lancet Public Health* 2020 May;5(5):e251-e252. [doi: [10.1016/s2468-2667\(20\)30060-8](https://doi.org/10.1016/s2468-2667(20)30060-8)]
37. Ma Q, Shan H, Zhang HL, Li GM, Yang RM, Chen JM. Potential utilities of mask-wearing and instant hand hygiene for fighting SARS-CoV-2. *J Med Virol* 2020 Sep 01;92(9):1567-1571 [FREE Full text] [doi: [10.1002/jmv.25805](https://doi.org/10.1002/jmv.25805)] [Medline: [32232986](https://pubmed.ncbi.nlm.nih.gov/32232986/)]
38. Zhai Z. Facial mask: A necessity to beat COVID-19. *Build Environ* 2020 May 15;175:106827 [FREE Full text] [doi: [10.1016/j.buildenv.2020.106827](https://doi.org/10.1016/j.buildenv.2020.106827)] [Medline: [32287995](https://pubmed.ncbi.nlm.nih.gov/32287995/)]
39. Fact check: Face masks being worn in China before COVID-19 doesn't prove they are ineffective. Reuters. 2020 Sep 4. URL: <https://www.reuters.com/article/uk-factcheck-china-mask-covid-idUSKBN25V1UR> [accessed 2020-09-26]
40. Cheng VC, Wong S, Chuang VW, So SY, Chen JH, Sridhar S, et al. The role of community-wide wearing of face mask for control of coronavirus disease 2019 (COVID-19) epidemic due to SARS-CoV-2. *J Infect* 2020 Jul;81(1):107-114 [FREE Full text] [doi: [10.1016/j.jinf.2020.04.024](https://doi.org/10.1016/j.jinf.2020.04.024)] [Medline: [32335167](https://pubmed.ncbi.nlm.nih.gov/32335167/)]
41. Catherine W. It's unfair to blame China for coronavirus pandemic, Lancet editor tells state media. *South China Morning Post*. 2020 May 2. URL: <https://www.scmp.com/news/china/science/article/3082606/its-unfair-blame-china-coronavirus-pandemic-lancet-editor-tells> [accessed 2020-05-08]
42. Zhang MW, Ho RC. Moodle: The cost effective solution for internet cognitive behavioral therapy (I-CBT) interventions. *THC* 2017 Feb 21;25(1):163-165. [doi: [10.3233/thc-161261](https://doi.org/10.3233/thc-161261)]
43. Soh HL, Ho RC, Ho CS, Tam WW. Efficacy of digital cognitive behavioural therapy for insomnia: a meta-analysis of randomised controlled trials. *Sleep Med* 2020 Nov;75:315-325. [doi: [10.1016/j.sleep.2020.08.020](https://doi.org/10.1016/j.sleep.2020.08.020)] [Medline: [32950013](https://pubmed.ncbi.nlm.nih.gov/32950013/)]
44. Simone L, Gnagnarella C. Differences Between Health Workers and General Population in Risk Perception, Behaviors, and Psychological Distress Related to COVID-19 Spread in Italy. *Front Psychol* 2020;11(2166):2166-2176 [FREE Full text] [doi: [10.3389/fpsyg.2020.02166](https://doi.org/10.3389/fpsyg.2020.02166)] [Medline: [33013555](https://pubmed.ncbi.nlm.nih.gov/33013555/)]

Abbreviations

DASS-21: Depression, Anxiety and Stress Scale–21 Items

IERS-R: Impact of Event Scale-Revised

PTSD: posttraumatic stress disorder

SARS: severe acute respiratory syndrome

Edited by G Eysenbach; submitted 08.02.21; peer-reviewed by B Van, G Bassi, PKH Mo; comments to author 10.03.21; revised version received 31.03.21; accepted 22.04.21; published 21.05.21.

Please cite as:

Wang C, López-Núñez MI, Pan R, Wan X, Tan Y, Xu L, Choo F, Ho R, Ho C, Aparicio García ME

The Impact of the COVID-19 Pandemic on Physical and Mental Health in China and Spain: Cross-sectional Study

JMIR Form Res 2021;5(5):e27818

URL: <https://formative.jmir.org/2021/5/e27818>

doi: [10.2196/27818](https://doi.org/10.2196/27818)

PMID: [33900933](https://pubmed.ncbi.nlm.nih.gov/33900933/)

©Cuiyan Wang, María Inmaculada López-Núñez, Riyu Pan, Xiaoyang Wan, Yilin Tan, Linkang Xu, Faith Choo, Roger Ho, Cyrus Ho, Marta E Aparicio García. Originally published in *JMIR Formative Research* (<https://formative.jmir.org>), 21.05.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Formative Research*, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Digital Phenotypes for Understanding Individuals' Compliance With COVID-19 Policies and Personalized Nudges: Longitudinal Observational Study

Ahmed Ibrahim^{1*}, MSc; Heng Zhang^{1*}, MSc; Sarah Clinch¹, PhD; Ellen Poliakoff¹, PhD; Bijan Parsia¹, PhD; Simon Harper¹, PhD

Department of Computer Science, The University of Manchester, Manchester, United Kingdom

*these authors contributed equally

Corresponding Author:

Heng Zhang, MSc

Department of Computer Science

The University of Manchester

LF7, Kilburn Building

Kilburn Building

Manchester, M13 9PL

United Kingdom

Phone: 44 7427630668

Email: heng.zhang@manchester.ac.uk

Abstract

Background: Governments promote behavioral policies such as social distancing and phased reopening to control the spread of COVID-19. Digital phenotyping helps promote the compliance with these policies through the personalized behavioral knowledge it produces.

Objective: This study investigated the value of smartphone-derived digital phenotypes in (1) analyzing individuals' compliance with COVID-19 policies through behavioral responses and (2) suggesting ways to personalize communication through those policies.

Methods: We conducted longitudinal experiments that started before the outbreak of COVID-19 and continued during the pandemic. A total of 16 participants were recruited before the pandemic, and a smartphone sensing app was installed for each of them. We then assessed individual compliance with COVID-19 policies and their impact on habitual behaviors.

Results: Our results show a significant change in people's mobility ($P < .001$) as a result of COVID-19 regulations, from an average of 10 visited places every week to approximately 2 places a week. We also discussed our results within the context of nudges used by the National Health Service in the United Kingdom to promote COVID-19 regulations.

Conclusions: Our findings show that digital phenotyping has substantial value in understanding people's behavior during a pandemic. Behavioral features extracted from digital phenotypes can facilitate the personalization of and compliance with behavioral policies. A rule-based messaging system can be implemented to deliver nudges on the basis of digital phenotyping.

(*JMIR Form Res* 2021;5(5):e23461) doi:[10.2196/23461](https://doi.org/10.2196/23461)

KEYWORDS

behavior; compliance; COVID-19; digital phenotyping; nudges; personalization; policy; sensor; smartphone

Introduction

Background

COVID-19 is a highly contagious disease with confirmed cases in more than 188 countries as between December 2019 and June 2020, resulting in a global pandemic [1]. To control the spread of COVID-19, governments have enforced behavioral policies,

such as stay-at-home and social distancing measures, which limit the usual patterns of human interaction [2,3]. The potential risk of problems with social isolation [4] complicates the implementation of these policies, which places an additional responsibility on governments to maintain mental health throughout the pandemic.

Currently, governments rely on communication campaigns to persuade people to adhere to COVID-19 behavioral policies and reduce disease spread. Health agencies, such as the National Health service (NHS) in the United Kingdom, design communication in a way that encourages the application of the promoted behaviors while avoiding problems related to social isolation. This approach to communications design employs behavioral insights derived from scientific studies to deliver behavioral guidance [5]. The communications resulting from this process are called “nudges” [6].

Despite the critical role of these campaigns in elevating community awareness, they are not designed to reflect differently when people exhibit different behavioral responses to the promoted procedures. Digital devices including smartphones can be used to recognize behavioral differences. Accordingly, communications can be personalized and contextualized on the basis of the individual’s behavior. Smartphones facilitate the capturing of behavioral features through the continuous and unobtrusive collection of sensor and interaction data; this process is known as “digital phenotyping.”

In this study, we show how an individual’s behavioral reactions to COVID-19 policies can be observed through digital phenotyping. Subsequently, we suggest a personalized way of delivering nudges designed around the individual’s reactions to the enforced regulations. We report 2 longitudinal studies that started before the outbreak of the pandemic to collect digital phenotypes. Our studies allow us to observe the impact on the overall behavior before and during the outbreak. Additionally, we observed the impact of COVID-19 on habitual behaviors and the uptake of new apps.

Our primary research contribution is the introduction of an approach that employs behavioral differences derived from digital phenotyping in the design of personalized nudges. Although we did not conduct an experiment to measure the real-time effects of personalized nudges, the proposed nudges conform to the general guidelines in behavioral science and are expected to improve individual compliance to them. Moreover, the development of mental health issues as a result of lockdown policies can be observed through digital phenotyping and better addressed through personalized nudges.

Related Work

With the popularity and evolution of personal electronic devices, people are producing an increasing number of digital footprints such as those generated through web-based communication and mobile device usage. These footprints can be linked and analyzed with clinical data to create an individualized, nuanced view of human disease, which is called a “digital phenotype” [7]. In 2015, a digital phenotype was defined by Jukka-Pekka Onnela as the “moment-by-moment quantification of the individual-level human phenotype in-situ using data from smartphones and other personal digital devices” [8]. Digital phenotyping has become one of the most innovative approaches to enhance health and wellness via human-computer interactions through digital technology.

Nowadays, smartphones have become the one of the ideal tools for digital phenotyping. Smartphones are the hub of personal communication, and almost everyone has a smartphone. Although smartphones are not specially designed for behavioral research, they can collect a large amount of related data directly and instantly with ecological validity. Social interaction on smartphones, including calls, messages, emails, and social media usage, can be captured without difficulty. Thus, social sensing could be less intrusive on smartphones than on any other device. Embedded multiple and power sensors also empower smartphones as an efficient tool to record the surrounding social context. For example, raw data from sensors such as microphones, the global positioning system (GPS), and accelerometers can be gathered and interpreted as conversation engagement, mobility patterns, and the number of encounters to infer social interaction occurring outside of smartphones. Thus, smartphones could be one of the most applicable ways of passive societal digital phenotyping.

Digital phenotyping on smartphones has been utilized in various fields, especially psychological and health-related studies. Abdullah et al [9] collected phone usage patterns to detect and predict discrepancies in sleep rhythms. Furthermore, LiKamWa et al [10] analyzed call, message, or email contacts and location clusters from smartphones to infer users’ daily mood. Farhan et al [11] combined the locations and activities from participants’ smartphones to predict depression. Boukhechba et al [12] explored the association of social anxiety with GPS and communication patterns. To confirm the findings and observations of passively collected smartphone data, all these studies asked for participants’ input through various means including interviews, focus groups, and questionnaires. All these studies claimed to have relatively high accuracy. Albeit with different aims, our study similarly implemented these smartphone monitoring technologies. We collected data before and during the COVID-19 lockdown, which provided us an opportunity to observe individual behavioral changes. We also conducted interviews with our study participants to verify our findings.

Methods

Methods Overview

We used behavioral indicators for the COVID-19 policies as proxies that would help us observe the adoption of the desired change by people. Our approach relies on transforming raw smartphone data collected longitudinally (ie, digital phenotypes) into behavioral features. Distance travelled and time spent at home by a person are examples of features derived from raw location data (ie, timestamped longitude and latitude attributes). The detection of behavioral indicators is achieved at the level of behavioral features rather than the raw data. This is because behavioral indicators are manifested at a higher level of human understanding expressible by those features. In the following section, we detail the behavioral features and their roles in recognizing the behavioral indicators of the proposed policies.

For this disease, transmitted through close contact, reducing the possibility of an uninfected person having physical contact with an infected person may be the only effective way to suppress

the transmission of the disease. Since the onset of the COVID-19 pandemic, governments worldwide enforced a series of behavioral policies based on this concept to control the spread of this highly infectious disease. For example, the government of the United Kingdom instructed individuals to stay home as much as possible, to limit contact with those from other households, and to maintain distance from others when stepping out of home (2 meters apart where possible) [13]. Other measures include school closures, working from home, cancellation of mass gatherings, and travel restrictions. These policies are referred to as “social distancing” or “physical distancing” policies.

Stay-at-Home Measures

Deriving behavioral indicators of social distancing from smartphone data was our primary consideration. There are some existing studies on the mobility responses to COVID-19; for instance, a previous study [14] analyzed public geolocated Twitter data to measure the travel behaviors of users. Allcott et al [15] combined surveys and GPS foot traffic patterns to observe partisan differences in social distancing. They reported a substantial reduction in the mobility of people in the United States, albeit with partisan gaps in beliefs and behavior. Similarly, we can expect that our participants should spend almost all their time at home and to limit the time and number of places when stepping out, which is usually only for essential shopping owing to the implementation of social distancing measures. These behavioral changes can be acquired from raw GPS data. Since participants' smartphones record latitude and longitude attributes continuously, their distance from home can always be calculated. Thus, we can determine the time and frequency of their trips outside of home.

Furthermore, social distancing measures can bring about adverse effects, especially on mental health. Some of these reported effects include stress, anxiety and depression, and panic [16]. To maintain mental well-being and while at home, people may find alternative methods of communication to replace their regular face-to-face interactions. Phone calls, messages, video chatting, and social media are possible substitutions people may choose; accordingly, a potential increase in the use of these communication methods is expected. With the various data sources, we could draw a comprehensive and personalized picture of how people react to the impact of COVID-19 restrictions.

Social Distancing Measures

Social distancing implies that people should meet fewer people than they would during normal times. Bluetooth signals are an effective reference for face-to-face interaction recorded on smartphones. Nowadays, almost everyone carries a smartphone, and almost every smartphone is equipped with Bluetooth technology, which scans surrounding signals and reports its identity continuously in a short range. Thus, every newly captured Bluetooth entry could potentially represent a new person in close proximity [17]. This technology has been widely used in the field to estimate face-to-face proximity [18]. Although it is not fully accurate because of the physical position

of the smartphone and surrounding environments, it can still provide a trend that people have less face-to-face interactions. Hence, owing to the social distancing policy, a reduction in the number of unique Bluetooth signals is expected. Theoretically, this would indicate whether our participants adhere to the rules of staying at home and avoiding others visiting their household.

Moreover, social distancing has also affected people when they go for essential shopping. Many grocery stores have a limited number of people in their branches and have introduced directional floor markings to help shoppers maintain a 2-meter distance from one another [19]. This policy could reduce the capacity of crowded grocery stores, and fewer people are expected to be in close proximity to our participants compared to the time before social distancing measures were implemented. Thus, from Bluetooth signals, we could expect a reduction in the number of unique devices from a single scan.

Experiments

We report results from 2 longitudinal studies conducted to gather smartphones' digital phenotypes. Both studies were underway prior to, and continued through, large-scale transmission of COVID-19 and associated social distancing behaviors.

Participants

The studies were reviewed and approved by the Department of Computer Science Ethics Committee at the University. A total of 16 participants were recruited (4 males and 4 females per experiment) through the university database and websites. The 2 experiments recruited individuals from different populations in the United Kingdom: (1) students and (2) patients with a diagnosis of Parkinson disease (aged 63-75 years).

The 2 studies used smartphones to capture data on the participants' activities. Both experiments rely on the same sensing platform.

Instrument

In this study, we used smartphones as independent sensing tools to retrieve participants' behavioral data. The AWARE sensing platform [20] and developed plug-ins were deployed on participants' smartphone as a monitor app. Under the approval of the ethics committee, different kinds of data, including calls, messages, social media app usage, smartphone usage, notifications, locations, Bluetooth signals, and Wi-Fi signals were collected passively. The content of sensitive communications, such as calls, messages, and conversations, was not recorded. All these data were processed to maintain the anonymity and confidentiality of all participants. All data sources are summarized in [Table 1](#).

Participants were asked to attend an introductory interview to obtain information on our study and to clarify any of their doubts. On obtaining formal approval from the participants, the AWARE app was installed on their smartphones. Participants were asked to keep the installed app running and use their phones as they normally do. An offline analysis was conducted on data synced with the backend AWARE server.

Table 1. Sources of the collected digital phenotypes with data descriptions.

Source	Data
Global positioning system	Location coordinates (longitude and latitude).
Weather	Temperature, humidity, pressure, wind speed, cloudiness, amount of rain and snow, and times of sunrise and sunset.
Apps	App usage data.
Notifications	All notifications generated by any app installed on the phone.
Screen usage	Screen interactions and visited websites.
Wi-Fi signals	Access points.
Bluetooth signals	Nearby devices.
Battery status	Charging status, charging start time, charging end time, discharging start time, and discharging end time.
Calls	Call types (outgoing, incoming, and missed calls) and times. Numbers are stored in an encrypted format.
Keyboard	Time and the typed letters (numbers and emails are replaced with asterisks).
Consent form	Name and signature.
Questionnaire	We asked about interests, which we inferred from the data.
Activity recognition	Tilting, running, being within a vehicle, walking, being on a bicycle, and traveling on foot.

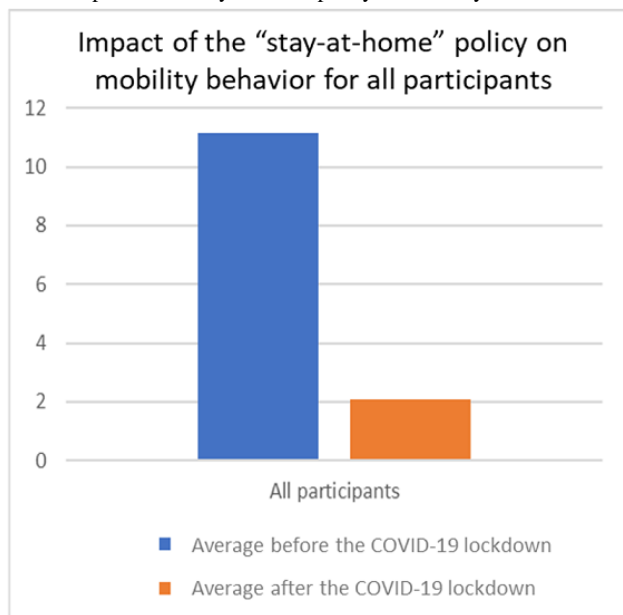
Results

Results Overview

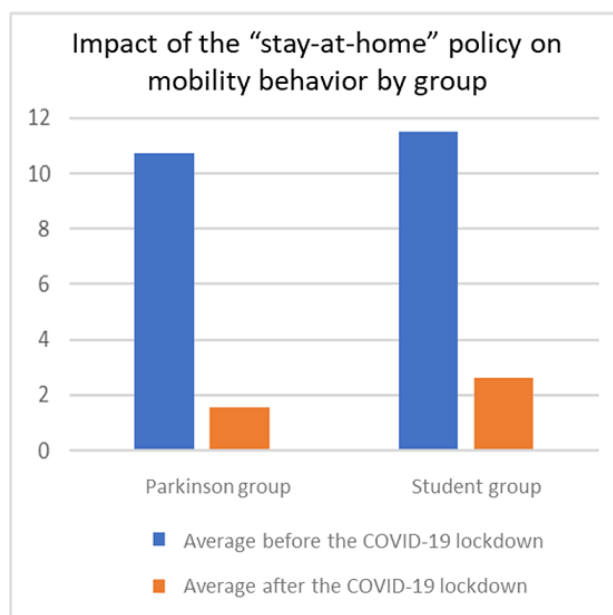
This section discusses the results obtained from the responses to the stay-at-home and social distancing policies. To show how digital phenotyping can help understand behavioral responses to these policies, we selected a prototypical participant who exemplified the general behavioral responses exhibited by all

participants, in each subsection, except for [Figure 1](#), which represents all participants. Our experiments started at different times; therefore, the lockdown timelines for each participant may differ. The behavioral responses to COVID-19 were captured despite the differences in the lockdown week. It was intended per our experimental design to have participants adhere to these policies at different times because participants were individually assessed, and no extrapolation among other participants was intended.

Figure 1. Impact of the stay-at-home policy on mobility behavior.



(A) Average number of visited places before and after the lockdown for all participants.



(B) Average number of visited places before and after the lockdown for each group.

Stay-at-Home Measures

Mobility patterns for participants in both experiments significantly decreased as a result of the compliance with the

stay-at-home policy ($P < .001$) ([Figure 1A](#)). Before the lockdown, the average number of places visited was slightly lesser among patients with Parkinson disease than among the students ([Figure 1B](#)). However, a patient with Parkinson disease and a student

may exhibit similar responses to the stay-at-home policy. Thus, individuals of the same group can exhibit a pattern that is different from the average behavior of their corresponding groups. Thus, individual analysis of digital phenotypes would help better understand people's compliance with the suggested policies.

Participants exhibited similar behavioral responses to COVID-19 regulations. We selected a participant who exemplifies the behavioral responses to present the results. We divided the participant's behavior window by week (Monday to Sunday), such that a whole cycle of a weekly social routine could be acquired. The stop point detection algorithms were applied for raw GPS data, such that the place of residence of the participant could be extracted. We used the algorithm proposed by Li et al [21] to extract stop points. The algorithm processes data points sequentially, and stop points are defined on the basis of predefined time and distance thresholds. Furthermore, we considered the location where participants spend most of their time of the day as their home. We used Foursquare [22] to determine the names of places, which allows for a better understanding of location semantics. By summing up the calculated results of the algorithm, the length of time participants spend at home and time spent by participants outside of home per week were obtained.

Another indicator is Bluetooth signals. As mentioned before, a scanned unique Bluetooth device could represent a person in close proximity. With everyone staying at home, fewer new identified Bluetooth entries were expected to be recorded. The time spent outside of home was usually below 30 minutes, but identified Bluetooth entries were all above 1000. To easily observe the similar trend of time spent outside of home and the number of new identified Bluetooth entries, we normalized the actual data so they can be plotted on the same graph. As illustrated in Figure 2, a clear boundary was observed, in that the participant went outside of home fewer times and presented decreased unique Bluetooth entries. Although fluctuations continue, the edge appeared around week 9; that is, March 15-22. This was the week before a lockdown was officially declared in the United Kingdom. Thus, it was observed that this participant perceived the stay-at-home policy and obeyed it objectively.

Figure 3 shows the impact of the "stay-at-home" policy on participant mobility. The figure represents the mobility behavior of participants who reside in the United Kingdom. Starting from week 12, the number of visited locations drastically decreased from an average of 7 locations to 2 locations. The 2 locations are the participant's home and a grocery store. To motivate this participant to comply with the stay-at-home policy, options for the delivery of grocery items or shopping times can be communicated.

Figure 2. Normalized unique Bluetooth signals and time spent out of home of a participant before and after the lock-down.

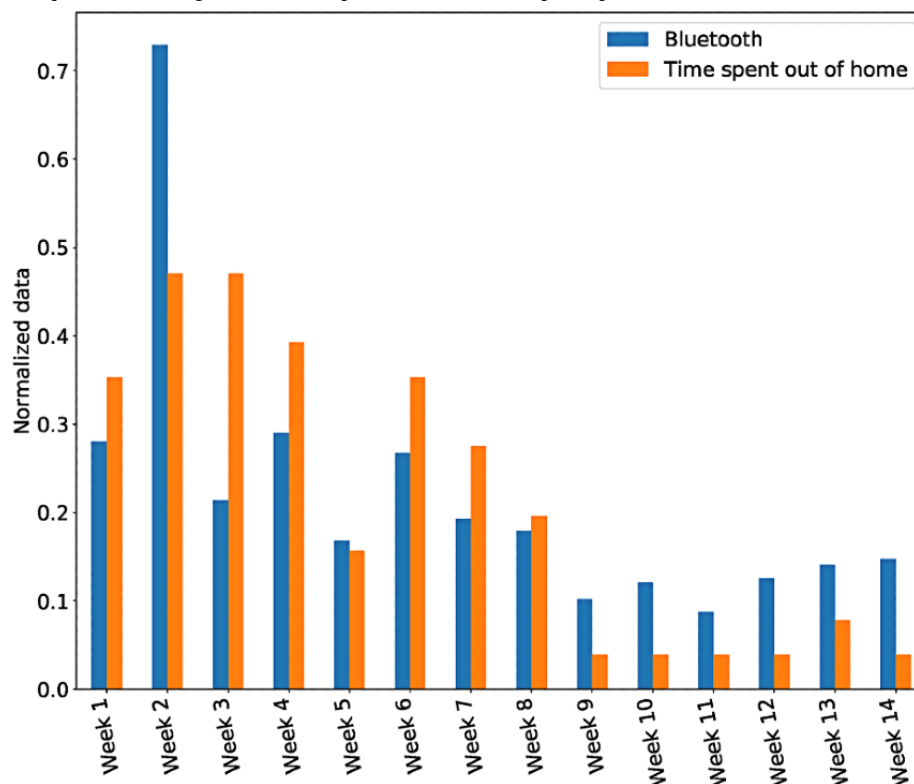
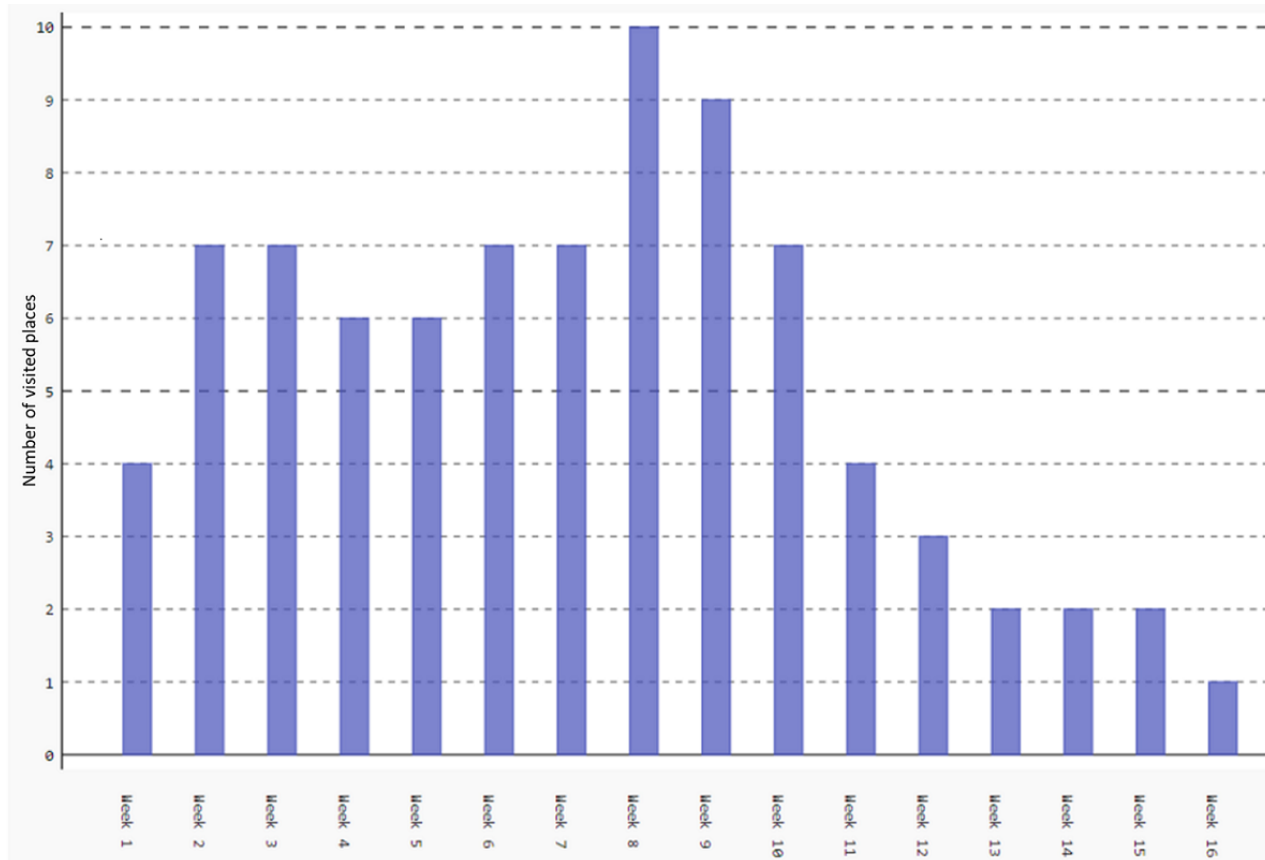


Figure 3. Location visited by a participant before and after the lockdown.

Social Distancing Measures

As described before, in accordance with the social distancing policy, people have to stay further away from each other than they would during normal times. Because of the capability of the Bluetooth technology, fewer scanned entries would be expected at a time. In this example, we also separated the data into natural weeks and combined all Bluetooth records within that week. Then, we divided this number by the total times for the scans to calculate the average 1-time Bluetooth discovery. As shown in Figure 4, the average 1-time Bluetooth entries decreased around week 9, which is the first week of the official lockdown in the United Kingdom. This potentially indicates that the participant maintained social distance with others and met fewer people during the lockdown.

The results of our experiment show that the participants complied with COVID-19 policies. Participants managed to

stay at home and adapt to the requested changes. However, to stay connected, the participant data show corresponding changes in app usage. The usage of social media apps, phone calls, and video conferences increased for most participants compared to the period before the lockdown. Figure 5 shows the app usage of a participant before and during the pandemic. Instagram was used the longest at 19.50 hours of usage, whereas the time spent on the Houseparty app was 9.27 hours. Values were normalized to easily observe the trend and be consistent with observations from other sources. The lockdown started during week 3. Consequently, the usage of apps, such as Facebook Messenger, WhatsApp, and Discord, has increased.

In contrast, 2 participants presented a decline in phone usage during the lockdown. When interviewed, the participants indicated that they started to use their personal computers and smart televisions more to accomplish the same tasks they previously did with smartphones.

Figure 4. Average 1-time Bluetooth entries before and after the lockdown.

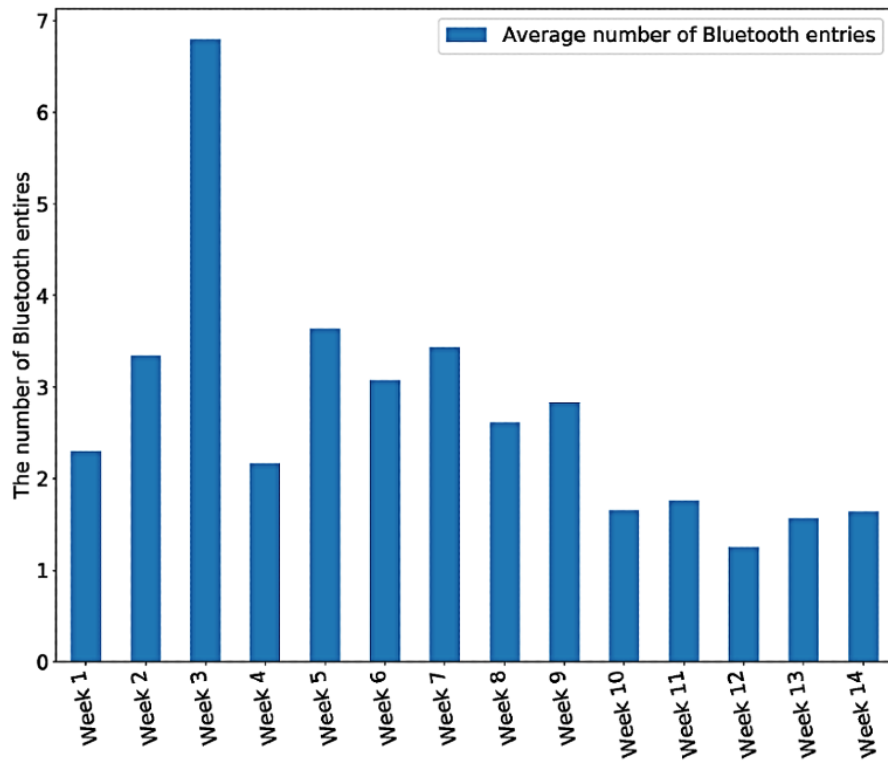
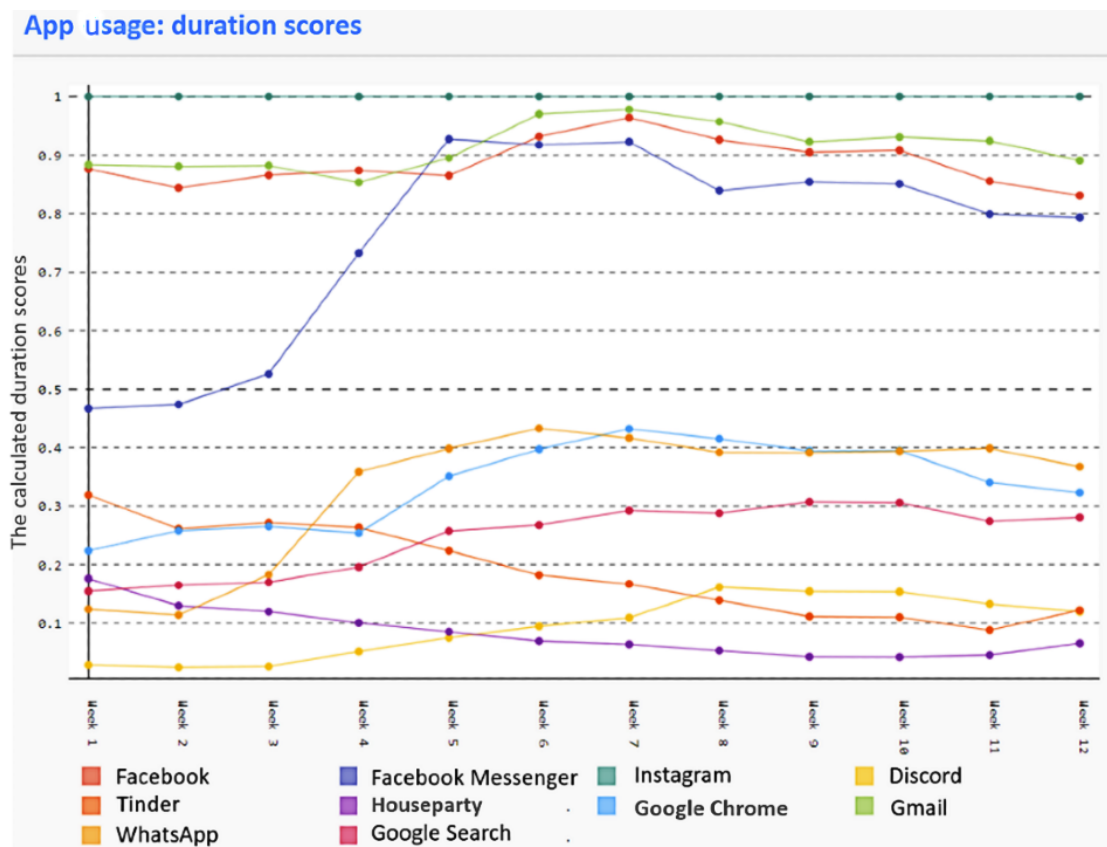


Figure 5. Normalized duration scores for a participant before and during the pandemic. The participant was enrolled in the first week of March and the lockdown started after the third week of data collection.



Discussion

Principal Findings

The reported results show that actionable information can be derived from digital phenotyping. The information derived from understanding participants' compliance, as well as the behavioral impact, can be used in personalized behavioral interventions. Behavioral nudges are used as an effective approach to promote behavioral changes. The NHS in the United Kingdom employs behavioral principles, such as reducing the cognitive load, to communicate nudges. We use actual text messages delivered by the NHS during the pandemic to demonstrate the potential benefit of personalization based on digital phenotyping. We show how a personalized understanding can be leveraged for more traction nudges and just-in-time intervention. The Behavioural Insights (BI) team [23] and the NHS have collaborated to nudge approximately 2 million people through text messages. The recipients of these nudges include people at the highest risk of developing critical complications should they contract the disease. The BI team employ the following behavioral principles to produce the content of a nudge (ie, the delivered text message).

1. Selection of the appropriate communication channel: since smartphone apps introduce multiple communication channels (eg, SMS, WhatsApp, and Messenger), personal preferences vary. The NHS and BI team have selected SMS as their preferred method on the basis of a study that shows that 85% of 600 participants do not mind receiving text messages on their personal devices from the NHS [24].

2. Signifying the key points: owing to the limitation of text messages, the NHS and BI team have to summarize extended guidelines into short messages. Accordingly, they designed messages such that the key ideas are prioritized.
3. Minimization of confusion and the cognitive load: the key ideas should be delivered in a language that is understandable by laypeople. Additionally, the messages should be clear to avoid confusion and misunderstanding that may quickly spread and negatively impact people.
4. Drawing on scientific behavioral findings: insights derived from behavioral and psychological studies are used to design nudges. For instance, it has been suggested that providing the rationale can help manage people's mental health when quarantined. Accordingly, the NHS and BI team comply with that when designing nudges.

These behavioral principles are population-based, which has been reflected on the content of the nudge. M1, M2, and M3 (Table 2) are examples of 3 nudges that are delivered in accordance with these principles. We hypothesize that digital phenotyping can better improve the content and delivery of these nudges through personalization. For instance, the predicate of M1 can be tailored in accordance with the participant's status as follows. We can predict whether or not a person lives alone from the digital phenotypes. Accordingly, 2 versions of the message can be prepared to deliver a personalized nudge. Versions can be tailored on the basis of the predicted status, age, or other demographics predictable through digital phenotyping.

Table 2. Text messages used by the National Health Service of the United Kingdom for nudging and our proposed personalization.

Code	Goal	Text message by the National Health Service of the United Kingdom	Personalization suggestions
M1	Nudge to establish social responsibility and stay connected	"If you live alone, text a friend or a family member to let them know you are following advice to stay at home until it is safer to mix with others. Plan to chat to someone over the phone at least once a day."	If a participant chats regularly or lives with others, do not send the message and prevent overmessaging.
M2	Nudge to maintain a normal routine and ease anxiety	"Try to stick as closely as you can to your typical daily routine."	If a participant frequents the cinemas, send the following message: "Watch a movie and try to stick as closely as you can to your typical daily routine."
M3	Nudge to preserve mental health	"Are there things you enjoy doing at home that you usually don't have time for?"	If a participant reports home activity, do not send the message and prevent overmessaging.

Digital phenotypes can also improve M2. For instance, an individual used to go to the cinema on Saturdays. Instead of delivering a general nudge about adhering to the typical routine, we can nudge the participant to watch a movie every Saturday during the pandemic. Thus, the typical routine can be embraced, and the delivery of the nudge can be contextualized (ie, just-in-time intervention). Adhering to typical routines can improve the mental health of individuals and reduce the negative impact of COVID-19 policies.

The information derived from digital phenotyping can also be used to prevent overmessaging. M1 encourages participants to chat with others to stay connected. If the derived data show that a participant regularly chats with others, there is no need to send M1. We speculate that crafting messages on the basis of both

data and behavioral principles as well as introducing fewer messages is expected to provide better results. However, actual field testing is required to scientifically measure the real effect of doing so.

Although our approach demonstrates a potential way of producing personalized nudges, it can be reflected in existing behavioral change frameworks such as the behavioral change wheel [25]. For instance, the framework of the behavioral change wheel identifies 3 main stages to the behavioral change: (1) understanding the behavior to be changed, (2) deciding on the intervention function, and (3) selecting the mode of delivery. We profile and understand the individuals' behaviors through digital phenotyping. Incentivization and persuasion are intervention functions that shape nudging [18]. Communication

as a delivery mode is then used to deliver text messages that nudge people to exhibit the desired behavior.

We are aware of the privacy concerns that may hinder the measurements and implementation of personalized nudges. However, apps can be designed in a way that allows people to partially share information in accordance with their needs. For instance, an individual may choose to share the location data only if diagnosed with COVID-19, to trace and limit the spread of the disease to others. Another individual may choose to share his/her data to receive personalized nudges that help him/her adhere to the daily routine (M2). Nevertheless, in these cases and others, personal behaviors are privately phenotyped, and it is up to the person whether or not to share the collected data. Alternatively, messages can be packaged with the app and delivered to participants on the basis of the outcome of a decision tree.

Stay-at-home, social distancing, and other policies are primarily behavioral measures aimed at changing individuals' behaviors to ensure that the risk of contracting the disease is reduced. From this standpoint, behavioral change frameworks (eg, nudging and the behavioral change wheel) can be relied upon to support the implementation of these behavioral policies. The use of digital phenotyping in activating these frameworks provides an opportunity to personalize the delivery of these policies on the basis of each individual's data. Individuals, institutions, and governments can benefit from such personalization in containing the spread of the virus. Governments may choose to develop apps that have behavioral policies implemented as built-in messages. The delivery of these messages is designed to adapt in accordance with the exhibited behaviors. Individuals who stayed at home (according to digital phenotyping) will not receive messages encouraging them to do so. This decision and others related to message delivery are made locally, on the individual's phone, without compromising his/her privacy. However, individuals who test positive can help governments reduce the potential impacts on others by voluntarily sharing their latest mobility behaviors.

Besides generating personalized nudges, digital phenotyping shows its capability to observe people's behavior on an individual level. In the context of the COVID-19 pandemic, digital phenotyping has great potential for various

implementations. Some of the COVID-19 tracking apps such as TraceTogether in Singapore and COVIDSafe in Australia have used Bluetooth technology embedded in smartphones as their primary contact tracing tool [26]. People are encouraged to install these apps so they can know if they have been in close contact with individuals who have tested positive for COVID-19. Institutions such as universities can implement digital phenotyping as innovative methods to study the traditional physiological or societal questions, since no face-to-face settlement is needed. Care facilities could also have digital phenotyping apps installed on their clients' smartphones, such that their issues can be noted without face-to-face reporting. Moreover, the large amount of personal and longitudinal digital phenotyping data could provide policymakers with a deeper understanding of the impact of COVID-19 on a sample of the population. This would shed light on how people actually react to these policies, rather than only determining the infection rate.

Conclusions

This study shows how digital phenotyping can be of value in understanding people's behavior during a pandemic. Behavioral features extracted from digital phenotypes represent the cornerstone that facilitates the personalization of and compliance with behavioral policies. We presented examples of using Bluetooth, GPS, and app usage data to analyze behavioral responses to COVID-19 policies. Additional sources can be further investigated, such as accelerometers and their role in understanding if people pause more to maintain safe distance.

To encourage the large-scale adaptation of digital phenotyping, governments can emphasize the potential benefits of public health and of maintaining mental health. To preserve privacy, an individual's data are stored locally, and he/she can make the ultimate decision on what to share and to whom the access is granted.

A rule-based messaging implementation can be used to deliver nudges on the basis of the analysis of digital phenotyping. In future studies, we intend to examine the impact of these suggested messages on a sample of the population to measure the impact of preventing overmessaging. Conducting a real-world experiment would also enable us to assess whether having more tailored messages would yield the expected benefits.

Conflicts of Interest

None declared.

References

1. Coronavirus. Wikipedia. URL: [http://en.wikipedia.org/w/index.php?title=Coronavirus%](http://en.wikipedia.org/w/index.php?title=Coronavirus%20) [accessed 2021-05-19]
2. Andersen M. Early Evidence on Social Distancing in Response to COVID-19 in the United States. SSRN Journal 2020 Apr 05 [FREE Full text] [doi: [10.2139/ssrn.3569368](https://doi.org/10.2139/ssrn.3569368)]
3. Guglielmo B, Nicola L, Mario M, Mirco T. Expectations, reference points, and compliance with COVID-19 social distancing measures. National Bureau of Economic Research. 2020. URL: https://www.nber.org/system/files/working_papers/w26916/w26916.pdf [accessed 2021-05-19]
4. Usher K, Bhullar N, Jackson D. Life in the pandemic: Social isolation and mental health. J Clin Nurs 2020 Aug;29(15-16):2756-2757. [doi: [10.1111/jocn.15290](https://doi.org/10.1111/jocn.15290)] [Medline: [32250493](https://pubmed.ncbi.nlm.nih.gov/32250493/)]

5. Bavel JJV, Baicker K, Boggio PS, Capraro V, Cichocka A, Cikara M, et al. Using social and behavioural science to support COVID-19 pandemic response. *Nat Hum Behav* 2020 May;4(5):460-471. [doi: [10.1038/s41562-020-0884-z](https://doi.org/10.1038/s41562-020-0884-z)] [Medline: [32355299](https://pubmed.ncbi.nlm.nih.gov/32355299/)]
6. Leonard TC, Richard H, Thaler, Cass R, Sunstein, Nudge: Improving decisions about health, wealth, and happiness. *Const Polit Econ* 2008 Aug 22;19(4):356-360. [doi: [10.1007/s10602-008-9056-2](https://doi.org/10.1007/s10602-008-9056-2)]
7. Jain SH, Powers BW, Hawkins JB, Brownstein JS. The digital phenotype. *Nat Biotechnol* 2015 May;33(5):462-463. [doi: [10.1038/nbt.3223](https://doi.org/10.1038/nbt.3223)] [Medline: [25965751](https://pubmed.ncbi.nlm.nih.gov/25965751/)]
8. Torous J, Kiang MV, Lorme J, Onnela JP. New Tools for New Research in Psychiatry: A Scalable and Customizable Platform to Empower Data Driven Smartphone Research. *JMIR Ment Health* 2016 May 05;3(2):e16 [FREE Full text] [doi: [10.2196/mental.5165](https://doi.org/10.2196/mental.5165)] [Medline: [27150677](https://pubmed.ncbi.nlm.nih.gov/27150677/)]
9. Abdullah S, Matthews M, Murnane EL, Gay G, Choudhury T. Towards circadian computing: "early to bed and early to rise" makes some of us unhealthy and sleep deprived. In: Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing. 2014 Presented at: The 2014 ACM Conference on Ubiquitous Computing; September 2014; Seattle, WA p. 673-684 URL: <https://dl.acm.org/doi/10.1145/2632048.2632100> [doi: [10.1145/2632048.2632100](https://doi.org/10.1145/2632048.2632100)]
10. LiKamWa R, Liu Y, Lane ND, Zhong L. MoodScope: building a mood sensor from smartphone usage patterns. In: Proceeding of the 11th annual international conference on Mobile systems, applications, and services. 2013 Presented at: The 11th Annual International Conference on Mobile Systems, Applications, and Services; June 2013; Taipei p. 465-466 URL: <https://dl.acm.org/doi/10.1145/2462456.2483967> [doi: [10.1145/2462456.2483967](https://doi.org/10.1145/2462456.2483967)]
11. Farhan AA, Yue C, Morillo R, Ware S, Lu J, Bi J, et al. Behavior vs. introspection: refining prediction of clinical depression via smartphone sensing data. 2016 Presented at: 2016 IEEE Wireless Health (WH); October 25-27, 2016; Bethesda, MD. [doi: [10.1109/wh.2016.7764553](https://doi.org/10.1109/wh.2016.7764553)]
12. Boukhechba M, Huang Y, Chow P, Fua K, Teachman BA, Barnes LE. Monitoring social anxiety from mobility and communication patterns. In: Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers. 2017 Presented at: The 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing; September 2017; Maui, HI p. 749-753. [doi: [10.1145/3123024.3125607](https://doi.org/10.1145/3123024.3125607)]
13. Coronavirus (COVID-19). Government of the United Kingdom. URL: <https://www.gov.uk/coronavirus> [accessed 2021-05-19]
14. Xu P, Dredze M, Broniatowski DA. The Twitter Social Mobility Index: Measuring Social Distancing Practices With Geolocated Tweets. *J Med Internet Res* 2020 Dec 03;22(12):e21499 [FREE Full text] [doi: [10.2196/21499](https://doi.org/10.2196/21499)] [Medline: [33048823](https://pubmed.ncbi.nlm.nih.gov/33048823/)]
15. Allcott H, Boxell L, Conway J, Gentzkow M, Thaler M, Yang D. Polarization and public health: Partisan differences in social distancing during the coronavirus pandemic. *J Public Econ* 2020 Nov;191:104254 [FREE Full text] [doi: [10.1016/j.jpubeco.2020.104254](https://doi.org/10.1016/j.jpubeco.2020.104254)] [Medline: [32836504](https://pubmed.ncbi.nlm.nih.gov/32836504/)]
16. Coping with Stress. Centers for Disease Control and Prevention. URL: <https://www.cdc.gov/coronavirus/2019-ncov/daily-life-coping/managing-stress-anxiety.html> [accessed 2021-05-19]
17. Liu S, Jiang Y, Striegel A. Face-to-Face Proximity Estimation Using Bluetooth On Smartphones. *IEEE Trans on Mobile Comput* 2014 Apr;13(4):811-823. [doi: [10.1109/TMC.2013.44](https://doi.org/10.1109/TMC.2013.44)]
18. Liu S, Striegel A. Accurate Extraction of Face-to-Face Proximity Using Smartphones and Bluetooth. 2011 Presented at: 20th International Conference on Computer Communications and Networks (ICCCN); July 31-August 4, 2011; Lahaina, HI. [doi: [10.1109/icccn.2011.6006081](https://doi.org/10.1109/icccn.2011.6006081)]
19. Tesco COVID-19 updates. Tesco. URL: <https://www.tesco.com/help/covid-19/> [accessed 2021-05-19]
20. Ferreira D, Kostakos V, Dey AK. AWARE: Mobile Context Instrumentation Framework. *Front ICT* 2015 Apr 20;2:6. [doi: [10.3389/fict.2015.00006](https://doi.org/10.3389/fict.2015.00006)]
21. Li Q, Zheng Y, Xie X, Chen Y, Liu W, Ma WY. Mining user similarity based on location history. In: Proceedings of the 16th ACM SIGSPATIAL international conference on Advances in geographic information systems. 2008 Presented at: 16th International Symposium on Advances in Geographic Information Systems; November 2008; Irvine, CA p. 1-10. [doi: [10.1145/1463434.1463477](https://doi.org/10.1145/1463434.1463477)]
22. Foursquare. URL: <https://foursquare.com/> [accessed 2021-05-19]
23. COVID-19. The Behavioural Insights Team. URL: <https://www.bi.team/our-work/covid-19/> [accessed 2021-05-19]
24. Burd H, Coleman C. Using behavioural insights to create a Covid-19 text service for the NHS. The Behavioural Insights Team. 2020 Apr 08. URL: <https://www.bi.team/blogs/using-behavioural-insights-to-create-a-covid-19-text-service-for-the-nhs/> [accessed 2021-05-19]
25. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. *Implement Sci* 2011 Apr 23;6:42 [FREE Full text] [doi: [10.1186/1748-5908-6-42](https://doi.org/10.1186/1748-5908-6-42)] [Medline: [21513547](https://pubmed.ncbi.nlm.nih.gov/21513547/)]
26. Culnane C, Leins K, Rubinstein BI. Tracking, tracing, trust: contemplating mitigating the impact of COVID-19 through technological interventions. *Med J Aust* 2020 Jul;213(2):94-94.e1 [FREE Full text] [doi: [10.5694/mja2.50683](https://doi.org/10.5694/mja2.50683)] [Medline: [32583470](https://pubmed.ncbi.nlm.nih.gov/32583470/)]

Abbreviations

BI: Behavioural Insights

GPS: global positioning system

NHS: National Health Service

Edited by G Eysenbach; submitted 12.08.20; peer-reviewed by G Deckard, S Badawy; comments to author 11.09.20; revised version received 04.11.20; accepted 17.03.21; published 27.05.21.

Please cite as:

Ibrahim A, Zhang H, Clinch S, Poliakoff E, Parsia B, Harper S

Digital Phenotypes for Understanding Individuals' Compliance With COVID-19 Policies and Personalized Nudges: Longitudinal Observational Study

JMIR Form Res 2021;5(5):e23461

URL: <https://formative.jmir.org/2021/5/e23461>

doi: [10.2196/23461](https://doi.org/10.2196/23461)

PMID: [33999832](https://pubmed.ncbi.nlm.nih.gov/33999832/)

©Ahmed Ibrahim, Heng Zhang, Sarah Clinch, Ellen Poliakoff, Bijan Parsia, Simon Harper. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 27.05.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Original Paper

Use of eHealth Platforms and Apps to Support Monitoring and Management of Home-Quarantined Patients With COVID-19 in the Province of Trento, Italy: App Development and Implementation

Lorenzo Gios^{1*}, PhD; Giulia Crema Falceri², MD; Stefano Micocci³, MSc; Luigi Patil², MSc; Sara Testa³, MSc; Simona Sforzin², MD; Ettore Turra², MSc; Diego Conforti⁴, MSc; Giulia Malfatti¹, PhD; Monica Moz¹, MD, MBA; Andrea Nicolini¹, MSc; Paolo Guarda⁵, PhD; Alessandro Bacchiega³, MSc; Carlo Mion³, MSc; Michele Marchesoni³, MSc; Rosa Maimone³, MSc; Pietro Benedetto Molini³, MSc; Alberto Zanella³, MSc; Venet Osmani^{3*}, PhD; Oscar Mayora-Ibarra³, PhD; Stefano Forti³, PhD

¹TrentinoSalute4.0 – Competence Center for Digital Health of the Province of Trento, Trento, Italy

²Azienda Provinciale per i Servizi Sanitari (APSS), Trento, Italy

³Fondazione Bruno Kessler, Trento, Italy

⁴Provincia Autonoma di Trento, Trento, Italy

⁵Facoltà di Giurisprudenza, Università degli Studi di Trento, Trento, Italy

*these authors contributed equally

Corresponding Author:

Lorenzo Gios, PhD

TrentinoSalute4.0 – Competence Center for Digital Health of the Province of Trento

Via Sommarive, 18. I-38123 POVO (TN)

Trento

Italy

Phone: 39 3408340665

Email: gios.lorenzo@gmail.com

Abstract

Background: Italy was the first country to largely experience the COVID-19 epidemic among other Western countries during the so-called first wave of the COVID-19 pandemic. Proper management of an increasing number of home-quarantined individuals created a significant challenge for health care authorities and professionals. This was especially true when considering the importance of remote surveillance to detect signs of disease progression and consequently regulate access to hospitals and intensive care units on a priority basis.

Objective: In this paper, we report on an initiative promoted to cope with the first wave of the COVID-19 epidemic in the Spring/Summer of 2020, in the Autonomous Province of Trento, Italy. A purposefully built app named TreCovid19 was designed to provide dedicated health care staff with a ready-to-use tool for remotely monitoring patients with progressive symptoms of COVID-19, who were home-quarantined during the first wave of the epidemic, and to focus on those patients who, based on their self-reported clinical data, required a quick response from health care professionals.

Methods: TreCovid19 was rapidly developed to facilitate the monitoring of a selected number of home-quarantined patients with COVID-19 during the very first epidemic wave. The app was built on top of an existing eHealth platform, already in use by the local health authority to provide home care, with the following functionalities: (1) to securely collect and link demographic and clinical information related to the patients and (2) to provide a two-way communication between a multidisciplinary health care team and home-quarantined patients. The system supported patients to self-assess their condition and update the multidisciplinary team on their health status. The system was used between March and June 2020 in the province of Trento.

Results: A dedicated multidisciplinary group of health care professionals adopted the platform over a period of approximately 3 months (from March-end to June 2020) to monitor a total of 170 patients with confirmed COVID-19 during home quarantine. All patients used the system until the end of the initiative. The TreCovid19 system has provided useful insights of possible viability and impact of a technological–organizational asset to manage a potentially critical workload for the health care staff involved in

the periodic monitoring of a relevant number of quarantined patients, notwithstanding its limitations given the rapid implementation of the whole initiative.

Conclusions: The technological and organizational model adopted in response to the COVID-19 pandemic was developed and finalized in a relatively short period during the initial few weeks of the epidemic. The system successfully supported the health care staff involved in the periodic monitoring of an increasing number of home-quarantined patients and provided valuable data in terms of disease surveillance.

(*JMIR Form Res* 2021;5(5):e25713) doi:[10.2196/25713](https://doi.org/10.2196/25713)

KEYWORDS

telemedicine; telemonitoring; quarantine management; COVID-19; connected care

Introduction

Since the first cases were reported in December 2019, the COVID-19 outbreak caused by the novel coronavirus SARS-CoV-2 has spread rapidly, creating an unprecedented challenge to health care systems worldwide [1].

The COVID-19 pandemic has considerably altered health care systems worldwide and put at risk their sustainability, while boosting the adoption of telemedicine, which has been designed to address the challenges related to patients with COVID-19 [2]. During the pandemic, medical institutions were rapidly facing a massive tsunami of patients requiring hospital treatment, with critical consequences in terms of health care staff workload and dwindling of medical care resources. To prevent the collapse of the global health care system, many countries have advocated for infected patients with mild symptoms to stay at home and self-quarantine [3]. However, it has been observed that the condition of some home-quarantined patients became severe or critical as the disease progressed [4], leading to a delay in timely treatment and hospitalization of these patients and, consequently, rapid deterioration or even death.

Italy was the first country to largely experience the COVID-19 epidemic among other Western countries, and the first country in Europe to impose a general lockdown in March 2020 so as to limit the spread of COVID-19 [5]. From an epidemiological perspective, the COVID-19 epidemic had spread at varying levels across the different regions of Italy with a significant geographic heterogeneity in terms of the number of cases and dynamics of the outbreak [6]. The first cases (n=4) officially reported in the Autonomous Province of Trento (PAT) were identified on March 3, 2020, whereas by the end of March, the number of individuals infected with COVID-19 increased to 2529. In terms of individuals infected per day, the peak of the first wave was reached on March 21, 2020, with 239 newly infected patients, whereas the highest number of daily deaths (n=18) was reached on March 30, 2020. The massive scale-up of infected patients exposed the provincial health care system to an urgent, wide, and rapid organizational and logistic rearrangement during the course of the very first epidemic wave.

In fact, within 4 weeks of the very first cases reported, an exponentially increasing number of patients needed monitoring, active support, and prompt hospitalization, and a significant proportion of patients required intensive care. The peak of the burden for health care services was reached on April 8, 2020, with 311 patients hospitalized in an infectious diseases

department, 77 patients managed in intensive care units, and 43 patients managed in high-intensity units. Only 3 months from the very onset of the pandemic, June 2, 2020, was the first day with 0 COVID-19 cases reported in the province of Trento, whereas the last infected patient was discharged from the intensive care unit on June 12, 2020.

The COVID-19 pandemic had suddenly exposed deficiencies in the whole health care system, revealing high uncertainty on key issues such as increasing difficulty in tracing the exact transmission route of the virus, inadequacy of the diagnostic testing system, and the lack of clear monitoring procedures in the case of home-quarantined patients and therapeutic approaches [7]. Ensuring proper management of an increasing number of home-quarantined individuals posed a demanding challenge to the health care authorities and professionals, considering the key role of strict monitoring in order to detect disease aggravation in view of prompt hospitalization and regulating access to hospitals and intensive care units only when needed.

Within this critical and rapidly changing scenario, there was a scattered phenomenon of swift and spontaneous—albeit valuable—attempts to adopt digital tools to support the health care system in dealing with this unexpected epidemic, particularly in the field of remote monitoring. Although some forms of remote monitoring of home-quarantined patients have been implemented in other countries during the very first stage of the pandemic [8,9], no such initiatives were undertaken in Italy. At the same time, the COVID-19 pandemic has been considered as a missed opportunity to improve telemedicine [6,10], which is still a scattered and embryonic phenomenon at the national level.

In this paper, we report on the initiative designed and implemented to cope with the first wave of the COVID-19 epidemic in Spring/Summer 2020 in the province of Trento, Italy. In the very first weeks after the outbreak, a telemedicine tool was purposefully developed to provide home-quarantined patients with COVID-19 with an app named TreCovid19. The app was linked to an already existing telemedicine system that was in use by nurses who provided home care. The app and the platform were set up within an extremely short period to enable an automated monitoring system supporting health care staff when dealing with an ever-increasing number of infected patients. This was achieved by merging organizational and technological components, that is, by embedding a telehealth service and related activities into the framework of health care

procedures put in place to face the epidemic's impact, particularly in terms of monitoring of home-quarantined patients.

The TreCovid19 tool was developed within an initiative promoted and coordinated by the Competence Centre on Digital Health of PAT, TrentinoSalute4.0 (TS4.0) [11].

Methods

Contextual Factors: Collaboration Among Health Care Bodies in Trentino

A key component of the initiative described in the present case study lies in a specific contextual factor characterizing the collaboration between health care bodies and policy and research stakeholders in the province of Trento. TS4.0 was formally established in 2016 with an Act of the local government as a partnership among three relevant stakeholders in Trentino: the Department of Health and Social Policies of PAT, the local Healthcare Trust of the Autonomous Province of Trento (Azienda Provinciale per i Servizi Sanitari [APSS]), and the Bruno Kessler Foundation (FBK)—a research entity with particular focus on the applicative dimension of technology in the field of digital health. This alliance has been established for strengthening cooperation among the three institutions and coordinating the eHealth agenda in Trentino; in May 2020, TS4.0 officially became a Joint Research Unit. Specific financial funding has been allocated by the province of Trento to support TS4.0 coordination activities, whereas a specific TS4.0 Steering Committee has been formed comprising representatives of the partner institutions (ie, PAT, APSS, and FBK). The Steering Committee is in charge of defining and approving the overall strategy of TS4.0 and prioritizing areas of action, as well as monitoring a smooth implementation of related activities. In recent years, this pre-existing framework of collaboration has provided a solid basis to a number of initiatives in the field of digital health (from piloting to the delivery of eHealth services), including telemedicine projects [12].

Organizational Response

In view of the first epidemic wave in the province of Trento, health care authorities set up a COVID-19 Special Unit, established to deliver a general strategy to cope with the pandemic and to ensure proper management of positive cases, from monitoring to hospitalization and treatment. This unit was strongly linked with PAT (in particular, with the Department of Health and Social Policies), as well as with the Directorate General of APSS, in order to ensure proper decision-making. The COVID-19 Special Unit was also linked with ad hoc contact points within different public health institutions at the local level (eg, hospitals and local districts). Patients with a probable or confirmed COVID-19 infection were reported to this unit through different channels, namely, the provincial registry of citizens with positive swabs, the prevention department, and via general practitioners (GPs).

The monitoring of home-quarantined patients with COVID-19 was considered a key component in the management of the epidemic, particularly in the first phase of the outbreak. This was despite the fact that it represented a challenging action considering the relative uncertainty about clinical manifestations

and related indicators in the early stage of the epidemic [7]. Therefore, within the COVID-19 Special Unit, a selected group of health care professionals were put in charge of monitoring patients at the provincial level, namely, 2 medical doctors and 2 nurse coordinators were in charge of managing and coordinating the monitoring activities and 13 nurses, 2 medical doctors (specialists), and 1 medical doctor from the Special Continuity Care Unit (so-called *Unità Speciali di Continuità Assistenziale* [USCA]) were in charge of performing the actual monitoring of COVID-19-positive cases. A total of nearly 80 health care professionals were also involved in the monitoring phase at the community level. The monitoring of home-quarantined patients was initially set up using the following procedure: medical doctors and nurses from the COVID-19 Special Unit phoned the home-quarantined patients twice a day (morning and afternoon) to collect information about their clinical status and progression of disease symptoms (if any). Data collected included the patients' self-reported body temperature, perceived pain, level of fatigue, dyspnoea, level of consciousness, and presence of deep vein thrombosis. This continuous and remote monitoring performed through periodical phone calls was set up to detect the progression of disease and to support direct referral to the GP for clinical assessment and access to the emergency room for evaluation with consequent hospitalization, if necessary.

Within a few weeks since the onset of the COVID-19 outbreak in the province, the rapidly increasing number of home-quarantined patients who needed monitoring created a demanding challenge among the health care authorities and professionals.

Selecting a Priority During the Very Start of the COVID-19 Outbreak

Since the onset of the COVID-19 outbreak in the province of Trento, TS4.0—as the leading player in the field of digital health within the province—was immediately given the mandate to facilitate the process of designing and delivering eHealth services to support a swift reaction to the pandemic. No specific extra-financial funding was acquired for this task to be implemented during the emergency period, as this has been considered part of the standard coordination activities of TS4.0.

Within this context, the TS4.0 Steering Committee identified the urgent need for core actions such as the rapid implementation of a digital solution to support health care staff in charge of monitoring patients with COVID-19, particularly those who are home-quarantined. This decision was the result of both internal consultations and a prompt negotiation with the COVID-19 Special Unit (the overview of the organizational asset is provided in [Multimedia Appendix 1](#)). The reasons behind this decision were both organizational and technological. From an organizational viewpoint, it was rapidly clear that the number of home-quarantined citizens was steadily increasing in the very first weeks since the onset of the epidemic and health care staff in charge of monitoring these patients were extremely exposed to overwork. Furthermore, because clinical management of the patients in the hospital was clearly a sole responsibility of the health care staff, the periodical collection of information on home-quarantined patients could be (partially) assigned to the

patients themselves, by and with reliable and ready-to-use self-care or self-monitoring solutions. From a technological viewpoint, it was considered that specific apps for telemonitoring and self-care of patients had already been piloted and used in the context of previous projects, and a technological tool in use for the management of integrated home care in Trentino was already available as part of the standard care (the so-called “@home platform”).

In terms of technological assets, a multidisciplinary approach was also adopted. A working group comprising medical doctors, nurses, information technology (IT) specialists, technologists, researchers, and project managers from the PAT, APSS, and FBK was set up under the coordination of TS4.0, with the aim of collecting clinical requirements and designing, testing, and implementing technological tools.

As previously mentioned, the project was based on an already existing partnership among the Department of Health and Social Policies of PAT, APSS, and FBK. This previous collaboration represented the ideal basis to quickly develop and deliver a ready-to-use tool for supporting health care staff in charge of monitoring patients. Clinical colleagues from APSS, including members of the COVID-19 Special Unit, were responsible for setting up the clinical assumptions and criteria behind the technological asset, whereas IT professionals from APSS and FBK were in charge of developing the app and linking the system with the already-existing telemedicine system, which was in use by the nurses who provided home care. PAT colleagues provided guidance and inputs in line with the provincial strategy implemented during the COVID-19 emergency period, particularly the strategy related to patient management.

Development of an App for Monitoring Home-Quarantined Patients With COVID-19

Overview

The aim of rapidly incorporating a technological and organizational approach toward COVID-19 monitoring in

Trentino was motivated by the need to support the health care staff in dealing with the sudden increase in the workload presented during the first stage of disease transmission.

To reach this objective, the TreCovid19 system was specifically set up to (1) regularly and automatically collect self-reported symptoms from home-quarantined patients with COVID-19 through a smartphone app; (2) translate subjective self-reports of these symptoms into numerical scales; and (3) allow a set of alerts based on specific cut-offs, periodically informing health care staff about the status of the patients and optimizing interventions and direct contacts if required.

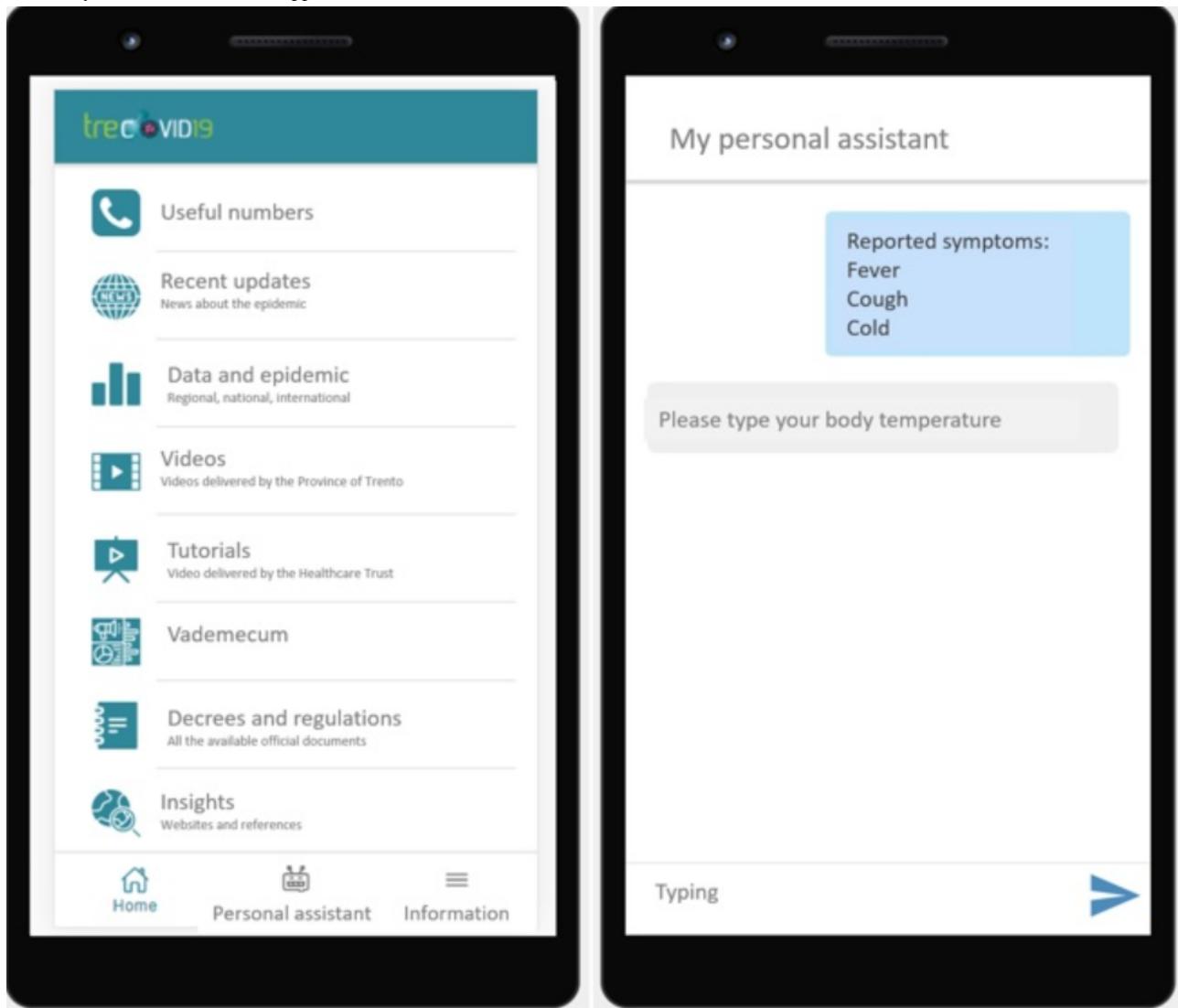
The mobile app was embedded into an already existing telemedicine platform, adopted by home care practitioners. The core approach adopted for this endeavor was to merge organizational and technological components by embedding a telehealth service (mainly supported by a dedicated app) into the already-existing framework of clinical procedures for monitoring of home-quarantined patients.

The system was rapidly developed by designing and delivering two components: (1) an app for patients to support daily self-collection of symptoms and (2) a dashboard for medical doctors and nurses in charge of monitoring patients.

App to Support Patients' Daily Self-Collection of Symptoms

The TreCovid19 app homepage has a number of functionalities that are divided into two main types (see [Figure 1](#)). The first category is related to providing official information about the COVID-19 pandemic and general information to cope with it, such as recent updates about the epidemic data, video tutorials delivered by APSS, tips and advice on specific safety procedures, information on regional and national decrees, and regulations related to the COVID-19 pandemic.

Figure 1. Layout of the TreCovid19 app and related functionalities.



The second category constituted the core tool for patients and is represented by an automated chatbot functionality. The app would periodically activate a specific chat with the patient, administering a set of items to gather self-reported data on their health status and related symptoms. The different pieces of information were collected twice a day (once in the morning and once in the afternoon) and communicated in real time to the central system. In case the level of self-reported symptoms exceeded the cut-offs set by the health care staff (see [Table 1](#)), a specific alarm was sent via email to the COVID-19 Special Unit for immediate (re)action.

Dashboard for Health Care Staff

The dashboard for the health care staff was based on the so-called @home system [13,14]. Such a system is constructed on top of the APSS technological tool currently in use for the management of integrated home care in Trentino. The @home application was successfully adopted to improve efficiency in managing the reporting process of all interventions performed by home care case managers. The adoption of the system reduced the workload for reports management, as well as reduced paperwork for nurses.

The dashboard for the health care staff was developed using a customer relationship management commercial tool, whereas nurses could use a dedicated tool during home visits. The dashboard uses a specific Platform as a Service (PaaS), which makes it possible to operate on isolated virtual environments.

The integration currently available in the @home platform allows a proper link with consent and privacy documentations available within the official document repository, in addition to a secure link and integration with the health register, the register of operators or health care staff, and the APSS notification system.

Setting up Clinical Requirements

To set up the COVID-19 quarantine monitoring tool (mobile app) for patients, a rapid review of the available scientific literature, technical documents, and reports was performed by the clinicians collaborating within the working group. A number of assessment scales were also reviewed by medical experts from the COVID-19 team to identify potential items to be included in the monitoring tool, such as standard validated scales [15]. A multidisciplinary approach was adopted, by setting up a working group comprising medical doctors, nurses, IT experts, and project managers, to set up an automated monitoring tool

to be used in the very first wave of the epidemic. The working group was constantly in contact with and reporting to the steering committee, to ensure an effective decision-making process. It should be underlined that the unexpectedness of the first COVID-19 outbreak and related organizational complications were detrimental to both the health care staff and IT staff in developing the system in a very short period. Other major challenges rose due to the relative uncertainty regarding clinical manifestations and related indicators of COVID-19 in the first few months of the epidemic outbreak, rendering it difficult to construct a stable and reliable list of indicators to be translated into the app functionalities.

Therefore, a pragmatic decision process was adopted by the core group of medical doctors in charge of managing the provincial COVID-19 task force. After a round of internal meetings, a core set of key indicators and related cut-offs were identified, as well as core IT functionalities. The proposal was negotiated and agreed with the Steering Committee group. In

addition, a shared decision was taken to group the participants into two categories according to their clinical status, as follows:

- Red group (Acv19): patients considered to be COVID-19 positive, determined based on one of the following criteria: reporting clinically relevant symptoms; a positive swab test result; or relevant clinical parameters based on radiological examination.
- Blue group (AIOcv19): cohabitants or family members living with a patient who has tested positive for COVID-19. The blue group was specifically initiated to monitor developments in the conditions of the cohabitants, specifically to signal their potential infection by the virus.

Table 1 details the key pieces of information that have been identified by the team of medical doctors and nurses in charge of monitoring the disease progression (the detailed questionnaire is provided in [Multimedia Appendix 2](#)). An automated alarm system was designed considering the shown variables and their cut-off values.

Table 1. Participants' grouping, variables, and cut-offs adopted for the TreCovid19 app.

Variables	Cut-off score or criterion	
	Red group (Acv19)	Blue group (AIOcv19)
Body temperature (°C)	≥39	≥37.5
Pain (Numerical Rating Scale)	≥4	≥4
Fatigue ^a	≥7	≥7
Peripheral oxygen saturation (SpO ₂ , %) ^b	≤95	N/A ^c
Dyspnoea	≥4	≥3
Level of consciousness	Confused or coma	Confused or coma
Deep vein thrombosis	Yes	Yes
Respiratory rate (per min)	≥22	≥22
Assumption of antipyretics	N/A	Yes

^aFatigue was measured based on a self-reported assessment scale ranging from 0 (not at all) to 10 (very much).

^bSpO₂ was considered only when oximeter was available.

^cN/A: not applicable.

Selecting the Patients

Considering the emergency context in which the telemonitoring system has been set up and delivered, researchers, health care staff, and managers decided to adopt a conservative approach. In line with a precautionary principle, only patients with relatively stable medical conditions were contacted and provided access to the app, thereby restricting the use of the app to a limited number of eligible patients. This was decided to (1) allow proper adoption of the system and related procedures in view of a potential scale-up of the initiative and (2) guarantee a controllable margin of safety for patients in this first phase of emergency.

Therefore, patients were included based on the following inclusion criteria: diagnosed with COVID-19 and home-quarantined, residents of the province of Trento, reporting relatively stable medical conditions, ability to use a smartphone or living with a cohabitant with a smartphone, and voluntary

participation. Medical doctors and nurses of the COVID-19 Special Unit were in charge of the assessment of the medical conditions, based on the experience they gained through managing the very first group of patients monitored until that time. Exclusion criteria were as follows: reporting severe medical conditions, belonging to vulnerable populations (eg, those with complex general and/or chronic health conditions), and reporting specific social and/or psychological needs (eg, anxiety management).

It should be underlined that, mainly because of emergency issues, neither the exact number of potential eligible patients nor precise clinical patterns were collected and registered. Only a limited number of patients were eligible mainly because of (1) the unexpectedness of the COVID-19 epidemic, (2) the exponential increase of the number of infected patients, and (3) the relative uncertainty about the clinical manifestations and related indicators of the infection. In addition, the app had been developed within a period of two weeks, to ensure the

availability of a monitoring tool to support health care staff in the shortest possible time; therefore, the app was adopted only for a convenient sample of patients. The decision of assigning the app to a patient was based on the clinical assessment performed by the health care staff in charge of monitoring the patients, following the available guidelines and experiences gained through the very first epidemic wave. Expanding the patient sample and extending the service to patients who were nonresidents of Trentino was not considered feasible.

Setting up the Procedural Flow

The procedure of the telemedicine system is presented below. Patients were contacted via phone by the COVID-19 Special Unit—the team in charge of monitoring COVID-19 cases. Members of the team assessed potential eligibility of patients via phone per the abovementioned criteria (see section “Selecting the Patients”).

After the eligibility for being enrolled in the project was confirmed, the participants were grouped into the two abovementioned categories, namely, Red group (Acv19, comprising patients considered to be COVID-19 positive) and Blue group (AIOcv19, comprising cohabitants or family members living with a patient with confirmed COVID-19). COVID-19-positive patients and their cohabitants were invited to access the mobile app TreCovid19. Activation of the app required patients to enter their health insurance code and fiscal code, matching with the health care platform. The app was then linked with the specific clinical profile of the patient.

Nurses and medical doctors were in charge of training and assisting patients in the process of downloading and activating the app, providing telephone support in case of issues related to installation and use. The participants underwent remote quarantine management monitoring, and the app was used to automatically prompt patients from both groups to fill in the requested information twice a day (morning and afternoon). Once an alert was received by the health care staff, the participant was contacted via phone by a trained operator (either a doctor or a nurse), to remotely assess the health status and the progression of the disease.

Two potential outcomes of the assessment were set up:

- Typing error: the alarm was triggered by incorrect data, entered by mistake. In this case, a manual correction of the parameter was performed to ensure accurate tracking and recording of data. Typing errors accounted for a very small part of the generated alarms.
- No typing error: the alarm was triggered by data correctly entered by the participants.

In the latter case, the health care staff performed a telephone-based in-depth assessment of patients' general conditions and previous parameters, resulting in one of the following scenarios: (1) continuous use of the monitoring system based on the app; (2) intensifying remote monitoring by adding periodical phone calls to the app usage; (3) direct referral to the GP for clinical assessment, in view of the need to incorporate additional interventions (eg, pulse oximeter delivery at home, home-visit scheduling, or access to the emergency room for evaluation with consequent hospitalization if necessary). In general, and based on the telephonic in-depth assessment, health care staff considered that the validity of the self-reported data was appropriate, with few exceptions related to the potential overestimation of some symptoms.

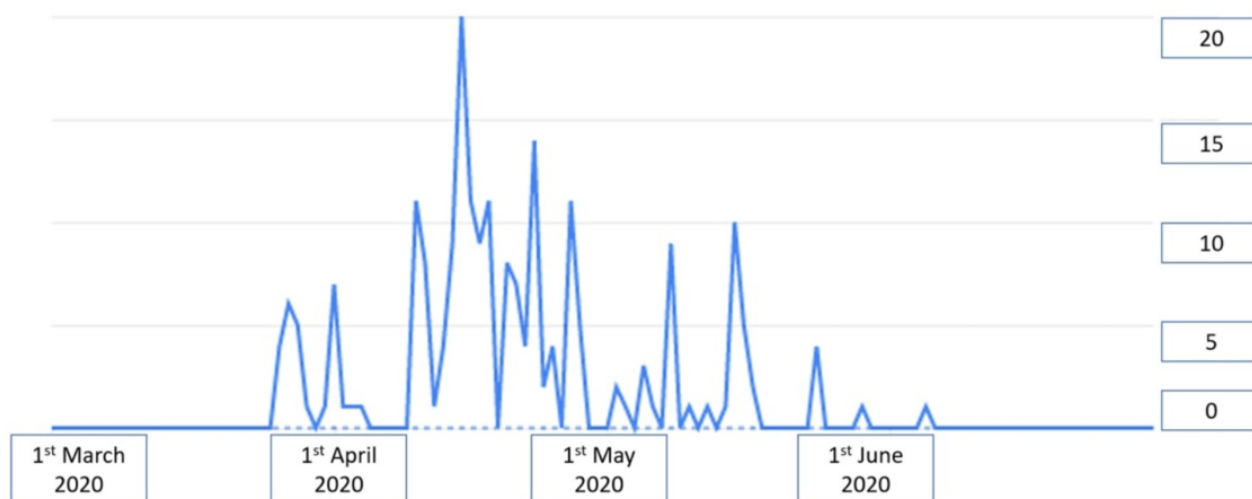
Ethical Issues

Dedicated information sheets and informed consent were already available for the telemedicine system in use. A specific information sheet was developed for patients when downloading and activating the TreCovid19 app.

To further improve safety and reliability of the system, a set of piloting phases were performed to ensure secure transmission of data and proper functioning of the alerting system and related cut-offs. Specific data to allow proper insights about the validity of self-reported data were not collected (eg, comparison between self-collected versus health care worker-collected data). Nevertheless, based on the information gathered through the telephone-based in-depth assessments, the validity of the data collected through the app was considered acceptable. Constant supervision of the system was ensured during the project piloting phase.

Results

The multidisciplinary group developed the technological platform for patient monitoring in time for its wide use during the very first months of the epidemic in the province of Trento. It should be highlighted that the first cases officially reported in the province ($n=4$) were identified on March 3, 2020, whereas at the end of March, the number of COVID-19 cases increased to 2529. The TreCovid19 platform was developed and made available in an extremely short period (approximately 2 weeks), which was possible owing to the existing infrastructure on which the new service had been constructed. This swift reaction enabled the monitoring of 170 home-quarantined patients with COVID-19 from March-end to June 2020 (see [Figure 2](#)). The large majority of patients monitored through the telemedicine system were enrolled in April.

Figure 2. Timeline of the piloting and enrolment phases of the project.

Of the 170 patients specifically selected for this piloting phase, 107 (62.9%) were assigned to the red group (Acv19), that is, COVID-19-positive patients in quarantine; the remaining 63 (37.1%) were assigned to the blue group (AIOcv19), that is, cohabitants or family members living with a COVID-19-positive patient. Of the enrolled patients, half were female (85/170, 50%) with a mean age of 37.83 (SD 15.54 years). The red group sample comprised 52 (48.6%) female and 55 (51.4%) male patients (age: mean 38.95, SD 14.98 years), whereas the blue group comprised 33 (52.4%) female and 30 (47.6%) male participants (age: mean 35.92, SD 16.26 years). With regard to the red group (Acv19), 2570 monitoring measurements were collected by the app (24 measurements per user on average), whereas 1057 sets of measurements were collected by the app (17 measurements per user on average) in the blue group (AIOcv19).

On the basis of the available data, none of the patients neither deteriorated nor needed prompt hospitalization. Once recovered, the patients were simply asked to uninstall the app, which was then unlinked with the @home system as well.

Discussion

Principal Findings

To our knowledge, the TreCovid19 app represents a unique example of swift design and delivery of a technological innovation supporting health care staff dealing with the monitoring of home-quarantined patients during the first COVID-19 wave in Italy and in other European countries as well. The uniqueness of this experience lies in the fact that (1) the app can be considered as one of the very first telemonitoring tools launched during the first COVID-19 wave [8,9] and (2) it was launched in Italy, one of the first Western countries to be significantly hit by the COVID-19 pandemic [6]. Despite the critical contextual situation, this project has successfully

achieved its goals owing to the two key strengths discussed below.

The first commendable strong point is in the presence of a pre-existing joint center for digital health at the province level. In fact, the TreCovid19 initiative was promoted and coordinated by the Competence Centre on Digital Health (TS4.0), a strategic alliance among the three core health stakeholders in the province, namely, PAT (Department of Health and Social Policies), APSS, and FBK. Despite the unexpectedness of the first COVID-19 outbreak and associated organizational complications, this collaboration represented a pivotal asset to promote swift development of the telemonitoring system within a considerably short period, promoting a prompt convergence of organizational, clinical, and technological competences within different institutions.

The second strong point is the integration of a specific telemedicine system (TreCovid19) within the health care platform already in use by thousands of citizens and an organizational asset, providing an immediately available tool for piloting a novel telemonitoring system. In fact, the pre-existing digital health infrastructures already in use for patients and the health care staff in the province of Trento [12,13,16] allowed an efficient development and delivery of a digital tool to tackle the epidemic, while the health care provincial institutions were under enormous pressure.

Limitations

The core limitation of this initiative is related to the swift and unstable scenario in which such a telemedicine system has been developed and adopted. First, the unexpectedness of the first COVID-19 outbreak and organizational complications associated with it have exposed the health care and IT staff to a challenging scenario while developing the system within a very limited timespan.

Second, the relative uncertainty about the clinical manifestations and related indicators of COVID-19 made it difficult to construct a stable and reliable list of indicators and triggers to be translated into the app functionalities, at least in the initial months of the epidemic. Furthermore, no specific clinical, automated, or laboratory-based indicators were considered to triangulate the different pieces of self-reported information collected in the preliminary phase of this project. Because of the emergency situation and the pressing need for implementing an immediate action, the project team used a rough validation strategy based only on the completion of the requirements from the clinical team over a thorough trial analysis. For the same reason, the collection of specific data on this process was not considered critical in this phase—again, because the urgency of having a platform immediately usable for the emergency purpose was the core priority considering the increasing number of patients to be monitored. In addition, clinical evidence, as well as well-structured guidelines about specific cut-offs, were not always available when the system was designed.

This explains why researchers, health care staff, and health managers decided to adopt a conservative approach in line with a precautionary principle, and to enroll in the project a limited and convenient sample of patients with relatively stable medical conditions. This was decided to (1) allow proper evaluation of the system and related procedures and (2) guarantee a controllable margin of safety for the patients in this first phase. As a result, it was not possible to obtain a larger sample size and, therefore, to test the app and its functionalities among a larger audience, providing further validity and robustness of these findings.

The system has not been promoted as standard procedure nor scaled considering several organizational, technological, and contextual factors of emergency during the very first wave of the COVID-19 epidemic. At the same time, proper qualitative and quantitative assessments are foreseen to explore the organizational and contextual factors (eg, digital literacy, internal procedures, clinical requirements) that can potentially contribute to or hamper larger adoption of the system. Likewise, an improved co-design of the app and an update of the clinical information and related cut-offs could further improve the usability of the system. In fact, a larger implementation of a reliable telemedicine system to support patients monitoring could be particularly important during the fall/winter season of 2020 and the first part of 2021, when a relevant increase in the number of home-quarantined patients is more likely to occur.

Lessons Learned: Do's and Don'ts

When health care institutions are facing a *health tsunami* of sorts as the one we all experienced during the first wave of the COVID-19 pandemic, there is a clear need for a rapid reaction and swift delivery of viable procedures and tools to tackle that calamity. On the basis of this experience, we have identified some core do's and don'ts learned through this initiative.

The first issue is related to the swift identification of the key priority to be addressed, among the different urgent issues to cope with. In our case, monitoring of home-quarantined patients with COVID-19 has been immediately considered a key component in the management of the epidemic, particularly in

the first phase of the outbreak, also with a view of preventing overflowing of patients into the intensive care wards. Even if this initiative represented a challenging action considering the relative uncertainty about the clinical manifestations and related indicators in the early stage of the epidemic, developing a dedicated telemedicine system seemed to be a vital action to support the health care staff in delivering efficient health care services despite the huge increase of infected patients. This prompt identification of a list of priorities to be addressed (in this case, telemonitoring) could occur only if teamwork and well-organized collaboration among key stakeholders are already in place.

The second issue is related to the ability to select pre-existing infrastructures, adapt them in light of new contextual factors, and deliver the service within a reasonable timeframe. The provision of a telemonitoring system was essential if it was possible in the shortest timeframe possible, considering the rapid increase of the COVID-19 epidemic that was putting the health care provincial institutions under enormous pressure. Developing a new tool for monitoring COVID-19-positive patients from scratch would have resulted in an impossible and detrimental task.

In other words, the result achieved in the province of Trento could be linked to specific contextual dynamics, that we consider to be even more essential factors in the framework of sudden outbreaks and public health calamities; these include the availability of a pre-established collaboration, pre-existing technological infrastructures, and a multidisciplinary approach.

The first essential factor was the availability of a former collaboration in place among the key health organizations within the province (ie, the partnership among APSS, APT, and FBK). This was likely the key factor of success of this very initiative. In the context of a pressing pandemic, such prevailing teamwork and the availability of pre-existing tools and technological infrastructures have represented a solid foundation for coordinating a complex task such as the swift delivery of an ad hoc telemonitoring system to tackle a relatively unknown epidemic. The TS4.0 alliance has also enabled the rapid adaptation of already available eHealth platforms that have been speedily converted into a ready-to-use tool to tackle the sudden COVID-19 outbreak.

An additional factor of success was the multidisciplinary approach adopted. The harmonization between technological assets and organizational procedures, as well as putting together the clinical know-how, public health expertise, and IT knowledge, was possible only because of a multifaceted and integrative working method. The fact that medical doctors, nurses, project managers, and IT technicians were already exposed to long-term cooperation in the field of digital health resulted in a prompt and smooth cooperation within a critical context due to the COVID-19 pandemic.

Finally, we would like to further emphasize the importance of the abovementioned factors in view of potential, future scenarios similar to the COVID-19 outbreak. First, this experience has underlined the clear need for establishing solid collaboration among key health, policy, and research organizations at the local level, ideally by establishing specific joint centers for

digital health. If already present, these alliances must be strengthened considering their potential pivotal role in case of tsunamis in the field of public health. In fact, this approach could guarantee the availability of an organizational asset to enable a prompt and swift reaction to emergencies. Second, promoting multidisciplinary collaboration and mutual inspiration among IT experts, public health managers, and health care staff could represent a vital long-term investment to ensure smooth convergence of different stakeholders in emergency circumstances. Lastly, the design of technological and digital health infrastructures with high levels of flexibility and adaptability could also be a strategy to pursue, to make these infrastructures flexible in case of emergency.

Lessons Learned: Evaluation and Assessment

We discuss here a number of lessons learned through this initiative in terms of evaluation and assessment, albeit time constraints led to the lack of a robust methodological asset for evaluating and testing the system.

First, there is a clear need to develop an approach for assessing and validating technological tools developed in a strict period of time for emergency purposes, such as the one described in the present paper. Evaluating to what extent the harmonization between technological assets and organizational procedures is put in place should be a key action, even when a rapid reaction to the emergency leads to a swift adaptation and implementation of available technologies to address the containment of the health care crisis.

Second, in ideal conditions, the adoption of robust study designs should be considered, even within a context of emergency such as the COVID-19 outbreak. Setting up a clear study design (eg, case-control design) within this first phase of the project could have further improved the interpretation of outcome and results of this initiative. For instance, the idea of including a control group was discussed, but the emergency prevented the project team to design a suitable research plan, as the core objective was to deliver a ready-to-use system to support the health care staff in the very early phase of the epidemic outbreak. More advanced strategies for implementing robust study designs for

reacting to emergencies should be considered in future crisis scenarios.

Third, the opportunity of a broader psychological-sociological assessment of the experiences of family members and patients monitored at home, as well as the experience of the health care staff adopting the technological tool, could have represented an added value in evaluating the initiative. Moreover, a cost-effectiveness analysis of the entire initiative would be a critical, albeit complex, task to further assess the viability and sustainability of the technological-organizational asset.

Conclusions

To conclude, we would like to highlight that the rapid onset of the COVID-19 pandemic urgently called upon the need for swift changes in health care provisioning, as well as the need for a rapid “design-to-deliver” approach such that specific solutions that can support the management of a rapidly increasing number of patients can be made immediately available to health care staff. Within this scenario, monitoring of home-quarantined patients with COVID-19 was a key component, as it constituted a challenging but important action to deliver efficient health care services and to control patients’ status and related hospitalization levels.

This paper describes how we managed to develop and deliver, in short time, an eHealth tool to assist health care staff in coping with an inflow of home-quarantined patients with COVID-19, even in the period of a severe and unpredictable outbreak.

The TreCovid19 system has provided useful insights of possible viability and impact of a technological-organizational asset to manage a potentially critical workload for health care staff involved in the periodic monitoring of a relevant number of quarantined patients, notwithstanding its limitations due to the rapid implementation of the whole initiative. TreCovid19 presents high potential to further support the local health care system when facing higher peaks of the COVID-19 epidemic or future health care emergencies. With this perspective, further optimization of the system, its potential extension to larger groups of home-quarantined patients in the Italian province of Trento, and a robust validation assessment of the entire model might further increase its applicability.

Acknowledgments

The authors would like to thank all TreCovid19 app users, including patients and health care staff, as well as all the colleagues from different institutions and departments contributing to the TreCovid19 project in such a difficult period as the one we all faced during the first COVID-19 outbreak.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Organizational diagram depicting the structural asset and relations among the different institutions involved in the initiative. [\[PNG File, 218 KB - formative_v5i5e25713_app1.png\]](#)

Multimedia Appendix 2

Detailed questionnaire (Italian version with English translation) adopted for the TreCovid-19 app.

[DOCX File , 20 KB - [formative_v5i5e25713_app2.docx](#)]

References

1. Olsen SJ, Chen MY, Liu YL, Witschi M, Ardoin A, Calba C, European COVID-19 Work Group. Early introduction of severe acute respiratory syndrome coronavirus 2 into Europe. *Emerg Infect Dis* 2020 Jul;26(7):1567-1570 [FREE Full text] [doi: [10.3201/eid2607.200359](#)] [Medline: [32197059](#)]
2. Hollander JE, Carr BG. Virtually perfect? Telemedicine for Covid-19. *N Engl J Med* 2020 Apr 30;382(18):1679-1681. [doi: [10.1056/nejmp2003539](#)]
3. Home care for patients with COVID-19 presenting with mild symptoms and management of their contacts. World Health Organization. URL: [https://www.who.int/publications/i/item/home-care-for-patients-with-suspected-novel-coronavirus-\(ncov\)-infection-presenting-with-mild-symptoms-and-management-of-contacts](https://www.who.int/publications/i/item/home-care-for-patients-with-suspected-novel-coronavirus-(ncov)-infection-presenting-with-mild-symptoms-and-management-of-contacts) [accessed 2021-05-07]
4. WHO Coronavirus Disease (COVID-19) Dashboard. World Health Organization. URL: <https://covid19.who.int/> [accessed 2021-05-07]
5. Guzzetta G, Riccardo F, Marziano V, Poletti P, Trentini F, Bella A, COVID-19 Working Group, et al. Impact of a nationwide lockdown on SARS-CoV-2 transmissibility, Italy. *Emerg Infect Dis* 2021 Jan;27(1):267-270 [FREE Full text] [doi: [10.3201/eid2701.202114](#)] [Medline: [33080168](#)]
6. Remuzzi A, Remuzzi G. COVID-19 and Italy: what next? *Lancet* 2020 Apr 11;395(10231):1225-1228 [FREE Full text] [doi: [10.1016/S0140-6736\(20\)30627-9](#)] [Medline: [32178769](#)]
7. Acuti Martellucci C, Flacco ME, Cappadona R, Bravi F, Mantovani L, Manzoli L. SARS-CoV-2 pandemic: An overview. *Adv Biol Regul* 2020 Aug;77:100736 [FREE Full text] [doi: [10.1016/j.jbior.2020.100736](#)] [Medline: [32773099](#)]
8. Xu H, Huang S, Qiu C, Liu S, Deng J, Jiao B, et al. Monitoring and management of home-quarantined patients with COVID-19 using a WeChat-based telemedicine system: Retrospective cohort study. *J Med Internet Res* 2020 Jul 02;22(7):e19514 [FREE Full text] [doi: [10.2196/19514](#)] [Medline: [32568727](#)]
9. Timmers T, Janssen L, Stohr J, Murk JL, Berrevoets MAH. Using eHealth to support COVID-19 education, self-assessment, and symptom monitoring in the Netherlands: Observational study. *JMIR Mhealth Uhealth* 2020 Jun 23;8(6):e19822 [FREE Full text] [doi: [10.2196/19822](#)] [Medline: [32516750](#)]
10. Omboni S. Telemedicine during the COVID-19 in Italy: A missed opportunity? *Telemed J E Health* 2020 Aug;26(8):973-975 [FREE Full text] [doi: [10.1089/tmj.2020.0106](#)] [Medline: [32324109](#)]
11. Mayora Ibarra O, Forti S, Conforti D, Tessari P, Testa S. Trentino salute 4.0 - the creation of a competence center on digital health integrating policy, healthcare trust and research in Trentino territory. *Int J Integr Care* 2019 Aug 08;19(4):56 [FREE Full text] [doi: [10.5334/ijic.s3056](#)]
12. Eccher C, Gios L, Zanutto A, Bizzarri G, Conforti D, Forti S. TreC platform. An integrated and evolving care model for patients' empowerment and data repository. *J Biomed Inform* 2020 Feb;102:103359 [FREE Full text] [doi: [10.1016/j.jbi.2019.103359](#)] [Medline: [31917253](#)]
13. Fontanari A, Zambiasi P, Turra E, Guarrera GM. Come valutare le app di sanità digitale: il caso della soluzione @home per l'Assistenza Domiciliare. Article in Italian. *eHealth* 2019;70:39-42 [FREE Full text]
14. Manchar A, Chouhan A. Salesforce CRM: A new way of managing customer relationship in cloud environment. : IEEE; 2017 Nov Presented at: Second International Conference on Electrical, Computer and Communication Technologies (ICECCT); February 22-24, 2017; Coimbatore, India p. 1-4. [doi: [10.1109/ICECCT.2017.8117887](#)]
15. Karcioğlu O, Topacoglu H, Dikme O, Dikme O. A systematic review of the pain scales in adults: Which to use? *Am J Emerg Med* 2018 Apr;36(4):707-714. [doi: [10.1016/j.ajem.2018.01.008](#)] [Medline: [29321111](#)]
16. Osmani V, Forti S, Mayora O, Conforti D. Challenges and opportunities in evolving TreC personal health record platform. New York: Association for Computing Machinery; 2017 May Presented at: PervasiveHealth '17: 11th EAI International Conference on Pervasive Computing Technologies for Healthcare; May 2017; Barcelona, Spain p. 288-291. [doi: [10.1145/3154862.3154910](#)]

Abbreviations

APSS: Azienda Provinciale per i Servizi Sanitari
FBK: Bruno Kessler Foundation
GP: general practitioner
IT: information technology
PaaS: Platform as a Service
PAT: Autonomous Province of Trento
TS4.0: TrentinoSalute4.0
USCA: Unità Speciali di Continuità Assistenziale

Edited by G Eysenbach; submitted 12.11.20; peer-reviewed by H Mehdizadeh, R Halkes; comments to author 02.12.20; revised version received 22.12.20; accepted 13.04.21; published 31.05.21.

Please cite as:

Gios L, Crema Falceri G, Micocci S, Patil L, Testa S, Sforzin S, Turra E, Conforti D, Malfatti G, Moz M, Nicolini A, Guarda P, Bacchiega A, Mion C, Marchesoni M, Maimone R, Molini PB, Zanella A, Osmani V, Mayora-Ibarra O, Forti S

Use of eHealth Platforms and Apps to Support Monitoring and Management of Home-Quarantined Patients With COVID-19 in the Province of Trento, Italy: App Development and Implementation

JMIR Form Res 2021;5(5):e25713

URL: <https://formative.jmir.org/2021/5/e25713>

doi: [10.2196/25713](https://doi.org/10.2196/25713)

PMID: [33909586](https://pubmed.ncbi.nlm.nih.gov/33909586/)

©Lorenzo Gios, Giulia Crema Falceri, Stefano Micocci, Luigi Patil, Sara Testa, Simona Sforzin, Ettore Turra, Diego Conforti, Giulia Malfatti, Monica Moz, Andrea Nicolini, Paolo Guarda, Alessandro Bacchiega, Carlo Mion, Michele Marchesoni, Rosa Maimone, Pietro Benedetto Molini, Alberto Zanella, Venet Osmani, Oscar Mayora-Ibarra, Stefano Forti. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 31.05.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org

<https://www.jmirpublications.com/>