

---

# JMIR Formative Research

---

Impact Factor (2022): 2.2

Volume 6 (2022), Issue 10 ISSN 2561-326X Editor in Chief: Gunther Eysenbach, MD, MPH, FACMI

---

## Contents

### Original Papers

- Using Personalized Avatars as an Adjunct to an Adult Weight Loss Management Program: Randomized Controlled Feasibility Study ([e36275](#))  
Maria Horne, Maryann Hardy, Trevor Murrells, Hassan Ugail, Andrew Hill. . . . . 6
- Mining the Gems of a Web-Based Mindfulness Intervention: Qualitative Analysis of Factors Aiding Completion and Implementation ([e37406](#))  
Muskan Yadav, Sandra Neate, Craig Hassed, Richard Chambers, Sherelle Connaughton, Nupur Nag. . . . . 19
- Adapting Child Health Knowledge Translation Tools for Use by Indigenous Communities: Qualitative Study Exploring Health Care Providers' Perspectives ([e36353](#))  
Sarah Elliott, Jason Kreutz, Kelsey Wright, Sherri Di Lallo, Shannon Scott, Lisa Hartling. . . . . 28
- Acceptability and Usability of a Reward-Based Mobile App for Opioid Treatment Settings: Mixed Methods Pilot Study ([e37474](#))  
Steven Proctor, Khary Rigg, Allen Tien. . . . . 41
- Bolstering the Business Case for Adoption of Shared Decision-Making Systems in Primary Care: Randomized Controlled Trial ([e32666](#))  
JoAnn Sperl-Hillen, Jeffrey Anderson, Karen Margolis, Rebecca Rossom, Kristen Kopski, Beth Averbeck, Jeanine Rosner, Heidi Ekstrom, Steven Dehmer, Patrick O'Connor. . . . . 48
- Formative Evaluation of the Acceptance of HIV Prevention Artificial Intelligence Chatbots By Men Who Have Sex With Men in Malaysia: Focus Group Study ([e42055](#))  
Mary Peng, Jeffrey Wickersham, Frederick Altice, Roman Shrestha, Iskandar Azwa, Xin Zhou, Mohd Halim, Wan Ikhtiaruddin, Vincent Tee, Adeeba Kamarulzaman, Zhao Ni. . . . . 62
- Factors Related to Smoking and Perceptions of a Behavioral Counseling and Messenger Service-Delivered Smoking Cessation Intervention for People With HIV in China: Qualitative Study ([e35923](#))  
Shanyin Yang, Jiegang Huang, Li Ye, Jianyan Lin, Zhiman Xie, Baodong Guo, Yanjun Li, Bingyu Liang, Zhigang Zheng, Karsten Lunze, Abu Abdullah, Hao Liang, Lisa Quintiliani. . . . . 77
- The Triage Capability of Laypersons: Retrospective Exploratory Analysis ([e38977](#))  
Marvin Kopka, Markus Feufel, Felix Balzer, Malte Schmieding. . . . . 90
- Online Health Information Seeking and eHealth Literacy Among Spanish Language-Dominant Latino Adults Receiving Care in a Community Clinic: Secondary Analysis of Pilot Randomized Controlled Trial Data ([e37687](#))  
Enmanuel Chavarria, Shannon Christy, Han Feng, Hongyu Miao, Rania Abdulla, Liliana Gutierrez, Diana Lopez, Julian Sanchez, Clement Gwede, Cathy Meade. . . . . 99

<b>Evaluating the Feasibility of a Digital Therapeutic Program for Patients With Cancer During Active Treatment: Pre-Post Interventional Study (e39764)</b>	
G Gudmundsson, Judit Mészáros, Ágústa Björnsdóttir, María Ámundadóttir, Gudrun Thorvardardóttir, Erna Magnúsdóttir, Halla Helgadóttir, Saemundur Oddsson. . . . .	115
<b>mHealth Technologies for Managing Problematic Pornography Use: Content Analysis (e39869)</b>	
Nathan Henry, Liesje Donkin, Matt Williams, Mangor Pedersen. . . . .	129
<b>An Evaluation of a Mobile App for Chronic Low Back Pain Management: Prospective Pilot Study (e40869)</b>	
Jonathan Browne, Michael Vaninetti, David Giard, Konstantinos Kostas, Ankur Dave. . . . .	143
<b>Developing Population Health Surveillance Using mHealth in Low-Resource Settings: Qualitative Assessment and Pilot Evaluation (e36260)</b>	
Natalie Benda, Sakie Zawtha, Kyrie Anderson, Mohit Sharma, Phoe Lin, Beichoza Zawtha, Ruth Masterson Creber. . . . .	152
<b>Evidence for Telemedicine’s Ongoing Transformation of Health Care Delivery Since the Onset of COVID-19: Retrospective Observational Study (e38661)</b>	
Soumik Mandal, Batia Wiesenfeld, Devin Mann, Katharine Lawrence, Rumi Chunara, Paul Testa, Oded Nov. . . . .	164
<b>Impact of the COVID-19 Pandemic on the Implementation of Mobile Health to Improve the Uptake of Hydroxyurea in Patients With Sickle Cell Disease: Mixed Methods Study (e41415)</b>	
Sherif Badawy, Lisa DiMartino, Donald Brambilla, Lisa Klesges, Ana Baumann, Ebony Burns, Terri DeMartino, Sara Jacobs, Hamda Khan, Chinonyelum Nwosu, Nirmish Shah, Jane Hankins, Sickle Cell Disease Implementation Consortium. . . . .	179
<b>An Accessible Communication System for Population-Based Genetic Testing: Development and Usability Study (e34055)</b>	
Tara Coffin, Deborah Bowen, Elizabeth Swisher, Karen Lu, Nadine Rayes, Barbara Norquist, Stephanie Blank, Douglas Levine, Jamie Bakkum-Gamez, Gini Fleming, Olufunmilayo Olopade, Alan D’Andrea, Denise Nebgen, Christine Peterson, Mark Munsell, Kathleen Gavin, Rebecca Lechner, Jamie Crase, Deborah Polinsky, Iris Romero. . . . .	192
<b>Readiness for Telemedical Services in Patients With Cardiovascular Diseases: Cross-sectional Study (e33769)</b>	
Barbara Kala ska-Lukasik, Aleksandra Gładny, Tomasz Jadczyk, Monika Gruz-Kwapisz, Wojciech Wojakowski, Malgorzata Kowalska. . . . .	204
<b>A Smartphone Physical Activity App for Patients in Alcohol Treatment: Single-Arm Feasibility Trial (e35926)</b>	
Ana Abrantes, Lidia Meshesha, Claire E Blevins, Cynthia Battle, Clifford Lindsay, Eliza Marsh, Sage Feltus, Matthew Buman, Emmanuel Agu, Michael Stein. . . . .	213
<b>Shared Decision-Making for Drug-Drug Interactions: Formative Evaluation of an Anticoagulant Drug Interaction (e40018)</b>	
Ainhua Gomez Lumbreras, Thomas Reese, Guilherme Del Fiol, Malinda Tan, Jorie Butler, Jason Hurwitz, Mary Brown, Kensaku Kawamoto, Henrik Thiess, Maria Wright, Daniel Malone. . . . .	227
<b>A Cognitive Behavioral Therapy Chatbot (Otis) for Health Anxiety Management: Mixed Methods Pilot Study (e37877)</b>	
Yenushka Goonesekera, Liesje Donkin. . . . .	240
<b>Social Media Communication and Network Correlates of HIV Infection and Transmission Risks Among Black Sexual Minority Men: Cross-sectional Digital Epidemiology Study (e37982)</b>	
Lindsay Young, Jack Tang, Yuanfeixue Nan. . . . .	253
<b>Mental Illness Concordance Between Hospital Clinical Records and Mentions in Domestic Violence Police Narratives: Data Linkage Study (e39373)</b>	
George Karystianis, Rina Cabral, Armita Adily, Wilson Lukmanjaya, Peter Schofield, Iain Buchan, Goran Nenadic, Tony Butler. . . . .	266

<b>Internet-Delivered Self-help for Adults With ADHD (MyADHD): Usability Study (e37137)</b> Robin Kenter, Adrian Schønning, Yavuz Inal. . . . .	280
<b>The Development and Evaluation of “Life Age”—a Primary Prevention and Population-Focused Risk Communication Tool: Feasibility Study With a Single-Arm Repeated Measures Design (e37385)</b> Adeogo Olusan, Suzanne Barr, Mark Cobain, Holly Whelan. . . . .	289
<b>An Automated Surveillance System (SurCeGGID) for the French Sexually Transmitted Infection Clinics: Epidemiological Monitoring Study (e31136)</b> Ndeindo Ndeikoundam Ngangro, Corinne Pioche, Sophie Vaux, Delphine Viriot, Julien Durand, Bénédicte Berat, Mohammed Hamdaoui, Florence Lot. . . . .	296
<b>User Behavior of a Publicly Available, Free-to-Use, Self-guided mHealth App for Depression: Observational Study in a Global Sample (e35538)</b> Langting Su, Page Anderson. . . . .	306
<b>Measuring the Usability of eHealth Solutions for Patients With Parkinson Disease: Observational Study (e39954)</b> Jonas Bendig, Anja Spanz, Jana Leidig, Anika Frank, Marcus Stahr, Heinz Reichmann, Kai Loewenbrück, Björn Falkenburger. . . . .	316
<b>Changes in Food Insecurity Among Individuals Using a Telehealth and Nutrition Platform: Longitudinal Study (e41418)</b> Shivani Bakre, Benjamin Shea, Kaylee Ortega, Jared Scharen, Jason Langheier, Emily Hu. . . . .	327
<b>Colocating Teleophthalmology Within Primary Care Settings to Improve Access to Diabetic Retinopathy Screening: Retrospective Descriptive Evaluation (e17838)</b> Tiffany Wandy, Shuja Rayaz, Jenna Brager, Michael Kiritsy, Elizabeth Offermann, Daniel Durand. . . . .	339
<b>A Digital Smartphone-Based Self-administered Tool (R+ Dietitian) for Nutritional Risk Screening and Dietary Assessment in Hospitalized Patients With Cancer: Evaluation and Diagnostic Accuracy Study (e40316)</b> Zhiwen Long, Shan Huang, Jie Zhang, Deng Zhang, Jun Yin, Chengyuan He, Qinqiu Zhang, Huilin Xu, Huimin He, Ho Sun, Ke Xie. . . . .	345
<b>Relationships Between Blood Pressure Reduction, Weight Loss, and Engagement in a Digital App-Based Hypertension Care Program: Observational Study (e38215)</b> OraLee Branch, Mohit Rikhy, Lisa Auster-Gussman, Kimberly Lockwood, Sarah Graham. . . . .	360
<b>Screening for Generalized Anxiety Disorder From Acoustic and Linguistic Features of Impromptu Speech: Prediction Model Evaluation Study (e39998)</b> Bazen Teferra, Sophie Borwein, Danielle DeSouza, Jonathan Rose. . . . .	371
<b>Informing mHealth and Web-Based Eating Disorder Interventions: Combining Lived Experience Perspectives With Design Thinking Approaches (e38387)</b> Hannah Jarman, Siân McLean, Rachel Rodgers, Matthew Fuller-Tyszkiewicz, Susan Paxton, Beth O’Gorman, Emily Harris, Adrian Shatte, Katie Bishop, Tahlia Baumann, Danielle Mahoney, Melissa-Claire Daugelat, Zali Yager. . . . .	387
<b>A Video-Delivered Family Therapeutic Intervention for Perinatal Women With Clinically Significant Depressive Symptoms and Family Conflict: Indicators of Feasibility and Acceptability (e41697)</b> Fallon Cluxton-Keller, Mark Hegel. . . . .	401
<b>Development of a Web-Based System to Report Medication-Related Adverse Effects: Design and Usability Study (e37605)</b> Renly Lim, Christopher Thornton, Jan Stanek, Lisa Ellett, Myra Thiessen. . . . .	411
<b>Information Resources Among Flemish Pregnant Women: Cross-sectional Study (e37866)</b> Dorien Lanssens, Inge Thijs, Pauline Dreesen, Ann Van Hecke, Pascal Coorevits, Gitte Gaethofs, Joyce Derycke, Inge Tency. . . . .	421

RWD-Cockpit: Application for Quality Assessment of Real-world Data (e29920)  
 Lmar Babrak, Erand Smakaj, Teyfik Agac, Petra Asprion, Frank Grimberg, Daan der Werf, Erwin van Ginkel, Deniz Tosoni, Ieuan Clay, Markus Degen, Dominique Brodbeck, Eriberto Natali, Erik Schkommodau, Enkelejda Miho. . . . . 430

Understanding Mental Health Apps for Youth: Focus Group Study With Latinx Youth (e40726)  
 Elena Agapie, Katherine Chang, Sneha Patrachari, Martha Neary, Stephen Schueller. . . . . 443

Long-term Follow-up of Patients With Hernia Using the Hernia-Specific Quality-of-Life Mobile App: Feasibility Questionnaire Study (e39759)  
 Ching-Shui Huang, Feng-Chuan Tai, Heng-Hui Lien, Jia-Uei Wong, Chi-Cheng Huang. . . . . 459

The Family Level Assessment of Screen Use–Mobile Approach: Development of an Approach to Measure Children’s Mobile Device Use (e40452)  
 Oriana Perez, Anil Kumar Vadathya, Alicia Beltran, R Barnett, Olivia Hinderer, Tatyana Garza, Salma Musaad, Tom Baranowski, Sheryl Hughes, Jason Mendoza, Ashutosh Sabharwal, Ashok Veeraraghavan, Teresia O’Connor. . . . . 468

A Tailored Gender-Sensitive mHealth Weight Loss Intervention (I-GENDO): Development and Process Evaluation (e38480)  
 Magdalena Pape, Tanja Färber, Caroline Seiferth, Tanja Roth, Stefanie Schroeder, Joerg Wolstein, Stephan Herpertz, Sabine Steins-Loeber. . . . . 4

From Real-world Individuals’ Data to National Health Indicators: Multiphase Pilot Study in Gabon (e35176)  
 Aimé Koumamba, Edgard Ngoungou, Jean Engohang-Ndong, Euloge Ibinga, Raymond Ondzigue Mbenga, Gayo Diallo. . . . . 497

Digital Peer-Supported Self-Management Intervention Codesigned by People With Long COVID: Mixed Methods Proof-of-Concept Study (e41410)  
 Hayley Wright, Andrew Turner, Stuart Ennis, Carol Percy, Garry Loftus, Wendy Clyne, Gabriela Matouskova, Faith Martin. . . . . 511

The First Asynchronous Online Evidence-Based Medicine Course for Syrian Health Workforce: Effectiveness and Feasibility Pilot Study (e36782)  
 Yazan Kenjrawi, Mayssoon Dashash. . . . . 526

Mobile App to Enhance Patient Activation and Patient-Provider Communication in Major Depressive Disorder Management: Collaborative, Randomized Controlled Pilot Study (e34923)  
 Maggie McCue, Christopher Blair, Ben Fehnert, James King, Francesca Cormack, Sara Sarkey, Anna Eramo, Christopher Kabir, Rasha Khatib, David Kemp. . . . . 536

Telehealth Movement-to-Music to Increase Physical Activity Participation Among Adolescents With Cerebral Palsy: Pilot Randomized Controlled Trial (e36049)  
 Byron Lai, Laura Vogtle, Raven Young, Mary Craig, Yumi Kim, Marissa Gowe, Erin Swanson-Kimani, Drew Davis, James Rimmer. . . . . 549

Social Determinants of Health and Diabetes-Related Distress in Patients With Insulin-Dependent Type 2 Diabetes: Cross-sectional, Mixed Methods Approach (e40164)  
 Natalie Levy, Agnes Park, Daniela Solis, Lu Hu, Aisha Langford, Binhuan Wang, Erin Rogers. . . . . 565

Separating Features From Functionality in Vaccination Apps: Computational Analysis (e36818)  
 George Shaw Jr, Devaki Nadkarni, Eric Phann, Rachel Sietaty, Madeleine Ledenyi, Razaan Abnowf, Qian Xu, Paul Arredondo, Shi Chen. . . . . 5

Social Media Mining of Long-COVID Self-Medication Reported by Reddit Users: Feasibility Study to Support Drug Repurposing (e39582)  
 Jonathan Koss, Sabine Bohnet-Joschko. . . . . 588

Collective Value Promotes the Willingness to Share Provaccination Messages on Social Media in China: Randomized Controlled Trial ( <a href="#">e35744</a> ) Chunye Fu, Xiaokang Lyu, Mingdi Mi. ....	601
COVID-Washing in US Food and Beverage Marketing on Twitter: Content Analysis ( <a href="#">e37642</a> ) Krystle Tsai, Omni Cassidy, Josh Arshonsky, Sara Bond, Inés Del Giudice, Marie Bragg. ....	612
Population Health Metrics During the Early Stages of the COVID-19 Pandemic: Correlative Pilot Study ( <a href="#">e40215</a> ) Marie Severson, David Cassada, Victor Huber, Daniel Snow, Lisa McFadden. ....	623
Feasibility of At-Home Serial Testing Using Over-the-Counter SARS-CoV-2 Tests With a Digital Smartphone App for Assistance: Longitudinal Cohort Study ( <a href="#">e35426</a> ) Carly Herbert, John Broach, William Heetderks, Felicia Qashu, Laura Gibson, Caitlin Pretz, Kelsey Woods, Vik Kheterpal, Thejas Suvarna, Christopher Nowak, Peter Lazar, Didem Ayturk, Bruce Barton, Chad Achenbach, Robert Murphy, David McManus, Apurv Soni. ....	636
Anxiety, Post-COVID-19 Syndrome-Related Depression, and Suicidal Thoughts and Behaviors in COVID-19 Survivors: Cross-sectional Study ( <a href="#">e36656</a> ) Sean Woodward, Sumra Bari, Nicole Vike, Shamal Lalvani, Khrystyna Stetsiv, Byoung Kim, Leandros Stefanopoulos, Nicos Maglaveras, Hans Breiter, Aggelos Katsaggelos. ....	646
COVID-Bot, an Intelligent System for COVID-19 Vaccination Screening: Design and Development ( <a href="#">e39157</a> ) Chinedu Okonkwo, Lateef Amusa, Hossana Twinomurinzi. ....	665
Sociodemographics and Transdiagnostic Mental Health Symptoms in SOCIAL (Studies of Online Cohorts for Internalizing Symptoms and Language) I and II: Cross-sectional Survey and Botometer Analysis ( <a href="#">e39324</a> ) Lorenzo Lorenzo-Luaces, Jacqueline Howard, Andy Edinger, Harry Yan, Lauren Rutter, Danny Valdez, Johan Bollen. ....	678
Examining the Twitter Discourse on Dementia During Alzheimer's Awareness Month in Canada: Infodemiology Study ( <a href="#">e40049</a> ) Juanita-Dawne Bacsu, Allison Cammer, Soheila Ahmadi, Mehrnoosh Azizi, Karl Grewal, Shoshana Green, Rory Gowda-Sookchoff, Corinne Berger, Sheida Knight, Raymond Spiteri, Megan O'Connell. ....	697

## Corrigenda and Addenda

Correction: Using Wake-Up Tasks for Morning Behavior Change: Development and Usability Study ( <a href="#">e42926</a> ) Kyue Oh, Jisu Ko, Jaemyung Shin, Minsam Ko. ....	676
---	-----

Original Paper

# Using Personalized Avatars as an Adjunct to an Adult Weight Loss Management Program: Randomized Controlled Feasibility Study

Maria Horne<sup>1\*</sup>, BA(Hons), MA, PhD; Maryann Hardy<sup>2\*</sup>, BSc(Hons), MSc, PhD; Trevor Murrells<sup>3\*</sup>, BSc, MSc; Hassan Ugail<sup>4\*</sup>, BSc, PGCE, PhD; Andrew John Hill<sup>5\*</sup>, BSc, PhD

<sup>1</sup>Faculty of Medicine and Health, School of Healthcare, University of Leeds, Leeds, United Kingdom

<sup>2</sup>Faculty of Health Studies, School of Allied Health Professions and Midwifery, University of Bradford, Bradford, United Kingdom

<sup>3</sup>Florence Nightingale Faculty of Nursing, Midwifery and Palliative Care, King's College London, London, United Kingdom

<sup>4</sup>University of Bradford, Centre for Visual Computing, Bradford, United Kingdom

<sup>5</sup>Institute of Health Sciences, School of Medicine, University of Leeds, Leeds, United Kingdom

\* all authors contributed equally

**Corresponding Author:**

Maria Horne, BA(Hons), MA, PhD

Faculty of Medicine and Health

School of Healthcare

University of Leeds

Baines Wing

Leeds, LS2 9JT

United Kingdom

Phone: 44 113 343 ext 1230

Email: [M.Horne@leeds.ac.uk](mailto:M.Horne@leeds.ac.uk)

## Abstract

**Background:** Obesity is a global public health concern. Interventions rely predominantly on managing dietary intake and increasing physical activity; however, sustained adherence to behavioral regimens is often poor. The lack of sustained motivation, self-efficacy, and poor adherence to behavioral regimens are recognized barriers to successful weight loss. Avatar-based interventions achieve better patient outcomes in the management of chronic conditions by promoting more active engagement. Virtual representations of *self* can affect real-world behavior, acting as a catalyst for sustained weight loss behavior.

**Objective:** We evaluated whether a personalized avatar, offered as an adjunct to an established weight loss program, can increase participant motivation, sustain engagement, optimize service delivery, and improve participant health outcomes.

**Methods:** A feasibility randomized design was used to determine the case for future development and evaluation of avatar-based technology in a randomized controlled trial. Participants were recruited from general practitioner referrals to a 12-week National Health Service weight improvement program. The main outcome measure was weight loss. Secondary outcome measures were quality-of-life and self-efficacy. Quantitative data were subjected to descriptive statistical tests and exploratory comparison between intervention and control arms. Feasibility and acceptability were assessed through interviews and analyzed using framework approach. Health Research Authority ethics approval was granted.

**Results:** Overall, 10 men (n=7, 70% for routine care and avatar and n=3, 30% for routine care) and 33 women (n=23, 70% for intervention and n=10, 30% for routine care) were recruited. Participants' initial mean weight was greater in the intervention arm than in the routine care arm (126.3 kg vs 122.9 kg); pattern of weight loss was similar across both arms of the study in T0 to T1 period but accelerated in T1 to T2 period for intervention participants, suggesting that access to the self-resembling avatar may promote greater engagement with weight loss initiatives in the short-to-medium term. Mean change in participants' weight from T0 to T2 was 4.5 kg (95% CI 2.7-6.3) in the routine care arm and 5.3 kg (95% CI 3.9-6.8) in the intervention arm. Quality-of-life and self-efficacy measures demonstrated greater improvement in the intervention arm at both T1 (105.5 for routine care arm and 99.7 for intervention arm) and T2 (100.1 for routine care arm and 81.2 for intervention arm). Overall, 13 participants (n=11, 85% women and n=2, 15% men) and two health care professionals were interviewed about their experience of using the avatar program.

**Conclusions:** Participants found using the personalized avatar acceptable, and feedback reiterated that seeing a future *self* helped to reinforce motivation to change behavior. This feasibility study demonstrated that avatar-based technology may successfully promote engagement and motivation in weight loss programs, enabling participants to achieve greater weight loss gains and build self-confidence.

**Trial Registration:** ISRCTN Registry 17953876; <https://doi.org/10.1186/ISRCTN17953876>

(*JMIR Form Res* 2022;6(10):e36275) doi:[10.2196/36275](https://doi.org/10.2196/36275)

## KEYWORDS

avatar; feasibility; obesity; weight loss; weight management

## Introduction

### Background

Obesity is a global health concern [1] and a health priority in the United Kingdom [2,3]. It is associated with a range of increased health risks including diabetes (type II), heart disease, cancer, depression, and mental health issues [4], affecting societal engagement and quality of life [5]. It has also been identified as a mortality and morbidity risk factor for COVID-19 [6]. In addition to affecting the health of individuals, obesity increases demand for, and complexity of, health care and is estimated to cost the UK National Health Service (NHS) £5 billion (US \$5.739 billion) annually with a projected increase of £1.9 billion to £2 billion (US \$2.08 billion to US \$2.19 billion) per annum by 2030 without intervention [7]. Consequently, in terms of individual, societal, and economic well-being and sustainability, developing effective interventions to reduce population obesity is imperative.

Obesity interventions in adults rely predominantly on managing dietary intake and increasing physical activity [8]. Systematic reviews have evaluated the effectiveness of interventions, and multicomponent interventions have been acknowledged to be more effective for weight loss than single-component approaches [9,10]. Common barriers to successful weight loss among adults with obesity have been identified as lack of sustained motivation and poor adherence to behavioral regimens [11,12]. However, interest in the use of digital technologies to support health behavior change is growing [13,14].

Technology-based interventions for weight loss have shown promise in the short term [15,16]; however, the potential of computer engineering has not been fully realized when designing weight loss interventions [17]. It has been argued that new multicomponent (integrated) electronic platforms integrating education with individually tailored weight loss programs, including the promotion of autonomous motivation, self-efficacy, self-regulation, and positive body image, may present a way forward for wide-scale weight loss solutions and obesity management [18-20]. An area that has significant potential to be exploited as a catalyst for weight loss behavior modification within health care is virtual reality (VR) and the creation of personal avatars [15]. VR enables people to experience an alternate visual reality, often through an avatar (computerized representation of self). Avatar technology is well established within the computer gaming industry, and third-person perspective of self within a VR setting has been shown to promote emotional engagement [21,22].

On the basis of the social cognitive theory by Bandura [23] (behavior learning through observation of models) and self-perception theory by Bem [24] (inference of own attitudes by observing self from a third-party perspective), studies have demonstrated a link between virtual representations of self and real-world behavior and attitudes, which is considered as a consequence of individuals identifying with their avatar [25,26], particularly when the avatar is self-resembling [27], and experience of presence within the computer-mediated environment [28].

The potential of VR as an intervention has been explored across a spectrum of health and well-being conditions [25,26,29,30], and it has been recommended as a potential behavior modification tool for tackling obesity [31,32]. Studies have shown that within the experimental environment, an individual's behavior conforms to their digital self-representation (Proteus effect) [33-35]. Importantly, observing a self-resembling avatar modeling an activity (eg, exercise) within a VR environment, as opposed to a generic avatar, differentially positively influences real-world behavior, thus increasing actual engagement with the activity [29,36]. This connectedness between the actual and digital self has been further explored using aging algorithms to present individuals with personalized virtual futures influenced by choices made today [37,38], and the results suggest that observation of the future self within a web-based environment may vicariously reinforce today's desirable behaviors and attitudes.

Transforming the appearance of self through an avatar is a particularly powerful motivational application, but there has been limited translation of this technology into health care service delivery. So far, no identified study has directly applied the technology to the clinical setting as an adjunct to an existing obesity intervention or overtly included people with obesity within the study sampling frame.

### Aims

We proposed to develop a self-resembling avatar generation program to allow individuals to explore potential health futures related to obesity, including observing their self-avatar gain and lose weight and information on the associated health risks at BMI boundaries. This feasibility study tested the avatar design capabilities; clinical application; and underlying premise that this technology, provided as an adjunct to an existing weight loss program, will positively influence participant motivation.

## Methods

### Study Design

This was a 2-phase, sequential, mixed methods feasibility study, incorporating a feasibility randomized controlled trial (RCT) with a subsequent qualitative component, to determine the case for, and the parameters of, a future RCT. CONSORT-EHEALTH (Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth) guidelines were used to report the study ([Multimedia Appendix 1](#)).

### Ethics Approval

Ethics approval for the study was obtained from the University of Bradford and Health Research Authority (research ethics committee reference 18/NE/0286), and it was registered with the International Standard Randomized Control Trial Number (17953876). All participants were screened for suitability, by a registered medical clinician specializing in obesity, and psychological vulnerability. All participants provided written informed consent, and a process of medical referral and reporting of any potential harm as a result of participation in the study was established.

### Participants

Participants were patients with obesity or those who were overweight from within West Yorkshire, England, referred to the Mid Yorkshire Hospitals NHS Trust weight improvement service (March 2019 to December 2019). Inclusion criteria were the following: (1) adults aged 18 to 65 years, (2) BMI >30 kg/m<sup>2</sup>, (3) referred from GP, and (4) no known comorbidities or medical treatment that may influence dietary intake or weight loss achievement (eg, type I diabetes or current medication with weight gain as a known side effect).

Exclusion criteria were the following: (1) pregnant women, (2) children, (3) older individuals (aged >65 years), (4) BMI >45 kg/m<sup>2</sup>, and (5) people considered to be psychologically vulnerable.

A parallel group approach was undertaken with participants assigned to routine weight management program (routine care) or routine weight management program and avatar (intervention) and treated according to group assignment.

### Recruitment

Potential participants were identified by the weight management service clinical team (Mid Yorkshire Hospitals NHS Trust) following initial assessment for suitability to participate in a weight management program and study participation screening. Potential participants were given an information pack, including invitation letter; information sheet; permission-to-contact form; and postage-paid envelope to read, complete, and return after the appointment. The forms were returned to the weight management service, and a secure database of potential participants was created by the NHS research administrator assigned to this project. The researcher confirmed with the administrator on a weekly basis about those patients who had returned the permission-to-contact forms and telephoned those

patients interested in participating to describe the study further and confirm their willingness to participate. A cluster randomization approach was adopted to allocate patients to routine care or routine care and avatar based on the geographic location of the weight management service intervention.

A subsample of participants within both arms who completed the 12-week weight management program were invited to participate in individual interviews, 3 to 6 months after commencement of the intervention, and share their experiences of using the avatar program.

Health care professionals (HCPs) involved in delivering the weight loss program and who had experience of using the avatar program were invited to participate in a semistructured interview at the conclusion of the study's data collection phase. This was done to explore their experiences of using the avatar program as an adjunct to routine care.

### Intervention

The avatar VR program (creation and display) was developed through an iterative design process with service users and HCPs. Service user volunteers were recruited from the University of Bradford service user group, and HCP participants were recruited from the Mid Yorkshire Hospitals NHS Trust weight management service. Feedback at each stage of development ensured that the final avatar program design was optimized before commencing the feasibility study.

The definition of avatar adopted for this study was a web-based 3D entity whose appearance and body proportions resembled that of the user. In this sense, the avatar was a close representation or a *digital clone* of the individual, allowing the individual to relate to it in terms of appearance and body proportions.

The avatar was developed using standard 3D modeling technology. Standard human 3D shapes were developed using Maya 3D modeling software (Autodesk). This 3D model is flexible, permitting body parts (eg, size of the lower abdomen) to be changed using simple numbers or parameters. Therefore, the parameterized 3D model can be efficiently manipulated using a handful of parameter values determined by physical measurements and photographs of the individual acquired using a tablet device (iPad).

The weight management routine care program operated over 12 weeks. Patients who were allocated to receive the intervention (routine care and avatar) attended the weight improvement service for their initial appointment (T0). Additional time was allocated for this appointment to meet the researcher and for baseline anthropomorphic measurements and digital photographs to be acquired. These parameters were inputted into the avatar creation program while participants completed three questionnaires—(1) Impact of Weight on Quality of Life-lite [39], (2) Weight Efficacy Lifestyle-Short Form questionnaire [40], and (3) EQ-5D-5L [41]—to monitor changes in well-being during the study period and identify the best tool to be adopted in a future RCT, if warranted. Once the avatar was generated (<5 minutes after data input), the participant was introduced to their avatar and program operation and given a unique username and password to allow them to



access their avatar on the web via computer or tablet. Participants were encouraged to explore how increasing or decreasing weight affects avatar appearance, considering the 5% and 10% weight loss goals promoted by the weight management routine care program alongside the BMI boundary health risks as their avatar increased and decreased in weight.

Follow-up appointments with the researcher coincided with weight management service routine care appointments at 1 month (T1; immediate effect) and 3 months (T2; short-term effect after program completion). A further appointment to meet with the researcher was planned for 6 months (T3; 3 months after program completion—moderate-term progress). At each of these appointments, the researcher collected the completed surveys and inputted any change in weight into the avatar program. The changes in avatar appearance were presented and discussed with the participants at each researcher appointment, and participants were encouraged to use the avatar to visualize progress.

### Routine Care

Routine care was the standard 12-week weight management program without the provision of an avatar. Participants allocated to receive routine care were provided the same number of appointments with the weight improvement service as those allocated to the intervention group and met the researcher to complete questionnaires and acquire body measurements, photographs, and weight change data at the same points as those participating in the intervention arm. However, although routine care participants underwent the same process of meeting the researcher and replicating the activity of the intervention, visualization of, and access to, their avatar was withheld until the final weight management service appointment (12 weeks).

### Outcome Measures

The primary outcome measure was weight loss based on body weight (kg) and calculated BMI ( $\text{kg}/\text{m}^2$ ). Assessment of weight loss was planned to be conducted at baseline and after 3 months and 6 months.

The secondary outcome measures were uptake and continuation rates and changes in self-perceived quality of life and self-efficacy.

### Sample Size

The target sample size for the feasibility study trial was between 30 and 60 patients, overall [42]. The sample size for interviews with service users and HCPs was based on data saturation, with 10 to 16 interviews anticipated [43].

### Data Analysis

Data analysis focused on describing the key feasibility outcomes using descriptive statistical tests, including calculation of means and frequencies. Mean (SD and 95% CI) and median (IQR) values were reported by the study arm (routine care and routine care and avatar) for the primary outcomes of weight (kg) and BMI and secondary effectiveness outcome data collected through the questionnaires. A generalized linear repeated measures mixed effects model for weight (kg) was used to account for the discrete timing of the follow-up assessments and adjust for baseline measures (T0). The aim of that model was to obtain an estimate of variability (residual SD) for weight change and to estimate the mean change in weight (95% CI) from baseline (T0) to T1 and T2, which will then be used to inform the sample size calculation for a future definitive trial, if warranted.

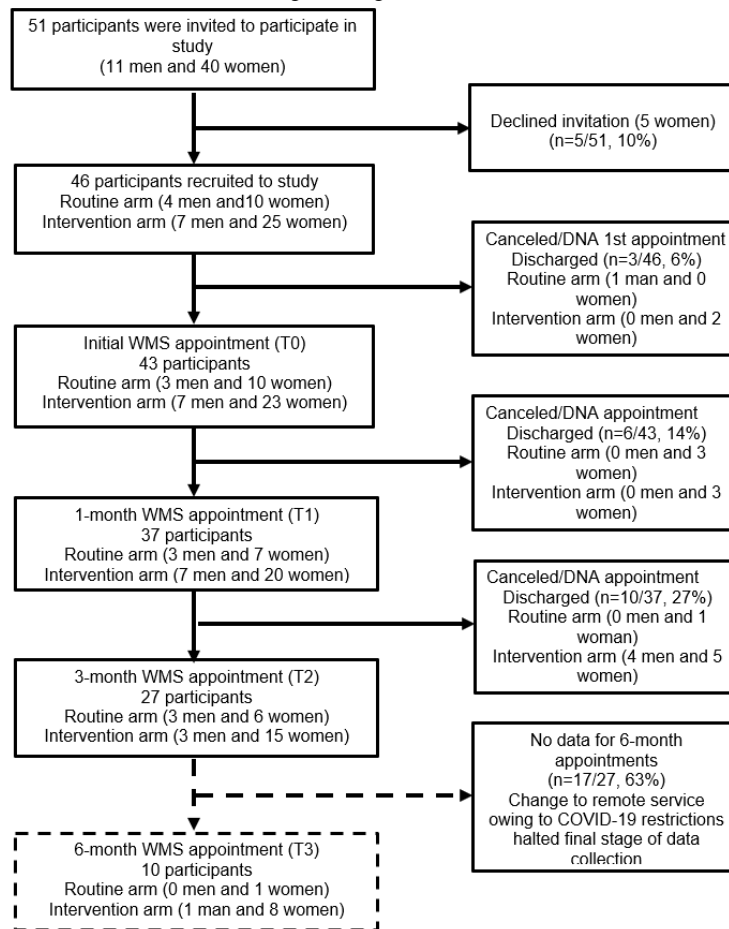
Qualitative data were analyzed using framework approach [44]. The first author coded all the transcripts. Then, the codes were discussed with the research team, who together developed the indexing scheme that was used by the first author to chart data. These charts were shared with others in the team to explore and interpret the data together and decide the final themes. Consensus on themes and subthemes was reached through discussion.

## Results

### Recruitment

The number of potential participants who met the recruitment criteria was less than that expected. This was mainly owing to the greater than expected proportion of referrals to the weight management service related to people with  $\text{BMI} > 45 \text{ kg}/\text{m}^2$ . In addition, restrictions and changes in service delivery from face-to-face to telephone appointments as a result of the COVID-19 pandemic influenced the final stages of data collection. Therefore, T3 (6 months) data were not collected for most participants, and analysis was restricted to T0 to T2.

Recruitment was conducted between March 2019 and December 2019. During this time, 51 potential participants were identified as being suitable for participation in the study, and information about the study was provided. Of these 51 participants, 5 (10%) declined participation and 3 (6%) did not attend the initial appointment and were discharged from the weight management service (Figure 1).

**Figure 1.** Participant flowchart. DNA: did not attend; WMS: weight management service.

## Participant Characteristics

Baseline data were collected for 84% (43/51) of the participants. A total of 10 men were recruited into the study ( $n=7$ , 70% into the routine care and avatar arm and  $n=3$ , 30% into the routine care arm). At the beginning of the program, their mean age was 50 years 10 months (range 39-63 years) and mean weight was 140.2 (range 114.6-173.2) kg. A total of 33 women were enrolled into the study at T0 ( $n=23$ , 70% into the intervention arm and  $n=10$ , 30% into the routine care arm). At the beginning of the program, their mean age was 40 years 1 month (range 21-57 years) and mean weight was 120.5 (range 89.6-156.4) kg. Mean BMI was 43.3 (range 38.6 to 49.5)  $\text{kg}/\text{m}^2$ .

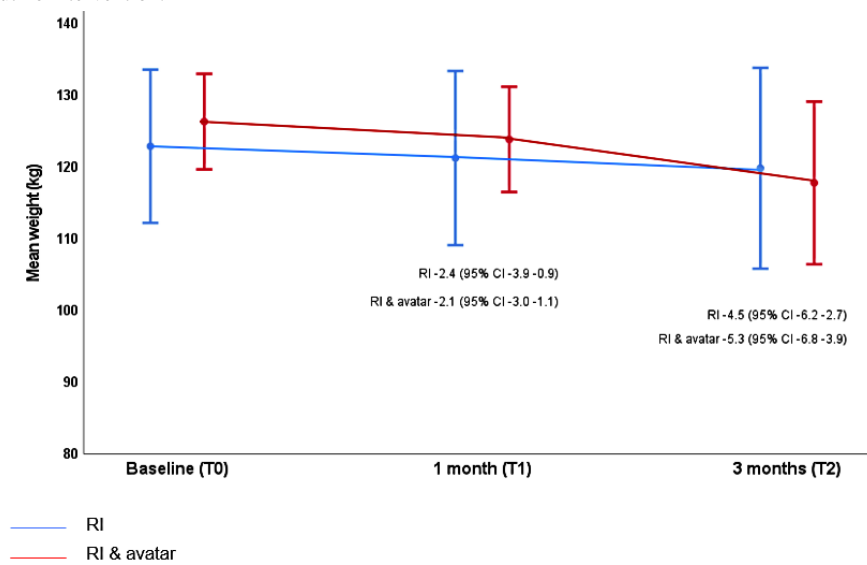
## Weight

The initial mean weight of participants was greater in the intervention arm than in the routine care arm (126.3 kg vs 122.9

kg), but the pattern of weight loss was similar across both arms of the study in the T0 to T1 period. During T1 to T2, weight loss accelerated in the intervention arm, suggesting that access to the self-resembling avatar may promote great engagement with weight loss initiatives in the short to medium term (Figure 2).

The mean change in weight of the participants from T0 to T2 was 4.5 kg (95% CI 2.7-6.3) in the routine care arm and 5.3 kg (95% CI 3.9-6.8) in the intervention arm. The initial mean BMI of participants was greater in the intervention arm than in the routine care arm (44  $\text{kg}/\text{m}^2$  vs 42.4  $\text{kg}/\text{m}^2$ ). However, reflecting the great weight loss in the intervention arm, the mean BMI of participants in both groups was identical at T2 (BMI 41.1  $\text{kg}/\text{m}^2$ ).

**Figure 2.** Weight (kg) by visit (cross-sectional means and SDs). Note: mean person-level change from baseline (95% CI) is shown for both arms at 1 month and 3 months. RI: routine intervention.



### Quality of Life and Self-efficacy

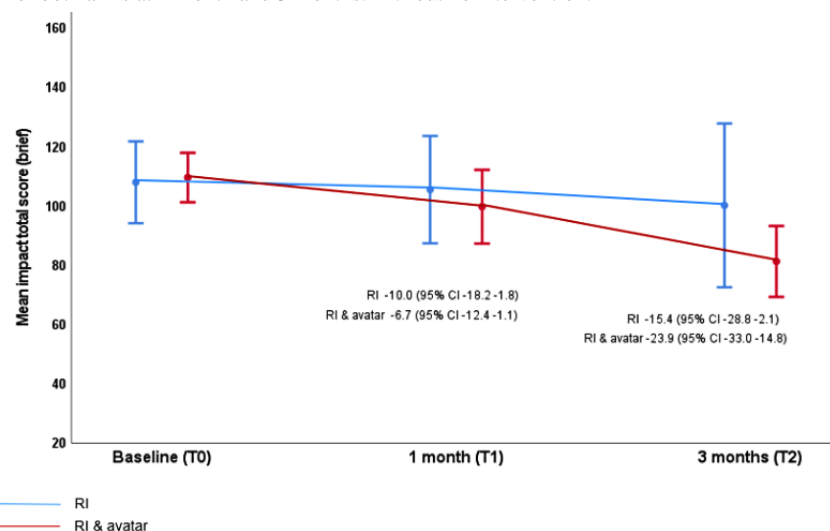
Quality of life and self-efficacy showed improvements, as seen in participants' scores across both arms of the study, but greater improvement was noted within the intervention arm (Figure 3).

Mean initial scores were comparable across the arms (107.9 in the routine care arm and 109.5 in the intervention arm), but participants in the intervention arm reported greater improvement at both T1 (105.5 in the routine care arm and 99.7 in the intervention arm) and T2 (100.1 in the routine care arm and 81.2 in the intervention arm). Improvement in participants' self-assessment of health at the time of appointment was also noted over the study time frame for participants in both arms of the study using EQ-5D-5L Visual Analog Scale (VAS) score; however, once again, great improvement was noted within the intervention arm. Mean VAS scores for the routine arm were 49.1 (T0), 50.2 (T1), and 52.8 (T2), whereas for the intervention

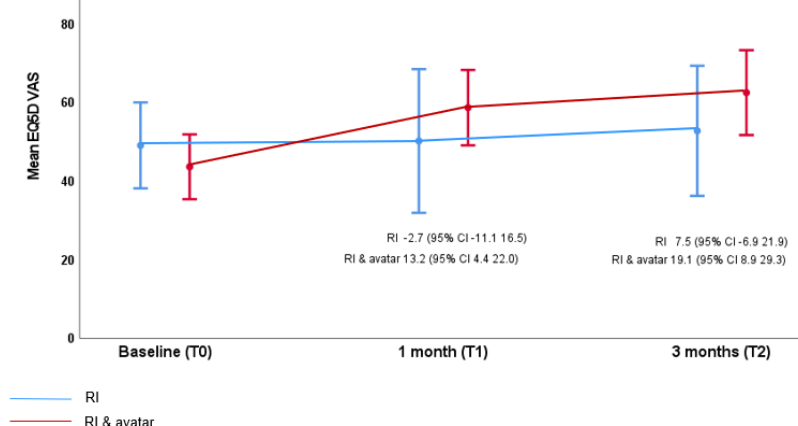
arm, mean VAS scores were 43.6 (T0), 58.7 (T1), and 62.5 (T2, Figure 4).

Having access to and visualizing changes in the self-resembling avatar improved weight loss motivation and perception of self-well-being. In contrast, the Weight Efficacy Lifestyle Questionnaire identified great improvements in participant self-belief in controlling eating behaviors in the routine care arm, with mean improvement in score between T0 and T2 of 17.6 in the routine arm and 12.1 in the intervention arm. It is unclear why the results from this questionnaire contrast with the findings of the other assessments. This may reflect the wording of statements, purpose of questionnaire, or interpretation by participants. However, it is important to note that the use of a single questionnaire alone will not allow for understanding of the complex motivations, enablers of, and barriers to weight loss.

**Figure 3.** Impact of Weight on Quality of Life-lite total score (brief) by visit (cross-sectional means and SDs). Note: mean person-level change from baseline (95% CI) is shown for both arms at 1 month and 3 months. RI: routine intervention.



**Figure 4.** EQ5D VAS by visit (cross-sectional means/SDs). Note: mean person-level change from baseline (95% CI) is shown for both arms at 1 month and 3 months. RI: routine intervention.



### Experiences of Using a Web-Based Avatar

In total, 13 participants (n=11, 85% women and n=2, 15% men) and two HCPs volunteered to be interviewed about their experiences of using a web-based avatar on a weight loss program. Regarding acceptability, most participants found using the personalized avatar as an acceptable and positive experience:

*Oh, I think the avatar's brilliant. It's really good to be able to see where you were and then to slowly see the progress because you can't always see it when you are looking in the mirror...I get comments; people go have you lost some weight? And I'm like oh yeah just a bit, but I can't see that I have lost weight but when you look at that [Avatar] you can see that even just a bit of weight makes that much difference. [MF45; female participant]*

However, a few participants found the initial viewing of the avatar as uncomfortable, but reported that the appearance of their avatar was a “reality check” of their current selves, which provided the added motivation to lose weight:

*...It was all a bit shocking to be fair [seeing avatar for first time] ...I don't look at myself very much in the mirror... [MM47; male participant]*

*It gives you an objective look on how others would see you. [MF13; female participant]*

*Devastated - that I had let myself get like that so...yeah...embarrassed...it spurred me on to get rid of all that round my belly...I know...as I get thinner I will get fitter and get rid of this lung issue and then I will get to where I want to be and I will go to the gym more and even do more at home. I will definitely walk the dog more; I'm trying to walk the dog faster so I am exerting myself more...so yeah, seeing that end result is brilliant to [compare with] where I was. [MF23; female participant]*

Visual perception of the personalized avatar assisted some participants to visualize a potential future reality:

*It helps me to visualize what I could, will be...what I am going to look like. [MF17; female participant]*

Participants also reported added psychological benefits, such as increased motivation and improved self-esteem and self-confidence:

*I need the visual to keep me motivated. [MF22; female participant]*

*Helps your self-esteem...when you can see it rather than someone just telling you. [MF39; female participant]*

*I feel like I am better equipped to kind of like be in control of myself. [MF02; female participant]*

Participants reported no concerns regarding the questionnaires used, despite some questions being personal and related to physical intimacy:

*There are some personal questions talking about intimacy and things, its important though because it is a part...you know that is...when you put weight on you feel undesirable especially when your partner is thin as a rake and...you weren't, you looked completely different when you met and you have been together you know...donkeys years you know, it is a big part of it, but I think it is important, you have to highlight every area or else you won't address your issues will? [MF22; female participant]*

Another participant reported that the questionnaires were good at pinpointing the physical and emotional issues that people with obesity have. Another participant talked about the positive boost received when reflecting back on the scores from T0 to T2 and seeing positive change in the scores that further embedded the feeling of well-being and achievement:

*...I think they pinpointed all the physical and emotional issues that people who are overweight probably face or are worried about. So yes, it was good and it was good that I could give honest answers and really think about each question. [MF01; female participant]*

Both health care practitioners were dietitians or nutritionists, and both felt that the personalized avatar, as an adjunct to a weight management program, was a useful addition to the questionnaires available:

*I think it exceeded my expectations...because obviously she takes a photo of the patient and then it's got their face on it and obviously their body...I know it didn't have any clothing or anything like that [underwear only], but I don't think that really matters because they get to see the actual body without the clothing...I think that's better because they can see their skin and where they are actually carrying the weight...I think overall it has been beneficial to have it as part of the program. [HCP1; female HCP]*

*I don't think I expected so much emotion, not the extent where people needed a lot of comforting...I didn't expect that no...Um...but I think it was a positive thing generally, it was something people could use to kind of scale down...It massively helped...just by dropping the BMI by a little bit or moving the scale option to do that...You know I would always highlight that this wasn't just about aesthetics although, you know everyone kind of wants to be a smaller dress size, but it's not really about that, it's about your health. It was quite good to drop it and show if your BMI was lowered by this much, your risk of diabetes reduces by this much and I think that's when it kind of hits people because then they are working to more health-related goals. [HCP2; female HCP]*

The addition of a personalized avatar to the standard NHS program was reported as helpful in terms of patients being able to visualize *reality* and confront denial:

*I feel like the avatar has been a great help for the patients that were on the MotiVar study...a lot of the patients were able to see what they actually look like, because having worked in the service for a few years now...I do understand a lot of the patients don't look in the mirror for example or they tend to be in denial about their weight or...they might be confident to lose the weight but then they know they are going to struggle to keep it off and that tends to knock them back if they have done several diets through their lifetime...but I think seeing that reality where you know they get to see the avatar and with their face on it as well...where they don't have to look in the mirror but they see themselves on the screen. I think it is a reality check for a lot of them, but also think it's a lot more real? [HCP1; female HCP]*

*I felt like with the females...because they could visualize it...it became something a bit more real and people made comments that existing apps they had used weren't realistic...one that did actually speak about that and she had downloaded a few apps in the past, but it wasn't like, proportionate to her and you didn't have to take so many measurements it was just your height. [HCP2; female HCP]*

Both HCPs felt that, for most participants, visualizing the personalized avatar assisted with the motivation to take up and maintain the weight loss program and to visualize personal goals:

*...When they are starting something like this program, I think it (personalized avatar) definitely helps in terms of motivation and giving them that kick start that they need, but also it helps to visualize their goal as well because of what the avatar does offer...how to select the different weights and seeing what the goal weight is but again you can check what your 5% target is and then see the difference as well when the weight does change on the avatar. I think it's really good to see it. [HCP1; female HCP]*

Health care practitioners reported no difficulties in administering the questionnaires and felt that they were appropriate for assessing confidence during the weight management program.

## Discussion

### Principal Findings

This study aimed to investigate the feasibility, acceptability, and implementation of a randomized design to determine the case for future development and evaluation of avatar-based technology in an RCT.

Following the recommendations for feasibility studies [45], feasibility was assessed by examining recruitment and retention (attrition) rates, together with participants' and HCPs' experiences of the use of avatar-based technology as an adjunct to a weight management program. In addition, we reported the pre-post intervention effect sizes on primary (weight loss) and other outcomes.

Recruitment was found to be feasible, with more than half of the contacted people consenting to participate, and the target of recruiting at least 30 participants into the intervention arm during the 6-month recruitment period was met. Reasons for declining to participate in the study were disinterest in the study and lack of guarantee of being in the intervention arm. Overall, 7% (3/46) of the recruited participants failed to attend their initial appointment and were discharged from the weight management service. The number of individuals recruited into the intervention arm (30/43, 70%) suggested that a personalized avatar, as an adjunct to a weight management program, will be feasible to offer and will be of interest to participants. However, attrition from the program was relatively high (19/46, 41%).

Regarding acceptability, all study participants mentioned experiencing benefits from using the avatar-based technology. Most participants emphasized the importance of the avatar's appearance as a reality check, whereas HCPs felt that it was important for patients to visualize *reality* and confront denial.

Most participants in our study also reported added psychological benefits, such as increased motivation to lose weight and improved self-esteem and self-confidence.

Finally, although the feasibility study was not powered to detect a difference in weight, the study found 10% greater mean weight loss with routine care and avatar compared with routine care weight management program.

## Comparison With Previous Studies

Attrition and nonadherence are known to be common problems across all weight loss interventions, with reported mean attrition rates ranging from 10% to >80% [46-49]. Therefore, further studies are required to determine the role that avatar-based technology may have in reducing attrition and supporting engagement within a weight management intervention.

The importance of the avatar's appearance in terms of visualizing *reality* and confronting denial builds on, and is consistent with, studies suggesting that mental imagery can induce strong positive and negative effects [50,51]. For example, Ridgway and King [51] found that when participants viewed their 3D avatars, their overall body satisfaction and mood decreased compared with that observed after viewing the baseline reports. However, participants also reported wanting to engage in greater appearance management behaviors after viewing the avatar compared with their baseline reports, suggesting that the avatar stimulated a change in perception of self and behavior. Similarly, Park [50], investigating how body image discrepancy or body satisfaction leverages behavioral intention for weight regulation among nonclinical participants who were aged  $\geq 18$  years, found that after participating in a web-based avatar session, those who demonstrated high body dissatisfaction exhibited great intention to be involved in behavioral change to achieve healthy body weight. In contrast, increased body image discrepancy in itself did not trigger an intention to lose weight after web-based avatar experience [50], confirming the multifactorial nature of self-perception and motivation to change behavior.

Weight management services are generally poorly adhered to [48], perhaps as a consequence of high body dissatisfaction, stigma, and discrimination toward people who are obese [52]. Viewing one's anthropometric web-based avatar can affect the viewer's self-body perception through comparative evaluation of self-concepts and lead to self-acceptance [53]. The initial discomfort experienced by some study participants on viewing their avatar appeared to provide the trigger needed for some to motivate them to change their behavior and adhere to a weight management program and warrants further investigation.

Our findings around the added psychological benefits, for example, increased motivation to lose weight and improved self-esteem and self-confidence, builds on and is consistent with previous studies suggesting that the experience of using an avatar can increase self-efficacy [54-56]. However, the literature suggests that this experience does not consistently translate into real-world settings and target health behaviors [56], thus supporting the need for further studies in this area.

Weight loss was similar between T0 to T1 for both routine care and avatar arm and routine care arm. However, from T1 to T2, weight loss was greater in the routine care and avatar arm, suggesting that high level of motivation was maintained. The reason for this was not explored directly, but can be explained

by the placebo effect of participating in a research trial [57]. Participants in both arms knew that they were part of a research study and were meeting researchers and the weight management team; this may have had an equal impact on diet motivation [57].

A recent systematic review [15], which focused on the inclusion of avatar technology in weight loss interventions, identified greater weight loss and weight maintenance compared with routine interventions. However, the differences were not consistent in terms of statistical or clinical significance [15]. No full-powered trial has been undertaken, and, therefore, although the findings in terms of weight loss demonstrate potential, large studies are required to confirm the results.

## Limitations

The feasibility trial design had several strengths, and, importantly, this was the first study of its type, conducted in a health care setting with a GP referral weight management service. As this was a feasibility study, assessing the effectiveness of the intervention was not the primary aim. However, future studies should examine the effectiveness of the avatar within an RCT design and better understand the economic impact.

Limitations of this study include the small sample size and relatively high attrition. However, it still remains as one of the larger sample size studies reported so far. Despite these limitations, this feasibility study has illustrated that avatar-based technology may successfully promote engagement with and motivation to lose weight as part of a weight management program.

This feasibility study demonstrates the possibilities of using avatar-based technology to motivate engagement with a weight management program in the short term. The findings also suggest that avatar-based technology may support greater self-confidence, belief, and efficacy in weight loss ambitions.

## Conclusions

We investigated the feasibility, acceptability, and implementation of a randomized design to determine the case for future development and evaluation of avatar-based technology in an RCT. This feasibility study demonstrates the possibilities of using avatar-based technology to motivate engagement with a weight management program in the short term. The findings also suggest that avatar-based technology may support greater self-confidence, belief, and efficacy in weight loss ambitions. Overall, the trial design was found to be feasible and acceptable by participants and HCPs. Expansion of the inclusion criteria to include both primary care and local community weight management services may be beneficial. A full trial of the technology, taking into account participant feedback about the avatar technology design and accessibility, is also warranted.

## Acknowledgments

This study was funded by the National Institute for Health Research (Research for Patient Benefit program [project reference PB-PG-1215-20016]). The views expressed are those of the authors and not necessarily those of the National Institute for Health Research or Department of Health and Social Care.

## Conflicts of Interest

None declared.

## Multimedia Appendix 1

CONSORT-eHEALTH checklist (V 1.6.1).

[[PDF File \(Adobe PDF File\), 3387 KB - formative\\_v6i10e36275\\_app1.pdf](#)]

## References

1. Obesity. World Health Organization. URL: [https://www.who.int/health-topics/obesity#tab=tab\\_1](https://www.who.int/health-topics/obesity#tab=tab_1) [accessed 2022-07-08]
2. Healthy lives, healthy people: a call to action on obesity in England. Department of Health and Social Care. 2011 Oct 13. URL: <https://www.gov.uk/government/publications/healthy-lives-healthy-people-a-call-to-action-on-obesity-in-england> [accessed 2022-07-08]
3. Statistics on obesity, physical activity and diet, England, 2020. NHS Digital. 2020 May 5. URL: <https://digital.nhs.uk/data-and-information/publications/statistical/statistics-on-obesity-physical-activity-and-diet/england-2020> [accessed 2022-07-08]
4. McPherson K, Marsh T, Brown M. Tackling Obesities: Future Choices – Modelling Future Trends in Obesity & Their Impact on Health 2nd Edition. London, UK: HMSO; 2007.
5. Scarborough P, Bhatnagar P, Wickramasinghe KK, Allender S, Foster C, Rayner M. The economic burden of ill health due to diet, physical inactivity, smoking, alcohol and obesity in the UK: an update to 2006-07 NHS costs. *J Public Health (Oxf)* 2011 Dec 11;33(4):527-535. [doi: [10.1093/pubmed/fdr033](https://doi.org/10.1093/pubmed/fdr033)] [Medline: [21562029](https://pubmed.ncbi.nlm.nih.gov/21562029/)]
6. Mohammad S, Aziz R, Al Mahri S, Malik SS, Haji E, Khan AH, et al. Obesity and COVID-19: what makes obese host so vulnerable? *Immun Ageing* 2021 Jan 04;18(1):1 [FREE Full text] [doi: [10.1186/s12979-020-00212-x](https://doi.org/10.1186/s12979-020-00212-x)] [Medline: [33390183](https://pubmed.ncbi.nlm.nih.gov/33390183/)]
7. Wang YC, McPherson K, Marsh T, Gortmaker SL, Brown M. Health and economic burden of the projected obesity trends in the USA and the UK. *Lancet* 2011 Aug;378(9793):815-825. [doi: [10.1016/S0140-6736\(11\)60814-3](https://doi.org/10.1016/S0140-6736(11)60814-3)]
8. Bray GA, Ryan DH. Evidence-based weight loss interventions: individualized treatment options to maximize patient outcomes. *Diabetes Obes Metab* 2021 Feb 24;23 Suppl 1(S1):50-62. [doi: [10.1111/dom.14200](https://doi.org/10.1111/dom.14200)] [Medline: [32969147](https://pubmed.ncbi.nlm.nih.gov/32969147/)]
9. Abbott S, Smith E, Tighe B, Lycett D. Group versus one-to-one multi-component lifestyle interventions for weight management: a systematic review and meta-analysis of randomised controlled trials. *J Hum Nutr Diet* 2021 Jun 28;34(3):485-493. [doi: [10.1111/jhn.12853](https://doi.org/10.1111/jhn.12853)] [Medline: [33368624](https://pubmed.ncbi.nlm.nih.gov/33368624/)]
10. Madigan CD, Graham HE, Sturgiss E, Kettle VE, Gokal K, Biddle G, et al. Effectiveness of weight management interventions for adults delivered in primary care: systematic review and meta-analysis of randomised controlled trials. *BMJ* 2022 May 30;377:e069719 [FREE Full text] [doi: [10.1136/bmj-2021-069719](https://doi.org/10.1136/bmj-2021-069719)] [Medline: [35636762](https://pubmed.ncbi.nlm.nih.gov/35636762/)]
11. Chew HS, Gao Y, Shabbir A, Lim SL, Geetha K, Kim G, et al. Personal motivation, self-regulation barriers and strategies for weight loss in people with overweight and obesity: a thematic framework analysis. *Public Health Nutr* 2022 Feb 22;25(9):2426-2435. [doi: [10.1017/s136898002200043x](https://doi.org/10.1017/s136898002200043x)]
12. Coons MJ, Demott A, Buscemi J, Duncan JM, Pellegrini CA, Steglitz J, et al. Technology interventions to curb obesity: a systematic review of the current literature. *Curr Cardiovasc Risk Rep* 2012 Apr 2;6(2):120-134 [FREE Full text] [doi: [10.1007/s12170-012-0222-8](https://doi.org/10.1007/s12170-012-0222-8)] [Medline: [23082235](https://pubmed.ncbi.nlm.nih.gov/23082235/)]
13. Robinson A, Husband A, Slight R, Slight S. Digital technology to support lifestyle and health behaviour changes in surgical patients: systematic review. *BJS Open* 2021 Mar 05;5(2):zraa009 [FREE Full text] [doi: [10.1093/bjsopen/zraa009](https://doi.org/10.1093/bjsopen/zraa009)] [Medline: [33688953](https://pubmed.ncbi.nlm.nih.gov/33688953/)]
14. Taj F, Klein MC, van Halteren A. Digital health behavior change technology: bibliometric and scoping review of two decades of research. *JMIR Mhealth Uhealth* 2019 Dec 13;7(12):e13311 [FREE Full text] [doi: [10.2196/13311](https://doi.org/10.2196/13311)] [Medline: [31833836](https://pubmed.ncbi.nlm.nih.gov/31833836/)]
15. Horne M, Hill A, Murells T, Ugail H, Irving, Chinnadorai R, et al. Using avatars in weight management settings: a systematic review. *Internet Interv* 2020 Mar;19:100295 [FREE Full text] [doi: [10.1016/j.invent.2019.100295](https://doi.org/10.1016/j.invent.2019.100295)] [Medline: [31871900](https://pubmed.ncbi.nlm.nih.gov/31871900/)]
16. 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](https://pubmed.ncbi.nlm.nih.gov/21091286/)]
17. Goldstein CM, Goldstein SP, Thomas DM, Hoover A, Bond DS, Thomas JG. The behavioral intervention with technology for e-weight loss study (BITES): incorporating energy balance models and the bite counter into an online behavioral weight

- loss program. *J Technol Behav Sci* 2021 Jun 27;6(2):406-418 [FREE Full text] [doi: [10.1007/s41347-020-00181-4](https://doi.org/10.1007/s41347-020-00181-4)] [Medline: [35356149](https://pubmed.ncbi.nlm.nih.gov/35356149/)]
18. Stubbs R, Duarte C, Palmeira A, Sniehotta F, Horgan G, Larsen S, et al. Evidence-based digital tools for weight loss maintenance: the NoHoW project. *Obes Facts* 2021 Apr 29;14(3):320-333 [FREE Full text] [doi: [10.1159/000515663](https://doi.org/10.1159/000515663)] [Medline: [33915534](https://pubmed.ncbi.nlm.nih.gov/33915534/)]
  19. Wiechert M, Holzapfel C. Nutrition concepts for the treatment of obesity in adults. *Nutrients* 2021 Dec 30;14(1):169 [FREE Full text] [doi: [10.3390/nu14010169](https://doi.org/10.3390/nu14010169)] [Medline: [35011045](https://pubmed.ncbi.nlm.nih.gov/35011045/)]
  20. Moravcová K, Karbanová M, Bretschneider MP, Sovová M, Ožana J, Sovová E. Comparing digital therapeutic intervention with an intensive obesity management program: randomized controlled trial. *Nutrients* 2022 May 10;14(10):2005 [FREE Full text] [doi: [10.3390/nu14102005](https://doi.org/10.3390/nu14102005)] [Medline: [35631145](https://pubmed.ncbi.nlm.nih.gov/35631145/)]
  21. Schuurink EL, Toet A. Effects of third person perspective on affective appraisal and engagement: findings from SECOND LIFE. *Simul Gaming* 2010 Apr 14;41(5):724-742. [doi: [10.1177/1046878110365515](https://doi.org/10.1177/1046878110365515)]
  22. Lim S, Reeves B. Being in the game: effects of avatar choice and point of view on psychophysiological responses during play. *Media Psychol* 2009 Nov 23;12(4):348-370. [doi: [10.1080/15213260903287242](https://doi.org/10.1080/15213260903287242)]
  23. 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/)]
  24. Bem D. Self-perception theory. In: *Advances in Experimental Social Psychology*. New York: Academic Press; 1972.
  25. Wrzesien M, Rodríguez A, Rey B, Alcañiz M, Baños RM, Vara MD. How the physical similarity of avatars can influence the learning of emotion regulation strategies in teenagers. *Comput Human Behav* 2015 Feb;43:101-111. [doi: [10.1016/j.chb.2014.09.024](https://doi.org/10.1016/j.chb.2014.09.024)]
  26. Seitz C, Poyrazli S, Harrison M, Flickinger T, Turkson M. Virtual reality exposure therapy for military veterans with posttraumatic stress disorder: a systematic review. *New School Psychol Bulletin* 2014;11(1):14-29.
  27. Schultze U, Leahy MM. The avatar-self relationshipnacting presence in second life. ICIS proceedings. 2009. URL: <https://aisel.aisnet.org/icis2009/12> [accessed 2022-07-08]
  28. Parrish DE, Oxhandler HK, Duron JF, Swank P, Bordnick P. Feasibility of virtual reality environments for adolescent social anxiety disorder. *Res Soc Work Pract* 2016 Aug 03;26(7):825-835. [doi: [10.1177/1049731514568897](https://doi.org/10.1177/1049731514568897)]
  29. Bordnick PS, Carter BL, Traylor AC. What virtual reality research in addictions can tell us about the future of obesity assessment and treatment. *J Diabetes Sci Technol* 2011 Mar 01;5(2):265-271 [FREE Full text] [doi: [10.1177/193229681100500210](https://doi.org/10.1177/193229681100500210)] [Medline: [21527092](https://pubmed.ncbi.nlm.nih.gov/21527092/)]
  30. Behm-Morawitz E. Mirrored selves: the influence of self-presence in a virtual world on health, appearance, and well-being. *Comput Human Behav* 2013 Jan;29(1):119-128. [doi: [10.1016/j.chb.2012.07.023](https://doi.org/10.1016/j.chb.2012.07.023)]
  31. Thomas JG, Goldstein CM, Bond DS, Hadley W, Tuerk PW. Web-based virtual reality to enhance behavioural skills training and weight loss in a commercial online weight management programme: the experience success randomized trial. *Obes Sci Pract* 2020 Dec 27;6(6):587-595 [FREE Full text] [doi: [10.1002/osp4.451](https://doi.org/10.1002/osp4.451)] [Medline: [33354337](https://pubmed.ncbi.nlm.nih.gov/33354337/)]
  32. Al-Rasheed A, Alabdulkreem E, Alduailij M, Alduailij M, Alhalabi W, Alharbi S, et al. Virtual reality in the treatment of patients with overweight and obesity: a systematic review. *Sustainability* 2022 Mar 11;14(6):3324. [doi: [10.3390/su14063324](https://doi.org/10.3390/su14063324)]
  33. Yee N, Bailenson J. The Proteus effect: the effect of transformed self-representation on behavior. *Human Comm Res* 2007 Jul;33(3):271-290. [doi: [10.1111/j.1468-2958.2007.00299.x](https://doi.org/10.1111/j.1468-2958.2007.00299.x)]
  34. Yee N, Bailenson JN, Ducheneaut N. The Proteus effect. *Commun Res* 2009 Jan 22;36(2):285-312. [doi: [10.1177/0093650208330254](https://doi.org/10.1177/0093650208330254)]
  35. Fox J, Bailenson J, Binney J. Virtual experiences, physical behaviors: the effect of presence on imitation of an eating avatar. *Presence Teleoperators Virtual Environ* 2009 Aug 01;18(4):294-303. [doi: [10.1162/pres.18.4.294](https://doi.org/10.1162/pres.18.4.294)]
  36. Song H, Kim J, Kwon RJ, Jung Y. Anti-smoking educational game using avatars as visualized possible selves. *Comput Human Behav* 2013 Sep;29(5):2029-2036. [doi: [10.1016/j.chb.2013.04.008](https://doi.org/10.1016/j.chb.2013.04.008)]
  37. Hershfield HE, Goldstein DG, Sharpe WF, Fox J, Yeykelis L, Carstensen LL, et al. Increasing saving behavior through age-progressed renderings of the future self. *J Market Res* 2019 Mar 06;48(SPL):S23-S37. [doi: [10.1509/jmkr.48.spl.s23](https://doi.org/10.1509/jmkr.48.spl.s23)]
  38. Kim Y, Sundar SS. Visualizing ideal self vs. actual self through avatars: impact on preventive health outcomes. *Comput Human Behav* 2012 Jul;28(4):1356-1364. [doi: [10.1016/j.chb.2012.02.021](https://doi.org/10.1016/j.chb.2012.02.021)]
  39. Kolotkin RL, Ervin CM, Meincke HH, Højbjerg L, Fehnel SE. Development of a clinical trials version of the Impact of Weight on Quality of Life-Lite questionnaire (IWQOL-Lite Clinical Trials Version): results from two qualitative studies. *Clin Obes* 2017 Oct 22;7(5):290-299 [FREE Full text] [doi: [10.1111/cob.12197](https://doi.org/10.1111/cob.12197)] [Medline: [28544443](https://pubmed.ncbi.nlm.nih.gov/28544443/)]
  40. Ames GE, Heckman MG, Grothe KB, Clark MM. Eating self-efficacy: development of a short-form WEL. *Eat Behav* 2012 Dec;13(4):375-378. [doi: [10.1016/j.eatbeh.2012.03.013](https://doi.org/10.1016/j.eatbeh.2012.03.013)] [Medline: [23121791](https://pubmed.ncbi.nlm.nih.gov/23121791/)]
  41. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res* 2011 Dec;20(10):1727-1736 [FREE Full text] [doi: [10.1007/s11136-011-9903-x](https://doi.org/10.1007/s11136-011-9903-x)] [Medline: [21479777](https://pubmed.ncbi.nlm.nih.gov/21479777/)]
  42. Teare MD, Dimairo M, Shephard N, Hayman A, Whitehead A, Walters SJ. Sample size requirements to estimate key design parameters from external pilot randomised controlled trials: a simulation study. *Trials* 2014 Jul 03;15(1):264 [FREE Full text] [doi: [10.1186/1745-6215-15-264](https://doi.org/10.1186/1745-6215-15-264)] [Medline: [24993581](https://pubmed.ncbi.nlm.nih.gov/24993581/)]



43. Guest G, Bunce A, Johnson L. How many interviews are enough? *Field Methods* 2016 Jul 21;18(1):59-82. [doi: [10.1177/1525822X05279903](https://doi.org/10.1177/1525822X05279903)]
44. Ritchie J, Spencer L, O'Connor W. Carrying out qualitative analysis. In: *Qualitative Research Practice*. London: SAGE Publications; 2003.
45. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M, Medical Research Council Guidance. Developing and evaluating complex interventions: the new Medical Research Council guidance. *BMJ* 2008 Sep 29;337(sep29 1):a1655 [FREE Full text] [doi: [10.1136/bmj.a1655](https://doi.org/10.1136/bmj.a1655)] [Medline: [18824488](https://pubmed.ncbi.nlm.nih.gov/18824488/)]
46. Franz MJ, VanWormer JJ, Crain AL, Boucher JL, Histon T, Caplan W, et al. Weight-loss outcomes: a systematic review and meta-analysis of weight-loss clinical trials with a minimum 1-year follow-up. *J Am Diet Assoc* 2007 Oct;107(10):1755-1767. [doi: [10.1016/j.jada.2007.07.017](https://doi.org/10.1016/j.jada.2007.07.017)] [Medline: [17904936](https://pubmed.ncbi.nlm.nih.gov/17904936/)]
47. LeBlanc ES, Patnode CD, Webber EM, Redmond N, Rushkin M, O'Connor EA. Behavioral and pharmacotherapy weight loss interventions to prevent obesity-related morbidity and mortality in adults: updated evidence report and systematic review for the US preventive services task force. *JAMA* 2018 Sep 18;320(11):1172-1191. [doi: [10.1001/jama.2018.7777](https://doi.org/10.1001/jama.2018.7777)] [Medline: [30326501](https://pubmed.ncbi.nlm.nih.gov/30326501/)]
48. Dombrowski SU, Knittle K, Avenell A, Araújo-Soares V, Snihotta FF. Long term maintenance of weight loss with non-surgical interventions in obese adults: systematic review and meta-analyses of randomised controlled trials. *BMJ* 2014 May 14;348(may14 6):g2646 [FREE Full text] [doi: [10.1136/bmj.g2646](https://doi.org/10.1136/bmj.g2646)] [Medline: [25134100](https://pubmed.ncbi.nlm.nih.gov/25134100/)]
49. Moroshko I, Brennan L, O'Brien P. Predictors of dropout in weight loss interventions: a systematic review of the literature. *Obes Rev* 2011 Nov;12(11):912-934. [doi: [10.1111/j.1467-789X.2011.00915.x](https://doi.org/10.1111/j.1467-789X.2011.00915.x)] [Medline: [21815990](https://pubmed.ncbi.nlm.nih.gov/21815990/)]
50. Park J. The effect of virtual avatar experience on body image discrepancy, body satisfaction and weight regulation intention. *Cyberpsychology J Psychosocial Res Cyberspace* 2018 Jul 25;12(1) [FREE Full text] [doi: [10.5817/cp2018-1-3](https://doi.org/10.5817/cp2018-1-3)]
51. Ridgway J. Before and after avatar exposure: the impact of body scanning technology on body satisfaction, mood, and appearance management. *Clothing Textiles Res J* 2017 Dec 26;36(2):91-103. [doi: [10.1177/0887302X17749924](https://doi.org/10.1177/0887302X17749924)]
52. Puhl RM, Heuer CA. Obesity stigma: important considerations for public health. *Am J Public Health* 2010 Jun;100(6):1019-1028. [doi: [10.2105/ajph.2009.159491](https://doi.org/10.2105/ajph.2009.159491)]
53. Park J, Ogle JP. How virtual avatar experience interplays with self-concepts: the use of anthropometric 3D body models in the visual stimulation process. *Fash Text* 2021 Jul 05;8(1):1-24. [doi: [10.1186/s40691-021-00257-6](https://doi.org/10.1186/s40691-021-00257-6)]
54. Navarro J, Cebolla A, Llorens R, Borrego A, Baños RM. Manipulating self-avatar body dimensions in virtual worlds to complement an internet-delivered intervention to increase physical activity in overweight women. *Int J Environ Res Public Health* 2020 Jun 05;17(11):4045 [FREE Full text] [doi: [10.3390/ijerph17114045](https://doi.org/10.3390/ijerph17114045)] [Medline: [32517117](https://pubmed.ncbi.nlm.nih.gov/32517117/)]
55. Napolitano MA, Hayes S, Russo G, Muresu D, Giordano A, Foster GD. Using avatars to model weight loss behaviors: participant attitudes and technology development. *J Diabetes Sci Technol* 2013 Jul 01;7(4):1057-1065 [FREE Full text] [doi: [10.1177/193229681300700430](https://doi.org/10.1177/193229681300700430)] [Medline: [23911189](https://pubmed.ncbi.nlm.nih.gov/23911189/)]
56. Johnston JD, Massey AP, Devaneaux CA. Innovation in weight loss programs: a 3-dimensional virtual-world approach. *J Med Internet Res* 2012 Sep 20;14(5):e120 [FREE Full text] [doi: [10.2196/jmir.2254](https://doi.org/10.2196/jmir.2254)] [Medline: [22995535](https://pubmed.ncbi.nlm.nih.gov/22995535/)]
57. Linde K, Fässler M, Meissner K. Placebo interventions, placebo effects and clinical practice. *Philos Trans R Soc Lond B Biol Sci* 2011 Jun 27;366(1572):1905-1912 [FREE Full text] [doi: [10.1098/rstb.2010.0383](https://doi.org/10.1098/rstb.2010.0383)] [Medline: [21576148](https://pubmed.ncbi.nlm.nih.gov/21576148/)]

## Abbreviations

**CONSORT-EHEALTH:** Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth

**HCP:** health care professional

**NHS:** National Health Service

**RCT:** randomized controlled trial

**VAS:** Visual Analog Scale

**VR:** virtual reality

*Edited by A Mavragani; submitted 11.01.22; peer-reviewed by J Wright, B Chaudhry; comments to author 25.05.22; revised version received 18.07.22; accepted 23.07.22; published 05.10.22.*

*Please cite as:*

*Horne M, Hardy M, Murrells T, Ugail H, Hill AJ*

*Using Personalized Avatars as an Adjunct to an Adult Weight Loss Management Program: Randomized Controlled Feasibility Study*  
*JMIR Form Res* 2022;6(10):e36275

URL: <https://formative.jmir.org/2022/10/e36275>

doi: [10.2196/36275](https://doi.org/10.2196/36275)

PMID: [36197703](https://pubmed.ncbi.nlm.nih.gov/36197703/)

©Maria Horne, Maryann Hardy, Trevor Murrells, Hassan Ugail, Andrew John Hill. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 05.10.2022. 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

# Mining the Gems of a Web-Based Mindfulness Intervention: Qualitative Analysis of Factors Aiding Completion and Implementation

Muskan Yadav<sup>1</sup>, MD; Sandra Neate<sup>2</sup>, MBBS; Craig Hassed<sup>3</sup>, MBBS; Richard Chambers<sup>3</sup>, DPsych; Sherelle Connaughton<sup>3</sup>, PhD; Nupur Nag<sup>2</sup>, PhD

<sup>1</sup>Department of General Practice, Monash University, Melbourne, Australia

<sup>2</sup>Neuroepidemiology Unit, Melbourne School of Population and Global Health, The University of Melbourne, Melbourne, Australia

<sup>3</sup>Monash Centre for Consciousness and Contemplative Studies, Monash University, Melbourne, Australia

**Corresponding Author:**

Nupur Nag, PhD

Neuroepidemiology Unit

Melbourne School of Population and Global Health

The University of Melbourne

Level 3, 207 Bouverie Street

Melbourne, 3010

Australia

Phone: 61 383447944

Email: [nnag@unimelb.edu.au](mailto:nnag@unimelb.edu.au)

## Abstract

**Background:** Digital health interventions provide a cost effective and accessible means for positive behavior change. However, high participant attrition is common and facilitators for implementation of behaviors are not well understood.

**Objective:** The goal of the research was to identify elements of a digital mindfulness course that aided in course completion and implementation of teachings.

**Methods:** Inductive thematic analysis was used to assess participant comments regarding positive aspects of the online mindfulness course Mindfulness for Well-being and Peak Performance. Participants were aged 18 years and older who had self-selected to register and voluntarily completed at least 90% the course. The course comprised educator-guided lessons and discussion forums for participant reflection and feedback. Participant comments from the final discussion forum were analyzed to identify common themes pertaining to elements of the course that aided in course completion and implementation of teachings.

**Results:** Of 3355 course completers, 283 participants provided comments related to the research question. Key themes were (1) benefits from the virtual community, (2) appeal of content, (3) enablers to participation and implementation, and (4) benefits noted in oneself. Of subthemes identified, some, such as community support, variety of easily implementable content, and free content access, align with that reported previously in the literature, while other subthemes, including growing together, repeating the course, evidence-based teaching, and immediate benefits on physical and mental well-being, were novel findings.

**Conclusions:** Themes identified as key elements for aiding participant completion of a mindfulness digital health intervention and the implementation of teachings may inform the effective design of future digital health interventions to drive positive health behaviors. Future research should focus on understanding motivations for participation, identification of effective methods for participant retention, and behavior change techniques to motivate long-term adherence to healthy behaviors.

(*JMIR Form Res* 2022;6(10):e37406) doi:[10.2196/37406](https://doi.org/10.2196/37406)

**KEYWORDS**

digital intervention; health education; mindfulness; online learning; behavior change; mental health; mental well-being; physical well-being; meditation; health promotion; digital health; eHealth; thematic analysis; attrition; participation; involvement; engagement; attitude; perspective; patient education; e-learning; user feedback

## Introduction

Web-based educational interventions have the potential to be implemented for disease prevention, positive behavior change, and self-managed well-being; they are scalable and can be disseminated widely [1-4]. In addition, the digital format enables participants to schedule and plan their learning, provides relative anonymity for open questioning and avoidance of perceived stigmatization for topics such as mental health, and minimizes physical accessibility and financial barriers [5,6]. During the COVID-19 pandemic, restrictions on in-person interactions were introduced to minimize virus spread. Physical barriers to education and demand for self-help courses led to a rapid change and increase in digital communication and learning [7]. Moreover, a global rise in mental health conditions resulting from the pandemic [8,9] prompts an immediate need for efficacious interventions that may increase resilience and well-being.

Mindfulness is a form of mental training that has been effectively used to improve health outcomes in clinical and nonclinical populations [10-12]. Defined as the nonjudgmental awareness of the environment, one's body, and one's own thoughts and feelings in the present moment, it can be developed through a range of formal (mindfulness meditation), informal (being mindful in daily life), and cognitive (reflecting on various aspects of mindfulness—for example, acceptance and letting go) practices. Meta-analyses indicate that mindfulness-based interventions (MBIs) facilitate healthy lifestyle change [3]; reduce stress, depression, and anxiety; and improve resilience, cognitive function, sleep, and general well-being [13].

Digital MBIs aim to enable health-based behavior change through real-time or asynchronous participation in self-guided or facilitator-supported content. MBIs have traditionally been enacted in-person where direct, personal, and reflective quality reinforces learning; these factors may be lacking in a digital format, bringing to question the effectiveness of digital MBIs. Studies directly comparing online versus in-person delivery of MBIs have reported similar benefits on measured health outcomes [14], and many studies have reported on the effectiveness of online delivery to reduce perceived stress and improve mindfulness, mental health-related outcomes, and well-being [15]. Key success components of web-based interventions include program content, multimedia resources, interactive online activities, and guidance or supportive feedback [16], while challenges include selecting and understanding the audience, effective learning, and defining learning objectives [1,17]. High attrition rates and low adherence are common. In a systematic review of studies of smartphone apps for depressive symptoms, attrition was near 50% but lower in studies offering human feedback and in-app mood monitoring [18]. Low adherence is associated with reduced intervention efficacy and possibly related to a combination of characteristics of the condition addressed by the intervention, the user, and the intervention itself [4,19]. Understanding components that aid in course completion and implementation of learned practices may inform methods to minimize some of the challenges in web-based educational interventions for positive behavior change.

Mindfulness for Well-being and Peak Performance (MWPP) is a massive online open course (MOOC) which, based on user reviews on Class Central, has been ranked in the top MOOCs globally since the year it began. A study of 2105 participants who completed MWPP found significant improvement in mindfulness, perceived stress, and work engagement [20]. Given the popularity and efficacy of MWPP, we qualitatively analyzed participants' free text comments on the course's final learner forum asking "What were the gems of the course?" to explore the positive aspects of the course. Having explored all forum comments, two research questions arose: "What course components aided course completion and implementation of learned practices?" and "How did participants learn mindfulness, and what outcomes resulted from this?" In this study, the research question regarding factors assisting course completion and implementation of learnings was explored.

## Methods

### Digital Mindfulness Intervention

The MWPP course was developed in 2015 and delivered digitally thrice annually until the end of 2021 by CH and RC. The course is hosted on the FutureLearn platform and open for free registration globally. Registered participants were sent weekly email reminders during the 4-week period prompting them to complete course modules.

The course recommended a commitment of 3 hours per week to cover topics that build skills and knowledge progressively. Content was delivered live in English via short videos, curated articles, downloadable guided meditations, quizzes on core content, links to further resources, moderated discussion forums, and weekly feedback videos. Discussion forums, based on weekly topics and designed to encourage self-reflective learning and discussion between participants, were facilitated by SC and another mentor with the support of educators CH and RC. Content for weekly feedback videos was based on prevalent participant questions and insights from discussion forums. At the end of the 4-week course, participants had an additional 2 weeks to complete the modules, without live mentoring. The course was free for this 6-week period with an optional one-time fee charged for unlimited access to content thereafter.

### Participants and Data Collection

Participants self-enrolled in the 16th run of the MWPP course, delivered from March to April 2020. Of 23,932 enrolled participants, 18,080 participants began the course and 3335 (18.4%) completed at least 90% of the modules. Of MWPP course completers, 527 responded to the final discussion forum questions "Have you joined this course before, and if so, what brings you back?" and "What are the gems for you from this course?" (data from responses to the latter were deidentified and extracted for analysis).

### Ethics Approval

This study was approved by the Monash University Human Research Ethics Committee (ID 18105).

### Analysis

Participants’ comments were analyzed using inductive thematic analysis, which identifies, analyses, and reports patterns within qualitative data in a structured format [21]. This form of analysis was selected given the content of the data varied and thematic analysis can detect patterns that facilitate understanding. Responses to the question of interest were securely exported into an Excel (Microsoft Corp) document.

The analysis team comprised MY, NN, and SN. All 527 participant responses to the final discussion forum question were independently reviewed in an iterative process to aid familiarization with the data and develop initial codes (NN and SN). Comments in languages other than English and from educators and mentors were excluded. From initial codes, NN and SN identified 2 primary and independent categories from the data: course-related (digital health interventions) codes versus value and impact-related codes. These categories were derived without preconceived ideas of categories or theoretical perspectives. A total of 283 participant comments related to the first overarching category of digital health interventions. These 283 comments were reanalyzed and recoded, and themes that contained a central organizing concept were identified and named (NN and MY). Analysis of comments relating to the second category are reported elsewhere [22].

The data were reviewed to ensure themes were appropriately named and reflected the data. Detailed records of researcher meetings (NN and MY) were maintained by MY to enhance transparency. Themes are presented in the context of a digital health program evaluation framework [23] and subthemes illustrated with verbatim quotes to enable readers to assess the validity and fittingness of researcher interpretations [24].

## Results

### Participant Characteristics

Of the course completers, 15.8% (527/3335) responded to the final discussion forum question. Of the comments, 53.7%

(283/527) were relevant to the research question. Due to FutureLearn’s privacy policy, it was not possible to match comments to individual participant demographics.

### Qualitative Findings

Four themes and 21 subthemes were identified for components of the course that aided in course completion and implementation of teachings:

1. Benefits from the virtual community
  - Mentor/educator feedback
  - Shared experiences
  - Support
  - Growing together
2. Appeal of content
  - Variety of format
  - Optional content
  - Unrestricted access
  - Visual aids
  - Life examples
  - Ease of implementation
  - Delivery by educators
3. Enablers to participation and implementation
  - Own pace
  - Free access
  - Option to repeat course
  - Evidence-based teaching
  - New perspectives
4. Benefits noted in oneself
  - Knowledge gained
  - Stress management
  - Behavior change
  - Mind and body benefits
  - Increased ability to cope

Selected participant quotes embodying subthemes are presented in text, and a representative quote for each theme tabulated (Table 1).

**Table 1.** Themes and illustrative quotes.

Theme	Illustrative quote
Benefits from the virtual community	“Having the opportunity to exchange ideas with other learners helps a lot in the process, as you learn from them also and share ideas and feelings.”
Appeal of content	“I am very impressed with the course and extra PDF files and videos. From the first week, I was making changes in my life.”
Enablers to participation and implementation	“Mindfulness is so helpful right now. I had a sense of how that would be but it’s only when you actually study and practice mindfulness at a time when all around you is noise, panic, and confusion that you truly realize its importance.”
Benefits noted in oneself	“Once again, many thanks for this course as it’s helped me to ease my worries working with COVID-19 patients. I feel more relaxed and less anxious. Worry is a just a thought. With meditation, I can change my thoughts for a positive one. And being connected with the present helps me to perform better in my job.”

### Benefits from the Virtual Community

Participants appreciated and realized positive outcomes from feeling part of and interacting with members of the course’s virtual community. This included feeling guided and engaged due to mentor feedback, feeling supported, experiencing comradery with fellow participants, and growing by virtue of

reflective interactions with other learners. This theme highlights principles of acceptability and adoption as well as technology and function within the evaluation framework [23], with 4 subthemes identified.

### **Mentor and Educator Feedback**

Mentoring was valuable as it helped stay on the right path, with individual concerns addressed and attended to. Educator feedback also enhanced the ability of the course to be engaging as the weekly feedback videos imparted a sense of real-time interaction.

*As ever, the “live” aspect of this course keeps me engaged.... I look forward to Craig and Richard’s feedback videos.*

### **Shared Experiences**

Through the online interactions with fellow learners who shared similar views, participants felt connected. The course also provided a platform for participant challenges to be validated, and they could find solace in that fellow learners were experiencing the same.

*It’s been especially meaningful to connect with like-minded souls across the world, with all of us sharing awareness and compassion.*

### **Support**

The support provided by educators, mentors, and fellow learners seemed to facilitate the practice of mindfulness taught in the course and aid in other challenging life situations.

*Thank you to everyone here: educators, mentors, and learners. Your knowledge and comments help go on with my practice, go through difficult times, and situations in my life.*

### **Growing Together**

Participants reported that through gaining insight from reading others’ views on the discussion forum and by sharing and reflecting on their own experiences, their learning was enhanced.

*Being able to read and comment on the “forum” is a great plus. Having the opportunity to exchange ideas with other learners helps a lot in the process, as you learn from them also and share ideas and feelings.*

### **Appeal of Content**

Theme 2, appeal of content, additionally reflected principles of acceptability, adoption, technology, and function [23]. Within this theme, 7 subthemes were identified in which participants noted course content features that were appealing and facilitated engagement in practice and learning. These included multimodal content, optional additional learning materials, relatable life examples, and clear and humorous course delivery.

### **Variety of Format**

The multimodal nature of the resources helped sustain participant interest in the material, continue to complete the course, and implement mindfulness in practice.

*I am very impressed with the course and extra PDF files and videos. From the first week I was making changes in my life.*

### **Optional Content**

Emphasis on autonomy and ability to choose the amount, type, and extra content they engaged with was valued.

*I would say the entire course is a gem! The way it gave short, open insights with the possibility of expansion and a view to more learning (videos, articles, abstracts, meditations).*

### **Unrestricted Access**

Unrestricted access to course content throughout the duration of the course was considered appealing as it enhanced equitable access to all resources and satisfaction with the course and enabled participants to revisit lessons and thereby consolidate knowledge at a user-defined pace.

*The resources in this course and the way it is delivered has really been useful. I have been able to go back and revisit sections to consolidate my understanding.*

### **Visual Aids**

The provision of visual aids in the course improved engagement, accessibility, and understanding, especially for people who did not speak English as their first language.

*I am Spanish-speaking and being able to study in English, at my own pace, in a relaxed tone and with visual help in the form of videos is a great experience for me.*

### **Life Examples**

The types of examples and approaches provided by the educators were relatable to participants and led to feelings that the content was personalized and applicable to their own life.

*It felt like, at times, that this course was designed specifically for me. It felt like I was having a conversation about my life.*

### **Ease of Implementation**

Mindfulness practices were appealing as they were easily implementable. Practices varied in duration, and educators provided examples of mindfulness that were practical and applicable in daily life and activities. The variable duration of practices enhanced accessibility as participants could tailor practices to suit their circumstances.

*The practices are as short or as long as you want them to be giving you just enough time to change your mindset at that moment to help cope with stress or just the rest of your day.*

### **Delivery by Educators**

Participants appreciated the delivery of content through videos and feedback-aided learning, effective educator communication, and content delivered with humor.

*The educators are really good at communicating the material in an easy-to-understand way and make it interesting and easy to practice.*

## Enablers to Participation and Implementation

Theme 3 encompassed aspects of the course that facilitated continued participation in the course and implementation of mindfulness practice, aligning with principles of accessibility, acceptability, and adoption [23]. Five subthemes were identified: ability to work through content at one's own pace, free access to course content, option to repeat course for consolidation of learning, evidence-based teaching, and new perspectives of mindfulness gained.

### Own Pace

Self-paced learning was a key enabler for participation and thus acceptability of the course. The flexible form of learning enabled accessibility, removing potential barriers of different learning speeds and capabilities, time availability, and motivation.

*It was at times hard for me to keep up here. I started late, and then COVID-19 hit us all with a blast. But I've made it!*

### Free Access

Free access to the course was important for accessibility, particularly relevant and valued during times of hardship. During the COVID-19 pandemic, emotional and financial distress were widespread across societies; having free access to resources was an enabler for participation in the course and implementation of mindfulness practice.

*Thank you for making the resources available for us for free. A bit more than 4 weeks into quarantine, jobless, and lost a few family members, and I know this course helped me going through all of that.*

### Option to Repeat Course

Participants reported that reenrolling to repeat the course was useful for implementation of mindfulness practice as it allowed reestablishment and building of knowledge and skills learned previously and revived motivation for practice.

*I have done this course before. I find it really useful to refresh and motivate myself. As I cannot take it all in during the 4 weeks, a revisit allows me to remind myself of what I have learned previously and to build on my understanding.*

### Evidence-Based Teaching

Delivery of evidence-based content and knowledge allowed for increased acceptability of the practices, enhanced motivation to adopt mindfulness, and increased participant ease.

*The gem for me was the evidence re mindfulness practice and improved cognitive ability, improved health outcomes, and improved quality of life. This wealth of evidence is helpful to maintain motivation for mindfulness practice.*

### New Perspectives

Through learning and practicing the course content, participant perspectives on the value of mindfulness practice for embracing and managing life challenges were either reinforced or new perspectives revealed.

*Mindfulness is so helpful right now. I had a sense of how that would be but it's only when you actually study and practice mindfulness at a time when all around you is noise, panic, and confusion that you truly realize its importance.*

## Benefits Noted in Oneself

Theme 4 aligned with the evaluation framework principles of safety and quality, health outcomes, and acceptability and adoption [23]. The 5 subthemes revealed positive outcomes for improved knowledge, stress management, positive behavior change, and improved well-being.

### Knowledge Gained

Participants became aware of certain habitual behaviors. Through course explanations provided regarding these behaviors and why they may be ineffective, participants gained skills that they could use to recognize and change their behaviors.

*Uni task. This section was the one that resonated most with me. I was happy in the belief that I could multitask, so it was when Craig and Richard were explaining what actually happens that I recognized some of the issues and barriers I face on a regular basis. I am going to make an effort to focus on just one task at a time.*

### Stress Management

The course taught participants practical skills of how to be mindful in different situations and environments that could be applied long term. Their immediate implementation of mindfulness resulted in identifiable positive outcomes including managing stress, changing thoughts from negative to positive, and improved performance.

*Once again, many thanks for this course as it's helped me to ease my worries working with COVID-19 patients. I feel more relaxed and less anxious. Worry is a just a thought. With meditation I can change my thoughts for a positive one. And being connected with the present helps me to perform better in my job.*

### Behavioral Change

Participants acknowledged that being mindful in performing daily tasks and actions helped them to focus, understand and enjoy life, and achieve more. They noted changes in behavior with an ability to regulate demands on self and others, let go of negative emotions, and accept limitations to maintain well-being.

*Perception, letting go, acceptance, and presence of mind are now my mantra. I've learned to be less demanding of myself and others, this does not mean that I have low expectations, it simply signifies that I'm more realistic about what to expect. I've also learned to let go of those harmful feelings of rage, anger, and disappointment. Being mindful, for me, also taught me that that's ok to say no, to respect and accept my limitations, and to safeguard my well-being.*

### ***Mind and Body Benefits***

Daily mindfulness resulted in benefits to participants' mental and physical well-being, with many reporting feelings of positivity, as well as improved concentration, sleep, and fitness, reflecting improved health outcomes.

*This course has made me feel more positive, I am loving the meditations, and it has helped me improve my fitness and focus on the here and now.*

### ***Increased Ability to Cope***

Improved coping skills were described as a positive outcome, gained through learning concepts of perception, letting go, acceptance, and presence of mind, which could be understood and predicted to be applicable to manage strong emotions arising in the future.

*This course was timely for me as it helped me push through the challenge, progress, and pick me up during the pandemic shutdown and work from home experience.*

## ***Discussion***

### **Principal Findings**

High participant attrition and low adherence to practice are common in digital health interventions [4,18,19]. Our qualitative study of participant feedback from a web-based mindfulness course identified 4 key elements that aided course completion and implementation of teachings: (1) benefits from the virtual community, (2) appeal of content, (3) enablers to participation and (4) benefits noted in oneself. Together the themes embodied safety and quality, accessibility, acceptability and adoption, health outcomes, and technology and function aspects of an evaluation framework [23], highlighting the effectiveness of the MWPP as a digital health intervention.

### **Comparison to Prior Work**

Various challenges have been identified in creating effective digital behavior change interventions (DBCIs), particularly in relation to understanding and promoting engagement [17]. These challenges include developing DBCIs that are person-centered and iterative and meet user requirements and establishing what constitutes effective engagement for a DBCI [25]. Through course elements like mentoring and feedback videos, a very person-centered, iterative, and supportive approach for the individual and the learning community were important elements in the success and popularity of the MWPP course. Further, engagement with an online learning platform does not necessarily translate into engagement with behavior change, nor does it mean that short-term behavior changes as a part of the program will translate into longer term healthy behaviors [26]. Mixed method approaches including qualitative data provide important insights into why participants find a DBCI effective or ineffective [27].

Benefits gained from a virtual community were reported to be achieved through interaction with mentors and fellow learners in the discussion forums and through educator feedback. Our findings align with prior studies showing support provided by therapists and other participants are significant factors in the

experience of participating in MBIs [4]; support through existence of another person or social presence can influence accountability and thus adherence to the course [28]. Other qualitative studies on mindfulness reflect the notion of group support to be valuable [14], with theoretical research suggesting it is the strongest predictor of behavior change [28]. It is interesting to note that comments in this theme reflect perceived support, with research showing this perception enhances coping and self-esteem [28]. A sense of connectedness and validation of feelings by sharing experiences and gaining and reflecting on learning and practice together were novel concepts for online interventions. These suggest the value of group interaction and discussion for enhancing motivation for continued learning, consolidation of teachings, and confidence in implementation of practice.

The multimodal and optional course content were valued and conducive to course completion, in accord with previous studies showing that appeal of intervention features improves the quality of the user experience and has strong influence on initiating and sustaining user engagement [29]. Unlimited and free access to content, flexibility of participation in the learning materials, and access to the benefits of discussion forums and educator feedback were also identified as aiding in course completion. Previous research has shown time constraints to be one of the key barriers to engaging in MBIs, and reducing this through features such as optional additional resources grants autonomy over amount of time invested and is favorable to participants [30]. Novel aspects identified in course content were inclusion of visual aids, which enabled accessibility to nonnative English speakers, and easy-to-understand delivery of concepts by educators. This provides valuable insight for a means to mitigate other potential language, reading, and listening barriers. Content that aided in implementation of mindfulness practice were provision of multiple options and providing life examples with which learners could identify.

Enablers to participation and implementation included provision of evidence-based content, which has been shown in prior research to improve outcomes by enabling acceptance in academically minded groups [31]. Self-pacing was a valued aspect of this course, and this is supported by the theory of self-determination—namely, that self-efficacy and autonomy increase successful implementation [32]. Novel findings within this theme were participants reporting repetition of the course provided motivation and enabled reestablishment and maintenance of skills and participants preconceptions of mindfulness changing because of course content. These aspects may be important for sustained long-term adherence to behavior change.

Benefits noted in oneself identified an increase in self-regulation (consciously managing emotions and behaviors), which is a recognized change induced by mindfulness meditation [33]. There is strong evidence shown in the literature that self-regulation facilitates behavior change as it improves participants' innate motivation and focus, which can assist in retaining and efficiently learning material leading to lasting changes [33].



## Strengths and Limitations

Limitations of the study are acknowledged. First, of participants who commenced the MWPP course, only 18% completed at least 90%. Reasons for and timing of dropout from the course were not collected, although for MOOCs, an 18% completion rate is quite high. These details, as well as information on motivations for commencing the course, may provide further insight to key components for participant retention. Second, of the MWPP course completers, 16% contributed to the data. Our results may not be representative of all course completers; however, data from a large sample are not crucial for qualitative analyses [34]. Methods to increase participant feedback may include mandatory survey completion at intervals to progress to subsequent sections and financial incentives for survey completion. A third limitation noted is the free-text nature of the feedback question, which directed participants to specifically comment on positive aspects of the course. Reframing the feedback question to query positive and negative aspects as well as course components that facilitated and deterred completion and implementation of teachings may reveal greater depth of knowledge.

The strengths of our study include the data being drawn from a diverse and significant sample of engaged participants, many

of whom were experienced with the course. Thematic analysis allowed for a range of relevant issues to be highlighted.

## Future Directions

Future research should focus on understanding motivations for participation in mindfulness and identifying effective methods for participant retention and behavior change techniques to motivate long-term adherence to healthy behaviors.

This may include research on behavior change techniques to motivate long-term adherence of healthy behaviors. Together, these questions may address challenges in the design of effective digital interventions for self-managed improved well-being.

For future research, the established mindfulness techniques used in this course can be explored to allow for formal reporting and use in other behavior change interventions.

## Conclusion

Supportive community, appealing content, enablers to participation and implementation, and positive self-benefits were identified as core aspects aiding in completion of an online mindfulness course and implementation of teachings. These insights may aid future digital health intervention designers to optimize elements that lead to optimal learning and effective implementation of health behaviors.

## Acknowledgments

The authors gratefully acknowledge participant responders and the contribution of Liz Brnjak as a Mindfulness for Well-being and Peak Performance mentor. The authors received no financial support for the research. Salaries of SN and NN and partial fees for open access were funded by philanthropic gifts to the Neuroepidemiology Unit from Mr Wal Piscoitta and other anonymous donors.

## Authors' Contributions

NN and SN were responsible for conceptualization, methodology, and data curation. NN, MY, and SN performed the formal analysis. CH, RC, and SC conducted the investigation. CH provided resources. NN and MY wrote the original draft of the manuscript. NN, RC, CH, SN, and SC reviewed and edited the manuscript. NN was responsible for visualization, supervision, and project administration. All authors have read and agreed to the published version of the manuscript.

## Conflicts of Interest

CH and RC are educators and creators of the Mindfulness for Well-being and Peak Performance course, and SC facilitated the discussion forums.

## References

1. Mrazek AJ, Mrazek MD, Cherolini CM, Cloughesy JN, Cynman DJ, Gougis LJ, et al. The future of mindfulness training is digital, and the future is now. *Curr Opin Psychol* 2019 Aug;28:81-86. [doi: [10.1016/j.copsyc.2018.11.012](https://doi.org/10.1016/j.copsyc.2018.11.012)] [Medline: [30529975](https://pubmed.ncbi.nlm.nih.gov/30529975/)]
2. Taylor H, Strauss C, Cavanagh K. Can a little bit of mindfulness do you good? A systematic review and meta-analyses of unguided mindfulness-based self-help interventions. *Clin Psychol Rev* 2021 Nov;89:102078. [doi: [10.1016/j.cpr.2021.102078](https://doi.org/10.1016/j.cpr.2021.102078)] [Medline: [34537665](https://pubmed.ncbi.nlm.nih.gov/34537665/)]
3. Sancho M, De Gracia M, Rodríguez RC, Mallorquí-Bagué N, Sánchez-González J, Trujols J, et al. Mindfulness-based interventions for the treatment of substance and behavioral addictions: a systematic review. *Front Psychiatry* 2018;9:95 [FREE Full text] [doi: [10.3389/fpsy.2018.00095](https://doi.org/10.3389/fpsy.2018.00095)] [Medline: [29651257](https://pubmed.ncbi.nlm.nih.gov/29651257/)]
4. Murray E. Web-based interventions for behavior change and self-management: potential, pitfalls, and progress. *Med* 2012;1(2):e3 [FREE Full text] [doi: [10.2196/med20.1741](https://doi.org/10.2196/med20.1741)] [Medline: [25075231](https://pubmed.ncbi.nlm.nih.gov/25075231/)]
5. Christensen H. Increasing access and effectiveness: using the internet to deliver low intensity CBT. In: Bennett-Levy J, Richards D, Farrand P, Christensen H, Griffiths K, Kavanagh D, editors. *Oxford Guide to Low Intensity CBT Intervention*. Oxford: Oxford University Press; 2010.

6. Rochlen AB, Zack JS, Speyer C. Online therapy: review of relevant definitions, debates, and current empirical support. *J Clin Psychol* 2004 Mar;60(3):269-283. [doi: [10.1002/jclp.10263](https://doi.org/10.1002/jclp.10263)] [Medline: [14981791](https://pubmed.ncbi.nlm.nih.gov/14981791/)]
7. Mheidly N, Fares MY, Fares J. Coping with stress and burnout associated with telecommunication and online learning. *Front Public Health* 2020 Nov 11;8:574969 [FREE Full text] [doi: [10.3389/fpubh.2020.574969](https://doi.org/10.3389/fpubh.2020.574969)] [Medline: [33262967](https://pubmed.ncbi.nlm.nih.gov/33262967/)]
8. Santomauro D, Mantilla Herrera A, Shadid J, Zheng P, Ashbaugh C, Pigott D. Global prevalence and burden of depressive and anxiety disorders in 204 countries and territories in 2020 due to the COVID-19 pandemic. *Lancet* 2021 Nov 06;398(10312):1700-1712 [FREE Full text] [doi: [10.1016/S0140-6736\(21\)02143-7](https://doi.org/10.1016/S0140-6736(21)02143-7)] [Medline: [34634250](https://pubmed.ncbi.nlm.nih.gov/34634250/)]
9. Torales J, O'Higgins M, Castaldelli-Maia JM, Ventriglio A. The outbreak of COVID-19 coronavirus and its impact on global mental health. *Int J Soc Psychiatry* 2020 Jun 31;66(4):317-320. [doi: [10.1177/0020764020915212](https://doi.org/10.1177/0020764020915212)] [Medline: [32233719](https://pubmed.ncbi.nlm.nih.gov/32233719/)]
10. Liu Y, Li I, Hsiao F. Effectiveness of mindfulness-based intervention on psychotic symptoms for patients with schizophrenia: a meta-analysis of randomized controlled trials. *J Adv Nurs* 2021 Jun;77(6):2565-2580. [doi: [10.1111/jan.14750](https://doi.org/10.1111/jan.14750)] [Medline: [33450107](https://pubmed.ncbi.nlm.nih.gov/33450107/)]
11. Scott-Sheldon LAJ, Gathright EC, Donahue ML, Balletto B, Feulner MM, DeCosta J, et al. Mindfulness-based interventions for adults with cardiovascular disease: a systematic review and meta-analysis. *Ann Behav Med* 2020 Jan 01;54(1):67-73 [FREE Full text] [doi: [10.1093/abm/kaz020](https://doi.org/10.1093/abm/kaz020)] [Medline: [31167026](https://pubmed.ncbi.nlm.nih.gov/31167026/)]
12. Galante J, Friedrich C, Dawson AF, Modrego-Alarcón M, Gebbing P, Delgado-Suárez I, et al. Mindfulness-based programmes for mental health promotion in adults in nonclinical settings: a systematic review and meta-analysis of randomised controlled trials. *PLoS Med* 2021 Jan;18(1):e1003481 [FREE Full text] [doi: [10.1371/journal.pmed.1003481](https://doi.org/10.1371/journal.pmed.1003481)] [Medline: [33428616](https://pubmed.ncbi.nlm.nih.gov/33428616/)]
13. Goldberg SB, Riordan KM, Sun S, Davidson RJ. The empirical status of mindfulness-based interventions: a systematic review of 44 meta-analyses of randomized controlled trials. *Perspect Psychol Sci* 2021 Feb 16. [doi: [10.1177/1745691620968771](https://doi.org/10.1177/1745691620968771)] [Medline: [33593124](https://pubmed.ncbi.nlm.nih.gov/33593124/)]
14. Klatt M, Bawa R, Gabram O, Westrick A, Blake A. Synchronous mindfulness in motion online: strong results, strong attendance at a critical time for health care professionals (HCPs) in the COVID era. *Front Psychol* 2021 Aug 18;12:725810 [FREE Full text] [doi: [10.3389/fpsyg.2021.725810](https://doi.org/10.3389/fpsyg.2021.725810)] [Medline: [34489830](https://pubmed.ncbi.nlm.nih.gov/34489830/)]
15. Zhang Y, Xue J, Huang Y. A meta-analysis: internet mindfulness-based interventions for stress management in the general population. *Medicine (Baltimore)* 2020 Jul 10;99(28):e20493. [doi: [10.1097/MD.00000000000020493](https://doi.org/10.1097/MD.00000000000020493)] [Medline: [32664060](https://pubmed.ncbi.nlm.nih.gov/32664060/)]
16. Barak A, Klein B, Proudfoot JG. Defining internet-supported therapeutic interventions. *Ann Behav Med* 2009 Aug;38(1):4-17. [doi: [10.1007/s12160-009-9130-7](https://doi.org/10.1007/s12160-009-9130-7)] [Medline: [19787305](https://pubmed.ncbi.nlm.nih.gov/19787305/)]
17. 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/)]
18. Torous J, Lipschitz J, Ng M, Firth J. Dropout rates in clinical trials of smartphone apps for depressive symptoms: a systematic review and meta-analysis. *J Affect Disord* 2020 Feb 15;263:413-419. [doi: [10.1016/j.jad.2019.11.167](https://doi.org/10.1016/j.jad.2019.11.167)] [Medline: [31969272](https://pubmed.ncbi.nlm.nih.gov/31969272/)]
19. Ryan C, Bergin M, Wells JS. Theoretical perspectives of adherence to web-based interventions: a scoping review. *Int J Behav Med* 2017 Jul 20;25(1):17-29. [doi: [10.1007/s12529-017-9678-8](https://doi.org/10.1007/s12529-017-9678-8)] [Medline: [28730402](https://pubmed.ncbi.nlm.nih.gov/28730402/)]
20. Bartlett L, Buscot M, Bindoff A, Chambers R, Hassed C. Mindfulness is associated with lower stress and higher work engagement in a large sample of MOOC participants. *Front Psychol* 2021 Sep 10;12:724126 [FREE Full text] [doi: [10.3389/fpsyg.2021.724126](https://doi.org/10.3389/fpsyg.2021.724126)] [Medline: [34566805](https://pubmed.ncbi.nlm.nih.gov/34566805/)]
21. 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)]
22. Neate SL, Reece JC, Hassed C, Chambers R, Connaughton S, Nag N. A qualitative analysis of free text comments of participants from a massive open online mindfulness course. *Front Public Health* 2022 Aug 12;10:947898 [FREE Full text] [doi: [10.3389/fpubh.2022.947898](https://doi.org/10.3389/fpubh.2022.947898)] [Medline: [36033817](https://pubmed.ncbi.nlm.nih.gov/36033817/)]
23. Evaluation resource guide: allied health telehealth capacity building project. Queensland Health. 2016. URL: [https://www.health.qld.gov.au/data/assets/pdf\\_file/0024/451077/telehealth-evalguide.pdf](https://www.health.qld.gov.au/data/assets/pdf_file/0024/451077/telehealth-evalguide.pdf) [accessed 2022-09-14]
24. Beck CT. Qualitative research: the evaluation of its credibility, fittingness, and auditability. *West J Nurs Res* 1993 Apr;15(2):263-266. [doi: [10.1177/019394599301500212](https://doi.org/10.1177/019394599301500212)] [Medline: [8470379](https://pubmed.ncbi.nlm.nih.gov/8470379/)]
25. Yardley L, Morrison L, Bradbury K, Muller I. The person-based approach to intervention development: application to digital health-related behavior change interventions. *J Med Internet Res* 2015;17(1):e30 [FREE Full text] [doi: [10.2196/jmir.4055](https://doi.org/10.2196/jmir.4055)] [Medline: [25639757](https://pubmed.ncbi.nlm.nih.gov/25639757/)]
26. Mohr DC, Schueller SM, Montague E, Burns MN, Rashidi P. The behavioral intervention technology model: an integrated conceptual and technological framework for eHealth and mHealth interventions. *J Med Internet Res* 2014 Jun;16(6):e146 [FREE Full text] [doi: [10.2196/jmir.3077](https://doi.org/10.2196/jmir.3077)] [Medline: [24905070](https://pubmed.ncbi.nlm.nih.gov/24905070/)]
27. Morrison LG, Hargood C, Lin SX, Dennison L, Joseph J, Hughes S, et al. Understanding usage of a hybrid website and smartphone app for weight management: a mixed-methods study. *J Med Internet Res* 2014 Oct;16(10):e201 [FREE Full text] [doi: [10.2196/jmir.3579](https://doi.org/10.2196/jmir.3579)] [Medline: [25355131](https://pubmed.ncbi.nlm.nih.gov/25355131/)]

28. Santarossa S, Kane D, Senn CY, Woodruff SJ. Exploring the role of in-person components for online health behavior change interventions: can a digital person-to-person component suffice? *J Med Internet Res* 2018 Apr 11;20(4):e144 [FREE Full text] [doi: [10.2196/jmir.8480](https://doi.org/10.2196/jmir.8480)] [Medline: [29643048](https://pubmed.ncbi.nlm.nih.gov/29643048/)]
29. Short C, Rebar A, Plotnikoff R, Vandelanotte C. Designing engaging online behaviour change interventions: a proposed model of user engagement. *Euro Health Psychol.*: 17; 2015. URL: [https://www.ehps.net/ehp/index.php/contents/issue/download/ehp.v17.i1/pdf\\_43](https://www.ehps.net/ehp/index.php/contents/issue/download/ehp.v17.i1/pdf_43) [accessed 2022-09-14]
30. Banerjee M, Cavanagh K, Strauss C. A qualitative study with healthcare staff exploring the facilitators and barriers to engaging in a self-help mindfulness-based intervention. *Mindfulness (NY)* 2017 May 25;8(6):1653-1664 [FREE Full text] [doi: [10.1007/s12671-017-0740-z](https://doi.org/10.1007/s12671-017-0740-z)] [Medline: [29201248](https://pubmed.ncbi.nlm.nih.gov/29201248/)]
31. Moore S, Barbour R, Ngo H, Sinclair C, Chambers R, Auret K, et al. Determining the feasibility and effectiveness of brief online mindfulness training for rural medical students: a pilot study. *BMC Med Educ* 2020 Apr 06;20(1):104 [FREE Full text] [doi: [10.1186/s12909-020-02015-6](https://doi.org/10.1186/s12909-020-02015-6)] [Medline: [32252750](https://pubmed.ncbi.nlm.nih.gov/32252750/)]
32. Ng JYY, Ntoumanis N, Thøgersen-Ntoumani C, Deci EL, Ryan RM, Duda JL, et al. Self-determination theory applied to health contexts: a meta-analysis. *Perspect Psychol Sci* 2012 Jul;7(4):325-340. [doi: [10.1177/1745691612447309](https://doi.org/10.1177/1745691612447309)] [Medline: [26168470](https://pubmed.ncbi.nlm.nih.gov/26168470/)]
33. Schuman-Olivier Z, Trombka M, Lovas DA, Brewer JA, Vago DR, Gawande R, et al. Mindfulness and behavior change. *Harv Rev Psychiatry* 2020;28(6):371-394 [FREE Full text] [doi: [10.1097/HRP.0000000000000277](https://doi.org/10.1097/HRP.0000000000000277)] [Medline: [33156156](https://pubmed.ncbi.nlm.nih.gov/33156156/)]
34. Marks D, Yardley L. *Research Methods for Clinical and Health Psychology*, First Edition. London: SAGE Publications Ltd; 2004.

## Abbreviations

- DBCI:** digital behavior change intervention  
**MBI:** mindfulness-based intervention  
**MOOC:** massive open online course  
**MWPP:** Mindfulness for Well-being and Peak Performance

*Edited by A Mavragani; submitted 20.02.22; peer-reviewed by C Ryan, C Tsou; comments to author 23.05.22; revised version received 21.07.22; accepted 02.09.22; published 05.10.22.*

*Please cite as:*

*Yadav M, Neate S, Hased C, Chambers R, Connaughton S, Nag N*

*Mining the Gems of a Web-Based Mindfulness Intervention: Qualitative Analysis of Factors Aiding Completion and Implementation*  
*JMIR Form Res* 2022;6(10):e37406

URL: <https://formative.jmir.org/2022/10/e37406>

doi: [10.2196/37406](https://doi.org/10.2196/37406)

PMID: [36197709](https://pubmed.ncbi.nlm.nih.gov/36197709/)

©Muskan Yadav, Sandra Neate, Craig Hased, Richard Chambers, Sherelle Connaughton, Nupur Nag. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 05.10.2022. 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

# Adapting Child Health Knowledge Translation Tools for Use by Indigenous Communities: Qualitative Study Exploring Health Care Providers' Perspectives

Sarah A Elliott<sup>1,2\*</sup>, PhD; Jason Kreutz<sup>1\*</sup>, BSc, MPP; Kelsey S Wright<sup>1</sup>, MSc; Sherri Di Lallo<sup>3</sup>, BScN, MN, RN; Shannon D Scott<sup>4</sup>, PhD, RN; Lisa Hartling<sup>1</sup>, PhD

<sup>1</sup>Alberta Research Centre for Health Evidence, Department of Pediatrics, Faculty of Medicine and Dentistry, University of Alberta, Edmonton, AB, Canada

<sup>2</sup>Cochrane Child Health, University of Alberta, Edmonton, AB, Canada

<sup>3</sup>Stollery Awasisak Indigenous Health, Stollery Children's Hospital, Edmonton, AB, Canada

<sup>4</sup>Faculty of Nursing, University of Alberta, Edmonton, AB, Canada

\*these authors contributed equally

**Corresponding Author:**

Sarah A Elliott, PhD

Alberta Research Centre for Health Evidence, Department of Pediatrics

Faculty of Medicine and Dentistry

University of Alberta

ECHA 4-472, 11405 87 Ave NW

Edmonton, AB, T6G 1C9

Canada

Phone: 1 780 492 1241

Email: [se2@ualberta.ca](mailto:se2@ualberta.ca)

## Abstract

**Background:** Our research groups have developed a number of parental knowledge translation (KT) tools to help families understand common childhood illnesses and make informed decisions regarding when to seek urgent care. We have developed a series of videos to help parents understand how to manage common acute childhood illnesses at home and when to contact emergency health care services. It is unclear whether the videos in their current form and language are useful for a wider range of populations, including Indigenous groups.

**Objective:** The purpose of this study was to explore whether and understand how our KT tools could be adapted for use with Indigenous communities.

**Methods:** Health care providers (HCPs) serving Indigenous families in Alberta, Canada, were asked to review 2 of our KT tools (one on croup and one on acute otitis media), complete a demographic survey, and participate in a one-on-one semistructured interview. HCPs were asked to reflect on the usability of the KT tools within their practice and what cultural adaptation considerations they felt would be needed to develop KT tools that meet the needs of Indigenous clients. Audio recordings from the interviews were transcribed verbatim and analyzed for relevant themes using thematic analysis.

**Results:** A total of 18 HCPs (n=15, 83% women and n=3, 17% men) from various health professions (eg, physician, registered nurse, and licensed practical nurse) were interviewed. Of these 18 HCPs, 7 (39%) self-identified as Indigenous. Four overarching themes were identified as important when considering how to adapt KT tools for use by Indigenous communities: accessibility, reliability, KT design, and relationship building. Access to tangible resources and personal and professional connections were considered important. Accessibility affects the types of KT tools that can be obtained or used by various individuals and communities and the extent to which they can implement recommendations given in those KT tools. In addition, the extent to which users relate to the depictions and content within KT tools must be considered. The environments, portrayals of characters, and cultural norms and values presented within KT tools should be relevant to users to increase the reliability and uptake of recommendations. Most importantly, fostering genuine and sustainable relationships with users and communities is a vital consideration for KT tool developers.

**Conclusions:** These findings serve to cultivate a greater understanding of the various components that HCPs consider important when developing or culturally adapting KT tools for use by Indigenous families. This information will help support the effective

adaptation and distribution of KT tools for use by a broad audience. Careful consideration of the themes identified in this study highlights the importance of working together with the knowledge users (health care consumers) when developing KT tools.

(*JMIR Form Res* 2022;6(10):e36353) doi:[10.2196/36353](https://doi.org/10.2196/36353)

## KEYWORDS

knowledge translation; culture; Indigenous culture; child health; adaptation

## Introduction

### Background

Effective synthesis, dissemination, and implementation of health research findings is critical for informing patient decision-making, promoting the effective use of health care system resources, and, ultimately, improving health outcomes [1]. Over the past few years, this process known as knowledge translation (KT) has been increasingly directed toward those accessing health care (ie, patients and parents) to foster an active involvement in informed decision-making [2]. Our research group strives to improve health outcomes for children by facilitating the uptake of evidence by actively engaging health care providers (HCPs) and families in the codevelopment and dissemination of a variety of user-friendly KT tools. We integrate the best available evidence (systematic reviews [3]) with parent experiences (qualitative interviews [4]) to create a successful KT model centered within the knowledge-to-action [5] cycle. We are now exploring within local contexts how best to tailor our KT products to diverse audiences.

Our research team has developed a series of web-based videos to help parents understand how to manage common acute childhood illnesses at home and when to contact health care services. These videos have been reviewed by our pediatric parent advisory group [6]; vetted by multidisciplinary, pediatric emergency health care professionals; tested for usability by parents in remote, rural, and urban settings [7,8]; and released on the web for the English-speaking general population (freely accessible at ECHO research [9]). However, it is unclear whether these videos, in their current form, will be useful for a wider range of population.

The reliability and accessibility of KT tools require developers to understand the contexts and realities in which their viewers find themselves [5,10]. Although KT tools are meant to present health care information in user-friendly and relatable formats, they are often developed with the majority cultural communities in mind (eg, in North America, this is predominantly White, English-speaking populations). Culturally adapting KT tools for use by diverse groups offers a great potential for decreasing health disparities and increasing access to resources.

In Canada, Indigenous populations face a plethora of health disparities and unique barriers when accessing health information [11,12]. This can include a lack of access to health resources and the discontinuity of care for families, which can directly affect child health outcomes [13,14]. Unfortunately, most existing KT initiatives may not effectively engage Indigenous populations or incorporate Indigenous knowledge and practices [15-17]. In addition, there is a lack of information

in the literature on how best to implement, practice, and appraise KT initiatives for Indigenous populations [18].

### Objective

We sought to understand whether and how our KT tools could be adapted for use by Indigenous communities. Although frameworks to achieve cultural adaptation for health promotion programs and health interventions have been documented [19,20], guidance on how best to adapt KT products for use by culturally and linguistically diverse audiences is lacking. As the first step, we explored the opinions of HCPs who work with Indigenous communities regarding the applicability of 2 of our KT tools for use by the families they serve. We aimed to gather their perspectives on the usefulness and appropriateness of the tools as well as considerations for how best to adapt or develop KT tools in such a way that they are relatable and accessible to Indigenous communities.

## Methods

### Stakeholder Engagement

Within our province, there are 45 First Nations in 3 treaty areas, representing culturally distinct and traditionally underserved Indigenous communities [21]. We engaged with several local stakeholders (researchers, clinicians, program managers, and Indigenous community members) working in Indigenous health and research to understand what approach would work best. Those early conversations helped develop the research plan and facilitated subsequent engagement with HCPs who serve Indigenous families and communities in Alberta, Canada.

### Overview

Semistructured one-on-one web-based interviews were conducted with HCPs who serve Indigenous communities to discuss 2 existing child health KT videos (on croup and acute otitis media [AOM]; [Multimedia Appendix 1](#)). The videos were originally produced for the general English-speaking population. The participants were also asked to complete a short demographic survey.

### Sampling and Recruitment

Participants were eligible for the study if they were a practicing HCP serving Indigenous families in Alberta and could read and speak English. HCPs could be any health care personnel or professional interacting with Indigenous patients and families and providing care (eg, physician, nurse, or midwife). Participants were recruited from a variety of disciplines, including provincial health authorities and those working in not-for-profit organizations, local hospitals or care units, and Indigenous organizations, through a targeted recruitment effort using professional contacts and snowball sampling [22].

Purposive sampling was used to generate an in-depth understanding of the HCPs' reflections [23]. Sample size was determined by evidence of data saturation, which was monitored through concurrent analysis of the data to assess the comprehensiveness, variation, and richness of the interviews [24].

### **Ethics Approval**

Ethics approval for this study (Pro00085188) was received by the appropriate Institutional Review Board offices at the University of Alberta. Before the interviews, all participants were informed of the study and the study's purpose through recruitment emails during the screening for eligibility process and at the start of the interview sessions. All participants provided informed consent before any data were collected. The anonymity, confidentiality, and secure storage of the data were guaranteed to the participants in both the written consent form and before each interview and respected. A meal or electronic gift voucher to the value of CAD \$25 (US \$19) was provided to all the participants before commencing the interview in recognition of their time.

### **Study Components**

#### ***Demographic Questionnaire***

The participants completed a demographic questionnaire via a secure web-based survey platform, REDCap (Research Electronic Data Capture; Vanderbilt University) [25]. The questionnaire asked the participants about their clinical or community health care practice history and setting, their Indigenous affiliation, and standard demographic information such as age and sex.

#### ***Semistructured Interview***

Our research team initially developed the interview questions before field testing. The questions were then adapted over the course of 3 nonrecorded interviews with in-house volunteers (HCPs, researchers, and parents). The first part of the interview consisted of questions probing into the existing information sources and KT tools that were used by the HCPs and the patients they served. The second part of the interview consisted of questions specifically about the content and format of the croup and AOM videos. The semistructured nature of the interview schedule allowed flexibility for the interviewer (JK) to explore the most meaningful experiences for the participants. The participants were treated as experts of their own experience and had full autonomy over the information they shared.

Interviews were audio recorded and then transcribed verbatim by a third-party transcriptionist as data were collected. The transcripts were reviewed as the interviews were completed to ensure accurate interpretation and complete inclusion of the data collected in the interviews.

### **Data Analysis**

Data collection and analysis were conducted concurrently to facilitate a more focused and meaningful data collection. Data management and analysis were facilitated using NVivo software (version 12, QSR International). Transcripts were read several times, coded, and analyzed using thematic analysis to identify common themes across the interviews. The thematic analysis process followed was outlined by Braun and Clark [26] and entailed familiarization with the data, initial coding, grouping similar codes together, and the development of themes and subthemes. Themes were then named and described. The interviewer (JK) regularly debriefed the study team members SAE and LH. Iterative data collection and analysis continued until data saturation was reached.

A line-by-line approach was used during the coding process. An inductive approach was used where codes emerged from the data [27]. Data were coded and sorted based on the themes that developed throughout the analysis. A focused coding approach in which similar codes were grouped together was used to elucidate patterns during data analysis. Analytical rigor and trust were promoted through continual communication with the research team and a detailed study log. Interview recordings and detailed field notes promoted confirmability of the findings. Field notes promoted reflexivity, allowing for the acknowledgment of bias, transferability, dependability, credibility, and inductive research praxis [28].

A second researcher (SAE or KSW) re-examined the transcripts and the codes to mitigate interpretive bias. All concerns were discussed until a consensus was reached.

The description and interpretation of themes were then reviewed by a fourth research team member who self-identified as Indigenous (SDL) to ensure that the language, meaning, and context of the findings were appropriately reflective of Indigenous cultures and the needs faced by Indigenous families within the scope of this study.

## **Results**

### **Study Participants**

A total of 18 HCPs from various health professions (eg, physician, registered nurse, and licensed practical nurse) participated in the study. Participant demographics are presented in [Table 1](#). All the participants were HCPs who served Indigenous families or communities; 83% (15/18) of the participants self-identified as female, and 39% (7/18) of the participants self-identified as Indigenous. The HCPs served Indigenous patients in rural communities (4/18, 22%), urban communities (6/18, 33%), or a mixture of both (8/18, 44%). The participants had, on average, 11 (SD 9) years of experience working with Indigenous communities.

**Table 1.** Demographic characteristics of the study participants (N=18).

Variable	Participants, n (%)
<b>Sex</b>	
Male	3 (17)
Female	15 (83)
<b>Age (years)</b>	
21-30	5 (28)
31-40	3 (17)
41-50	3 (17)
51-60	4 (22)
31-70	3 (17)
<b>Indigenous descent</b>	
First Nations	4 (22)
Métis	3 (17)
Inuit	0 (0)
Non-Indigenous	11 (61)
<b>Level of education</b>	
Bachelor of Arts (BA)	3 (17)
Master of Science (MSc); Doctor of Philosophy (PhD)	2 (11)
Nurse Practitioner (NP); Registered Nurse (RN)	5 (28)
Doctor of Medicine (MD)	8 (44)
<b>Practice setting</b>	
Private or community clinic	5 (28)
Hospital	8 (44)
Mixed or other	5 (28)
<b>Locale</b>	
Rural	4 (22)
Urban	6 (33)
Both	8 (44)
<b>Working with Indigenous populations (years)</b>	
<5	7 (39)
5-10	4 (22)
10-20	2 (11)
≥20	5 (28)

<sup>a</sup>BA: Bachelor of Arts.

<sup>b</sup>MSc: Master of Science.

<sup>c</sup>PhD: Doctor of Philosophy.

<sup>d</sup>NP: Nurse Practitioner.

<sup>e</sup>RN: Registered Nurse.

<sup>f</sup>MD: Doctor of Medicine.

### Considerations for Adapting and Developing KT Tools for Indigenous Communities

The participants spoke about various considerations needed during the development or adaptation of KT tools to ensure they were relevant to the Indigenous communities they served. In

total, 4 main themes were identified, and a summary of these themes can be found in [Table 2](#). The participants described how addressing (1) accessibility, (2) relatability, (3) KT design, and (4) relationship building could produce more effective, culturally relevant, and meaningful KT tools.

**Table 2.** Descriptions and quotes of the identified themes and subthemes.

Themes and subthemes	Descriptions	Theme exemplars
<b>Accessibility</b>		
<ul style="list-style-type: none"> <li>Tangible resources</li> <li>Relationships</li> </ul>	<ul style="list-style-type: none"> <li>Tangible resources refer to the physical resources, or a lack thereof, that a caregiver or community has access to. This can include access to the internet, clean water, medication, medical devices, and transportation.</li> <li>Relationships refer to having consistent or easy access to HCPs<sup>a</sup>, services, and community connections.</li> </ul>	<ul style="list-style-type: none"> <li>“...if you were looking at antibiotics and like try to look at an antibiotic that while maybe if they’re like extremely rural and in the bush and they don’t have access to a refrigerator like you wouldn’t want to give them something that had to be refrigerated, right, like trying to take into consideration those little the tiny details in regards to lifestyle, where they’re from, what kind of life they live I guess”</li> <li>“not a lot of people have Wi-Fi because we’re very rural”</li> <li>“...often times families don’t have thermometers at home”</li> <li>“...accessing healthcare isn’t as easy, especially if they don’t have a physician or a nurse in their community...”</li> <li>“You have to weigh the pros and cons about going [to seek care]. Is it that severe that I have to go and wake up a nurse in the middle of the night? I think talking to rural communities and seeing their perspective on things and getting like an idea from a kid about how they actually access the healthcare services”</li> </ul>
<b>Relatability</b>		
<ul style="list-style-type: none"> <li>Environmental contexts</li> <li>Individual representation</li> <li>Culture</li> </ul>	<ul style="list-style-type: none"> <li>Environmental contexts refer to the need for relatable scenery and environments, such as depicting characters from a remote community.</li> <li>Individual representation refers to having characters that resemble the target audience in terms of appearance, attire, and customs.</li> <li>Culture refers to integrating cultural practices and dynamics into KT<sup>b</sup> tools and ensuring that these tools are culturally sensitive.</li> </ul>	<ul style="list-style-type: none"> <li>“...tailoring it more towards Indigenous culture, different languages, and incorporating Indigenous art might be helpful...”</li> <li>“I thought was not very Indigenous-friendly based on the lifestyle and the settings that most First Nations and Métis and Inuit people live”</li> <li>“Wow, I thought the second video actually looked maybe more visually that people would be an Indigenous kind of people. Maybe that would be more relatable if the people looked more like you. The other piece on that is that you might have a grandmother or somebody involved in the picture, which might also make it a little bit more relatable...”</li> <li>“...a lot of people have like Indigenous bedding like so if you had a baby...it would be nice to have like a baby wrapped in a special blanket things like that that are like very obviously Indigenous”</li> <li>“Yeah, like a blanket on a wall. Like a lot of people have those really nice Pendleton blankets on the wall on their like office or like of their home”</li> </ul>
<b>KT design</b>		
<ul style="list-style-type: none"> <li>Format</li> <li>Plain language</li> <li>Translations</li> </ul>	<ul style="list-style-type: none"> <li>Format refers to the medium of a KT tool and the benefits and drawbacks that come with it; for example, video versus paper-based resource.</li> <li>Plain language includes avoiding medical jargon, using clear and simple wording, using active voice, avoiding excessive text, and using lower reading levels.</li> <li>Translations are important for ensuring that Indigenous patients and communities that speak Indigenous languages are able to use KT tools.</li> </ul>	<ul style="list-style-type: none"> <li>“...the more concise and brief, the more easily received...”</li> <li>“Videos are very short and concise and like even with the new videos that are coming out with AHS in regards to emergency and like their kiddos at the hospital those are very clear and concise and are easy to kind of absorb and understand...”</li> <li>“I think there’s many resources that could be dwindled down on their information or make it more like clear and concise so it’s easy for them to kind of absorb and not feel frightened about the information...”</li> <li>“They should really be striving to make resources that are adapted to their patients if they can because not all parents are willing to read a 30-page document”</li> <li>“I definitely think that translating them into languages like Cree or Dené or Stoney if we’re talking all through Alberta not just Northern Alberta and Blackfoot all of the languages I think would be very, very valuable”</li> <li>“...a lot of the Elders, a lot of grandparents, a lot of the people that would be relied upon to have this information or who would best have this information are more comfortable in their Indigenous language. I mean, the younger people, definitely a lot of them are more comfortable”</li> </ul>



Themes and subthemes	Descriptions	Theme exemplars
<b>Relationship building</b>		
<ul style="list-style-type: none"> <li>Understanding community needs</li> <li>Consultations while adapting tools</li> <li>Spreading information</li> </ul>	<ul style="list-style-type: none"> <li>Relationship building with community members, Elders, and HCPs is necessary to understand community needs using a bottom-up approach.</li> <li>Consultations with Indigenous community stakeholders are needed throughout the development and adaptation of KT tools to ensure that their needs are appropriately met.</li> <li>Connections with Indigenous stakeholders, HCPs, and community members are vital for distributing KT tools and facilitating the uptake of health information.</li> </ul>	<ul style="list-style-type: none"> <li>“Definitely consulting in the community with Elders to learn about their values and their cultural practices, just being very aware of that”</li> <li>“...we really need their [Elders] perspective and understanding of their resources when we’re creating these things and advising them...”</li> <li>“...But you have to work with the community. It can’t be done at the community, it needs to be done with the communities”</li> <li>“...I’m connected to my First Nation’s Facebook page and it’s like by far the easiest way to disseminate like videos, educational videos...”</li> <li>“I thought like Facebook would probably be like the easiest option (for dissemination) as well as just connecting with the community nurses”</li> </ul>

<sup>a</sup>HCP: health care provider.

<sup>b</sup>KT: knowledge translation.

## Accessibility

### Overview

HCPs described how the accessibility of KT tools and the resources mentioned within the tools varied greatly within the Indigenous communities they serve. HCPs often spoke of tangible resources (physical resources) that a family or community may or may not have access to as well as the relationships needed to have a consistent or an easy access to health care resources.

### Tangible Resources

HCPs discussed access to tangible resources as a barrier to viewing KT tools and relating to the content within each tool. Access to internet, equipment, and health centers were mentioned as barriers to Indigenous community members seeking health information. Often, KT tools are disseminated on the web, which may not reach those without adequate internet connection. An HCP described the situation as follows:

*Not a lot of people have Wi-Fi because we’re very rural, and I’m not sure what the cost of Wi-Fi is out there but not very many people have it and there’s often people who don’t have phone data.* [Participant\_018]

This lack of internet reliability presents a challenge for the dissemination of and access to web-based KT tools.

Along with access to web-based resources, HCPs described the limited access their patients and some community members had to health centers. The representation of resources within the tools themselves was also discussed. The participants described the resources depicted in KT tools to which Indigenous community members might not have access. An HCP explained as follows:

*I put out a doctor’s post about hand washing and then the Facebook comments were like well, it’d be nice if we had something to wash our hands with because at that time there was no hand sanitizer in the stores*

*and they just kind of expected us to have a stock to give out.* [Participant\_010]

Another HCP mentioned the following:

*I think often times families don’t have thermometers at home, but I also see the value in explaining the difference between a low fever and a high fever. I don’t know so maybe you could just say if you have a thermometer it’s useful to see if it’s a low fever or a high fever.* [Participant\_016]

These relatively small components of KT tool could impose a barrier for the user in understanding important content or acting upon specific recommendations. Similarly, another HCP described access to transportation as an issue, “...the Inuit people I work with don’t have cars.” One of the participants later went on to say, “There’s a variety of things that [these KT tools] have made assumptions about parents, that doesn’t always fit” (participant\_007). As this HCP mentioned, finding a good fit with the message and audience is important for translating relevant evidence to various communities.

### Relationships

When searching for health information, HCPs voiced several difficulties their community members faced when finding and receiving appropriate health care services. It was often described as “weighing up the pros and cons about going to seek care” (participant\_008). Several HCPs mentioned that knowledgeable and accessible health care workers were already overwhelmed within the communities, so patients might not want to further burden them with requests or feel that such requests would put a strain on the provider-patient relationship. An HCP described their patient’s difficult decision-making, “Is it that severe that I have to go and wake up a nurse in the middle of the night?” (participant\_008). This notion of balancing the barriers and benefits of seeking care was mentioned as particularly difficult for those in rural communities where access to care is limited. An HCP also suggested that researchers and practitioners would better understand Indigenous health care needs by “talking to rural communities and seeing their perspective on things” (participant\_005).

## Relatability

### Overview

Beyond assessing the resources in a given community, HCPs discussed how adapting KT tools to demonstrate relatable environmental contexts, individual representations, and the culture of a community can improve the relevance of the tools for Indigenous community members.

### Environmental Contexts

HCPs described how familiar scenery could provide a more relatable tool for the intended community. For some communities residing in terrains difficult to navigate, physicians do not just wait in a clinic to see their patients. An HCP described their situation around traveling to visit the patients as follows:

*Very often especially even in the summer with the storms, flights are restricted or cancelled. So, the ability to get in and out and right now it's being served by a group of three flying docs who have been kind of commuting up. [Participant\_005]*

The importance of tailoring visuals for increased relatability within the given community was also highlighted through the interviews. Upon viewing one of the KT tools, an HCP said the following:

*I thought it was not very Indigenous-friendly based on the lifestyle and the settings that most First Nations and Métis and Inuit people live, it's more middle class. [Participant\_013]*

Furthermore, they explained the importance of representing the community setting for an individual to relate the content to their own experience. An HCP mentioned that this was perhaps why information was disseminated more broadly in a recent health campaign on social media as follows:

*Yeah, or like I said, Indigenize it up. I know pretty clearly on my Facebook because I'm Cree, when Alberta Health Services posted that poster translated in Cree for washing hands, I saw that a lot more of my own relatives and a lot more on my Facebook share that than the poster of wash your hands. [Participant\_012]*

### Individual Representation

Similarly, in terms of visual representation, HCPs described how animated characters in the KT tools should look similar to the target communities. A self-identified Indigenous HCP mentioned the following:

*I do think having representative characters would be a good starting point. Skin tone and skin coloration...It would be more relatable if people looked more like you... [Participant\_015]*

Including HCPs that resemble Indigenous community members was also mentioned as a way to foster comfort and inclusion. Others discussed considering family structures, inclusion of grandparents as primary caregivers, and representative clothing animations as key considerations for improving relatability.

## Culture

The idea of representing culture was a major discussion item throughout the interviews. HCPs felt that to engage with Indigenous communities, these KT tools required visual representations of the culture and ways of being that are unique to Indigenous peoples. "Indigenize it up" was the advice given by an HCP. More specifically, other HCPs suggested including items such as a Métis sash, traditional medicines, beadwork, and artwork on the walls to improve the relatability of the video:

*Of course, like braids would be great and then yeah, bedding. Like a lot of people have like Indigenous bedding like so if you had a baby like it would be nice to have like a baby wrapped in a special blanket things like that that are like very obviously Indigenous...like physicians or doctors or nurses they have those nice beaded lanyards that are nice. Those are just like some ideas how to indigenize the space. [Participant\_011]*

Cultural considerations as a whole were suggested as essential in producing relatable material for any specific community, not just for Indigenous peoples.

## KT Tool Design

### Overview

In addition to questions regarding content, the participants were encouraged to comment on the design features of the KT tools and were asked for their opinions on the format of the tools and their preferences. Although design preferences were mixed, HCPs suggested that those developing or adapting KT tools should consider the appropriateness of format, depth of language, and translation of language when catering to Indigenous communities.

### Format

Rather than one specific format recommendation, HCPs suggested that understanding the accessibility barriers would better inform the development of KT tool formatting. An HCP said the following:

*There's often people that don't have phone data so I think there is still a role for physical pamphlets and handouts. [Participant\_018]*

HCPs also suggested that the best format for KT tools is dependent on the context of the target audience.

### Plain Language

HCPs recognized that a subsection of those they work with does not have English or health literacy, which would enable them to understand the health information in these KT tool formats. An HCP voiced this by saying the following:

*We get a lot of patients especially from up north where English is not their first language or they're not comfortable with English and like in a conversation you can slow down, you can change your words to make them simpler or like cue in what they understand but you can't do that in a video. [Participant\_010]*

In addition, while discussing the use of various KT tools more broadly, more verbose documents were identified as a barrier to effectively communicating with patients and families:

*Often times, they're really text heavy, they're written at a higher level of reading, and have a lot of medical jargon.* [Participant\_012]

These sorts of documents would not successfully inform the target audience and may even isolate users; therefore, HCPs suggested using plain language and avoiding jargon when creating or adapting KT tools for use by Indigenous groups.

### Translations

When considering language, the idea of translating the tools into Indigenous dialects was discussed. HCPs offered suggestions of “translating them into languages like Cree or Dené or Stoney” but recognized the multitude of Indigenous languages spoken across Canada. Once again, HCPs urged that tool development match the unique needs of the community first, whether that involves translating the tools into a specific language or making use of subtitles or syllabics.

### Relationship Building

#### Overview

The most commonly mentioned theme through these discussions was the importance of building trusting relationships with community members to provide relevant information. HCPs believed that understanding community needs, consulting with community members throughout the KT process, and disseminating information in meaningful ways would promote effective practice in developing and adapting KT tools for use by Indigenous communities.

#### Understanding Community Needs

HCPs voiced the need for relationship building in Indigenous communities to inform content development in KT efforts. Understanding resources, services, and health provider access as well as common health concerns would require conversations with community members, which HCPs mentioned as being essential for creating meaningful products. An HCP said the following:

*You have to work with the community. It can't be done at the community, it needs to be done with the communities.* [Participant\_006]

This collaboration was suggested not only for assessing needs but also for the continued adaptation and revision of KT tools.

#### Consultations

HCPs suggested that while adapting tools by fostering relationships with Indigenous community members, those facilitating KT efforts could seek relevant thoughts and opinions and adjust their products accordingly. Rather than simply improving the quality, an HCP said that community collaboration is a necessary component of KT efforts:

*I think we really need their perspective and understanding of their resources when we're creating these things and advising them.* [Participant\_009]

### Spreading Information

HCPs said that by understanding context, resources, and overall culture, one could plan an appropriate method of delivering health information to each community. An HCP recommended using social media to spread information:

*I'm connected to my First Nation's Facebook page and it's by far the easiest way to disseminate educational videos.* [Participant\_003]

Other HCPs had specific recommendations of where they have found success in dissemination but overall suggested consulting with community members throughout the KT process.

### Discussion

#### Principal Findings

Cultural adaptation has the potential to enhance the acceptability, uptake, and adherence to evidence-based information. Where culture represents a dynamic set of norms, values, and practices in a social sphere, cultural adaptation focuses on modifying certain aspects to enhance the fit for an individual's cultural values, preferences, and norms [29]. Incorporating Indigenous knowledge into KT efforts has the potential to create “new knowledge, policies, and practices to address health issues in communities” [30]. Unfortunately, KT has received relatively little attention in Indigenous health contexts.

As a starting point to understanding how KT tools could be adapted or tailored to meet the needs of Indigenous communities in Alberta, we assessed the value of 2 of our existing child health KT tools from the perspectives of HCPs. From our thematic analysis, 4 key themes were identified by HCPs to consider when adapting or developing child health KT tools use by for Indigenous families: accessibility, relatability, KT design, and relationship building. Here, we present important considerations for researchers and health professionals developing or adapting KT tools for use by Indigenous populations.

#### Resources

Tangible resources, such as access to internet, being able to afford certain medications, or having access to clean water or transportation, were key elements to consider when portraying an event or message within the KT tools. In general, remote and rural communities have greater accessibility barriers than their urban counterparts and often face challenges in relation to transportation infrastructure and internet connectivity [31-34]. In addition, a community's ability to leverage resources, services, and providers affects the relevance and applicability of the recommendations made in a KT tool. Researchers and developers should ensure that KT tools are tailored to the intended population or community with specific accessibility barriers in mind so that the users can implement the recommendations effectively.

Accessibility barriers can also impact whether users are able to access or view KT tools. For instance, caregivers residing in remote locales with sparse internet connectivity may be unable to use web-based KT tools. Although these interviews did not reveal a preference for a particular KT format, the ideal KT format is something which is easily accessed and used by the

target population. In this case, social media platforms such as Facebook were identified as an extremely effective method for distributing web-based tools to populations with internet access and have been successfully used previously to share health information with First Nations communities within our province [35].

### Language

To ensure widespread accessibility and usability, the translation of KT tools into various Indigenous languages and dialects for specific communities and linguistic groups should be considered. However, a deep linguistic approach, as referred to by Resnicow et al [36], should be used. This involves ensuring that texts are culturally relevant rather than simply providing a direct translation. Direct language translations can often lack cultural nuances; therefore, high quality translations by Indigenous translators are needed.

A community-level collaboration approach should be adopted to support the development of community-owned products that incorporate cultural sensitivity and unique linguistic needs to improve their uptake and use within the community.

### Cultural Awareness and Understanding Diversity

Incorporating different aspects of a group's culture *appropriately* into KT tools can foster engagement and increase knowledge, especially if shared through locally developed and contextualized ways [17,18]. Incorporating Indigenous cultural components, such as the medicine wheel, or practices, such as smudging, may help users to better connect with KT tools. An important caveat here is that Indigenous peoples and cultures are incredibly diverse [37], and as such, it is extremely challenging to create content that is relevant to all Indigenous groups within a single tool. However, culture has been successfully incorporated within KT efforts in the past. For instance, Laird et al [38] used a culturally safe KT approach to help Indigenous Australian caregivers manage protracted bacterial bronchitis in their children. Another study aimed to improve dementia health literacy in Canadian Indigenous communities by developing culturally appropriate fact sheets in collaboration with Indigenous organizations and Elders [39]. However, each Indigenous community has a unique culture and history; therefore, a successful approach to KT for one community may not be relevant to another. By actively engaging the community of interest, stakeholders could shape the information design and content to suit the community's needs and preferences.

Ensuring that KT tools, specifically those that use visuals or graphics, incorporate relevant and appropriate depictions of environments and characters is important. Scenic or environmental portrayals must be relatable to KT users. Likewise, portrayals of characters within visual or story-based KT tools must also bear resemblance to the target users for optimal reception. This could involve having characters in a visual KT tool that have a similar skin complexion to that of the target audience and similar clothing items based on cultural or environmental factors or integrating customs and practices specific to that group [40,41].

Notably, those developing KT tools should avoid stereotypical or inaccurate representations of Indigenous peoples or cultures by effectively engaging with and involving the communities and tailoring the tools appropriately.

### Importance of Relationships

Relationship building is important throughout all aspects of the KT and dissemination process and is integral for working with Indigenous communities in a respectful and meaningful way.

Engaging with and building long-lasting mutual relationships with Indigenous communities, stakeholders, and Elders has been shown to be extremely important for working with Indigenous populations in a multitude of capacities and contexts [42-44]. A scoping review highlighted the importance of the contributions of the Indigenous Elders to well-being at the individual and community levels and the need to include Indigenous Elders in consultation and relationship-building efforts [45].

Relationship building entails creating long-term, sustainable partnerships with community members, stakeholders, and health professionals. Understanding community needs is required from a holistic and grassroots approach, which should only be achieved through genuine reciprocal relationships. Relationship building is also necessary to facilitate consultations with Indigenous community members, Elders, and health providers to adapt KT tools for use by their respective communities. KT tools should be codeveloped with the members of a given community to ensure that they are culturally sensitive, relevant to their needs, and accessible.

### Limitations and Future Directions

While reflecting on the enduring power dynamics of researchers engaging with Indigenous peoples [46], we believed that recruiting Indigenous community members required a sensitive process of meaningful reciprocity that was not feasible for this study. In consultation with Indigenous community representatives, it was suggested that HCPs who work with Indigenous families were well situated to comment on the available health care resources for the Indigenous communities they serve. The interviewed HCPs also provided important considerations on how to engage Indigenous patients and communities to develop relevant and impactful KT tools.

An important limitation to address is the biases that may arise from the type of qualitative research conducted in this study. Specifically, interpretive biases could arise during the interview or the data analysis and interpretation stages. To mitigate the impact of any interpretive biases, a second researcher (KSW) reviewed the data and brought forth any concerns for discussion with the primary author to reach a resolution. The findings of this study were also reviewed by one of our team members (SDL) who self-identified as Indigenous, further mitigating any interpretive biases.

Despite having a good representation from HCPs working directly with Indigenous families and communities, only 39% (7/18) of our participants self-identified as Indigenous. Even though the themes present across the interviews were similar, it is important to keep in mind that the lived experiences of

HCPs who self-identify as Indigenous are distinct from those that are not. Future studies may benefit from incorporating a higher proportion of HCPs self-identifying as Indigenous or aim to recruit exclusively Indigenous participants. This may prove difficult given that Indigenous peoples have historically been underrepresented in health care-related fields [47]. It is important to recognize that given the small number of Indigenous people involved in the study, the findings are not generalizable to the broader population of Indigenous people. In addition, these findings only represent the perspectives of the HCP, and future studies should expand beyond consulting only HCPs and include the health consumers or knowledge users of the KT tool (Indigenous families). We recognize that this study is just a first step; consultations with Indigenous community members, Elders, Indigenous leaders, and other stakeholders are needed to fully understand how to best develop and adapt KT tools for use by Indigenous communities in accordance with their unique needs. Furthermore, we recognize that this study was conducted within a local context in Canada, and the findings might not be representative of all Indigenous populations, not just within Alberta but across Canada (and elsewhere). However, the croup video usability was completed at the Stanton Territorial Hospital in Yellowknife, Northwest Territories, Canada, which has a significant population of Dené people; Portage La Prairie, Manitoba, Canada, which has a significant population of Anishinaabe or Ojibwe, Cree, and Dakota or Sioux people; and the Stollery Children's Hospital

(Edmonton, Alberta, Canada). The AOM video usability was conducted in 3 rural emergency departments in Nova Scotia; the Royal University Hospital in Saskatoon, Saskatchewan, Canada; and Leduc Community Hospital, Alberta, Canada. Therefore, it is important to acknowledge that the perspectives of various Indigenous communities and Indigenous peoples may differ from the themes we have derived in this study. Future KT tool development studies may also benefit from usability testing that occurs across an increased number of locales, particularly those with significant populations of Indigenous peoples.

## Conclusions

Developing or adapting KT products that are relevant to different cultural groups is important to enhance the reach of knowledge and ultimately improve health outcomes. On the basis of the interviews with HCPs, accessibility, relatability, KT design, and relationship building were perceived by HCPs as important considerations for adapting and developing KT tools that would meet the needs of the Indigenous communities they serve. This study is a critical first step in identifying how to adapt existing KT tools for use by culturally and linguistically diverse communities in general. This study demonstrated the importance of engaging end users (health care consumers) in the development and critique of KT products. Including Indigenous perspectives in the continuing refinement of KT tool content and delivery will increase the relevance and accessibility of the messaging for future KT efforts in this unique population.

---

## Acknowledgments

This research was funded in part by the generous support of the Stollery Children's Hospital Foundation through the Women and Children's Health Research Institute and the Alberta Strategy for Patient Oriented Research (SPOR) SUPPORT Unit Knowledge Translation Platform, which is funded by Alberta Innovates and the Canadian Institutes of Health Research. LH is supported by a Canada Research Chair in Knowledge Synthesis and Translation. SDS is supported by a Canada Research Chair in Knowledge Translation in Child Health. LH and SDS are distinguished researchers with the Stollery Science Lab, supported by the Stollery Children's Hospital Foundation. JK was funded by the Alberta Innovates Summer Research Studentship program. The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

The authors extend their sincere gratitude to the health care providers (HCPs) who generously volunteered their time to contribute to this study. The authors would also like to thank the Alberta SPOR SUPPORT Unit Knowledge Translation Platform research staff (Gabrielle Zimmermann, Denise Thomson, Stephanie Brooks, and Devonne Brandys) and an external consultant (Aretha Greatrix) who were involved in the initial planning of this work.

---

## Data Availability

These are available upon reasonable request from the corresponding author.

---

## Conflicts of Interest

None declared.

---

## Multimedia Appendix 1

Examples of knowledge translation tools used.

[[PDF File \(Adobe PDF File\), 349 KB - formative\\_v6i10e36353\\_app1.pdf](#)]

---

## References

1. Chapman E, Haby MM, Toma TS, de Bortoli MC, Illanes E, Oliveros MJ, et al. Knowledge translation strategies for dissemination with a focus on healthcare recipients: an overview of systematic reviews. *Implement Sci* 2020 Mar 04;15(1):14 [FREE Full text] [doi: [10.1186/s13012-020-0974-3](https://doi.org/10.1186/s13012-020-0974-3)] [Medline: [32131861](https://pubmed.ncbi.nlm.nih.gov/32131861/)]
2. Albrecht L, Scott SD, Hartling L. Knowledge translation tools for parents on child health topics: a scoping review. *BMC Health Serv Res* 2017 Sep 29;17(1):686 [FREE Full text] [doi: [10.1186/s12913-017-2632-2](https://doi.org/10.1186/s12913-017-2632-2)] [Medline: [28962637](https://pubmed.ncbi.nlm.nih.gov/28962637/)]
3. Gates A, Gates M, Vandermeer B, Johnson C, Hartling L, Johnson DW, et al. Glucocorticoids for croup in children. *Cochrane Database Syst Rev* 2018 Aug 22;8(8):CD001955 [FREE Full text] [doi: [10.1002/14651858.CD001955.pub4](https://doi.org/10.1002/14651858.CD001955.pub4)] [Medline: [30133690](https://pubmed.ncbi.nlm.nih.gov/30133690/)]
4. Scott SD, Hartling L, O'Leary KA, Archibald M, Klassen TP. Stories – a novel approach to transfer complex health information to parents: a qualitative study. *Arts Health* 2012 Jun;4(2):162-173. [doi: [10.1080/17533015.2012.656203](https://doi.org/10.1080/17533015.2012.656203)]
5. Graham ID, Logan J, Harrison MB, Straus SE, Tetroe J, Caswell W, et al. Lost in knowledge translation: time for a map? *J Contin Educ Health Prof* 2006;26(1):13-24. [doi: [10.1002/chp.47](https://doi.org/10.1002/chp.47)] [Medline: [16557505](https://pubmed.ncbi.nlm.nih.gov/16557505/)]
6. Hartling L, Elliott SA, Buckreus K, Leung J, Scott SD. Development and evaluation of a parent advisory group to inform a research program for knowledge translation in child health. *Res Involv Engagem* 2021 Jun 07;7(1):38 [FREE Full text] [doi: [10.1186/s40900-021-00280-3](https://doi.org/10.1186/s40900-021-00280-3)] [Medline: [34099069](https://pubmed.ncbi.nlm.nih.gov/34099069/)]
7. Le A, Hartling L, Scott SD. The development and usability testing of digital knowledge translation tools for parents of children with acute otitis media. *medRxiv* 2021 Jul 3. [doi: [10.1101/2021.06.29.21259431](https://doi.org/10.1101/2021.06.29.21259431)]
8. Scott S, Le A, Hartling L. Developing and testing an arts-based, digital knowledge translation tool for parents about childhood croup. *medRxiv* 2021 Jun 4. [doi: [10.1101/2021.06.03.21257424](https://doi.org/10.1101/2021.06.03.21257424)]
9. ECHO - Evidence in Child Health to Enhance Outcomes. URL: <https://www.echokt.ca/> [accessed 2022-09-22]
10. Kothari A, Wathen CN. Integrated knowledge translation: digging deeper, moving forward. *J Epidemiol Community Health* 2017 Jun;71(6):619-623. [doi: [10.1136/jech-2016-208490](https://doi.org/10.1136/jech-2016-208490)] [Medline: [28298415](https://pubmed.ncbi.nlm.nih.gov/28298415/)]
11. Kim PJ. Social determinants of health inequities in indigenous Canadians through a life course approach to colonialism and the residential school system. *Health Equity* 2019 Jul 25;3(1):378-381 [FREE Full text] [doi: [10.1089/heap.2019.0041](https://doi.org/10.1089/heap.2019.0041)] [Medline: [31346558](https://pubmed.ncbi.nlm.nih.gov/31346558/)]
12. Nesdole R, Voigts D, Lepnurm R, Roberts R. Reconceptualizing determinants of health: barriers to improving the health status of First Nations peoples. *Can J Public Health* 2014 May 09;105(3):e209-e213 [FREE Full text] [doi: [10.17269/cjph.105.4308](https://doi.org/10.17269/cjph.105.4308)] [Medline: [25165841](https://pubmed.ncbi.nlm.nih.gov/25165841/)]
13. Greenwood ML, de Leeuw SN. Social determinants of health and the future well-being of Aboriginal children in Canada. *Paediatr Child Health* 2012 Aug;17(7):381-384 [FREE Full text] [Medline: [23904782](https://pubmed.ncbi.nlm.nih.gov/23904782/)]
14. Horrill T, McMillan DE, Schultz AS, Thompson G. Understanding access to healthcare among Indigenous peoples: a comparative analysis of biomedical and postcolonial perspectives. *Nurs Inq* 2018 Jul;25(3):e12237 [FREE Full text] [doi: [10.1111/min.12237](https://doi.org/10.1111/min.12237)] [Medline: [29575412](https://pubmed.ncbi.nlm.nih.gov/29575412/)]
15. Smylie J, Martin CM, Kaplan-Myrth N, Steele L, Tait C, Hogg W. Knowledge translation and indigenous knowledge. *Int J Circumpolar Health* 2004;63 Suppl 2:139-143. [doi: [10.3402/ijch.v63i0.17877](https://doi.org/10.3402/ijch.v63i0.17877)] [Medline: [15736639](https://pubmed.ncbi.nlm.nih.gov/15736639/)]
16. Estey E, Kmetc A, Reading J. Knowledge translation in the context of Aboriginal health. *Can J Nurs Res* 2008 Jun;40(2):24-39. [Medline: [18714896](https://pubmed.ncbi.nlm.nih.gov/18714896/)]
17. Estey EA, Kmetc AM, Reading JL. Thinking about aboriginal KT: learning from the Network Environments for Aboriginal Health Research British Columbia (NEARBC). *Can J Public Health* 2010;101(1):83-86 [FREE Full text] [doi: [10.1007/bf03405569](https://doi.org/10.1007/bf03405569)] [Medline: [20364545](https://pubmed.ncbi.nlm.nih.gov/20364545/)]
18. Morton Ninomiya ME, Atkinson D, Brascoupe S, Firestone M, Robinson N, Reading J, et al. Effective knowledge translation approaches and practices in Indigenous health research: a systematic review protocol. *Syst Rev* 2017 Feb 20;6(1):34 [FREE Full text] [doi: [10.1186/s13643-017-0430-x](https://doi.org/10.1186/s13643-017-0430-x)] [Medline: [28219446](https://pubmed.ncbi.nlm.nih.gov/28219446/)]
19. 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/)]
20. Kreuter MW, Lukwago SN, Bucholtz RD, Clark EM, Sanders-Thompson V. Achieving cultural appropriateness in health promotion programs: targeted and tailored approaches. *Health Educ Behav* 2003 Apr;30(2):133-146. [doi: [10.1177/1090198102251021](https://doi.org/10.1177/1090198102251021)] [Medline: [12693519](https://pubmed.ncbi.nlm.nih.gov/12693519/)]
21. Cardinal JC, Schopflocher DP, Svenson LW, Morrison KB, Laing L. First nations in Alberta: a focus on health service use. Alberta Health and Wellness. Edmonton, Canada: University of Alberta; 2004. URL: <https://tinyurl.com/y58mz6pn> [accessed 2020-08-14]
22. Johnson TP. Snowball sampling: introduction. In: *Wiley StatsRef: Statistics Reference Online*. Hoboken, NJ, USA: John Wiley & Sons; 2014.
23. Palinkas LA, Horwitz SM, Green CA, Wisdom JP, Duan N, Hoagwood K. Purposeful sampling for qualitative data collection and analysis in mixed method implementation research. *Adm Policy Ment Health* 2015 Sep;42(5):533-544 [FREE Full text] [doi: [10.1007/s10488-013-0528-y](https://doi.org/10.1007/s10488-013-0528-y)] [Medline: [24193818](https://pubmed.ncbi.nlm.nih.gov/24193818/)]
24. Guest G, Namey E, Chen M. A simple method to assess and report thematic saturation in qualitative research. *PLoS One* 2020 May 5;15(5):e0232076 [FREE Full text] [doi: [10.1371/journal.pone.0232076](https://doi.org/10.1371/journal.pone.0232076)] [Medline: [32369511](https://pubmed.ncbi.nlm.nih.gov/32369511/)]

25. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, REDCap Consortium. The REDCap consortium: building an international community of software platform partners. *J Biomed Inform* 2019 Jul;95:103208 [FREE Full text] [doi: [10.1016/j.jbi.2019.103208](https://doi.org/10.1016/j.jbi.2019.103208)] [Medline: [31078660](https://pubmed.ncbi.nlm.nih.gov/31078660/)]
26. 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)]
27. Nowell LS, Norris JM, White DE, Moules NJ. Thematic analysis: striving to meet the trustworthiness criteria. *Int J Qual Methods* 2017 Oct 02;16(1):160940691773384. [doi: [10.1177/1609406917733847](https://doi.org/10.1177/1609406917733847)]
28. Guba EG, Lincoln YS. Paradigmatic controversies, contradictions, and emerging confluences. In: Denzin NK, Lincoln YS, editors. *The SAGE Handbook of Qualitative Research*. 3rd edition. Thousand Oaks, CA, USA: Sage Publications; 2005:191-215.
29. Cabassa LJ, Baumann AA. A two-way street: bridging implementation science and cultural adaptations of mental health treatments. *Implement Sci* 2013 Aug 19;8:90 [FREE Full text] [doi: [10.1186/1748-5908-8-90](https://doi.org/10.1186/1748-5908-8-90)] [Medline: [23958445](https://pubmed.ncbi.nlm.nih.gov/23958445/)]
30. Ellison C. Indigenous Knowledge and Knowledge Synthesis Translation and Exchange (KSTE). University of Alberta. 2014. URL: <https://www.nccih.ca/docs/context/RPT-IndigenousKnowledgeKSTE-Ellison-EN.pdf> [accessed 2020-10-02]
31. Hardin HK, Alchami H, Lee D, Jones MS. Unmet health need and perceived barriers to health care among adolescents living in a rural area. *Child Health Care* 2021;50(1):108-123 [FREE Full text] [doi: [10.1080/02739615.2020.1833333](https://doi.org/10.1080/02739615.2020.1833333)] [Medline: [33424079](https://pubmed.ncbi.nlm.nih.gov/33424079/)]
32. Ansari Z. A review of literature on access to primary health care. *Aust J Prim Health* 2007;13(2):80-95. [doi: [10.1071/py07026](https://doi.org/10.1071/py07026)]
33. Velaga NR, Beecroft M, Nelson JD, Corsar D, Edwards P. Transport poverty meets the digital divide: accessibility and connectivity in rural communities. *J Transp Geography* 2012 Mar;21:102-112 [FREE Full text] [doi: [10.1016/j.jtrangeo.2011.12.005](https://doi.org/10.1016/j.jtrangeo.2011.12.005)]
34. Ali-Hassan H, Eloulabi R, Keethakumar A. Internet non-use among Canadian indigenous older adults: Aboriginal Peoples Survey (APS). *BMC Public Health* 2020 Oct 15;20(1):1554 [FREE Full text] [doi: [10.1186/s12889-020-09659-5](https://doi.org/10.1186/s12889-020-09659-5)] [Medline: [33059658](https://pubmed.ncbi.nlm.nih.gov/33059658/)]
35. Molyneaux H, O'Donnell S, Kakekaspan C, Walmark B, Budka P, Gibson K. Social media in remote first nation communities. *Can J Commun* 2014 Jun 01;39(2):275-288. [doi: [10.22230/cjc.2014v39n2a2619](https://doi.org/10.22230/cjc.2014v39n2a2619)]
36. Resnicow K, Baranowski T, Ahluwalia JS, Braithwaite RL. Cultural sensitivity in public health: defined and demystified. *Ethn Dis* 1999;9(1):10-21. [Medline: [10355471](https://pubmed.ncbi.nlm.nih.gov/10355471/)]
37. France MH, del Carmen Rodríguez M, Hett GG. *Diversity, Culture and Counselling: A Canadian Perspective*. 3rd edition. Edmonton, Canada: Brush Education; 2021.
38. Laird P, Walker R, Lane M, Totterdell J, Chang AB, Schultz A. Recognition and management of protracted bacterial bronchitis in Australian aboriginal children: a knowledge translation approach. *Chest* 2021 Jan;159(1):249-258. [doi: [10.1016/j.chest.2020.06.073](https://doi.org/10.1016/j.chest.2020.06.073)] [Medline: [32673622](https://pubmed.ncbi.nlm.nih.gov/32673622/)]
39. Webkamigad S, Warry W, Blind M, Jacklin K. An approach to improve dementia health literacy in indigenous communities. *J Cross Cult Gerontol* 2020 Mar;35(1):69-83 [FREE Full text] [doi: [10.1007/s10823-019-09388-2](https://doi.org/10.1007/s10823-019-09388-2)] [Medline: [31853783](https://pubmed.ncbi.nlm.nih.gov/31853783/)]
40. Castro FG, Barrera Jr M, Holleran Steiker LK. Issues and challenges in the design of culturally adapted evidence-based interventions. *Annu Rev Clin Psychol* 2010;6:213-239 [FREE Full text] [doi: [10.1146/annurev-clinpsy-033109-132032](https://doi.org/10.1146/annurev-clinpsy-033109-132032)] [Medline: [20192800](https://pubmed.ncbi.nlm.nih.gov/20192800/)]
41. Spanhel K, Balci S, Feldhahn F, Bengel J, Baumeister H, Sander LB. Cultural adaptation of Internet- and mobile-based interventions for mental disorders: a systematic review. *NPJ Digit Med* 2021 Aug 25;4(1):128 [FREE Full text] [doi: [10.1038/s41746-021-00498-1](https://doi.org/10.1038/s41746-021-00498-1)] [Medline: [34433875](https://pubmed.ncbi.nlm.nih.gov/34433875/)]
42. Boyd B, Loreface S. Understanding consultation and engagement of Indigenous Peoples in resource development: a policy framing approach. *Can Public Admin* 2018 Dec 16;61(4):572-595 [FREE Full text] [doi: [10.1111/capa.12301](https://doi.org/10.1111/capa.12301)]
43. Kwiatkowski R. Indigenous community based participatory research and health impact assessment: a Canadian example. *Environ Impact Assess Rev* 2011 Jul;31(4):445-450 [FREE Full text] [doi: [10.1016/j.eiar.2010.02.003](https://doi.org/10.1016/j.eiar.2010.02.003)]
44. Berg K, McLane P, Eshkakogan N, Mantha J, Lee T, Crowshoe C, et al. Perspectives on indigenous cultural competency and safety in Canadian hospital emergency departments: a scoping review. *Int Emerg Nurs* 2019 Mar;43:133-140. [doi: [10.1016/j.ienj.2019.01.004](https://doi.org/10.1016/j.ienj.2019.01.004)] [Medline: [30733006](https://pubmed.ncbi.nlm.nih.gov/30733006/)]
45. Viscogliosi C, Asselin H, Basile S, Borwick K, Couturier Y, Drolet M, et al. Importance of Indigenous elders' contributions to individual and community wellness: results from a scoping review on social participation and intergenerational solidarity. *Can J Public Health* 2020 Oct;111(5):667-681 [FREE Full text] [doi: [10.17269/s41997-019-00292-3](https://doi.org/10.17269/s41997-019-00292-3)] [Medline: [32109314](https://pubmed.ncbi.nlm.nih.gov/32109314/)]
46. Nicholls R. Research and Indigenous participation: critical reflexive methods. *Int J Soc Res Methodol* 2009 Apr;12(2):117-126. [doi: [10.1080/13645570902727698](https://doi.org/10.1080/13645570902727698)]
47. Curtis E, Wikaire E, Stokes K, Reid P. Addressing indigenous health workforce inequities: a literature review exploring 'best' practice for recruitment into tertiary health programmes. *Int J Equity Health* 2012 Mar 15;11:13 [FREE Full text] [doi: [10.1186/1475-9276-11-13](https://doi.org/10.1186/1475-9276-11-13)] [Medline: [22416784](https://pubmed.ncbi.nlm.nih.gov/22416784/)]

---

**Abbreviations****AOM:** acute otitis media**HCP:** health care provider**KT:** knowledge translation**REDCap:** Research Electronic Data Capture

---

*Edited by A Mavragani; submitted 07.06.22; peer-reviewed by MO Lucha-López, D Viswanath, S Shibasaki; comments to author 30.08.22; revised version received 02.09.22; accepted 06.09.22; published 05.10.22.*

*Please cite as:*

*Elliott SA, Kreutz J, Wright KS, Di Lallo S, Scott SD, Hartling L*

*Adapting Child Health Knowledge Translation Tools for Use by Indigenous Communities: Qualitative Study Exploring Health Care Providers' Perspectives*

*JMIR Form Res 2022;6(10):e36353*

*URL: <https://formative.jmir.org/2022/10/e36353>*

*doi: [10.2196/36353](https://doi.org/10.2196/36353)*

*PMID: [36197711](https://pubmed.ncbi.nlm.nih.gov/36197711/)*

©Sarah A Elliott, Jason Kreutz, Kelsey S Wright, Sherri Di Lallo, Shannon D Scott, Lisa Hartling. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 05.10.2022. 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

# Acceptability and Usability of a Reward-Based Mobile App for Opioid Treatment Settings: Mixed Methods Pilot Study

Steven L Proctor<sup>1,2,3</sup>, PhD; Khary K Rigg<sup>4</sup>, PhD; Allen Y Tien<sup>5</sup>, MD, MHS

<sup>1</sup>PRO Health Group, Miami Beach, FL, United States

<sup>2</sup>Thriving Mind South Florida, Miami, FL, United States

<sup>3</sup>Department of Psychiatry and Behavioral Health, Herbert Wertheim College of Medicine, Florida International University, Miami, FL, United States

<sup>4</sup>Department of Mental Health Law and Policy, University of South Florida, Tampa, FL, United States

<sup>5</sup>Medical Decision Logic, Inc, Baltimore, MD, United States

**Corresponding Author:**

Steven L Proctor, PhD

Thriving Mind South Florida

7205 Corporate Center Drive

Suite 200

Miami, FL, 33126

United States

Phone: 1 305 858 3335

Email: [sproctor@thrivingmind.org](mailto:sproctor@thrivingmind.org)

## Abstract

**Background:** Contingency management is an evidence-based yet underutilized approach for opioid use disorder (OUD). Reasons for limited adoption in real-world practice include ethical, moral, and philosophical concerns regarding use of monetary incentives, and lack of technological innovation. In light of surging opioid overdose deaths, there is a need for development of technology-enabled solutions leveraging the power of contingency management in a way that is viewed by both patients and providers as acceptable and feasible.

**Objective:** This mixed methods pilot study sought to determine the perceived acceptability and usability of PROCare Recovery, a reward-based, technology-enabled recovery monitoring smartphone app designed to automate contingency management by immediately delivering micropayments to patients for achieving recovery goals via smart debit card with blocking capabilities.

**Methods:** Participants included patients receiving buprenorphine for OUD (n=10) and licensed prescribers (n=5). Qualitative interviews were conducted by 2 PhD-level researchers via video conferencing to explore a priori hypotheses. Thematic analysis of interviews was conducted and synthesized into major themes.

**Results:** Participants were overwhelmingly in favor of microrewards (eg, US \$1) to incentivize treatment participation (up to US \$150 monthly). Participants reported high acceptability of the planned debit card spending restrictions (blocking cash withdrawals and purchases at bars or liquor stores, casinos or online gambling). Quantitative data revealed a high level of perceived usability of the PROCare Recovery app.

**Conclusions:** Patients and providers alike appear receptive to microfinancial incentives in standard OUD treatment practices. Further pilot testing of PROCare is underway to determine acceptability, feasibility, and preliminary effectiveness in a rigorous randomized controlled trial.

(*JMIR Form Res* 2022;6(10):e37474) doi:[10.2196/37474](https://doi.org/10.2196/37474)

**KEYWORDS**

opioids; contingency management; mHealth; digital health; mobile app; innovation; opioid use disorder; recovery; acceptability

## Introduction

America's escalating opioid overdose crisis requires innovative solutions. Over 100,000 people died of drug overdose in 2021 in the United States—the majority of which involved opioids

[1]. Contingency management (ie, rewarding people, often with monetary incentives, for achieving recovery goals) is an effective, evidence-based intervention for opioid use disorder (OUD) backed by decades of research [2] but remains highly underutilized. Motivational incentives are rarely used in

real-world clinical practice due to several concerns, including most notably, the apparent lack of innovation, as well as moral, philosophical, ethical, and economic concerns, and even federal rules meant to prevent illegal inducements in health care [3]. Traditional contingency management protocols have become rudimentary, outdated, and onerous in the current digital era (eg, requirement for in-person appointments, use of a “prize bowl” filled with slips of paper), necessitating novel, technology-enabled solutions to facilitate widespread adoption. Many accepted contingency management procedures reward drug-free urinalysis screens exclusively, and there is only a low chance that the desired behavior will actually be rewarded in the commonly used probabilistic “prize-based” procedure in which patients earn draws from a prize bowl containing slips of paper with either no monetary value or a low-value prize when the target behavior is exhibited. This raises the common complaint that contingency management is a “game of chance” due to the lack of immediate and consistent meaningful reinforcement that is required for lasting behavior change.

Treatment programs may understandably voice concerns about increased costs associated with providing monetary rewards given that contingency management is often an “add-on” to usual care (ie, adjunctive intervention). However, research shows that contingency management, when combined with medication treatment for OUD, demonstrated the largest cost-savings relative to other evidence-based interventions for OUD, including medication alone [4]. A recent study examining the net impact of a digital therapeutic delivering contingency management via mobile app on medical costs due to hospital-based encounters and procedures among a sample of patients treated with buprenorphine for OUD documented that the medical cost reduction in patients using the app relative to those receiving standard care offset the cost of the digital therapeutic itself, thereby resulting in a net cost-savings of US \$720 per patient [5].

Although there have been a number of recent strides in coverage for contingency management, including pilot programs in several US states, many commercial and government insurers remain slow to cover contingency management. There may also be legal concerns about whether the use of monetary incentives violates federal and state law because it could be considered unlawful to give money to patients who are enrolled in federally or state-funded health plans or programs. However, recent guidance from the Office of Inspector General (OIG) in the form of an advisory legal opinion (OIG Advisory Opinion No. 22-04) in March 2022 approved the use of a digital contingency management program using smartphone and smart debit card technology, which could clear the way for wider use of similar programs in routine treatment settings. The OIG has also dispelled the oft-stated assumption that the OIG bans contingency management incentives with a monetary value greater than US \$75 [6].

Accumulating evidence suggests smartphone ownership, although certainly not universal, is no longer the barrier it once was [7-9]. In light of the increasing penetration of smartphone users, and the fact that many patients already leverage technology in all facets of their lives, reward-based apps have the potential to bring contingency management into the hands

of more people receiving treatment for OUD. As a number of mobile apps begin to emerge in the treatment of OUD [10], there remain concerns about their quality, safety, potential efficacy, and availability of empirical evidence supporting their use for populations with OUD. Findings from a recent review study conducted to characterize currently available smartphone apps for the prevention, management, and treatment of opioid use, misuse, and related harm found that few apps meet basic quality standards [11], and even fewer reward-based apps have published peer-reviewed evidence regarding patient and provider perspectives on acceptability to inform uptake in real-world treatment settings. All currently available opioid-related apps were identified via web scraping of data from the Google Play and Apple App Store using the following keywords: *opioid use disorder, opioid abuse opioid misuse, opioid addiction, prescription opioid misuse, prescription opioid abuse, opioid abuse treatment, opioid abuse intervention, opioid abuse therapy, opioid abuse management, and opioid addiction recovery*. Of the 619 apps identified by the researchers, 59 apps met basic criteria for quality assessment, and only 1 app satisfied all standards on the screener for quality, as assessed by the American Psychiatric Association’s App Evaluation Model [12], which addresses the most fundamental questions to ask when considering using a digital health app (eg, *Does the app have a clinical/recovery foundation relevant to your intended use? Is there evidence of specific benefit from academic institutions, end user feedback, or research studies?*). Further work is warranted to fill this gap in technological solutions for OUD recovery management.

Guided by the Innovation Corps methodology, which uses the Lean Launchpad approach to developing hypotheses, then rapidly moving to continuous customer discovery with the aim of translating hypotheses into facts [13], this study evaluated PROCare Recovery, a multiplatform (iOS and Android), reward-based recovery management mobile app for patients receiving medication treatment for OUD. PROCare uses the power of motivational incentives and self-monitoring to reward people for achieving their recovery goals and engaging in their treatment plan. Development of PROCare was supported by a Small Business Innovation Research (SBIR) grant from the National Institute on Drug Abuse (NIDA). With an evidence-based reward system, people recovering from OUD can earn rewards for taking their medication as directed, attending appointments, taking routine self-report surveys to track recovery progress, completing science-backed learning modules in the psychoeducational library, as well as engaging in other recovery-oriented activities within the app. PROCare is a recovery management tool for people receiving medication treatment for OUD who are currently enrolled in outpatient treatment under the supervision of a clinician, and allows for many aspects of contingency management to be fully or partially automated, thereby addressing common logistical barriers to implementation. Patients have the opportunity to earn both monetary and nonmonetary rewards for completing various activities within the app as well as elements of their care plan such as Rapid Daily Check-In surveys (assessing craving, motivation, etc); more comprehensive monthly assessments (health care utilization, occupational functioning, quality of life, etc); taking their medication as directed; and accessing resources

in the psychoeducational library to help educate, support, and encourage patients. Automated delivery of monetary rewards is achieved by depositing money to a pre-paid debit card with the option to apply spending restrictions. The “smart” debit card allows card administrators (eg, treatment program staff) to toggle specified blocking capabilities on/off to prevent cash withdrawals or purchases at identified high-risk vendors (eg, bars, liquor stores, casinos). Medication adherence and treatment engagement are translated to tangible financial rewards. With nonmonetary rewards, patients earn “credits” toward their “Degree(s) in Recovery” (associate, bachelor’s, master’s, and doctorate of recovery). Patients immediately earn monetary and non-monetary rewards for certain activities (eg, daily check-ins), whereas other activities first require verification before rewards are released. Medication adherence is confirmed by way of a combination of self-report and verification via urinalysis, and attendance at individual and group therapy appointments is confirmed via GPS location verification.

In light of surging opioid overdose deaths, there is a need for development of technology-enabled solutions leveraging the power of contingency management in a way that is viewed by

both patients and providers as acceptable and feasible. This mixed methods study sought to determine the perceived acceptability and usability of PROCare Recovery, a reward-based, recovery monitoring smartphone app designed to automate contingency management by immediately delivering micropayments (eg, US \$1) to patients for achieving recovery goals via smart debit card with blocking capabilities.

## Methods

### Recruitment

The current study was conducted as part of a phase I SBIR grant from the NIDA to build and test a reward-based recovery management smartphone app (PROCare Recovery) for patients with OUD receiving medication treatment. Participants were recruited from an addiction treatment system in South Florida. The study sample included licensed providers actively prescribing buprenorphine (n=5), and patients currently receiving buprenorphine for OUD (n=10). Participants had an average age of 41.93 years, and were predominately White, male, and employed full-time (Table 1).

**Table 1.** Sample demographic characteristics (N=15).

Variable	Value, n (%)
<b>Sex</b>	
Male	9 (60)
Female	6 (40)
<b>Ethnicity</b>	
Hispanic	3 (20)
Non-Hispanic	12 (80)
<b>Race</b>	
White	13 (86.7)
Asian	1 (6.7)
Other	1 (6.7)
<b>Marital status</b>	
Single	8 (53.3)
Married	3 (20)
Divorced	2 (13.3)
Widowed	1 (6.7)
Separated	1 (6.7)
<b>Employment status</b>	
Full-time	11 (73.3)
Part-time	2 (13.3)
Unemployed	2 (13.3)

### Measures

Individual in-depth qualitative interviews were guided by the Innovation Corps methodology given its emphasis on the Lean Launchpad approach to developing hypotheses, then rapidly moving to continuous customer discovery with the aim of translating hypotheses into facts [13]. A Project Advisory Board

of addiction treatment industry experts, researchers, clinicians, administrators, people in recovery from opioid addiction, and family members of people with OUD, varying in background and expertise, was formed to inform the scientific and strategic direction of the phase I SBIR project. Following Project Advisory Board input with developing a list of hypotheses about

the problem under study—here specifically regarding planned features and components of PROCare Recovery—semistructured interviews were conducted by two PhD-level researchers (authors SLP, KKR). In-depth interviews were selected for their ability to capture individual experiences and elicit detailed, granular responses. Participants were also administered the 10-item System Usability Scale (SUS) [14] as a quantitative measure of perceived PROCare usability (scores can range from 0 to 100 with >68 considered above-average). The SUS is a valid and reliable measure commonly used for global assessments of systems usability to evaluate a wide variety of products (mobile apps, software, websites, etc), including studies evaluating patient- and provider-facing mobile apps in addiction treatment settings [15-19]. The SUS can be accessed online [20].

### Procedure

After informed consent was obtained, participants were administered the semistructured interview, including a series of open-ended questions exploring a priori hypotheses, but space was made for the emergence of unanticipated themes allowing participants to fully describe their perceptions. Participants were asked for their thoughts on the acceptability of contingency management for routine use in addiction treatment settings in general, and perceived usability of the PROCare Recovery app in particular. Participants were informed that PROCare was a reward-based, technology-enabled recovery monitoring smartphone app delivering micropayments via smart debit card (True Link), with blocking capabilities for achieving recovery goals. Participants were shown the app and presented prompts with actual interview questions on the screen covering various topics and app features (ie, contingency management, outcomes monitoring, daily check-in, smart debit card system, psychoeducational library, appointment scheduling, target behaviors for rewards, and progress charting). Examples of questions included: (1) *What expectations do you have about an app designed to monitor outcomes for patients?* (2) *Do you have any reservations about the PROCare app?* (3) *What do you think of the features of the app?*

During the rewards section of the interview, participants were first asked how open they were to the idea of giving people small amounts of money (ranging from US \$1 to US \$5) for achieving their treatment goals, before being asked how much money in particular they thought it would take to properly incentivize participation each month. Participants were then asked a series of questions focused on the use of the True Link smart debit card system to facilitate delivery of monetary rewards. After participants provided their initial impressions regarding the smart debit card system, they were asked how important it was that such a debit card have blocking capabilities to prevent specified purchases. Participants offered their thoughts on what types of stores, businesses, or spending categories they believed should be automatically blocked for all card users in addiction treatment settings. Participants were also asked to comment on how they thought patients would feel about restrictions being placed on how they could spend the money earned from the PROCare app; that is, whether they believed such a strategy would be viewed as understandable or offensive to most patients receiving addiction treatment.

Examples of questions included: (1) *What are your thoughts on the True Link smart debit card system?* (2) *How do you feel about restrictions being placed on how patients can use the money?* (3) *How receptive do you think patients will be to the non-monetary rewards?* At the conclusion of the interview, participants were administered the SUS to quantify perceived usability of PROCare. Interviews lasted approximately 60 minutes and were conducted via secure video platform (Zoom). Participants received a US \$50 gift card for their participation. Interviews were recorded and automatically transcribed by Zoom. Transcripts were coded and thematically analyzed using Quirkos software by a PhD-level researcher with specialized training in mixed methods [21].

### Ethical Considerations

All study procedures were approved by the Medical Decision Logic, Inc Institutional Review Board (IRB00001558).

### Results

In general, our qualitative analysis revealed participants were very excited about the idea of an app that can help them in their recovery. Qualitative responses were consistent with SUS scores showing an exceptionally high level of perceived usability, as evidenced by a mean participant usability score of 92.2 (range 72.5-100). It was clear that participants perceived a need for a mobile app-based program and that PROCare Recovery seemed to be filling this void. Using such an app was only viewed as beneficial, and no major concerns about using the app were voiced.

When participants were asked for their thoughts about contingency management as delivered by PROCare, patients and providers were overwhelmingly supportive. There were no objections to giving patients small amounts of money for reaching their recovery goals. In fact, the idea of paying patients was generally described as “innovative” and “smart,” with very little concern being voiced. With regard to apparent benefits of using monetary rewards, one patient stated:

*A lot of addicts will use the app just for the money, but it might save their life in the process.*

The micropayment model in particular was described as “wise” because it did not give the patient “all the money at once.” The planned maximum monthly reward limit of US \$150 was viewed as a reasonable amount for patients to earn monthly, as one patient explained:

*People in recovery are broke. [US \$150/month] is just the right amount. Not too much to hurt themselves. Small, which is good for cigarettes, food, and other things.*

Further analysis revealed that participants believed the True Link debit card system was an appropriate safeguard. This was viewed as a superior option to giving patients actual cash, with some even describing the debit card system as a “necessity”. The blocking capabilities were seen as a crucial piece to the debit card system and several suggestions were made regarding which stores, businesses, and activities should be blocked (eg,

cash withdrawals, bars, liquor stores, casinos or online gambling, strip clubs).

## Discussion

### Principal Findings

This study provides important qualitative and quantitative evidence supporting the use of technology-enabled contingency management in real-world opioid treatment settings using a mobile app delivering rewards via smart debit card. Responses from in-depth interviews revealed that not only are patients themselves receptive to the idea of rewarding patients for achieving treatment goals, but providers also reported seeing value in the use of monetary incentives when proper safeguards are in place (blocking capabilities preventing ATM cash withdrawals, etc). Patients and providers both expressed enthusiasm for the micropayment model in which patients can earn small amounts of money (typically around US \$1) for actively participating in their treatment (attending individual therapy appointments, producing negative urinalysis drug screens, attending community-based mutual-help support groups such as Narcotics Anonymous, etc), with the opportunity to earn a maximum monthly amount of US \$150. Of particular interest, patients were in favor of the blocking capabilities, reporting that such spending restrictions “made perfect sense,” especially early in the recovery process. In addition to the qualitative findings, the observed SUS score of 92.2 far exceeds the industry standard of 68 indicating a high level of perceived usability of the PROCare Recovery app.

### Comparison With Prior Work

There has been rapid progress on the innovation front in recent years with respect to emerging mobile app technologies leveraging contingency management in the treatment of OUD [22-24]. The current findings contribute to the extant literature on promising reward-based mobile apps for OUD treatment populations and extend prior work by providing empirical evidence on patient and provider perspectives on the acceptability of providing motivational incentives in routine clinical practice. Several studies [25-27] have identified concerns shared by patients as well as providers tasked with implementing contingency management, including the overreliance on abstinence, fairness, perceived power imbalance,

and how incentives will be spent. One of the most commonly identified concerns about contingency management is how patients will use the money earned (ie, “giving people ‘extra’ money at a vulnerable point in their treatment pathway may do more harm than good”) [27,28]. The current study findings, however, demonstrate that contingency management protocols using “micro” rewards and a mobile app with accompanying “smart” debit card technology—where blocking capabilities and spending restrictions can be put in place—is viewed by both patients and providers as an appropriate safeguard and a critical piece to any reward system.

### Limitations

This study has several limitations, including most notably, a relatively small sample size, thereby limiting the generalizability of the findings. Although a small sample is generally acceptable for a focused, qualitative pilot study designed to inform preliminary technology development, conclusions drawn from such a small sample require replication in a larger scale study. The study sample was also predominately White and did not include any Black, American Indian, or Pacific Islander participants. Further research with a larger, more racially diverse population is warranted. Finally, although this study was able to collect useful data on perceived feasibility of PROCare by way of in-depth qualitative interviews with both patients and providers, a logical next step for future work is to assess for additional indicators of feasibility, including patient access to smartphone technology, level of digital health literacy, and comfort with technology.

### Conclusions

Notwithstanding sample size limitations, current findings suggest strong preliminary evidence that both patients and providers alike appear to be in favor of contingency management for OUD, particularly when using a micropayment model delivered via mobile app with smart debit card and spending parameters for monetary rewards earned. Although patients and providers reported a high level of perceived usability of the PROCare Recovery app, further testing in a large-scale randomized controlled trial is necessary to determine preliminary effectiveness of the PROCare Recovery app as a solution to enhance medication and care plan adherence, and ultimately improve outcomes for OUD treatment populations.

### Conflicts of Interest

Research reported in this publication was supported by the National Institute on Drug Abuse of the National Institutes of Health under Award Number R43DA051298. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. SLP and AYT are Dual Principal Investigators on a Small Business Innovation Research (SBIR) grant from the National Institute on Drug Abuse to develop and test the PROCare Recovery app. SLP is the original creator of the PROCare Recovery app, developed in collaboration with AYT, President of Medical Decision Logic, Inc (mdlogix), the company that received SBIR funding for app development.

### References

1. Ahmad F, Cisewski J, Rossen LM, Sutton P. Provisional drug overdose death counts. National Center for Health Statistics. 2022. URL: <https://www.cdc.gov/nchs/nvss/vsrr/drug-overdose-data.htm> [accessed 2022-09-29]

2. Bolívar HA, Klemperer EM, Coleman SRM, DeSarno M, Skelly JM, Higgins ST. Contingency management for patients receiving medication for opioid use disorder: a systematic review and meta-analysis. *JAMA Psychiatry* 2021 Oct 01;78(10):1092-1102 [FREE Full text] [doi: [10.1001/jamapsychiatry.2021.1969](https://doi.org/10.1001/jamapsychiatry.2021.1969)] [Medline: [34347030](https://pubmed.ncbi.nlm.nih.gov/34347030/)]
3. Proctor SL. Rewarding recovery: the time is now for contingency management for opioid use disorder. *Ann Med* 2022 Dec;54(1):1178-1187 [FREE Full text] [doi: [10.1080/07853890.2022.2068805](https://doi.org/10.1080/07853890.2022.2068805)] [Medline: [35471190](https://pubmed.ncbi.nlm.nih.gov/35471190/)]
4. Fairley M, Humphreys K, Joyce VR, Bounthavong M, Trafton J, Combs A, et al. Cost-effectiveness of treatments for opioid use disorder. *JAMA Psychiatry* 2021 Jul 01;78(7):767-777 [FREE Full text] [doi: [10.1001/jamapsychiatry.2021.0247](https://doi.org/10.1001/jamapsychiatry.2021.0247)] [Medline: [33787832](https://pubmed.ncbi.nlm.nih.gov/33787832/)]
5. Velez FF, Colman S, Kauffman L, Anastassopoulos K, Murphy S, Maricich Y. Real-world changes in US health system hospital-based services following treatment with a prescription digital therapeutic for opioid use disorder. *Hosp Pract (1995)* 2021 Dec;49(5):341-347. [doi: [10.1080/21548331.2021.1956256](https://doi.org/10.1080/21548331.2021.1956256)] [Medline: [34275401](https://pubmed.ncbi.nlm.nih.gov/34275401/)]
6. Federal Register – The Daily Journal of the United States Government. Medicare and State Health Care Programs: Fraud and Abuse; Revisions to Safe Harbors Under the Anti-Kickback Statute, and Civil Monetary Penalty Rules Regarding Beneficiary Inducements (Final Rule); Document citation: 85 FR 77684. 2020 Dec 2. URL: <https://www.federalregister.gov/documents/2020/12/02/2020-26072/medicare-and-state-health-care-programs-fraud-and-abuse-revisions-to-safe-harbors-under-the> [accessed 2022-09-29]
7. Pew Research Center. Mobile Technology and Home Broadband 2021. 2021 Jun 3. URL: <https://www.pewresearch.org/internet/2021/06/03/mobile-technology-and-home-broadband-2021/> [accessed 2022-09-29]
8. Ashford RD, Lynch K, Curtis B. Technology and social media use among patients enrolled in outpatient addiction treatment programs: cross-sectional survey study. *J Med Internet Res* 2018 Mar 06;20(3):e84 [FREE Full text] [doi: [10.2196/jmir.9172](https://doi.org/10.2196/jmir.9172)] [Medline: [29510968](https://pubmed.ncbi.nlm.nih.gov/29510968/)]
9. Jennings L, Lee N, Shore D, Strohminger N, Allison B, Conserve DF, et al. U.S. minority homeless youth's access to and use of mobile phones: implications for mhealth intervention design. *J Health Commun* 2016 Jul;21(7):725-733. [doi: [10.1080/10810730.2015.1103331](https://doi.org/10.1080/10810730.2015.1103331)] [Medline: [27232544](https://pubmed.ncbi.nlm.nih.gov/27232544/)]
10. Nuamah J, Mehta R, Sasangohar F. Technologies for opioid use disorder management: mobile app search and scoping review. *JMIR Mhealth Uhealth* 2020 Jun 05;8(6):e15752 [FREE Full text] [doi: [10.2196/15752](https://doi.org/10.2196/15752)] [Medline: [32501273](https://pubmed.ncbi.nlm.nih.gov/32501273/)]
11. Vilardaga R, Fisher T, Palenski PE, Kumaresan V, Mannelli P, Sweitzer MM, et al. Review of popularity and quality standards of opioid-related smartphone apps. *Curr Addict Rep* 2020 Dec 10;7(4):486-496 [FREE Full text] [doi: [10.1007/s40429-020-00344-6](https://doi.org/10.1007/s40429-020-00344-6)] [Medline: [33777644](https://pubmed.ncbi.nlm.nih.gov/33777644/)]
12. The App Evaluation Model. American Psychiatric Association. 2018. URL: <https://www.psychiatry.org/psychiatrists/practice/mental-health-apps/app-evaluation-model> [accessed 2022-09-29]
13. Nnakwe CC, Cooch N, Huang-Saad A. Investing in academic technology innovation and entrepreneurship: moving beyond research funding through the NSF I-CORPS program. *Technol Innov* 2018 Jun 10;19(4):773-786. [doi: [10.21300/19.4.2018.773](https://doi.org/10.21300/19.4.2018.773)]
14. Brooke J. SUS: A “quick and dirty” usability scale. In: *Usability Evaluation In Industry*. London, UK: Taylor & Francis; 1996.
15. Curtis AC, Satre DD, Sarovar V, Wamsley M, Ly K, Satterfield J. A mobile app to promote alcohol and drug SBIRT skill translation among multi-disciplinary health care trainees: Results of a randomized controlled trial. *Subst Abus* 2022 Nov 11;43(1):13-22. [doi: [10.1080/08897077.2019.1686723](https://doi.org/10.1080/08897077.2019.1686723)] [Medline: [31710269](https://pubmed.ncbi.nlm.nih.gov/31710269/)]
16. Malte CA, Dulin PL, Baer JS, Fortney JC, Danner AN, Lott AMK, et al. Usability and acceptability of a mobile app for the self-management of alcohol misuse among veterans (Step Away): pilot cohort study. *JMIR Mhealth Uhealth* 2021 Apr 08;9(4):e25927 [FREE Full text] [doi: [10.2196/25927](https://doi.org/10.2196/25927)] [Medline: [33830064](https://pubmed.ncbi.nlm.nih.gov/33830064/)]
17. Miguel A, Smith C, Perea N, Johnson K, McDonell M, McPherson S. Development of Automated Reinforcement Management System (ARMS): protocol for a phase i feasibility and usability study. *JMIR Form Res* 2021 Jul 19;5(7):e25796 [FREE Full text] [doi: [10.2196/25796](https://doi.org/10.2196/25796)] [Medline: [34279238](https://pubmed.ncbi.nlm.nih.gov/34279238/)]
18. O'Grady M, Kapoor S, Gilmer E, Neighbors C, Conigliaro J, Kwon N, et al. Point-of-care mobile application to guide health care professionals in conducting substance use screening and intervention: a mixed-methods user experience study. *ACI Open* 2019 Mar 27;03(01):e1-e12. [doi: [10.1055/s-0039-1684002](https://doi.org/10.1055/s-0039-1684002)]
19. Waselewski ME, Flickinger TE, Canan C, Harrington W, Franklin T, Otero KN, et al. A mobile health app to support patients receiving medication-assisted treatment for opioid use disorder: development and feasibility study. *JMIR Form Res* 2021 Mar 23;5(2):e24561 [FREE Full text] [doi: [10.2196/24561](https://doi.org/10.2196/24561)] [Medline: [33620324](https://pubmed.ncbi.nlm.nih.gov/33620324/)]
20. System Usability Scale (SUS). usability.gov. URL: <https://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html> [accessed 2022-09-30]
21. 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)]
22. Velez FF, Colman S, Kauffman L, Ruetsch C, Anastassopoulos K. Real-world reduction in healthcare resource utilization following treatment of opioid use disorder with reSET-O, a novel prescription digital therapeutic. *Expert Rev Pharmacoecon Outcomes Res* 2021 Mar 04;21(1):69-76. [doi: [10.1080/14737167.2021.1840357](https://doi.org/10.1080/14737167.2021.1840357)] [Medline: [33146558](https://pubmed.ncbi.nlm.nih.gov/33146558/)]

23. DeFulio A, Rzeszutek MJ, Furgeson J, Ryan S, Rezania S. A smartphone-smartcard platform for contingency management in an inner-city substance use disorder outpatient program. *J Subst Abuse Treat* 2021 Jan;120:108188. [doi: [10.1016/j.jsat.2020.108188](https://doi.org/10.1016/j.jsat.2020.108188)] [Medline: [33298295](https://pubmed.ncbi.nlm.nih.gov/33298295/)]
24. Velez FF, Colman S, Kauffman L, Ruetsch C, Anastassopoulos K, Maricich YA. Comparison of healthcare resource utilization between patients who engaged or did not engage with a prescription digital therapeutic for opioid use disorder. *Clinicoecon Outcomes Res* 2021;13:909-916 [FREE Full text] [doi: [10.2147/CEOR.S334274](https://doi.org/10.2147/CEOR.S334274)] [Medline: [34754205](https://pubmed.ncbi.nlm.nih.gov/34754205/)]
25. Gagnon M, Guta A, Payne A. "Setting people up for success and then failure" - health care and service providers' experiences of using prize-based contingency management. *Subst Abuse Treat Prev Policy* 2020 Sep 14;15(1):71 [FREE Full text] [doi: [10.1186/s13011-020-00316-z](https://doi.org/10.1186/s13011-020-00316-z)] [Medline: [32928275](https://pubmed.ncbi.nlm.nih.gov/32928275/)]
26. Gagnon M, Payne A, Guta A. What are the ethical implications of using prize-based contingency management in substance use? A scoping review. *Harm Reduct J* 2021 Aug 04;18(1):82 [FREE Full text] [doi: [10.1186/s12954-021-00529-w](https://doi.org/10.1186/s12954-021-00529-w)] [Medline: [34348710](https://pubmed.ncbi.nlm.nih.gov/34348710/)]
27. Sinclair J, Burton A, Ashcroft R, Priebe S. Clinician and service user perceptions of implementing contingency management: a focus group study. *Drug Alcohol Depend* 2011 Dec 01;119(1-2):56-63 [FREE Full text] [doi: [10.1016/j.drugalcdep.2011.05.016](https://doi.org/10.1016/j.drugalcdep.2011.05.016)] [Medline: [21680110](https://pubmed.ncbi.nlm.nih.gov/21680110/)]
28. Cameron J, Ritter A. Contingency management: perspectives of Australian service providers. *Drug Alcohol Rev* 2007 Mar;26(2):183-189. [doi: [10.1080/09595230601184653](https://doi.org/10.1080/09595230601184653)] [Medline: [17364854](https://pubmed.ncbi.nlm.nih.gov/17364854/)]

## Abbreviations

**NIDA:** National Institute on Drug Abuse  
**OIG:** Office of Inspector General  
**OD:** opioid use disorder  
**SBIR:** Small Business Innovation Research  
**SUS:** System Usability Scale

*Edited by A Mavragani; submitted 22.02.22; peer-reviewed by K Langdon, KL Mauco; comments to author 09.05.22; revised version received 30.07.22; accepted 12.09.22; published 05.10.22.*

*Please cite as:*

*Proctor SL, Rigg KK, Tien AY*

*Acceptability and Usability of a Reward-Based Mobile App for Opioid Treatment Settings: Mixed Methods Pilot Study*

*JMIR Form Res* 2022;6(10):e37474

URL: <https://formative.jmir.org/2022/10/e37474>

doi: [10.2196/37474](https://doi.org/10.2196/37474)

PMID: [36197705](https://pubmed.ncbi.nlm.nih.gov/36197705/)

©Steven L Proctor, Khary K Rigg, Allen Y Tien. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 05.10.2022. 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

# Bolstering the Business Case for Adoption of Shared Decision-Making Systems in Primary Care: Randomized Controlled Trial

JoAnn M Sperl-Hillen<sup>1,2</sup>, MD; Jeffrey P Anderson<sup>3</sup>, SCD, MPH; Karen L Margolis<sup>1</sup>, MD, MPH; Rebecca C Rossom<sup>1</sup>, MD, MSCR; Kristen M Kopski<sup>4</sup>, MD; Beth M Averbeck<sup>5</sup>, MD; Jeanine A Rosner<sup>5</sup>, RN; Heidi L Ekstrom<sup>1,2</sup>, MA; Steven P Dehmer<sup>1</sup>, PhD; Patrick J O'Connor<sup>1,2</sup>, MD, MA, MPH

<sup>1</sup>HealthPartners Institute, Bloomington, MN, United States

<sup>2</sup>Research Department, HealthPartners Center for Chronic Care Innovation, Bloomington, MN, United States

<sup>3</sup>Genesis Research, Hoboken, NJ, United States

<sup>4</sup>Medica Health Plan, Minneapolis, MN, United States

<sup>5</sup>HealthPartners, Bloomington, MN, United States

**Corresponding Author:**

JoAnn M Sperl-Hillen, MD

Research Department

HealthPartners Center for Chronic Care Innovation

8170 33rd Ave S

Bloomington, MN, 55425

United States

Phone: 1 952 967 5009

Email: [JoAnn.M.SperlHillen@HealthPartners.Com](mailto:JoAnn.M.SperlHillen@HealthPartners.Com)

## Abstract

**Background:** Limited budgets may often constrain the ability of health care delivery systems to adopt shared decision-making (SDM) systems designed to improve clinical encounters with patients and quality of care.

**Objective:** This study aimed to assess the impact of an SDM system shown to improve diabetes and cardiovascular patient outcomes on factors affecting revenue generation in primary care clinics.

**Methods:** As part of a large multisite clinic randomized controlled trial (RCT), we explored the differences in 1 care system between clinics randomized to use an SDM intervention (n=8) versus control clinics (n=9) regarding the (1) likelihood of diagnostic coding for cardiometabolic conditions using the 10th Revision of the International Classification of Diseases (ICD-10) and (2) current procedural terminology (CPT) billing codes.

**Results:** At all 24,138 encounters with care gaps targeted by the SDM system, the proportion assigned high-complexity CPT codes for level of service 5 was significantly higher at the intervention clinics (6.1%) compared to that in the control clinics (2.9%), with  $P<.001$  and adjusted odds ratio (OR) 1.64 (95% CI 1.02-2.61). This was consistently observed across the following specific care gaps: diabetes with glycated hemoglobin A<sub>1c</sub> (HbA<sub>1c</sub>)>8% (n=8463), 7.2% vs 3.4%,  $P<.001$ , and adjusted OR 1.93 (95% CI 1.01-3.67); blood pressure above goal (n=8515), 6.5% vs 3.7%,  $P<.001$ , and adjusted OR 1.42 (95% CI 0.72-2.79); suboptimal statin management (n=17,765), 5.8% vs 3%,  $P<.001$ , and adjusted OR 1.41 (95% CI 0.76-2.61); tobacco dependency (n=7449), 7.5% vs. 3.4%,  $P<.001$ , and adjusted OR 2.14 (95% CI 1.31-3.51); BMI >30 kg/m<sup>2</sup> (n=19,838), 6.2% vs 2.9%,  $P<.001$ , and adjusted OR 1.45 (95% CI 0.75-2.8). Compared to control clinics, intervention clinics assigned ICD-10 diagnosis codes more often for observed cardiometabolic conditions with care gaps, although the difference did not reach statistical significance.

**Conclusions:** In this randomized study, use of a clinically effective SDM system at encounters with care gaps significantly increased the proportion of encounters assigned high-complexity (level 5) CPT codes, and it was associated with a nonsignificant increase in assigning ICD-10 codes for observed cardiometabolic conditions.

**Trial Registration:** ClinicalTrials.gov NCT 02451670; [https://clinicaltrials.gov/ct2/show/NCT\\_02451670](https://clinicaltrials.gov/ct2/show/NCT_02451670)

(JMIR Form Res 2022;6(10):e32666) doi:[10.2196/32666](https://doi.org/10.2196/32666)



**KEYWORDS**

clinical decision support; primary care; ICD-10 diagnostic coding; CPT levels of service; shared decision-making

## Introduction

Care delivery systems are increasingly considering an array of software products that promote clinical decision support (CDS), care efficiency, and shared decision-making (SDM) in primary care environments. CDS uses computable biomedical information, person-specific data, and inferencing mechanisms to generate helpful information to clinicians, patients, and care teams, as care is being delivered with the objective of reducing errors and adverse events and promoting best practices [1]. CDS can also be used to generate SDM interfaces to facilitate patient engagement and help patients make choices, incorporate personal preferences, and help them prioritize clinical recommendations and decisions [2]. Key features shown to improve the success of SDM products include incorporating them into clinician workflows without disruption, delivery at the right time in the clinical encounter to influence decision-making, and provision of SDM output to patients as well as clinicians [3]. We developed an SDM system that involves patient-centered CDS and a workflow that presents clinicians and patients with printed information about chronic care gaps in low- and high-literacy formats and prioritizes care options based on potential benefits to the individual early in primary care encounters. We have shown that an SDM system with these features achieves high clinician satisfaction rates and sustainable high SDM use, improves glucose and blood pressure (BP) control in patients with diabetes mellitus (DM), lowers 10-year cardiovascular (CV) risk in patients without DM or heart disease, and positively influences the frequency and quality of SDM [4-6].

However, many care systems are operating with tight budgets and facing difficult choices regarding adoption of SDM due to the cost of implementation and maintenance of SDM technology [3,7]. Very few studies have assessed the cost-effectiveness of SDM, and those that have invariably adopt the societal or health insurer perspective. In 1 such study, Gilmer et al [8] estimated the base-case incremental cost-effectiveness of implementing a clinical decision-making system used in SDM for patients with DM at US \$3017 per quality adjusted life year gained [8]. This amount was considered cost-effective by usual standards from a payer perspective [9]. However, most of the cost burden for implementing SDM falls on the care delivery system rather than the payer, and the lack of data needed to estimate the impact of SDM implementation on care delivery system revenue is often cited as a major barrier to adoption [7,10].

The objective of this analysis was to evaluate the impact of an SDM system on diagnostic coding and billing at primary care encounters because these factors substantially impact revenue generation for a care delivery system and can ultimately influence the case for SDM adoption. In a care delivery model that relies on “fee for service (FFS)” reimbursement, it is important that care systems are able to capture billing codes that reflect the extent to which SDM might increase the amount of time, number of clinical issues addressed, and complexity of medical decision-making at patient encounters [11]. In today’s

emerging transition to value-based care agreements, accurate and complete diagnostic coding is related to risk-adjusted reimbursement for the population served. In most health care settings today, both adequate billing service levels and accurate coding of conditions are necessary to optimize revenue and optimally manage the health care needs of the patients and populations they serve [11].

The SDM system studied in this analysis included no specific components to encourage diagnostic coding or influence billing codes. However, it directed clinician attention to care gaps related to diabetes and uncontrolled CV risk factors, and it would be expected to indirectly influence diagnostic coding and billing. Therefore, in an exploratory analysis conducted as part of a multisite randomized controlled trial (RCT) to evaluate the quality impact of the SDM system on patients with high CV risk and serious mental illness, we assessed the effect of using SDM at intervention clinics within 1 medical group on (1) rates of diagnostic coding for SDM-related clinical domains based on the 10th Revision of the International Statistical Classification of Diseases and Related Health Problems (ICD-10) [12] and (2) current procedural terminology (CPT) codes in the level of service used for billing at clinics using the evidence-based SDM system [13].

## Methods

### Study Design and Study Population

This analysis occurred as part of a larger multisite clinical trial (trial registration: NCT 02451670) funded by the National Institute of Mental Health that developed, implemented, and evaluated an SDM system for adults with serious mental illness (SMI), such as schizophrenia, schizoaffective disorder, or bipolar disorder, who die on average 17 years earlier than the rest of their birth cohort, primarily due to CV disease [14]. The objective of the study was to determine if an SDM system targeting reversible CV risk factors would lower the reversible 10-year CV risk for patients with SMI over 12-18 months. The study showed that the rate of increase in the total modifiable CV risk was 4% lower among intervention patients compared to the control, emphasizing the value of using the SDM system for the prompt management of modifiable CV risk factors in the SMI population [15]. Of the 3 participating medical groups, 1 was used for this exploratory analysis of SDM impact on billing and coding. The SDM implementation at this site also included patients with diabetes, CV disease, and high CV risk in addition to those with SMI. For this analysis, we explored the impact of the SDM system on diagnosis and CPT coding at all encounters of adult patients with diabetes, SMI, CV disease, or high reversible CV risk, plus suboptimal control of 1 or more major CV risk factors. The specific inclusion and exclusion criteria are described in more detail below. In this medical group, 17 primary care clinics were randomly assigned to receive (n=8) or not receive (n=9) the SDM system beginning March 15, 2017. The control clinics were scheduled to receive the SDM system 18 months later, in September 2018. The clinic

randomization was conducted using a computer-generated random allocation sequence while ensuring a balance in terms of the clinic size and percentage of patients with Medicaid insurance. Clinic names were concealed until intervention assignment.

Inclusion and exclusion criteria for study analysis eligibility were determined for each study-eligible patient by SDM algorithms at the start of every primary care encounter and included the following.

### **Inclusion Criteria**

These criteria include an office encounter in a primary care department with a patient aged 18 to 75 years and one of the following two clinical criteria:

1. The presence of DM, CV disease, or SMI and not meeting evidence-based goals for one or more of the following major CV risk factors were considered: statin use [16], BP [17], glycemic control [18], weight (BMI > 25 kg/m<sup>2</sup>) [19], tobacco cessation [20], and aspirin use, if indicated [21,22].
2. The reversible 10-year CV risk score was greater than 10% (without DM, CVD, or SMI identified). The reversible CV risk score was the sum of the amount of 10-year CV risk attributable to each of the above risk factors that could potentially be eliminated if the patient were to achieve the guideline-recommended clinical goal. For weight, the reversible risk was the amount of reversible CV risk attributable to a drop of 3 units in the BMI (kg/m<sup>2</sup>), which is equivalent to a 10- to 20-pound weight loss for most individuals.

### **Exclusion Criteria**

ICD-10 visit codes and problem list codes were used to exclude all patients with one or more of the following conditions: hospice or nursing home, active cancer, current or recent pregnancy, and cognitive impairment.

### **Intervention Description**

The evidence-based SDM system directs patient and clinician attention to a patient's personalized care priorities at the point of care. A custom routine programmed in the electronic health record (EHR) gathers key clinical data and securely exchanges it with a web service, where algorithms are applied to identify and prioritize evidence-based CV risk factor care improvement options for patients and clinicians. The algorithms use published risk-prediction equations [23-25] to calculate an individual's

reversible CV risk potential and then prioritize out-of-control CV risk factors in a list form from the most to the least likely to lower CV risk if successfully addressed. Care suggestions include personalized pharmacologic and lifestyle treatment options that account for the patient's current therapy, most recent status of clinical values, distance from clinical goals, and relevant comorbidities. The SDM also provides safety alerts (eg, for drug contraindications and interactions), screening and monitoring reminders, and suggestions for appropriate follow-up intervals [5].

The SDM system is automatically triggered at adult visits when clinic rooming staff enter any BP value into the EHR, as is the case in over 95% of all primary care clinic visits. When web-based clinical algorithms identify a patient who meets the study eligibility criteria, a flag is returned to the EHR that triggers an EHR best practice advisory (BPA) pop-up inviting clinic rooming staff to open (1 click) and print (1 click) the SDM tools for patients and clinicians in intervention clinics. Having paper interfaces available to clinicians and patients at the beginning of the encounter was key to the SDM process and workflow, and using rooming staff to print the interfaces was key to ensure high use rates and exposure to the SDM tools.

### **Printed SDM Interfaces for Patients and Clinicians**

To meet a wide range of health literacy needs, the printed SDM tools included a more detailed "clinician-oriented" decision support interface as well as a companion low-literacy "patient-oriented" interface. The interfaces went through multiple iterations based on feedback received from clinicians and patients during the study. Figure 1 shows the version of the SDM system for a synthetic patient.

The lay/patient version is printed by the rooming nurse and given to the patient to review while waiting in the exam room for the provider, with the following message: "*If you act on the things with high priority or needs attention, you may be able to reduce your danger of a stroke or heart attack. Talk to your doctor about things you can do.*"

The professional/clinician version in Figure 2 is printed by the rooming nurse and placed on the exam room door for rapid review by the provider just before the visit. Uncontrolled CV risk factors are prioritized by the potential absolute risk reduction that may be achieved by managing those risk factors. The data presented in Figure 2 are obtained from web service interfaces for synthetic patients and are not from actual patients.

**Figure 1.** Example of the printed shared decision-making interfaces for patients. A<sub>1c</sub>: glycated hemoglobin; BP: blood pressure; LDL: low-density lipoprotein.

WIZARD®

[Suggestions](#)
[FAQ](#)

Clinical Priorities
Mayo Statin Tool

PRIORITY WIZARD®

[Provider](#)
[Patient](#)
[Print](#)











**TALK TO YOUR DOCTOR ABOUT HOW YOU CAN IMPROVE YOUR HEALTH**

Start the conversation! Use the priorities below as a guide to take action to better your health.

High Priority

Needs Attention


Doing well

 <b>CHOLESTEROL</b>	 Your LDL: 106	A cholesterol lowering drug called a statin may be beneficial for you. Talk to your doctor.
 <b>BLOOD SUGAR</b>	 Your A1C: 7.4	<b>Your Goal: A1C less than 7</b> You may be due for an A1c test. It would be a good idea for you to see a diabetes educator or dietitian.
 <b>WEIGHT</b>	 Your Weight : 220	For support with weight management contact: HP Nutrition Services (952-967-7616), or visit <a href="http://www.healthpartners.com/healthyliving">www.healthpartners.com/healthyliving</a> , or call your clinic.
 <b>BLOOD PRESSURE</b>	 Your Blood Pressure:(119/75)	<b>Ideal BP &lt; 120/80</b>
 <b>TOBACCO</b>	 Your Blood Pressure:(119/75)	<b>Good work maintaining your health!</b>

How are you liking the Wizard tool? Please take a quick survey to provide us feedback.  
[www.healthpartners.com/wizardsurvey](http://www.healthpartners.com/wizardsurvey)


2019 © HealthPartners

**Figure 2.** Example of the printed shared decision-making interfaces for clinicians. A<sub>1c</sub>: glycated hemoglobin; ALT: alanine amino-transferase; BP: blood pressure; CV: cardiovascular; HDL: high-density lipoprotein; LDL: low-density lipoprotein; TC: total cholesterol; TRIG: triglycerides; UMACR: urine microalbumin to creatinine ratio.



Suggestions
FAQ

Clinical Priorities
Mayo Statin Tool



Provider
Patient
Print

**Relevant Conditions:** Diabetes

**10-year Cardiovascular Risk :** 3.4% (Risk of stroke or heart attack over the next 10 years)

<b>#1 LIPID</b>	<p><b>Goal:</b> Consider statin initiation.</p> <p><b>Treatment Considerations</b></p> <ul style="list-style-type: none"> <li>Statin initiation or intensification is recommended due to diabetes and CV risk. Many experts recommend moderate therapy.</li> </ul> <p><b>Other Alerts</b></p> <ul style="list-style-type: none"> <li>Baseline ALT measurement is recommended by many experts prior to statin therapy initiation.</li> </ul>	<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr><th colspan="3">Results</th></tr> </thead> <tbody> <tr><td>LDL (mg/dl)</td><td>106</td><td>10/11/19</td></tr> <tr><td>HDL (mg/dl)</td><td>54</td><td>10/11/19</td></tr> <tr><td>TRIG (mg/dl)</td><td>204</td><td>10/11/19</td></tr> <tr><td>TC (mg/dl)</td><td>201</td><td>10/11/19</td></tr> </tbody> </table>	Results			LDL (mg/dl)	106	10/11/19	HDL (mg/dl)	54	10/11/19	TRIG (mg/dl)	204	10/11/19	TC (mg/dl)	201	10/11/19	No Medications
Results																		
LDL (mg/dl)	106	10/11/19																
HDL (mg/dl)	54	10/11/19																
TRIG (mg/dl)	204	10/11/19																
TC (mg/dl)	201	10/11/19																
<b>#2 GLYCEMIC CONTROL</b>	<p><b>Goal:</b> A1C &lt;= 6.9</p> <p><b>Other Alerts</b></p> <ul style="list-style-type: none"> <li>Need updated A1c to guide recommendations.</li> <li>Consider monthly visits and/or interim phone calls until A1c goal achieved.</li> <li>Urinary albumin excretion test (e.g. UMACR) may be due.</li> <li>Diabetes educator and/or dietitian support is suggested.</li> </ul>	<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr><th colspan="3">Results</th></tr> </thead> <tbody> <tr><td>A1C</td><td>7.4</td><td>8/16/18</td></tr> </tbody> </table>	Results			A1C	7.4	8/16/18	<p><b>Medications</b></p> <ul style="list-style-type: none"> <li>Insulin Aspart Soln Pen-injector 100 Unit/ML</li> <li>Insulin Glargine Soln Pen-Injector 100 Unit/ML</li> </ul>									
Results																		
A1C	7.4	8/16/18																
<b>BMI (WEIGHT)</b>	<p><b>Treatment Considerations</b></p> <ul style="list-style-type: none"> <li>Discuss advantages of reducing weight by 10-20 lbs. Potential actions are listed on patient interface.</li> </ul>	<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr><th colspan="3">Results</th></tr> </thead> <tbody> <tr><td>Weight(lbs)</td><td>220</td><td>10/11/19</td></tr> <tr><td>BMI</td><td>29.84</td><td>10/11/19</td></tr> </tbody> </table>	Results			Weight(lbs)	220	10/11/19	BMI	29.84	10/11/19	No Medications						
Results																		
Weight(lbs)	220	10/11/19																
BMI	29.84	10/11/19																
<b>RELEVANT INFORMATION AND RECOMMENDATIONS</b>	<p><b>BLOOD PRESSURE</b></p> <ul style="list-style-type: none"> <li>Consider adding hypertension to the patient problem list based on meeting guidelines from the American Heart Association.</li> <li>A referral for ambulatory BP monitoring can be helpful to guide diagnostic and treatment decisions.</li> </ul> <p><b>TOBACCO</b></p> <ul style="list-style-type: none"> <li>Passive smoking exposure is noted. Suggest smoke free home and auto.</li> </ul> <p><b>ASPIRIN</b></p> <ul style="list-style-type: none"> <li>Aspirin is not recommended for primary prevention for adults at low to moderate CV risk (10-year CV risk less than 10%).</li> </ul>	<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr><th colspan="3">Results</th></tr> </thead> <tbody> <tr><td>BP (mm Hg)</td><td>119/75</td><td>10/11/19</td></tr> <tr><td>Last BP (mm Hg)</td><td>125/90</td><td>5/16/19</td></tr> <tr><td>Smoking Status/Review Date</td><td>PASSIVE</td><td>10/11/19</td></tr> <tr><td>Smokeless Tobacco</td><td>NOT ASKED</td><td>10/11/19</td></tr> </tbody> </table>	Results			BP (mm Hg)	119/75	10/11/19	Last BP (mm Hg)	125/90	5/16/19	Smoking Status/Review Date	PASSIVE	10/11/19	Smokeless Tobacco	NOT ASKED	10/11/19	No medications
Results																		
BP (mm Hg)	119/75	10/11/19																
Last BP (mm Hg)	125/90	5/16/19																
Smoking Status/Review Date	PASSIVE	10/11/19																
Smokeless Tobacco	NOT ASKED	10/11/19																

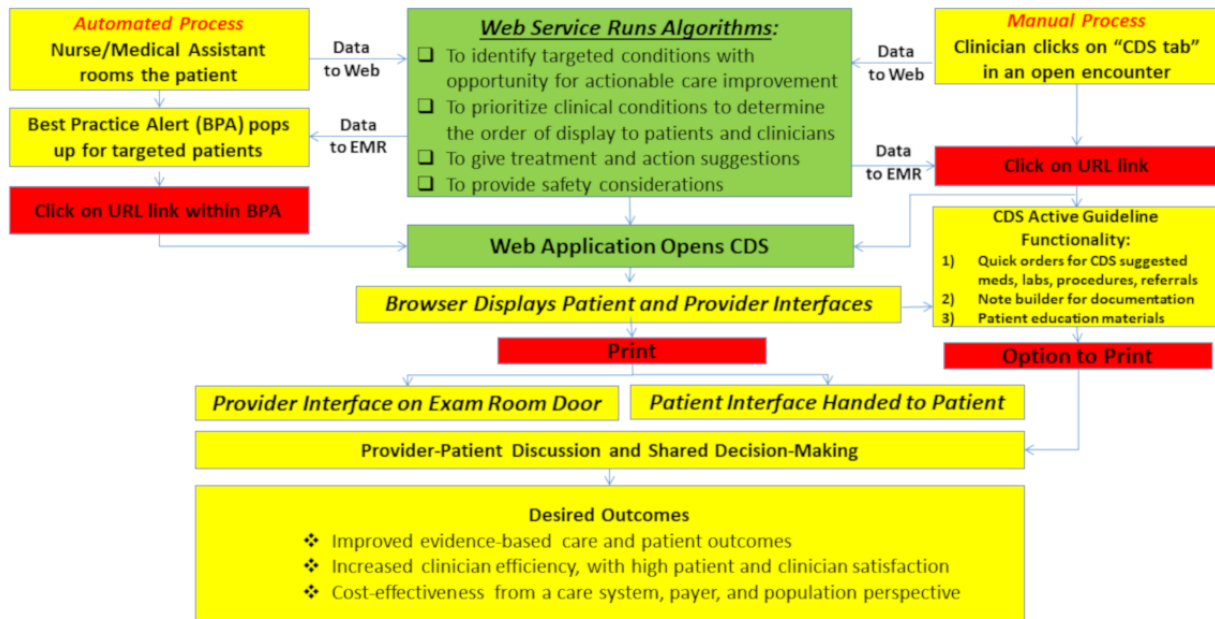
**Disclaimer:** The Priority Wizard® suggestions are based on electronically available data and are not intended to be a substitute for clinical judgment. Alternative actions to those that Wizard suggest may be indicated. Exercise independent clinical judgment, review allergies, and follow product labelling instructions before choosing Wizard prescribing suggestions.

### Recommended Workflow

The automated BPA pop-up for targeted patient encounters in the recommended workflow shown in Figure 3 was for rooming staff to print and give the patient-oriented interface to patients while they waited for the clinician, a design that promoted engagement and improved efficiency when making important decisions for care priorities. The “clinician-oriented” interface was given to clinicians before the encounter to review patient CV risk factor–related clinical priorities. Clinicians at intervention clinics also could manually view the SDM within the EHR for any adult patient (independent of study eligibility

or CV risk) from an SDM activity tab visible in all open encounters. Later in the study, when the SDM was opened from this tab, the SDM display included active guideline features that facilitated quick orders for medications, labs, procedures, and referrals based on recommendation options generated by the SDM algorithms. Rooming staff in the control clinics did not receive the BPA and clinicians could not display the SDM tools. Providers in both intervention and control clinics could use a “smart dot phrase” within encounter notes to summarize and document the patient’s 10-year CV risk score and CV risk factors not at goal.

**Figure 3.** Workflow for shared decision-making use in primary care encounters. CDS: clinical decision support; EMR: electronic medical record.

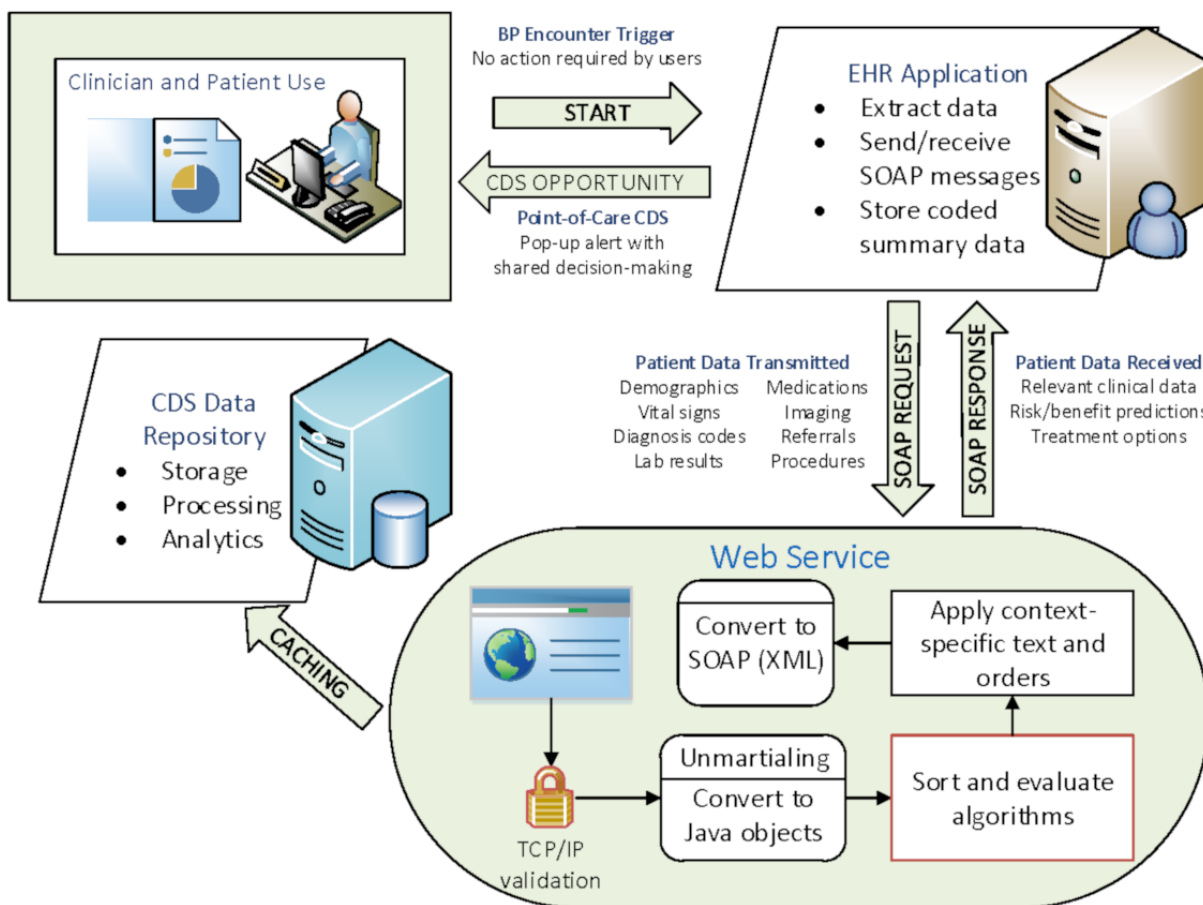


### Technical Description of the SDM System Functionality

The SDM system shown in Figure 4 consists of three interconnected components: (1) a custom routine to extract data from the EHR, (2) web services running on server clusters that process algorithms, and (3) a website that displays the SDM patient and clinician interfaces. The first component of the SDM system involves installation of a program using a Massachusetts General Hospital Utility Programming System (MUMPS) routine in Epic's database management system (Epic Systems Corporation) called Chronicles. When the BP is entered at the encounter, it triggers the MUMPS routine to extract all the data needed to run the algorithms and packages the information into a Simple Object Access Protocol messaging request that uses text in the XML format. Epic's Interconnect Infrastructure is used to connect to the web service over https that contains a unique web service call identifier. The web service then processes the data, runs algorithms, stores the unique call identifier, and returns results to the EHR. The EHR code then

processes the response and extracts and saves relevant pieces of information into discrete data fields. For targeted patients with care gaps identified within the web service response, the BPA contains a URL link to the web application that displays the SDM tools. When the rooming staff or clinician clicks on the URL link containing the unique patient identifier, the patient's personalized SDM tools are displayed in real time within the EHR browser. With 1 additional click, the tools can also be printed for patients and clinicians to use in SDM. To the end user, the process to display the SDM takes less than 2 seconds and appears to be entirely integrated within the EHR experience. All data are exchanged via transport layer security with extra layers of security enforced via exchange of unique identifiers and IP address authorization. The SDM system currently uses Epic web services where possible, and these can be replaced with Fast Healthcare Interoperability Resources (FHIR) as FHIR features mature and offer improved interoperability with other EHR systems and software.

**Figure 4.** Technology behind the shared decision-making system. BP: blood pressure; CDS: clinical decision support; EHR: electronic health record; SOAP: Simple Object Access Protocol; TCP: Transmission Control Protocol.



**Training**

Intervention clinic staff were offered a 1-hour luncheon training to introduce the SDM tools and learn the recommended workflow before the SDM system became available on March 15, 2017. Over the intervention period, nursing leaders at intervention clinics were given monthly reports of how often the tool was being printed for the target population. From our previous implementation experience, these monthly process measurements were essential for promoting and maintaining high SDM use rates. With this implementation process and workflow, the SDM was printed for 75% of eligible encounters on average after the first month, and these rates were sustained for the duration of the intervention.

**Analysis**

On entering the BP for all encounters, data exchange with the SDM web service occurred in both the SDM system intervention and control clinics for analysis purposes and was saved in a data repository, but the SDM tools were displayed only at the SDM system intervention clinics. We evaluated all eligible patient encounters occurring from March 15, 2017, to December 31, 2017, in the intervention and control clinics. Encounter-level data from the SDM repository were later merged with data extracted from the EHR (Epic Clarity) [26], which included the ICD-10 visit diagnostic codes and CPT level of service for the same encounters. The objective of this analysis was to evaluate differences between the intervention and control clinics in terms

of the following: (1) likelihood of ICD-10 diagnostic coding for DM (E10-E11), hypertension (I10-I16), hyperlipidemia (E78), obesity (E66), smoking (F17 and Z72), and CVD (G45, I20-25, I63-70, and I74); and (2) CPT billing codes documented by clinicians as straightforward (level 2, CPT 99212), low complexity (level 3, CPT 99213), moderate complexity (level 4, CPT 99214), or high complexity (level 5, CPT 99215).

Documentation of the CPT levels of service was done by clinicians based on the intensity and complexity of medical decision-making at encounters using the recommended criteria related to the nature and number of clinical problems, amount and complexity of the data reviewed, and the risk of morbidity and mortality to the patient [13]. If counseling or coordination of care accounts for more than 50% of the visit, the CPT service level can be based on the length of the visit as well [13].

Descriptive summaries of diagnostic and billing codes were tabulated, including frequencies, means, 95% CIs, and percentiles of the continuous distribution, where applicable. We used the Fisher exact test for unadjusted comparisons between intervention and control clinics. Generalized linear mixed regression models (with a binomial distribution and logit link) were used for covariate adjustment and random intercepts to account for clustering at the provider and clinic levels. All modeling results reported here are adjusted for the age (continuous), gender (female/male), and race (White, Black, and other/unknown) of the patients. The P values reported are 2-sided. Analyses were conducted using SAS (version 9.4, SAS

Institute) and R (version 3.4.3, R Foundation for Statistical Computing).

### Ethics Approval

The study was reviewed in advance, approved, and monitored by the HealthPartners Institutional Review Board (IRB, reference: 13-154). The IRB approved waiver of written consent from participants.

## Results

### Analysis Population

During the 9.5-month evaluation period, 32,735 primary care encounters with 18,070 unique adult patients were identified.

**Table 1** summarizes the demographic characteristics of the eligible encounters. Approximately half of the encounters (16,335/32,735, 49.9%) were with female patients; the mean age was 58.1 years (slightly older for the intervention clinics, 59.0 years vs 57.2 years), with 75% (13,553/18,070) of patients being White. The mean BMI was 33.0 kg/m<sup>2</sup>. CVD was present in 20.4% (6678) of the encounters, and of the encounters in which CVD was not identified, the mean 10-year estimated CV risk [27] was 15.8%. Type 2 diabetes was identified in 69.7% (22,816), type 1 diabetes in 4.2% (1375), and hypertension in 67.7% (22,162) of the encounters.

**Table 1.** Demographic characteristics of patient encounters by clinic intervention status, 2017.

Characteristic	All encounters (N=32,735)	Control clinics (N=16,417)	Intervention clinics (N=16,318)	P value
<b>Age in years, n (%)</b>				<.001
18-39	2782 (8.5)	1478 (9)	1289 (7.9)	
40-49	4386 (13.4)	2529 (15.4)	1877 (11.5)	
50-59	8642 (26.4)	4515 (27.5)	4145 (25.4)	
60-69	11,163 (34.1)	5336 (32.5)	5826 (35.7)	
70-75	5761 (17.6)	2561 (15.6)	3182 (19.5)	
<b>Gender, n (%)</b>				.1
Female	16,335 (49.9)	8110 (49.4)	8208 (50.3)	
Male	16,400 (50.1)	8307 (50.6)	8110 (49.7)	
<b>Race, n (%)</b>				<.001
White	24,346 (75.2)	12,198 (74.3)	12,418 (76.1)	
Black	4223 (12.9)	1904 (11.6)	2333 (14.3)	
Other/unknown	3895 (11.9)	2331 (14.2)	1567 (9.6)	

### Diagnostic Coding

**Table 2** shows the proportion of encounters with each condition and care gap identified by the SDM system that included a corresponding visit diagnostic code. There was a significant increase in diagnostic coding for almost all the cardiometabolic

conditions that the SDM addressed. Odds ratios (ORs) from generalized linear mixed models were generally consistent with unadjusted comparisons; however, these estimates were not statistically significant, possibly due to a relatively small sample of clinics (N=17).

**Table 2.** Proportion of encounters with the 10th Revision of the International Classification of Diseases coding for clinical conditions and care gaps identified by shared decision-making algorithms based on intervention clinic status, 2017<sup>a</sup>.

Condition subgroup	Control clinics (N=16,417)	Intervention clinics (N=16,318)	P value	Adjusted odds ratio	95% CI
Diabetes (N=24,138), n (%)	17,186 (71.2)	18,055 (74.8)	<.001	1.04	0.8-1.35
Diabetes with HbA <sub>1c</sub> <sup>b</sup> above individualized goal (N=12,786), n (%)	9705 (75.9)	10,280 (80.4)	<.001	1.12	0.86-1.47
Diabetes with HbA <sub>1c</sub> ≥8% (N=8463), n (%)	6618 (78.2)	6931 (81.9)	<.001	1.15	0.89-1.49
Hypertension (N=22,127), n (%)	12,834 (58)	13,542 (61.2)	<.001	1.04	0.73-1.48
Hypertension with BP <sup>c</sup> above goal (N=8515), n (%)	5779 (68.1)	5926 (69.6)	.13	1.04	0.68-1.59
Suboptimal lipid management (N=17,765), n (%)	5330 (30)	5738 (32.3)	.002	1.01	0.67-1.51
ASCVD <sup>d</sup> (N=6679), n (%)	1656 (24.8)	1803 (27)	.041	1.18	0.93-1.51
Tobacco use (N=7449), n (%)	1952 (26.2)	2451 (32.9)	<.001	1.38	0.98-1.95
BMI≥30 kg/m <sup>2</sup> (N=19,838), n (%)	2956 (14.9)	3491 (17.6)	<.001	1.19	0.73-1.94

<sup>a</sup>Age (continuous), sex (female/male), and race (White/Black/other) were included in multivariable logistic regression models.

<sup>b</sup>HbA<sub>1c</sub>: glycated hemoglobin A<sub>1c</sub>.

<sup>c</sup>BP: blood pressure.

<sup>d</sup>ASCVD: atherosclerotic cardiovascular disease.

### CPT Levels of Service

**Table 3** shows the proportion of encounters coded as “5,” indicating a high-complexity CPT level (as opposed to other lower CPT levels 2-4), for selected subgroups of encounters with targeted care gaps, glycated hemoglobin A<sub>1c</sub> (HbA<sub>1c</sub>)>8%, hypertension with BP over goal, suboptimal lipid management, obesity, and tobacco dependency. There were higher proportions of encounters with high-complexity codes for all SDM-eligible encounters. This held true for subgroups of encounters with

specific identified conditions, including DM, hypertension, statin use not at goal, CVD, smoking, obesity, and high reversible risk (without DM or CVD). In generalized linear mixed models, accounting for clinic clustering and demographic factors, encounters were statistically significantly more likely to be coded as “5” (high complexity) in intervention clinics overall (OR 1.64, 95% CI 1.02-2.61), and in patients with DM and HbA<sub>1c</sub>>8% (OR 1.93, 95% CI 1.01-3.67), and active smokers (OR 2.14, 95% CI 1.31-3.51).

**Table 3.** Proportion of encounters coded as 5, indicating a high-complexity current procedural terminology level for clinical conditions identified by shared decision-making algorithms based on clinic intervention status, 2017<sup>a</sup>.

Condition subgroup	Control clinics (N=16,417)	Intervention clinics (N=16,318)	P value	Adjusted odds ratio	95% CI
All encounters (N=32,735), n (%)	949 (2.9)	1997 (6.1)	<.001	1.64	1.02-2.61
Diabetes (N=24,138), n (%)	724 (3)	1472 (6.1)	<.001	1.44	0.73-2.85
Diabetes with HbA <sub>1c</sub> <sup>b</sup> above individualized goal (N=12,786), n (%)	384 (3)	793 (6.2)	<.001	1.80	0.92-3.52
Diabetes with HbA <sub>1c</sub> ≥8% (N=8463), n (%)	288 (3.4)	609 (7.2)	<.001	1.93	1.01-3.67
Hypertension (N=22,127), n (%)	708 (3.2)	1416 (6.4)	<.001	1.36	0.7-2.62
Hypertension with BP <sup>c</sup> above goal (N=8515), n (%)	315 (3.7)	553 (6.5)	<.001	1.42	0.72-2.79
Suboptimal lipid management (N=17,765), n (%)	533 (3)	1030 (5.8)	<.001	1.41	0.76-2.61
ASCVD <sup>d</sup> (N=6679), n (%)	274 (4.1)	448 (6.7)	<.001	1.43	0.77-2.63
Tobacco use (N=7449), n (%)	253 (3.4)	559 (7.5)	<.001	2.14	1.31-3.51
BMI≥30 kg/m <sup>2</sup> (N=19,838), n (%)	575 (2.9)	1230 (6.2)	<.001	1.45	0.75-2.8

<sup>a</sup>Age (continuous), sex (female/male), and race (White/Black/other) were included in multivariable logistic regression models.

<sup>b</sup>HbA<sub>1c</sub>: glycated hemoglobin A<sub>1c</sub>.

<sup>c</sup>BP: blood pressure.

<sup>d</sup>ASCVD: atherosclerotic cardiovascular disease.



## Discussion

### Principal Findings and Implications

SDM has been recommended for years as a strategy to improve outcomes for patients with chronic disease [28]. However, in the context of other technical accomplishments of this decade, the adoption of SDM beyond rather simple process prompts and reminders has been incredibly slow. The reasons for slow uptake are numerous, including the challenges related to developing, maintaining, and updating SDM content; workflow constraints in busy health care settings; and lack of evidence, until recently, directly correlating SDM use with improved patient outcomes [29]. High-quality SDM has become more widely available to improve quality of care and promote evidence-based standards [30]. To avoid influencing patient and clinician behavior and medical decision-making in nonevidence-based ways, it is important that the SDM developers avoid financial conflicts of interest, use of medication brand names, and biases introduced through commercialization strategies. However, adoption and implementation of high-quality SDM can be costly to care systems, and almost no data are available to describe directly how SDM impacts coding and billing factors that can affect revenue generation.

The SDM system implemented for this study was developed at HealthPartners Institute through a series of federally funded research grants by a team with no financial conflicts of interest and with the main objective of improving patient outcomes. Previous versions of the SDM system had been proven to improve patient outcomes for targeted individuals with diabetes and several cardiometabolic conditions as well as SMI in RCTs [4,5,15,31]. It required minimal staff training for implementation. It has been integrated into external care systems through business associate agreements, service agreements, and assurance of secure data transfer between EHRs and the SDM web service. It currently requires about 4 to 6 months of commitment by the recipient organization to conduct data mapping, programming, and testing prior to implementation, although this work may be streamlined in the future with increasing EHR data interoperability and improved FHIR applications. This SDM is currently in use at all HealthPartners and Park Nicollet primary care clinic systems in Minnesota and Wisconsin and external care organizations in rural Minnesota and 10 other states through collaborative research agreements, with over 250,000 web service calls per month. Because SDM algorithms are maintained and the SDM system output is delivered through web-based functionality, there are no dissemination-related geographic limits or boundaries to overcome. When research projects have ended, the annual costs for keeping the SDM system clinically up to date for diabetes and cardiovascular conditions and maintaining required informatics technology have been modest (estimated at US \$200,000 annually) and have been shared by participating care systems. However, even for a clinically successful SDM technology such as this one, dissemination and scalability efforts have been hampered by the inability to demonstrate the value proposition to care delivery systems, with the typical entities deciding whether to adopt the SDM and pay for its integration and maintenance [1].

Although this analysis was exploratory in nature, it was rigorously conducted, and the findings were quite consistent across all coding variables assessed. The findings were compatible with what would be expected if the SDM increased clinician time and attention to important care gaps identified by the SDM. The increase in the appropriate CPT billing codes observed in SDM-targeted encounters is important to care delivery models that rely on FFS reimbursement to capture the increased time and level of medical decision-making at these encounters. Without this appropriate alignment of billing codes, provision of additional value in care may go inadequately incentivized due to reduced FFS reimbursement [32]. Providers tend to systematically underestimate the value of their medical decision-making, which can lead to reduced revenue to support high-quality care [11]. To the extent that using this SDM system increased the amount of time, number of clinical issues addressed, and complexity of medical decision-making at encounters, the increased levels of CPT coding observed are clinically justified and may, as in several of our published studies, improve patient health [11].

In many health systems, FFS reimbursement models are transitioning to or being blended with value-based models [32,33] driven by federal programs such as medical homes and accountable care organizations [33,34]. Value-based models are also being adopted by commercial lines of business, with nearly two-thirds of payments now based to some extent on value [35,36]. Many of these models use a risk adjustment factor (RAF) based on patterns of diagnostic coding to determine the amount of payment to appropriately care for patient populations, assigning a higher RAF to the care of more complex patient populations [33,37]. For many health care organizations, attention to diagnostic coding for risk adjustment has become a top priority to improve care and sustain appropriate reimbursement for the populations they serve under value-based agreements. Some care systems are implementing software programs to explicitly promote accurate diagnosis of chronic conditions [38] in conjunction with mechanisms to ensure sufficient clinical documentation of care to support the diagnostic codes [39,40]. This SDM system did not contain features to explicitly encourage increased diagnostic coding at visits. The observed increases in CPT coding for chronic conditions addressed by the SDM in intervention clinics were not enough to achieve statistical significance with the limited number of clinics included in the analysis (N=17), but further research is warranted, given the consistency of coding changes.

Previous work has established the clinical benefits of using this SDM system, but fostering the business case for implementation and maintenance is critical to scalability and broad dissemination of SDM technology. For this SDM system, implementation and maintenance costs (excluding research-related costs) are known. A formal cost-effectiveness analysis demonstrated cost-effectiveness with likely cost savings to payers at scale [8], but more research is needed to demonstrate that the use of SDM systems does not negatively impact revenue generation. The data presented here demonstrate for the first time that outpatient SDM use at the point of care for patients with DM, SMI, and high CV risk increases high-complexity CPT level of service codes. It accomplishes this by broadening the clinical content

of the visit while guiding clinician and patient attention to specific evidence-based clinical actions with potential substantive benefits to a particular patient at the time of a clinical encounter.

These effects of the SDM system could improve short-term revenue generation for a care delivery system. For example, under the assumption that scheduling systems did not change and physician productivity was not affected, the estimated magnitude of revenue generation based on these CPT data for targeted encounters (34,300 SDM-eligible visits) would be approximately US \$63,919 over 12 months in a model that assumed that all encounters received published Medicare FFS CPT reimbursement rates (published 2018 revenue rates for CPT codes 99212, 99213, 99214, and 99215 are US \$45, 74, 109, and 148, respectively) [13]. In a pure value-based reimbursement model or a blended model with FFS, any positive impacts on diagnostic coding and quality of care would also be expected to increase revenue. As the clinical scope of SDM technology expands beyond CV domains, the extent of the SDM impact on a higher proportion of primary care visits could further enhance the SDM business case.

A number of factors limit the interpretation of the data we present here. First, the data were derived from a single care delivery system and should be replicated elsewhere and in larger studies. The SDM system we evaluated is now being used in many other care systems, which would enable such additional analyses. However, other SDM developers should consider assessing revenue impacts and the impact on physician productivity, such as the production of relative value units, which may vary according to payment models. Second, the value of the increased revenue is only justified by improved quality of clinical care delivered and improved clinical outcomes. Any increases in revenue from billing and coding changes due to the intervention would be in addition to what could be expected through higher incentive payments from better quality outcomes. We have shown in published clinical randomized trials that this SDM system improves clinical outcomes, but future investigations should jointly consider the

clinical and economic impacts of SDM technology from the point of view of both the payer and the care delivery system. Third, we focused on encounters made by patients who had the potential for diabetes improvement or CV risk reduction at their clinic visit. It remains to be seen if similar findings would apply to other chronic disease conditions or different acute or preventive care needs. Our results relate to CPT coding practices in a care delivery system that provides close oversight and routine audits of clinician coding practices to ensure accuracy and avoid fraud. Coding practices, and thus the impact of SDM systems on coding practices, may vary across care delivery organizations. The data used to for this analysis were collected before the 2021 changes to the Center for Medicare and Medicaid Services (CMS) Physician Fee Schedule to “reflect a broader administration-wide strategy to create a health care system that results in better accessibility, quality, affordability, empowerment, and innovation” [41]. Further studies are needed to assess the SDM impact with these newest CMS changes that attempt to simplify billing and coding for office-based services and compensate physicians for additional time spent with patients. Lastly, there are SDM-related factors that could positively influence revenue generation in any of these models, but they must be considered in the context of implementation and maintenance costs of SDM systems and the costs of promoting treatments that may be more intensive.

## Conclusions

This analysis demonstrates that use of an SDM system with proven clinical effectiveness was associated with significantly higher levels of CPT level of service 5 coding and with consistent but nonsignificant increases in ICD-10 coding at routine primary care encounters of patients with diabetes and uncontrolled cardiometabolic conditions. An appropriate shift in CPT coding was observed with a significantly increased proportion of encounters coded as high complexity for patients with poorly controlled diabetes and tobacco dependence. The study provides novel and important information that may inform the business decisions related to implementation of SDM technology to improve quality of care for targeted conditions.

---

## Acknowledgments

This work was funded by the National Institute of Mental Health (grant U19MH092201) for a study called “A Randomized Clinical Trial to Reduce Cardiovascular Risk in Adults with SMI.” JMS-H and PJO'C also received grant support from the National Institute of Diabetes and Digestive and Kidney Diseases (award P30DK092924) Health Delivery Systems Center for Diabetes Translational Research. The research received no funding from commercial sources. The contents of this article reflect the views of the authors and do not necessarily reflect the views of any specific care delivery system, or the position of the funding agency or the US government.

---

## Data Availability

The data sets used and analyzed in this study are available from the corresponding author on reasonable request.

---

## Authors' Contributions

JMS-H is a principal investigator and co-investigator on research studies and conceived the intervention, contributed to the study design, conducted data analysis, and drafted the paper. JPA conceived and designed analysis procedures, collected data, contributed data and analysis tools, performed data analysis, and wrote the methods section of the paper, along with being the lead statistician for data analysis. KLM contributed to data analysis, discussed results, and was involved in preparing the final manuscript. RCR

is the principal investigator for the main study of this analysis, involved in intervention implementation, data analysis, result interpretation, and preparing the final manuscript. KMK is a co-investigator and care system leader involved in implementation of the intervention, data analysis, result interpretation, and preparing the final manuscript. BMA is a care system leader who participated in intervention implementation, result interpretation, and final manuscript preparation. JAR is a care system leader involved in implementation of the intervention, result interpretation, and final manuscript preparation. HLE is the project leader for CDS implementation and research and was involved in data collection, result interpretation, and final manuscript preparation. SPD is an economist who contributed to data analysis and final manuscript preparation. PJO'C is a principal investigator and co-investigator on research studies that conceived the intervention, was involved in intervention implementation, contributed to the study design and implementation of the research, performed data analysis, and contributed significantly toward writing the manuscript.

## Conflicts of Interest

None declared.

## Multimedia Appendix 1

CONSORT-eHEALTH checklist (V 1.6.1).

[\[PDF File \(Adobe PDF File\), 1181 KB - formative\\_v6i10e32666\\_app1.pdf\]](#)

## References

1. Tchong JE, Bakken S, Bates DW, Bonner HIII, Gandhi TK, Josephs M, et al, editors. *Optimizing Strategies for Clinical Decision Support* summary of a Meeting Series. Washington, DC: National Academy of Medicine; 2017.
2. Beach MC, Sugarman J. Realizing shared decision-making in practice. *JAMA* 2019 Sep;322(9):811-812 [[FREE Full text](#)] [doi: [10.1001/jama.2019.9797](https://doi.org/10.1001/jama.2019.9797)] [Medline: [31343669](https://pubmed.ncbi.nlm.nih.gov/31343669/)]
3. Lobach D, Sanders GD, Bright TJ, Wong A, Dhurjati R, Bristow E, et al. Enabling health care decisionmaking through clinical decision support and knowledge management. *Evid Rep Technol Assess (Full Rep)* 2012 Apr(203):1-784. [Medline: [23126650](https://pubmed.ncbi.nlm.nih.gov/23126650/)]
4. O'Connor PJ, Sperl-Hillen JM, Rush WA, Johnson PE, Amundson GH, Asche SE, et al. Impact of electronic health record clinical decision support on diabetes care: a randomized trial. *Ann Fam Med* 2011 Jan;9(1):12-21 [[FREE Full text](#)] [doi: [10.1370/afm.1196](https://doi.org/10.1370/afm.1196)] [Medline: [21242556](https://pubmed.ncbi.nlm.nih.gov/21242556/)]
5. Sperl-Hillen JM, Crain AL, Margolis KL, Ekstrom HL, Appana D, Amundson G, et al. Clinical decision support directed to primary care patients and providers reduces cardiovascular risk: a randomized trial. *J Am Med Inform Assoc* 2018 Sep;25(9):1137-1146 [[FREE Full text](#)] [doi: [10.1093/jamia/ocy085](https://doi.org/10.1093/jamia/ocy085)] [Medline: [29982627](https://pubmed.ncbi.nlm.nih.gov/29982627/)]
6. Kharbanda EO, Asche SE, Sinaiko AR, Ekstrom HL, Nordin JD, Sherwood NE, et al. Clinical decision support for recognition and management of hypertension: a randomized trial. *Pediatrics* 2018 Feb;141(2):e20172954 [[FREE Full text](#)] [doi: [10.1542/peds.2017-2954](https://doi.org/10.1542/peds.2017-2954)] [Medline: [29371241](https://pubmed.ncbi.nlm.nih.gov/29371241/)]
7. Jacob V, Thota AB, Chattopadhyay SK, Njie GJ, Proia KK, Hopkins DP, et al. Cost and economic benefit of clinical decision support systems for cardiovascular disease prevention: a community guide systematic review. *J Am Med Inform Assoc* 2017 May;24(3):669-676 [[FREE Full text](#)] [doi: [10.1093/jamia/ocw160](https://doi.org/10.1093/jamia/ocw160)] [Medline: [28049635](https://pubmed.ncbi.nlm.nih.gov/28049635/)]
8. Gilmer TP, O'Connor PJ, Sperl-Hillen JM, Rush WA, Johnson PE, Amundson GH, et al. Cost-effectiveness of an electronic medical record based clinical decision support system. *Health Serv Res* 2012 Dec;47(6):2137-2158 [[FREE Full text](#)] [doi: [10.1111/j.1475-6773.2012.01427.x](https://doi.org/10.1111/j.1475-6773.2012.01427.x)] [Medline: [22578085](https://pubmed.ncbi.nlm.nih.gov/22578085/)]
9. Braithwaite RS, Meltzer DO, King JT, Leslie D, Roberts MS. What does the value of modern medicine say about the \$50,000 per quality-adjusted life-year decision rule? *Med Care* 2008 Apr;46(4):349-356. [doi: [10.1097/MLR.0b013e31815c31a7](https://doi.org/10.1097/MLR.0b013e31815c31a7)] [Medline: [18362813](https://pubmed.ncbi.nlm.nih.gov/18362813/)]
10. Lyman JA, Cohn WF, Bloomrosen M, Detmer DE. Clinical decision support: progress and opportunities. *J Am Med Inform Assoc* 2010 Sep;17(5):487-492 [[FREE Full text](#)] [doi: [10.1136/jamia.2010.005561](https://doi.org/10.1136/jamia.2010.005561)] [Medline: [20819850](https://pubmed.ncbi.nlm.nih.gov/20819850/)]
11. How ICD-10 impacts revenue cycle management. *Healthcare Innovation*. 2012. URL: <https://www.hcinnovationgroup.com/home/article/13004951/how-icd10-impacts-revenue-cycle-management> [accessed 2022-09-21]
12. *International Statistical Classification of Diseases and Related Health Problems, 10th Revision, Fifth Edition, 2016*. Geneva, Switzerland: World Health Organization; 2015.
13. Medicare fee for office visit CPT codes, CPT code 99213, 99214, 99203. Medicare Payment and Reimbursement. 1999. URL: <http://www.medicarepaymentandreimbursement.com/p/medicare-fee-schedule-for-office-visit.html> [accessed 2022-09-21]
14. Morbidity and mortality in people with serious mental illness. National Association of State Mental Health Program Directors Medical Directors. Alexandria, VA: NASMHPD Medical Directors Council Publications and Reports; 2006. URL: [https://www.nasmhpd.org/sites/default/files/Mortality%20and%20Morbidity%20Final%20Report%208.18.08\\_0.pdf](https://www.nasmhpd.org/sites/default/files/Mortality%20and%20Morbidity%20Final%20Report%208.18.08_0.pdf) [accessed 2022-09-27]

15. Rossom RC, Crain AL, O'Connor PJ, Waring SC, Hooker SA, Ohnsorg K, et al. Effect of clinical decision support on cardiovascular risk among adults with bipolar disorder, schizoaffective disorder, or schizophrenia: a cluster randomized clinical trial. *JAMA Netw Open* 2022 Mar;5(3):e220202 [FREE Full text] [doi: [10.1001/jamanetworkopen.2022.0202](https://doi.org/10.1001/jamanetworkopen.2022.0202)] [Medline: [35254433](https://pubmed.ncbi.nlm.nih.gov/35254433/)]
16. US Preventive Services Task Force, Bibbins-Domingo K, Grossman DC, Curry SJ, Davidson KW, Epling JW, et al. Statin use for the primary prevention of cardiovascular disease in adults: US Preventive Services Task Force Recommendation Statement. *JAMA* 2016 Nov;316(19):1997-2007. [doi: [10.1001/jama.2016.15450](https://doi.org/10.1001/jama.2016.15450)] [Medline: [27838723](https://pubmed.ncbi.nlm.nih.gov/27838723/)]
17. Whelton PK, Carey RM, Aronow WS, Casey DE, Collins KJ, Dennison Himmelfarb C, et al. 2017 ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PCNA guideline for the prevention, detection, evaluation, and management of high blood pressure in adults: a report of the American College of Cardiology/American Heart Association Task Force on clinical practice guidelines. *Hypertension* 2018 Jun;71(6):e13-e115 [FREE Full text] [doi: [10.1161/HYP.0000000000000065](https://doi.org/10.1161/HYP.0000000000000065)] [Medline: [29133356](https://pubmed.ncbi.nlm.nih.gov/29133356/)]
18. American Diabetes Association. 6. Glycemic targets: Standards of Medical Care in Diabetes-2018. *Diabetes Care* 2018 Jan;41(Suppl 1):S55-S64. [doi: [10.2337/dc18-S006](https://doi.org/10.2337/dc18-S006)] [Medline: [29222377](https://pubmed.ncbi.nlm.nih.gov/29222377/)]
19. Managing overweight and obesity in adults. National Heart Lung and Blood Institute. Washington, DC: National Institutes of Health; 2013. URL: <https://www.nhlbi.nih.gov/sites/default/files/media/docs/obesity-evidence-review.pdf> [accessed 2022-09-27]
20. Tobacco smoking cessation in adults, including pregnant persons: interventions. U.S. Preventive Services Task Force. Rockville, MD; 2021. URL: <https://www.uspreventiveservicestaskforce.org/Page/Document/RecommendationStatementFinal/tobacco-use-in-adults-and-pregnant-women-counseling-and-interventions1> [accessed 2019-04-08]
21. Aalbers J. Aspirin in primary prevention: USPSTF recommendations. *Cardiovasc J Afr* 2010 May;21(3):176 [FREE Full text] [Medline: [20532445](https://pubmed.ncbi.nlm.nih.gov/20532445/)]
22. Rothwell PM, Cook NR, Gaziano JM, Price JF, Belch JFF, Roncaglioni MC, et al. Effects of aspirin on risks of vascular events and cancer according to bodyweight and dose: analysis of individual patient data from randomised trials. *Lancet* 2018 Aug;392(10145):387-399 [FREE Full text] [doi: [10.1016/S0140-6736\(18\)31133-4](https://doi.org/10.1016/S0140-6736(18)31133-4)] [Medline: [30017552](https://pubmed.ncbi.nlm.nih.gov/30017552/)]
23. Stone NJ, Robinson JG, Lichtenstein AH, Bairey Merz CN, Blum CB, Eckel RH, American College of Cardiology/American Heart Association Task Force on Practice Guidelines. 2013 ACC/AHA guideline on the treatment of blood cholesterol to reduce atherosclerotic cardiovascular risk in adults: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. *Circulation* 2014 Jun;129(25 Suppl 2):S1-45 [FREE Full text] [doi: [10.1161/01.cir.0000437738.63853.7a](https://doi.org/10.1161/01.cir.0000437738.63853.7a)] [Medline: [24222016](https://pubmed.ncbi.nlm.nih.gov/24222016/)]
24. Stevens RJ, Kothari V, Adler AI, Stratton IM, United Kingdom Prospective Diabetes Study (UKPDS) Group. The UKPDS risk engine: a model for the risk of coronary heart disease in Type II diabetes (UKPDS 56). *Clin Sci (Lond)* 2001 Dec;101(6):671-679. [Medline: [11724655](https://pubmed.ncbi.nlm.nih.gov/11724655/)]
25. Gaziano TA, Young CR, Fitzmaurice G, Atwood S, Gaziano JM. Laboratory-based versus non-laboratory-based method for assessment of cardiovascular disease risk: the NHANES I Follow-up Study cohort. *Lancet* 2008 Mar;371(9616):923-931 [FREE Full text] [doi: [10.1016/S0140-6736\(08\)60418-3](https://doi.org/10.1016/S0140-6736(08)60418-3)] [Medline: [18342687](https://pubmed.ncbi.nlm.nih.gov/18342687/)]
26. Epic. Verona, WI: Epic Systems Corporation URL: <https://www.epic.com/software> [accessed 2022-09-21]
27. Arnett DK, Blumenthal RS, Albert MA, Buroker AB, Goldberger ZD, Hahn EJ, et al. 2019 ACC/AHA guideline on the primary prevention of cardiovascular disease: a report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. *J Am Coll Cardiol* 2019 Sep;74(10):e177-e232 [FREE Full text] [doi: [10.1016/j.jacc.2019.03.010](https://doi.org/10.1016/j.jacc.2019.03.010)] [Medline: [30894318](https://pubmed.ncbi.nlm.nih.gov/30894318/)]
28. Institute of Medicine, Committee on Improving the Patient Record. In: Dick RS, Steen EB, Detmer DE, editors. *The Computer-Based Patient Record: An Essential Technology for Health Care, Revised Edition*. Washington, DC: National Academies Press; 1997.
29. Bright TJ, Wong A, Dhurjati R, Bristow E, Bastian L, Coeytaux RR, et al. Effect of clinical decision-support systems: a systematic review. *Ann Intern Med* 2012 Jul;157(1):29-43 [FREE Full text] [doi: [10.7326/0003-4819-157-1-201207030-00450](https://doi.org/10.7326/0003-4819-157-1-201207030-00450)] [Medline: [22751758](https://pubmed.ncbi.nlm.nih.gov/22751758/)]
30. Clinical decision support. Agency for Healthcare Research and Quality. Rockville, MD; 2013. URL: <https://www.ahrq.gov/cpi/about/otherwebsites/clinical-decision-support/index.html> [accessed 2022-09-20]
31. Gold R, Larson AE, Sperl-Hillen JM, Boston D, Shepler CR, Heintzman J, et al. Effect of clinical decision support at community health centers on the risk of cardiovascular disease: a cluster randomized clinical trial. *JAMA Netw Open* 2022 Feb;5(2):e2146519 [FREE Full text] [doi: [10.1001/jamanetworkopen.2021.46519](https://doi.org/10.1001/jamanetworkopen.2021.46519)] [Medline: [35119463](https://pubmed.ncbi.nlm.nih.gov/35119463/)]
32. Conrad DA. The theory of value-based payment incentives and their application to health care. *Health Serv Res* 2015 Dec;50 Suppl 2:2057-2089 [FREE Full text] [doi: [10.1111/1475-6773.12408](https://doi.org/10.1111/1475-6773.12408)] [Medline: [26549041](https://pubmed.ncbi.nlm.nih.gov/26549041/)]
33. What are the value-based programs? Centers for Medicare & Medicaid Services. URL: <https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Value-Based-Programs/Value-Based-Programs.html> [accessed 2022-09-22]
34. Burwell SM. Setting value-based payment goals--HHS efforts to improve U.S. health care. *N Engl J Med* 2015 Mar;372(10):897-899. [doi: [10.1056/NEJMp1500445](https://doi.org/10.1056/NEJMp1500445)] [Medline: [25622024](https://pubmed.ncbi.nlm.nih.gov/25622024/)]

35. Masterson L. Payers moving to value-based care faster than expected. HEALTHCARE DIVE. 2018. URL: <https://www.healthcaredive.com/news/payers-moving-to-value-based-care-faster-than-expected/525900/> [accessed 2022-09-21]
36. Guram JS, Moffit RE. The Medicare advantage success story--looking beyond the cost difference. *N Engl J Med* 2012 Mar;366(13):1177-1179. [doi: [10.1056/NEJMp1114019](https://doi.org/10.1056/NEJMp1114019)] [Medline: [22356263](https://pubmed.ncbi.nlm.nih.gov/22356263/)]
37. Mandal AK, Tagomori GK, Felix RV, Howell SC. Value-based contracting innovated Medicare advantage healthcare delivery and improved survival. *Am J Manag Care* 2017 Feb;23(2):e41-e49 [FREE Full text] [Medline: [28245661](https://pubmed.ncbi.nlm.nih.gov/28245661/)]
38. Gosden T, Forland F, Kristiansen IS, Sutton M, Leese B, Giuffrida A, et al. Capitation, salary, fee-for-service and mixed systems of payment: effects on the behaviour of primary care physicians. *Cochrane Database Syst Rev* 2000 Jul(3):CD002215. [doi: [10.1002/14651858.CD002215](https://doi.org/10.1002/14651858.CD002215)] [Medline: [10908531](https://pubmed.ncbi.nlm.nih.gov/10908531/)]
39. Rudman WJ, Eberhardt JS, Pierce W, Hart-Hester S. Healthcare fraud and abuse. *Perspect Health Inf Manag* 2009 Sep;6:1g [FREE Full text] [Medline: [20169019](https://pubmed.ncbi.nlm.nih.gov/20169019/)]
40. Hill C, Hunter A, Johnson L, Coustasse A. Medicare fraud in the United States: Can it ever be stopped? *Health Care Manag (Frederick)* 2014 Jul;33(3):254-260. [doi: [10.1097/HCM.000000000000019](https://doi.org/10.1097/HCM.000000000000019)] [Medline: [25068880](https://pubmed.ncbi.nlm.nih.gov/25068880/)]
41. Final policy, payment, and quality provisions changes to the medicare physician fee schedule for calendar year 2021. Centers for Medicare & Medicaid Services. Baltimore, MD; 2020. URL: <https://www.cms.gov/newsroom/fact-sheets/final-policy-payment-and-quality-provisions-changes-medicare-physician-fee-schedule-calendar-year-1> [accessed 2022-09-18]

## Abbreviations

- BPA:** best practice advisory
- BP:** blood pressure
- CDS:** clinical decision support
- CMS:** Center for Medicare and Medicaid Services
- CPT:** current procedural terminology
- CV:** cardiovascular
- DM:** diabetes mellitus
- EHR:** electronic health record
- FFS:** fee for service
- FHIR:** Fast Healthcare Interoperability Resources
- HbA<sub>1c</sub>:** glycated hemoglobin
- ICD-10:** 10th Revision of the International Classification of Diseases
- MUMPS:** Massachusetts General Hospital Utility Programming System
- RAF:** risk adjustment factor
- RCT:** randomized controlled trial
- SDM:** shared decision-making
- SMI:** serious mental illness

*Edited by A Mavragani; submitted 05.08.21; peer-reviewed by D Newman, J Pevnick; comments to author 26.01.22; revised version received 27.07.22; accepted 23.08.22; published 06.10.22.*

*Please cite as:*

*Sperl-Hillen JM, Anderson JP, Margolis KL, Rossom RC, Kopski KM, Averbek BM, Rosner JA, Ekstrom HL, Dehmer SP, O'Connor PJ*

*Bolstering the Business Case for Adoption of Shared Decision-Making Systems in Primary Care: Randomized Controlled Trial*

*JMIR Form Res* 2022;6(10):e32666

URL: <https://formative.jmir.org/2022/10/e32666>

doi: [10.2196/32666](https://doi.org/10.2196/32666)

PMID: [36201392](https://pubmed.ncbi.nlm.nih.gov/36201392/)

©JoAnn M Sperl-Hillen, Jeffrey P Anderson, Karen L Margolis, Rebecca C Rossom, Kristen M Kopski, Beth M Averbek, Jeanine A Rosner, Heidi L Ekstrom, Steven P Dehmer, Patrick J O'Connor. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 06.10.2022. 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

# Formative Evaluation of the Acceptance of HIV Prevention Artificial Intelligence Chatbots By Men Who Have Sex With Men in Malaysia: Focus Group Study

Mary L Peng<sup>1</sup>, BA; Jeffrey A Wickersham<sup>2,3,4</sup>, PhD; Frederick L Altice<sup>2,3,4,5</sup>, MD; Roman Shrestha<sup>3,4,6</sup>, PhD; Iskandar Azwa<sup>4,7</sup>, MD; Xin Zhou<sup>2</sup>, PhD; Mohd Akbar Ab Halim<sup>4</sup>, BSc; Wan Mohd Ikhtiaruddin<sup>4</sup>, MSc; Vincent Tee<sup>4</sup>, MD; Adeeba Kamarulzaman<sup>2,4,7</sup>, MD; Zhao Ni<sup>3,8</sup>, BMedSci, RN, PhD

<sup>1</sup>Social and Behavioral Sciences Department, Yale School of Public Health, New Haven, CT, United States

<sup>2</sup>Section of Infectious Disease, Department of Internal Medicine, Yale School of Medicine, New Haven, CT, United States

<sup>3</sup>Center for Interdisciplinary Research on AIDS (CIRA), Yale University, New Haven, CT, United States

<sup>4</sup>Centre of Excellence for Research in AIDS (CERiA), Faculty of Medicine, University of Malaya, Kuala Lumpur, Malaysia

<sup>5</sup>Division of Epidemiology of Microbial Diseases, Yale School of Public Health, New Haven, CT, United States

<sup>6</sup>Department of Allied Health Sciences, University of Connecticut, Storrs, CT, United States

<sup>7</sup>Department of Medicine, Infectious Disease Unit, Faculty of Medicine, Kuala Lumpur, Malaysia

<sup>8</sup>School of Nursing, Yale University, New Haven, CT, United States

**Corresponding Author:**

Zhao Ni, BMedSci, RN, PhD

School of Nursing

Yale University

400 West Campus Drive

New Haven, CT, 06477

United States

Phone: 1 203 737 3039

Email: [zhao.ni@yale.edu](mailto:zhao.ni@yale.edu)

## Abstract

**Background:** Mobile technologies are being increasingly developed to support the practice of medicine, nursing, and public health, including HIV testing and prevention. Chatbots using artificial intelligence (AI) are novel mobile health strategies that can promote HIV testing and prevention among men who have sex with men (MSM) in Malaysia, a hard-to-reach population at elevated risk of HIV, yet little is known about the features that are important to this key population.

**Objective:** The aim of this study was to identify the barriers to and facilitators of Malaysian MSM's acceptance of an AI chatbot designed to assist in HIV testing and prevention in relation to its perceived benefits, limitations, and preferred features among potential users.

**Methods:** We conducted 5 structured web-based focus group interviews with 31 MSM in Malaysia between July 2021 and September 2021. The interviews were first recorded, transcribed, coded, and thematically analyzed using NVivo (version 9; QSR International). Subsequently, the unified theory of acceptance and use of technology was used to guide data analysis to map emerging themes related to the barriers to and facilitators of chatbot acceptance onto its 4 domains: performance expectancy, effort expectancy, facilitating conditions, and social influence.

**Results:** Multiple barriers and facilitators influencing MSM's acceptance of an AI chatbot were identified for each domain. *Performance expectancy* (ie, the perceived usefulness of the AI chatbot) was influenced by MSM's concerns about the AI chatbot's ability to deliver accurate information, its effectiveness in information dissemination and problem-solving, and its ability to provide emotional support and raise health awareness. Convenience, cost, and technical errors influenced the AI chatbot's *effort expectancy* (ie, the perceived ease of use). Efficient linkage to health care professionals and HIV self-testing was reported as a *facilitating condition* of MSM's receptiveness to using an AI chatbot to access HIV testing. Participants stated that *social influence* (ie, sociopolitical climate) factors influencing the acceptance of mobile technology that addressed HIV in Malaysia included privacy concerns, pervasive stigma against homosexuality, and the criminalization of same-sex sexual behaviors. Key design strategies that could enhance MSM's acceptance of an HIV prevention AI chatbot included an anonymous user setting; embedding

the chatbot in MSM-friendly web-based platforms; and providing user-guiding questions and options related to HIV testing, prevention, and treatment.

**Conclusions:** This study provides important insights into key features and potential implementation strategies central to designing an AI chatbot as a culturally sensitive digital health tool to prevent stigmatized health conditions in vulnerable and systematically marginalized populations. Such features not only are crucial to designing effective user-centered and culturally situated mobile health interventions for MSM in Malaysia but also illuminate the importance of incorporating social stigma considerations into health technology implementation strategies.

(*JMIR Form Res* 2022;6(10):e42055) doi:[10.2196/42055](https://doi.org/10.2196/42055)

## KEYWORDS

artificial intelligence; chatbot; HIV prevention; implementation science; men who have sex with men; MSM; mobile health design; mHealth design; unified theory of acceptance and use of technology; mobile phone

## Introduction

### Background

During the past 3 decades, great progress has been made worldwide in HIV prevention, including HIV testing. However, the HIV epidemic in key populations such as men who have sex with men (MSM) continues to grow worldwide in the setting of stigma and discrimination [1]. Malaysia has one of the fastest-growing HIV epidemics among MSM in Southeast Asia [2]. Surveillance data in 2019 suggest that over 1 in 5 (21.6%) MSM in Malaysia live with HIV [3]. HIV testing guidelines recommend high-risk MSM to be tested every 3 to 6 months [4,5]. However, MSM do not get tested optimally, with only 70.3% having ever been tested, 40.9% having been tested in the past year, and only 9.5% being tested more than once annually [6,7]. Moreover, in 2019, only 3% of MSM in Malaysia had adequate knowledge of HIV prevention, and only 36.7% reported ever having received HIV prevention services [3]. The low testing rates and inadequate prevention efforts are associated with high levels of stigma and discrimination against MSM, which are perpetuated in Malaysia, where same-sex sexual behaviors are criminalized by both secular and Shariah law [8]. To access HIV testing and prevention services, MSM in Malaysia must navigate individual and systemic barriers that perpetuate stigma and discrimination. Previous research has associated the heightened HIV burden among MSM in Malaysia with both etiological factors such as substance use during sexual encounters [9], which is linked to higher HIV transmission rates [10], and social factors such as perception of social stigma of pre-exposure prophylaxis use to prevent HIV [11], anxiety and fear for safety following the arrest of HIV prevention workers under legislation that codified sexual practices with persons of the same sex as criminal and unnatural [12], and health care workers' discriminatory intentions against MSM [13]. Mobile health (mHealth) strategies that use theory-guided behavior change interventions and address both individual and systematic health care challenges have the potential to overcome multilevel barriers to HIV testing and prevention [14].

### mHealth and Chatbots

Mobile technologies' increasing global penetration and capacity for swift information delivery and retrieval make mHealth a promising tool for HIV intervention, which requires timely monitoring of disease development at the population level as

well as early detection, effective prevention, and efficient diagnosis at the individual level. mHealth strategies involve several modalities, including apps that are installed on smartphones or chatbots that are embedded within existing apps or websites. Multiple studies have identified mHealth as an effective means of facilitating HIV testing and prevention. For example, the HIVSmart app showed high efficacy in detecting new HIV infections and increasing self-test referrals [15]. A recent randomized controlled trial of an app intervention for increasing access to HIV testing and care among young cisgender men and transgender women highlighted the capacity of mHealth technologies for effective HIV prevention and real-time intervention delivery [16]. A meta-analysis of 41 studies evaluating 28 evidence-based mHealth interventions to support HIV self-management further revealed that mHealth interventions significantly improved individual-level medication adherence, mental health, and social support [17]. Despite the ample literature on the effectiveness of mHealth interventions, these studies have primarily involved apps, and chatbots are absent from HIV testing and prevention strategies.

Chatbots have the potential to create new opportunities to promote HIV testing and prevention. Chatbots are software-based programs that imitate human conversational agents to interact with users [18]. Users often converse with a chatbot on mobile platforms such as phones, laptops, websites, apps, and SMS text messaging, where the chatbot provides the information and solutions demanded by users through text responses or other engaging formats such as pictures, videos, and audio. Following the advances in natural language processing (NLP) and artificial intelligence (AI), AI chatbots, broadly referring to computer programs with minimal design interfaces embedded with AI to simulate conversation with human users, have been developed as mHealth strategies to support patient care and identified as an effective intervention for increasing health information delivery and promoting physical activity and a healthy diet [19]. A 2021 systematic review of the application of AI chatbots in health care and oncology accentuated the potential of integrating AI chatbots into clinical practice to reduce costs, improve patient outcomes, and enhance health practitioners' work efficiencies [20]. Another recent review on the application of AI chatbots in digital mental health care suggested that, given AI chatbots' unique ability to learn from and interface with patients, understanding the individual and contextual factors that might affect the impact

of AI chatbots was particularly significant compared with conventional mHealth interventions [21]. Our study echoes this concern and aims to expand the understanding of AI chatbots' health care application by demonstrating how individual and social factors are intricately linked to the mechanisms that affect people's acceptance of AI chatbots as digital health interventions. Furthermore, in light of the onset of the COVID-19 pandemic, a recent review examined 61 chatbots used for pandemic public health response in 30 countries and highlighted chatbots' "scalability, wide accessibility, ease of use, and fast information dissemination" as prominent features that served the interest of public health while calling for discussions on sophisticated chatbot design synergies [22]. In this study, we hope to deepen people's understanding of AI chatbot design by illuminating user-centered features informed by multifaceted facilitators of and barriers to AI chatbot acceptance.

Although existing studies on the use of AI chatbots to improve health care outcomes encourage continuous investigation into the adoption of this technology, systematic discussions on contextual adaptations of AI chatbots to suit the unique dynamics of different populations, sociocultural contexts, and targeted health conditions remain scant. For example, a year-long prospective study of conversations between patients with breast cancer and an AI chatbot demonstrated the efficacy of using the chatbot to increase patients' knowledge of breast cancer, medication adherence, and satisfaction with health care support [23]. Specifically, this study highlighted patients' willingness to communicate sensitive and intimate information, such as their experiences of sexuality, with the AI chatbot, but whether patients' perception of cancer-related social stigma existed or affected chatbot acceptance remained unexplored [23]. When adapting an AI chatbot to different health conditions and target populations in different sociocultural contexts, the same level of acceptance should not be assumed ubiquitously without deliberating the cultural and disease contexts in which an AI chatbot might receive sensitive personal information. Whether the same level of willingness can be found among stigmatized populations who are subjected to potential negative sociopolitical consequences because of disclosures of sensitive personal information, exemplified by the MSM population in Malaysia, remains to be investigated. Thus, this study heeds the significance of cross-cultural technology adaption and situates the examination of AI chatbot acceptance in the careful consideration of person-disease-context dynamics among MSM in Malaysia. Another study examined the acceptance of using a tuberculosis prevention AI chatbot among people who were vulnerable to tuberculosis in South Korea and found that the chatbot facilitated the dissemination of desired tuberculosis care information to at-risk populations [24]. Although the researchers acknowledged the stigmatization of tuberculosis in South Korea, the interrogation of health stigma remained outside the systematic consideration of how multilevel social influence could affect technology acceptance [24]. Therefore, we hope to incorporate systemic stigma into the discourse of social influence that affects technology acceptance, highlighting the interconnection among AI chatbot acceptance, individual needs, and macrosocial determinants of health.

## Purposes of the Study

To develop culturally tailored AI chatbots to promote HIV testing and prevention among MSM in Malaysia, researchers must identify the facilitators and barriers of chatbots among MSM and features that would promote acceptance. Such findings can play a critical role in informing and optimizing the design of the AI chatbot that is currently being developed, which may in turn influence its uptake and use [25]. Overall, this study aimed to increase our understanding of developing culturally tailored AI chatbots as mHealth interventions in a discriminatory environment against MSM and lay the groundwork for the future development, implementation, and scale-up of an MSM-friendly AI chatbot in Malaysia to improve HIV testing and use of prevention services. Specifically, we explored what influenced MSM's acceptance of an AI chatbot designed to promote HIV testing and prevention using the unified theory of acceptance and use of technology (UTAUT).

## Conceptual Framework for Analysis

The UTAUT was used to guide the analysis of MSM's acceptance of AI chatbots designed to promote HIV testing and prevention in Malaysia. The UTAUT is an extension of the technology acceptance model (TAM) that is commonly used to evaluate factors predicting people's acceptance of new technologies. The TAM involves 2 constructs—perceived usefulness and perceived ease of use—as the primary predictors of users' acceptance of new technologies [26]. The TAM was later expanded to the UTAUT, which includes (1) performance expectancy (perceived usefulness), (2) effort expectancy (perceived ease of use), (3) social influence, and (4) facilitating conditions as empirically proven predictors of technology acceptance and use intention [27].

The UTAUT was used as the analytical framework for 3 reasons. First, we hoped to use the insights learned from this study to guide the design of an HIV prevention AI chatbot. The UTAUT's emphasis on users' perception of what makes a technological tool useful rather than what is assumed to be useful by researchers and product designers would allow us to design the AI chatbot from a user-centered perspective. Second, it was important to use a validated framework such as the UTAUT, which has predictive value for identifying effective users' acceptance of technology [28-33]. Third, although perceived self-efficacy, facilitating conditions, and system quality have been added to the UTAUT in evaluating health care technologies [34], there has been little investigation regarding the application of the UTAUT in HIV-specific health technology within cultural settings where stigma against HIV and MSM is pervasive. The standard definition of social influence in the UTAUT highlights an individual's family's, peers', and friends' influence on the individual's technology acceptance and behavioral intention [27] rather than the influence of macro sociocultural determinants. Given the prevalent antihomosexuality sentiment and the heightened stigma toward HIV-infected MSM in Malaysia, we sought to provide insights into the adaptation of the UTAUT in assessing culturally specific and HIV-specific mHealth acceptance by incorporating stigma into the domain of social influence in the current model.



## Methods

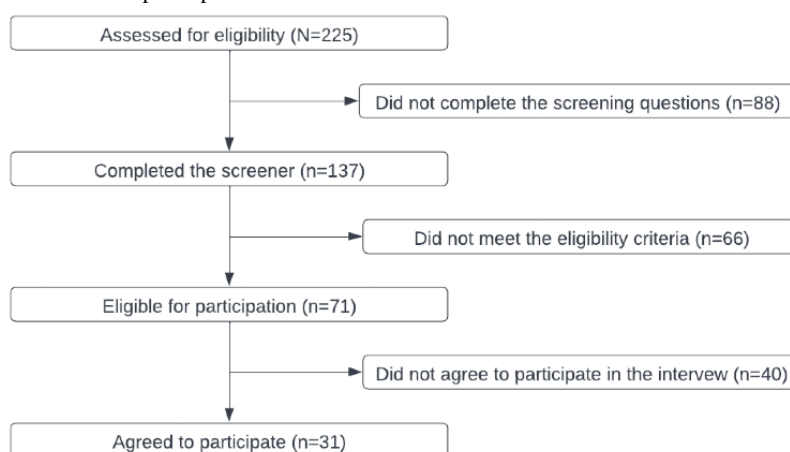
### Eligibility Criteria

The inclusion criteria for the participants were as follows: (1) being cisgender men, (2) being aged  $\geq 18$  years, self-reporting (3) condomless sex with another man in the past 6 months and (4) HIV-negative or unknown status, and (5) speaking Bahasa Malaysia or English.

### Recruitment

Figure 1 depicts findings from the recruitment strategy. The study setting was Malaysia, and the target population was MSM.

**Figure 1.** The screening process for research participants.



### Procedures for the Focus Group Interviews

After obtaining informed consent, eligible participants were assigned to different groups based on their available times. The focus group interviews were conducted using Zoom (Zoom Video Communications). To ensure anonymity, the participants were instructed to set their preferences for their name (pseudonym) and video use. An experienced qualitative interviewer (JAW) with expertise in HIV prevention with MSM in Malaysia led all the interviews with other members of the research team to observe and record them. A total of 5 focus group interviews (range 5-9 participants each) were conducted using a semistructured interview guide (Multimedia Appendix 1) between July 2021 and September 2021. The interview guide was developed by adapting the evidence-based 5-phase process by Kallio et al [35]. First, we identified the prerequisite for using focus group discussions by assigning user technology acceptance and perception as the predetermined thematic concerns of our interviews. Second, informed by the purpose of formulative evaluation and previous literature, we decided on our 3 main topics for discussion. *Topic 1: attitude toward HIV testing* aimed to generate insight into the health behavior change targeted by chatbot interventions. *Topic 2: chatbot* focused on participants' perceptions of the technology of interest. *Topic 3: social networking app* hoped to elicit user-centered suggestions regarding the platforms that should host MSM-friendly HIV prevention chatbots in response to the increasing mobilization of social media in the digital health industry [36]. Third, we formulated the interview guide by ensuring that the 3 topics proceeded logically from general to more specific questions,

A web-based screener was posted on MSM social networking apps, including Grindr, Hornet, and Blued, where MSM often meet to find sexual partners. Of the 225 people who initiated the web-based screener, 137 (60.9%) completed it, of whom 71 (51.8%) were eligible for participation. Of these 71 MSM, 31 (44%) agreed to participate in the focus group interviews and were enrolled. As part of the screening process, each participant was asked to complete a brief survey that included questions on demographic characteristics; HIV prevention practices; and self-reported measures of substance use, depression, and social media use.

leaving room for interviewers to probe unanticipated but relevant issues raised during the discussion [37]. Fourth, we pilot-tested the guide among research team members not involved in the development of the interview guide. Finally, we presented the complete guide to the entire research team for approval.

During the focus group discussion, participants first watched an introductory video (90 seconds) about AI chatbots and how they worked, followed by open-ended questions regarding the facilitators and barriers associated with using an AI chatbot to promote HIV testing and prevention. Sample questions included "What do you think would be helpful for AI-chatbots to promote HIV testing?" and "Where would you expect to find an AI-chatbot designed to promote HIV testing and prevention?" Themes identified during the interviews were probed for further details as they emerged. Each group interview lasted 85 minutes on average (SD 5; range 81-94 minutes).

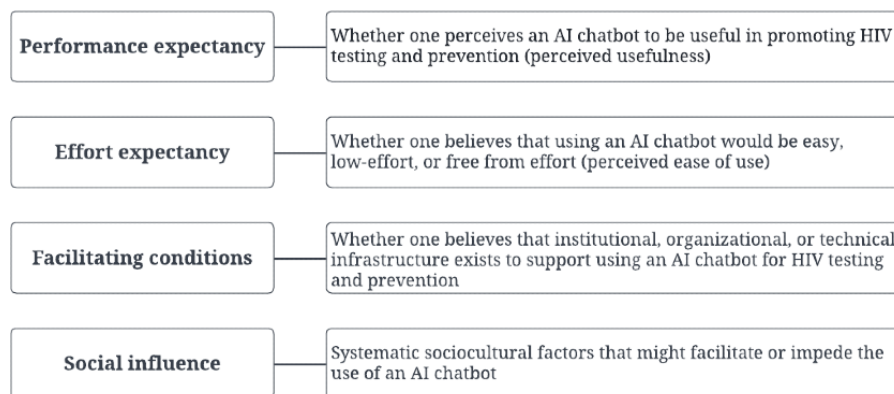
### Analysis

Descriptive statistics were calculated using SAS (version 9.4; SAS Institute) to summarize participants' demographic characteristics; HIV prevention practices; and self-reported measures of substance use, depression, and social media use characteristics. The interviews were recorded, transcribed verbatim, and analyzed using thematic analysis. Emerging themes were mapped onto the 4 constructs of the UTAUT. Thematic coding was conducted using NVivo (version 9; QSR International) by 4 researchers (ZN, MAAH, WMI, and VT) who independently completed the initial round of coding. Each coder identified the patterns and themes that emerged from the findings regarding the benefits, concerns, and desirable features

associated with using a chatbot for HIV testing and prevention. All coders collaboratively discussed their findings, resolved any discrepancies, and achieved saturation when no new themes were identified. The first author (MP) reviewed all the findings; identified similarities, differences, and redundancies among each coder's findings; and organized the themes and illustrative quotations according to the four constructs of the UTAUT (Figure 2): (1) performance expectancy (ie, perceived usefulness), (2) effort expectancy (ie, perceived ease of use), (3) facilitating conditions, and (4) social influence. Later, a retrospective assessment of thematic saturation was conducted by the 4 coders by comparing the study findings with the themes in the System Usability Scale, a simple, 10-item attitude Likert scale that provides a global view of subjective assessments of

usability [17]. We compared the findings with the System Usability Scale as it is a reliable and validated tool evaluating a wide range of technical themes, including a new technology's use frequency, complexity, technical issues, integration of various functions, and inconsistency. Any new themes appearing in the System Usability Scale were double-checked by the 4 coders in the transcripts to ensure that our results were comprehensive enough to guide a wide variety of mHealth interventions and products. To increase reliability, themes were reviewed after coding, where the final coded themes were returned to 6 interviewees for accuracy check and to identify if there were any discrepancies between our interpretations and their intended opinions. All participants concurred with our findings.

**Figure 2.** Summary of contextualized definitions of key terms. AI: artificial intelligence.



## Ethics Approval

This study was approved by the institutional review boards of Yale University (ID 2000025910) and University of Malaya (ID 202049-8488). Participants reviewed all study-related risks and benefits and provided signed informed consent before participation. All procedures performed involving human participants were in accordance with the ethical standards of the institutional or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Participants were paid 45 Malaysian ringgits (US \$10) after completing the study.

## Results

### Overview

Participants' demographic characteristics are presented in Table 1. The mean age of the participants was 30.6 (SD 6.4) years,

and they encompassed 4 ethnic groups: Malay (12/31, 39%), Chinese (16/31, 52%), Indian (1/31, 3%), and mixed (2/31, 6%). Among all participants, 48% (15/31) had never taken pre-exposure prophylaxis; 23% (7/31) had not been tested for HIV in the past 6 months; and 32% (10/31) had a PHQ-2 score >3, indicating a likelihood of depressive disorder.

Overall, most participants responded positively to using an AI chatbot for HIV testing and prevention. Multimedia Appendix 2 summarizes the key themes identified according to the 4 domains of the UTAUT with illustrative quotations. Figure 3 presents MSM's preferred features of the AI chatbot suggested by the participants.

**Table 1.** Participant characteristics (N=31).

Variable	Values
Age (years), mean (SD)	30.6 (6.4)
<b>Ethnicity, n (%)</b>	
Malay	12 (39)
Chinese	16 (52)
Indian	1 (3)
Mixed	2 (6)
<b>Had been tested for HIV, n (%)</b>	
In the last 3 months	12 (39)
In the last 4 to 6 months	12 (39)
In the last 7 to 9 months	2 (6)
In the last 10 to 12 months	5 (16)
<b>Had ever taken PrEP<sup>a</sup>, n (%)</b>	
Yes	16 (52)
No	15 (48)
<b>Taking PrEP (n=16), n (%)</b>	
Yes	11 (69)
No	5 (31)
<b>Any drug use (past 6 months), n (%)</b>	
Yes	7 (23)
No	24 (77)
<b>Substance used (past 6 months), n (%)<sup>b</sup></b>	
Alcohol	11 (35)
Crystal methamphetamine	4 (13)
Gamma hydroxybutyrate	2 (6)
Cigarettes	5 (16)
Other drugs	1 (3)
<b>PHQ-2<sup>c</sup> score<sup>d</sup>, n (%)</b>	
<3	21 (68)
≥3	10 (32)
<b>Smartphone system, n (%)</b>	
Android	15 (48)
iOS	16 (52)
<b>App used, n (%)<sup>e</sup></b>	
WhatsApp	29 (94)
Facebook	25 (81)
Instagram	27 (87)
Telegram	19 (61)
Grindr	15 (48)
Twitter	24 (77)
Blued	6 (19)
WeChat	8 (26)

Variable	Values
Discord	6 (19)
Jack'd	2 (6)
Scruff	2 (6)

<sup>a</sup>PrEP: pre-exposure prophylaxis.

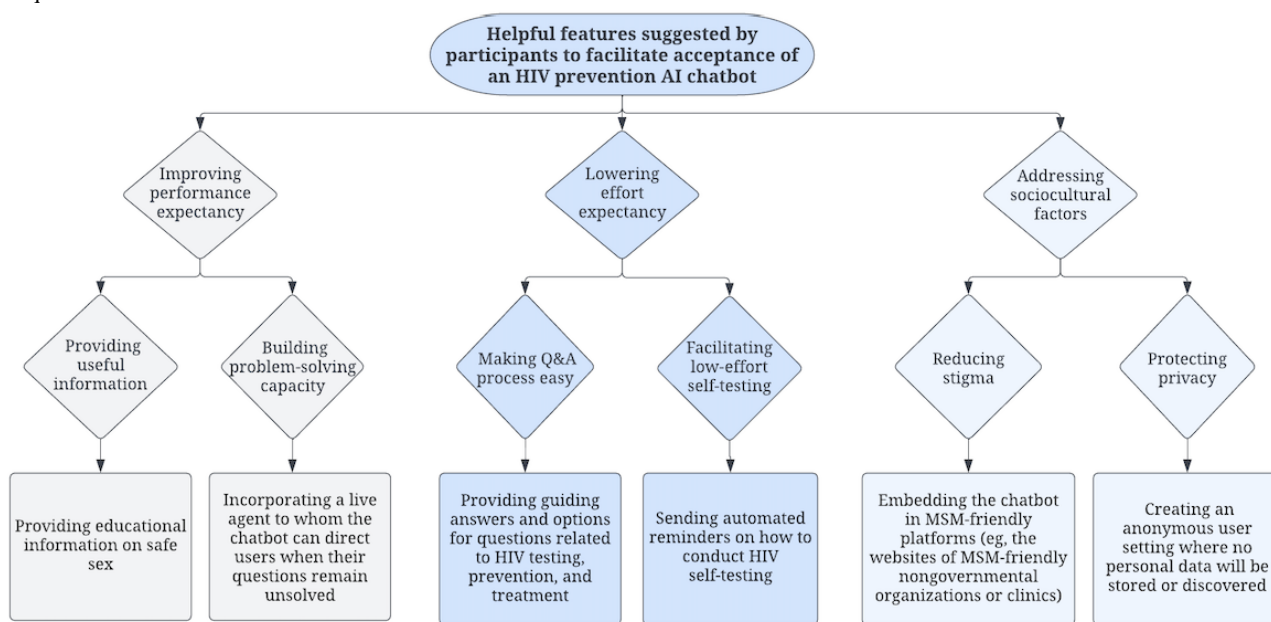
<sup>b</sup>Some participants used multiple drugs.

<sup>c</sup>PHQ-2: Patient Health Questionnaire-2.

<sup>d</sup>The PHQ-2 was used to screen for depression. Major depression was likely when the score was  $\geq 3$ .

<sup>e</sup>Some participants used multiple apps.

**Figure 3.** Helpful features for an HIV prevention artificial intelligence (AI) chatbot suggested by the participants. MSM: men who have sex with men; Q&A: question and answer.



## Performance Expectancy

### Overall Perception

Participants reported both positive and negative perceptions regarding AI chatbots' performance expectancy. They indicated that an AI chatbot's ability to disseminate valuable information, solve routine questions, and raise HIV prevention awareness positively affected whether they perceived the AI chatbot to be useful. They also described concerns with information accuracy and skepticism of AI chatbots' ability to provide emotional support and solve complex problems as the primary factors undermining performance expectancy.

### Contributors to Positive Performance Expectancy

An AI chatbot's capacity for information dissemination was a major contributor to positive performance expectancy. In total, 13% (4/31) of the participants explained that they perceived AI chatbots as useful "one-stop centers" for information with a simple user interface. A participant further elaborated on the relative advantage of using an AI chatbot to obtain information on HIV testing and prevention compared with canvassing multiple available sources and web-based platforms, which he described as a major obstacle to "getting tested often." Another participant highlighted his perception of using an AI chatbot as

a reliable source of information, stating that, compared with humans who might conceal information or twist facts, he perceived the information provided by AI chatbots as "factual" and more credible. Information regarding HIV awareness was also discussed. "Chatbot will be good for awareness purposes," as one participant stated regarding AI chatbots' usefulness in raising awareness for HIV testing and prevention. Nonetheless, another participant voiced skepticism, questioning that "how can chat robot send [relevant information to increase] awareness to the user?" In total, 13% (4/31) of the participants described that an AI chatbot was useful for providing solutions to users' routine questions regarding HIV. A participant emphasized that "[if] it's just a simple question, I believe the chatbot can help." For instance, another participant stated the following:

*...some people will ask simple questions like "where could I get tested" and the chatbot can totally help that by giving template answers.*

### Major Concerns

Despite some participants favoring AI chatbots as credible sources of information, a perceived lack of information accuracy and trustworthiness was reported by some as the primary reason that would hinder them from using AI chatbots. Specifically,

with regard to obtaining information on HIV self-testing kits from AI chatbots, a participant commented the following:

*I don't know if I can trust those...products are working or not.*

An overwhelming concern regarding AI chatbots' ability to provide emotional support and a personal touch to health care was reported by 19% (6/31) of the participants. As participants stated, because of chatbots' robotic nature, they "are not human alike," "will reply [to the user] in really regulated answers," and "there's no replacing the human element, especially for those first-timers [in HIV-testing]." A total of 10% (3/31) of the participants emphasized the importance of human touch, reiterating their preference for in-person health care services. For example, regarding the preference for HIV service modality, a participant commented the following:

*...at the end of the day, I think that most of them...wanted to reach out to talk to someone in person.*

The same sentiment was echoed by another participant, who highlighted the following:

*...sometimes you need another person in the same room offering you comfort and additional advice and support on what can be done and what needs to be done.*

Participants' perception of an AI chatbot's ability to solve complex health problems was another major source of negative performance expectancy. A total of 19% (6/31) of the participants perceived using a chatbot as sometimes frustrating and limited in providing useful solutions to users' complex questions and concerns. A participant elaborated that, in their previous experience using a chatbot, "it keeps being repetitive about the same kind of solution that I don't need." Another participant echoed the concern, stating the following:

*[Chatbots] always go back to the same question that cannot get my question answered so sometimes it's very frustrating.*

Some participants also highlighted their desire to interact with a real person. As a participant claimed, "if it's [the question] too complex, then I will prefer a human interaction." Among the concerned participants, one nevertheless mentioned that, for complicated questions, a chatbot could be useful in "filtering and narrowing down those inquiries."

### **Relevant Features Suggested by Participants**

In total, 10% (3/31) of the participants further discussed the features that they perceived to be critical in improving an AI chatbot's performance and usefulness. A participant suggested that the AI chatbot should provide educational information on safe sex, such as information on "chemsex," defined as "the use of mephedrone, crystal methamphetamine and gamma hydroxybutyrate/butyrolactone (GHB/GBL) to enable, enhance and prolong sexual interactions" [38], and "what to expect when they're going to try [chemsex]." Another participant suggested having real-time human representatives to whom the chatbot could direct users when their questions could not be adequately answered by the chatbot. He elaborated the following:

*...the chatbot cannot cater for every query or question that asked by the human, so I think it's good to have a counselor as to fulfill the other questions that's not being listed in the chatbot.*

To increase reliability, the final coded themes were returned to 19% (6/31) of the interviewees for accuracy check. All participants agreed on our findings, and 33% (2/6) provided valuable feedback worthy of mention. They pointed out that AI chatbots needed to be able to provide information in a neutral tone and should not strive to be or pretend to be humans, although being able to converse smoothly was important.

## **Effort Expectancy**

### **Overall Perception**

The participants reported both positive and negative perceptions regarding effort expectancy. The predominantly positive effort expectancy was associated with the convenience of using a chatbot. Major concerns included costs, technical difficulties, and the level of technical literacy required when using a chatbot.

### **Positive Contributors to Low Effort Expectancy**

Many participants expressed favorable attitudes toward the convenience and ease of using a chatbot. A participant stated the following:

*...in terms of convenience, I think it [an AI-chatbot] is brilliant, especially if you are planning to, you know, introduce it to the MSM community.*

The participants perceived AI chatbots as convenient tools, specifically for their 24/7 availability and their ability to provide immediate responses. Compared with human agents, from whom one "might have to wait for a reply," participants perceived an AI chatbot to be a more readily available source of help. A participant described the following:

*I think the reason why I said [it is] convenient [is] because usually with chatbots you get your answer almost immediately.*

Similarly, another participant elaborated the following:

*...it's just great because this hunting thoughts of whether you need to be checked up or not tends to come late at night when no personnel are actually available for you to speak to.*

Overall, low effort and ease of use were reported as the major contributors to positive effort expectancy and, consequently, the facilitators of participants' acceptance of AI chatbots designed to promote HIV testing and prevention.

### **Major Concerns**

Positive perceptions aside, some participants spoke about the potential cost, technical difficulties, and required technical literacy when using AI chatbots as sources of negative effort expectancy. A participant speculated that using an AI chatbot would be costly, maintaining that "I think it's going to be expensive, and I do not think that we should invest on it." Another perceived obstacle to AI chatbots' ease of use was the risk of technical breakdowns. In total, 6% (2/31) of the

participants reported the same concern, one stating the following:

*...sometimes it will break down. Then you have to do it again and again and again and again, and will be frustrated.*

The other participant expressed similar frustration with the following:

*...[imagine] your first attempt to just get HIV treatment...or PrEP...then suddenly a technical issue arises.*

The participants generally perceived AI chatbots as high-technology products, prompting 6% (2/31) of the participants to express concerns regarding the difficulty of navigating an AI chatbot for the older population. A participant conjectured that more effort would be required from older users who might not be technologically savvy, stating the following:

*...the younger generation definitely will find it easier to adapt...and as for the older generation...they still need to have someone to consult.*

The other participant raised the same concern, reiterating the potential difficulty to use an AI chatbot for HIV testing and prevention among the older generation when “things start too complicated.”

### **Relevant Features Suggested by Participants**

In terms of helpful features that could reduce use effort, some participants suggested that an AI chatbot should be designed to generate fast and seamless responses with guiding questions and options. A participant also envisioned that automatic information reminders on how to conduct HIV self-testing and notifications on progress throughout the testing process could enhance the AI chatbot's ease of use. For example, he stated the following:

*...you just need to enter something, [the chatbot will] give you the information and [to] book your parcel and to collect your parcel from the office.*

### **Facilitating Conditions**

Themes regarding both external and internal facilitating conditions emerged from participants' responses. External facilitating conditions pertained to the overall HIV virtual care infrastructure, whereas internal facilitating conditions included the users' attitudes toward HIV self-testing. Efficient linkage between users and health care professionals was a major external facilitating condition associated with the participants' acceptance of the AI chatbot. If an individual believed that there would be an integrated HIV telemedicine infrastructure and an effective channel of information relay established between the AI chatbot and health care services, they would be more inclined to use the chatbot. A participant added that the automated linkage to health care and physician referral from an AI chatbot would encourage chatbot use because the referral process would make them feel more comfortable compared with going straight to a clinical setting as the chatbot would have mentally prepared them to consult health care professionals. They stated the following:

*I would feel much more comfortable because the barrier is already broken when I checked with a chatbot, then I already let my guard down, then when it comes to the real person that can consult, then you feel much more open.*

Participants' attitudes toward self-testing were an important internal facilitating condition. On the one hand, most participants claimed that they would be encouraged to use an AI chatbot that could link them to HIV self-testing kits as they viewed HIV self-testing kits as a proprivacy, quick, and convenient testing method to learn their HIV status. By contrast, lack of self-testing awareness would undermine the incentives to use such a chatbot. As a participant reported, “we know that not everyone is aware of the self-testing kit,” and some participants “personally never heard of a self-testing kit.” Moreover, potential users' concerns about the accuracy and trustworthiness of HIV self-testing results could hamper the use of the AI chatbot to obtain HIV self-tests. As emphasized by a participant, he “would like to be assured by a professional that my test [result] is right.”

### **Social Influence**

#### **Overview**

Our participants also reported sociocultural factors that would facilitate or impede their acceptance of an AI chatbot, including stigma, fear of discrimination, privacy concerns, and drug-related legal concerns. Many participants mentioned the discriminatory cultural climate against MSM and the HIV stigmatization targeted at MSM in Malaysia:

*HIV is still a taboo [in Malaysia]*

*...there is huge stigma on the subject. People don't generally engage in the conversation about HIV.*

*...there are some judgments during those [HIV] checkups, and it deters me from any future visits.*

As a result, participants perceived an AI chatbot as a helpful learning and service tool in an environment with discrimination and stigmatization, believing that self-testing facilitated by the AI chatbot could help MSM avoid stigma and discrimination and, thus, increase HIV testing rates.

On the issue of avoiding discrimination and stigmatization, 10% (3/31) of the participants discussed the issue of HIV testing privacy, emphasizing that an AI chatbot could be a helpful means to protect their identity and privacy compared with routine clinical testing, especially for people “not willing to be tested by professionals.” A participant elaborated the following:

*...once told the nurse, the nurse told another nurse, and another nurse, and [eventually] everyone knew [I am] there for HIV testing.*

The benefit of using an AI chatbot for HIV self-testing in terms of safeguarding MSM's privacy was reiterated by another participant, who reported that he had to commute to another district far away from his current living address for HIV testing to protect his privacy. He stated the following:

*I think a chatbot is very helpful because when I took my HIV test, I was not checking at my area. I live at*

*Putrajaya, but I checked at Johor because I need to protect my privacy.*

Some participants further reported that, coupled with the fear of social stigma against HIV and MSM in Malaysia, they were concerned about the criminalization of substance use. As explained by a participant, they would be concerned about information censorship and unintentional engagement with illegal drug use if an AI chatbot was perceived as providing information on HIV-related or MSM sex-related drug use, which could in turn hamper their acceptance of the AI chatbot. Another participant further elaborated the following:

*I just feel that by having so much you know, putting up so much information out there, via chatbot I'm afraid that you know it seems that we are promoting of using drugs but, so we have to be quite careful with the information that we display out.*

### **Relevant Features Suggested by Participants**

Owing to the preponderant stigma associated with same-sex sexual behaviors and HIV in Malaysia, most MSM participants preferred having an AI chatbot embedded in MSM-friendly nongovernmental organizations' or clinics' websites. Participants believed that these platforms may offer greater privacy and a more secure virtual environment than a social networking app such as Facebook, Twitter, or Instagram, where users would be required to register using their personal information. Some participants also indicated that using MSM-friendly platforms could facilitate more effective offline linkage to HIV testing and prevention services as they would feel less weary about reaching out to health workers affiliated with MSM-friendly institutions. A participant further emphasized the importance of having an anonymous user setting, reiterating privacy concerns and stating that "people might not have the comfort to go to a person and let them know they want to do the testing."

## **Discussion**

### **Principal Findings**

In settings such as Malaysia, where same-sex sexual behaviors are illegal in secular and Shariah law, there are several key design and implementation factors that must be considered to improve the acceptability of AI chatbots that target HIV prevention in MSM. The UTAUT framework is an ideal heuristic for examining barriers and facilitators for new technologies; therefore, we analyzed the potential facilitators and barriers associated with MSM's acceptance of AI chatbots aimed at promoting HIV testing and prevention based on the UTAUT, categorizing participants' insights into performance expectancy, effort expectancy, facilitating conditions, and social influence. There were no additional or anomalous themes that stood incompatible or incongruent with the 4 aforementioned factors. The findings suggested that, with the right design, features, and platforms, the implementation of an AI chatbot to promote HIV testing and prevention could be acceptable to MSM, an at-risk population highly vulnerable to HIV infection in Malaysia.

The UTAUT states that positively perceived performance expectancy and perceived low effort to use facilitate use intention and technology acceptance [27]. On the basis of our results, an AI chatbot's perceived usefulness and ease of use stem from its ability to disseminate valuable HIV-related information, provide easy answers to users' questions, and raise HIV prevention awareness, and from its convenience and 24/7 availability. The participants placed a high value on an AI chatbot's ability to operationalize these features. In light of participants' opinions and direct feedback on preferred features, the design of an HIV prevention AI chatbot needs to (1) incorporate security and privacy settings and ensure anonymity to protect chatbot users; (2) have easily reachable human agents and a customer support section embedded in the AI chatbot interface; (3) have a learning center or information center embedded in the interface and consider a wide variety of educational, awareness-raising, and informative materials to be incorporated into the information center; (4) undergo meticulous model training to answer complex questions with high accuracy and clarity; (5) provide signposting questions to break down complex questions to facilitate complex problem-solving; (6) have a user-friendly and intuitive interface that does not undermine older users' ability to navigate the AI chatbot; (7) provide a navigation manual or chatbot function tour for newly registered users; and (8) have an efficient and responsive technical support sector where technical issues can be reported and investigated 24/7. As demonstrated in previous applications of the UTAUT in assessing technology acceptance, performance expectancy and effort expectancy had a significant impact on use intention, although performance expectancy is often reported as a stronger factor [39-43]. Coupled with these findings, the qualitative insights obtained from this study further demonstrate the necessity to ensure and continuously enhance an HIV prevention AI chatbot's user-centered designs grounded on users' feedback on the chatbot's performance and ease of use.

Findings from this study further expand the facilitating conditions that may encourage HIV prevention AI chatbot acceptance and use. Incorporating linkage to health care professionals and services should be considered in the design and implementation of the chatbot. Not only should an HIV prevention AI chatbot be designed as a one-stop center for information provision, but it should also be envisioned as a hub of health care connection that provides effective linkage and communication between health care consumers and health care providers. This finding accentuates the importance of socioecological and systematic thinking in the design of digital health tools. Health care services should ideally be treated as an ecosystem rather than as separate localities of services where one service locale readily affects the acceptance and use of another. Eliciting participation from local HIV clinics, HIV testing sites, and HIV health care providers; incorporating them as part of an HIV prevention AI chatbot's user-chatbot-provider ecosystem; and allowing the AI chatbot to readily transfer patient data upon request to health care services could prove crucial in improving the implementation outcomes of the chatbot, such as acceptability and sustainability.

It should also be noted that individual internal facilitating conditions, such as attitudes toward HIV self-testing, are

intricately shaped by external infrastructural conditions. If an AI chatbot intends to encourage HIV self-testing by linking users to HIV home testing services, the overall functionality of the self-testing service system and the information dissemination and educational infrastructure placed by HIV home test vendors would be a significant factor in shaping participants' attitudes toward HIV self-testing and, consequently, their attitudes toward the acceptance of the HIV prevention chatbot. Therefore, the chatbot development team could reach out to HIV home test vendors to encourage educational campaigns regarding their products, with an emphasis on the efficacy, transparency, and limitations of their HIV home testing kits. To help shape users' positive perception of self-testing, the AI chatbot could also incorporate a function that allows users to compare the quality of different self-testing kits, such as their price, sensitivity and specificity, positive and negative predictive values, and user reviews and satisfaction ratings. As Venkatesh et al [27] maintained, "facilitating conditions have a direct positive effect on intention to use." As supported by a previous mixed methods study on the acceptance of an informational antituberculosis chatbot among South Korean adults, facilitating conditions, specifically "the extent to which users think organizational and technical infrastructure exists to support the use of antituberculosis chatbots," showed a strong connection with the acceptance of the chatbot by patients with tuberculosis [24]. Given the findings of this study and the previous literature, the design of an HIV prevention AI chatbot should adopt systematic thinking by creating an efficacious user-centered chatbot HIV care service system and investing in improving facilitating conditions beyond focusing on the characteristics of the chatbot itself.

This study also provides insights into how social influence might affect the acceptance of an HIV prevention AI chatbot. Social influence is originally defined by Venkatesh et al [27] as "the degree to which an individual perceives that important others believe he or she should use the new system" in the UTAUT. The implication of this factor is that people's behavior is affected by and adjusted to others' perceptions of them. In this study, the definition of social influence goes beyond how one's behavioral intention is directly affected by important others' perceptions of them and focuses on how sociocultural climate and collective social perceptions, such as stigma, discrimination, and criminalization, affect individuals' behavior. The interviews revealed that the anti-lesbian, gay, bisexual, transgender, and queer culture and the taboo around HIV in Malaysia shape MSM's privacy concerns around HIV care and act as an impetus for MSM to resort to proprivacy technologies such as an HIV prevention AI chatbot, which would allow them to bypass human interaction in HIV testing and care services. Although previous studies attempting to empirically validate the UTAUT demonstrate that social influences have an inconsistent effect on technology acceptance across different settings [26], this study points to a new lens through which social influence can be examined. Instead of focusing purely on how individual behavior is affected by important others' perceptions of them, macrolevel social influences, such as cultural stigma, discrimination, social inequality, health justice, and identity oppression, can be further investigated as potential variables of social influence and, thus, technology acceptance. The

consideration of these systematic sociocultural factors could prove particularly salient for the design of mHealth interventions targeted at systematically disadvantaged at-risk populations who experience a high burden of social marginalization and health care inequality.

We offer 2 major suggestions for researchers worldwide when designing culturally tailored AI chatbots. First, the examination of the social determinants of the health outcome or disease that an AI chatbot aims to address should be an integral part of the design process. The understanding of the sociocultural mechanism that drives a specific population's behaviors and attitudes toward a health condition and health care can be used to decide what health care stakeholders an AI chatbot should aspire to connect to effectively encourage health-promoting actions among targeted populations. Different levels of stakeholders that can be connected to form a user-centered AI health care ecosystem include users, families, friends, health care providers, community health workers, nonprofit organizations, governmental agencies, policy makers, researchers, and community support networks. Second, although upstream social determinants such as health policies, discrimination and sociocultural stigma, and systematic health care inequality cannot be single-handedly solved by one mHealth technology, features such as empathetic NLP and human custom support agent empathy training that could assuage users' system-, stigma-, or discrimination-induced anxiety around a health condition can be incorporated into an AI chatbot to encourage positive attitudes toward health and care seeking. The latest publication by Rahmanti et al [44] on a chatbot designed with artificial empathy features for weight management demonstrated the promising impact of empathetic NLP on engendering long-term behavior change and fostering emotional and social support. Thus, we recommend the incorporation of artificial empathy into the design of AI chatbots that target stigma-ridden and anxiety-inducing health conditions in high-risk populations. Efforts should also be dedicated to training culturally and socially tailored empathetic responses in AI chatbot NLPs rather than assuming one-response-fits-all empathetic languages. Empathy training for health care workers has been described as of "unquestionable importance" [45]. In the virtual space, an AI chatbot's human support agents undertake the same tasks of user or patient communication and problem-solving as health care workers in the clinical setting. The final implementation and operationalization of the entire AI chatbot network, where the chatbot may refer users to human agents for further support, necessitates the consideration of empathy training not only within the chatbot programming but also within the human extension of the chatbot system. Anti-implicit bias training and compassionate care may be considered as part of the human support empathy training.

Finally, the UTAUT states that the effect of performance expectancy, effort expectancy, facilitating conditions, and social influence is "moderated by age, gender, experience and voluntariness of use" [27]. Although this study did not systematically explore the moderating effects of these factors, it indicated the potential disparity in the acceptance of an HIV prevention AI chatbot between younger and older users. The older population may expect a higher level of difficulty in using



the chatbot because of limited technological literacy. Future randomized controlled trials that examine the moderating effects of multi-sectional identities such as age, experience, socioeconomic background, and education level could prove valuable in demonstrating whether, given the same AI chatbot, use intention and acceptance of technology would differ among different populations.

Acceptability is a widely assessed implementation outcome, defined as the perception among implementation stakeholders that a given treatment, service, practice, or innovation is agreeable or satisfactory [46]. Both *acceptance* and *acceptability* were mentioned in this study, but the distinction between the 2 terms should be noted. We used *acceptance* as intended in the UTAUT, denoting users' intentions to use a technology [27]. This study specifically investigated the potential barriers and facilitators associated with users' intention to use AI chatbots for HIV testing and prevention. We used *acceptability* as intended in implementation science, denoting an implementation outcome [46]. Insights into what affects users' acceptance were used to discuss the implications of this study on the implementation outcome of the chatbot.

Overall, the study's findings illustrate the potential facilitators of and barriers to MSM's acceptance of an AI chatbot as a means to promote HIV testing and prevention, which are classified according to the 4 constructs of the UTAUT. Insights related to performance expectancy, effort expectancy, facilitating conditions, and social influences all suggest that an HIV prevention AI chatbot could have high acceptance among MSM in Malaysia. Our evidence speaks to the imperative of using user-centered design to enhance AI chatbots' convenience, availability, problem-solving capacity, and ability to provide valuable information and raise HIV prevention awareness. Conversely, negligence of these factors could impede user acceptance of the chatbot. Our findings also point to a greater chance of user acceptance if systematic and socioecological thinking is effectively incorporated into the design of AI chatbots, where linkage to health care services and vendor-consumer information dissemination are well supported by the functions and health care network incorporated into the chatbot. This study further provides hypotheses regarding the impact of health stigma and discrimination on HIV-related health technologies, which warrants further investigation and validation through randomized controlled trials and quantitative regression analysis. As the preimplementation formative research for an HIV prevention AI chatbot in Malaysia evolves, these insights will be used to inform the design of the AI chatbot and the testing, evaluation, and optimization of the implementation outcomes of the AI chatbot, such as appropriateness, feasibility, and sustainability, in the mid- and postimplementation stages of the AI chatbot development program.

### Limitations

Although this study contributes important knowledge to the understanding of the use of AI chatbots as an HIV prevention strategy, the following limitations must be acknowledged. First,

given the preimplementation phase and the explorative nature of this study, the findings neither statistically tested any hypothesis regarding user acceptance of AI chatbot technology nor confirmed the actual acceptability of an AI chatbot. Further quantitative or mixed methods research should be conducted to corroborate whether the reported factors are related to the actual acceptability and usability of an MSM-targeted AI chatbot in Malaysia. Second, all interviews were conducted in English, which may have limited the interpretation of the data. A large sample stratified by English and non-English speakers may reveal other facilitators and barriers associated with participants' acceptance of an AI chatbot. Nevertheless, as the aim of this study was not to test a hypothesis but rather to generate in-depth insights, this sample would not undermine the significance or validity of the participants' opinions [47]. Third, convenience sampling was used in the recruitment of interviewees. Although the recruited participants represented a wide range of demographic characteristics (shown in Table 1), thus mitigating the risk of overly narrow insights because of homogeneous population characteristics, random sampling could strengthen the study by incorporating a greater demographic diversity of participants, thus generating a greater breadth of insights. Furthermore, given the taboo culture around HIV and nonheterosexual orientation in Malaysia, participants might have withheld information regarding sensitive personal experiences and social issues regarding HIV care, which might otherwise provide more insights for this study. We also acknowledge the possibility of the misreporting of MSM identity among participants, although we believe this was highly improbable as participants voluntarily responded to our recruitment on social media and, in a sociopolitical environment that discriminates against MSM, the motivation to take on a false marginalized identity would be irrational. In addition, this study only obtained potential MSM users' perspectives. An investigation of health care providers' and other stakeholders' perspectives, such as community health workers and local policy stakeholders, could have provided a richer understanding of what might encourage or discourage the acceptance and use of an AI chatbot in Malaysia. Further research is warranted in this regard.

### Conclusions

In the absence of research that examines what influences the acceptance of AI chatbots as tools for HIV prevention in an anti-lesbian, gay, bisexual, transgender, and queer cultural setting, this study contributes to the understanding of what may affect users' acceptance of AI chatbots designed to promote HIV testing and prevention in Malaysia. The results of this study present the perceived benefits and concerns of using an AI chatbot among MSM and apply the UTAUT framework to understand the acceptance of health technology used for the prevention of stigma-ridden diseases. Not only can these findings inform the future design of AI chatbots aimed at promoting HIV testing and prevention, but they also provide direction for the integration of health stigma into the UTAUT's application in health technology settings.

## Acknowledgments

This work was supported by research grants from the National Institutes of Health (R21 TW011663 for ZN and FLA; R21 AI152927 for JAW, FLA, and ZN; and R21 TW011665 for RS) and the Fund for Lesbian and Gay Studies (FLAGS) Research Award at Yale University.

## Conflicts of Interest

None declared.

### Multimedia Appendix 1

Interview guide of the chatbot project—men who have sex with men.

[[PDF File \(Adobe PDF File\), 98 KB - formative\\_v6i10e42055\\_app1.pdf](#)]

### Multimedia Appendix 2

Participants' insights with illustrative quotes.

[[PDF File \(Adobe PDF File\), 148 KB - formative\\_v6i10e42055\\_app2.pdf](#)]

## References

1. HIV in Men Who Have Sex with Men. *The Lancet*. 2012 Jul 20. URL: <https://www.thelancet.com/series/hiv-in-men-who-have-sex-with-men> [accessed 2022-05-21]
2. Baral S, Turner RM, Lyons CE, Howell S, Honermann B, Garner A, et al. Population size estimation of gay and bisexual men and other men who have sex with men using social media-based platforms. *JMIR Public Health Surveill* 2018 Feb 08;4(1):e15 [FREE Full text] [doi: [10.2196/publichealth.9321](https://doi.org/10.2196/publichealth.9321)] [Medline: [29422452](https://pubmed.ncbi.nlm.nih.gov/29422452/)]
3. Suleiman A, Tze CP. Country Progress Report On HIV/AIDS 2019 Malaysia. HIV/STI/Hepatitis C Section of Ministry of Health, Malaysia. 2019. URL: [https://www.moh.gov.my/moh/resources/Penerbitan/Laporan/Umum/Report\\_GAM\\_2019\\_\(Final\).pdf](https://www.moh.gov.my/moh/resources/Penerbitan/Laporan/Umum/Report_GAM_2019_(Final).pdf) [accessed 2022-09-21]
4. Buchbinder SP, Liu AY. CROI 2019: advances in HIV prevention and plans to end the epidemic. *Top Antivir Med* 2019 Apr;27(1):8-25 [FREE Full text] [Medline: [31136999](https://pubmed.ncbi.nlm.nih.gov/31136999/)]
5. DiNenno EA, Prejean J, Irwin K, Delaney KP, Bowles K, Martin T, et al. Recommendations for HIV screening of gay, bisexual, and other men who have sex with men - United States, 2017. *MMWR Morb Mortal Wkly Rep* 2017 Aug 11;66(31):830-832 [FREE Full text] [doi: [10.15585/mmwr.mm6631a3](https://doi.org/10.15585/mmwr.mm6631a3)] [Medline: [28796758](https://pubmed.ncbi.nlm.nih.gov/28796758/)]
6. DiDomizio E. HIV-Syphilis Co-infection Among MSM in Malaysia. New Haven, CT, USA: School of Public Health, Yale University; 2019.
7. Ngadiman S, Suleiman A, Aziz MN, Yuswan F, Taib SM, Masnin H, et al. Malaysia National Strategic Plan for Ending AIDS 2016-2030. HIV/STI Section, Disease Control Division, Ministry of Health Malaysia. 2015. URL: <https://www.aidsdatahub.org/sites/default/files/resource/malaysia-national-strategic-plan-ending-aids-2016-2030.pdf> [accessed 2020-11-30]
8. Ellis-Petersen H. Malaysia accused of 'state-sponsored homophobia' after LGBT crackdown. *The Guardian*. 2018 Aug 22. URL: <https://www.theguardian.com/world/2018/aug/22/malaysia-accused-of-state-sponsored-homophobia-after-lgbt-crackdown> [accessed 2022-03-27]
9. Lim SH, Cheung DH, Guadamuz TE, Wei C, Koe S, Altice FL. Latent class analysis of substance use among men who have sex with men in Malaysia: findings from the Asian Internet MSM Sex Survey. *Drug Alcohol Depend* 2015 Jun 01;151:31-37. [doi: [10.1016/j.drugalcdep.2015.02.040](https://doi.org/10.1016/j.drugalcdep.2015.02.040)] [Medline: [25865907](https://pubmed.ncbi.nlm.nih.gov/25865907/)]
10. Purcell DW, Parsons JT, Halkitis PN, Mizuno Y, Woods WJ. Substance use and sexual transmission risk behavior of HIV-positive men who have sex with men. *J Subst Abuse* 2001;13(1-2):185-200. [doi: [10.1016/s0899-3289\(01\)00072-4](https://doi.org/10.1016/s0899-3289(01)00072-4)] [Medline: [11547619](https://pubmed.ncbi.nlm.nih.gov/11547619/)]
11. Bourne A, Cassolato M, Thuan Wei CK, Wang B, Pang J, Lim SH, et al. Willingness to use pre-exposure prophylaxis (PrEP) for HIV prevention among men who have sex with men (MSM) in Malaysia: findings from a qualitative study. *J Int AIDS Soc* 2017 Aug 02;20(1):21899 [FREE Full text] [doi: [10.7448/IAS.20.1.21899](https://doi.org/10.7448/IAS.20.1.21899)] [Medline: [28782336](https://pubmed.ncbi.nlm.nih.gov/28782336/)]
12. Poteat T, Diouf D, Drame FM, Ndaw M, Traore C, Dhaliwal M, et al. HIV risk among MSM in Senegal: a qualitative rapid assessment of the impact of enforcing laws that criminalize same sex practices. *PLoS One* 2011;6(12):e28760 [FREE Full text] [doi: [10.1371/journal.pone.0028760](https://doi.org/10.1371/journal.pone.0028760)] [Medline: [22194906](https://pubmed.ncbi.nlm.nih.gov/22194906/)]
13. Earnshaw VA, Jin H, Wickersham JA, Kamarulzaman A, John J, Lim SH, et al. Stigma toward men who have sex with men among future healthcare providers in Malaysia: would more interpersonal contact reduce prejudice? *AIDS Behav* 2016 Jan;20(1):98-106 [FREE Full text] [doi: [10.1007/s10461-015-1168-x](https://doi.org/10.1007/s10461-015-1168-x)] [Medline: [26324078](https://pubmed.ncbi.nlm.nih.gov/26324078/)]
14. Duthely LM, Sanchez-Covarrubias AP. Digitized HIV/AIDS treatment adherence interventions: a review of recent SMS/texting mobile health applications and implications for theory and practice. *Front Commun (Lausanne)* 2020 Nov;5:530164 [FREE Full text] [doi: [10.3389/fcomm.2020.530164](https://doi.org/10.3389/fcomm.2020.530164)] [Medline: [33644162](https://pubmed.ncbi.nlm.nih.gov/33644162/)]

15. Pai N, Esmail A, Saha Chaudhuri P, Oelofse S, Pretorius M, Marathe G, et al. Impact of a personalised, digital, HIV self-testing app-based program on linkages and new infections in the township populations of South Africa. *BMJ Glob Health* 2021 Aug;6(9):e006032 [FREE Full text] [doi: [10.1136/bmjgh-2021-006032](https://doi.org/10.1136/bmjgh-2021-006032)] [Medline: [34475026](https://pubmed.ncbi.nlm.nih.gov/34475026/)]
16. Wood OR, Garofalo R, Kuhns LM, Scherr TF, Zetina AP, Rodriguez RG, et al. A randomized controlled trial of an mHealth intervention for increasing access to HIV testing and care among young cisgender men and transgender women: the mLab App study protocol. *BMC Public Health* 2021 Oct 29;21(1):1959 [FREE Full text] [doi: [10.1186/s12889-021-12015-w](https://doi.org/10.1186/s12889-021-12015-w)] [Medline: [34715833](https://pubmed.ncbi.nlm.nih.gov/34715833/)]
17. Cooper V, Clatworthy J, Whetham J, EmERGE Consortium. mHealth interventions to support self-management in HIV: a systematic review. *Open AIDS J* 2017 Nov 21;11:119-132 [FREE Full text] [doi: [10.2174/1874613601711010119](https://doi.org/10.2174/1874613601711010119)] [Medline: [29290888](https://pubmed.ncbi.nlm.nih.gov/29290888/)]
18. Schachner T, Keller R, Wangenheim FV. Artificial intelligence-based conversational agents for chronic conditions: systematic literature review. *J Med Internet Res* 2020 Sep 14;22(9):e20701 [FREE Full text] [doi: [10.2196/20701](https://doi.org/10.2196/20701)] [Medline: [32924957](https://pubmed.ncbi.nlm.nih.gov/32924957/)]
19. Laranjo L, Dunn AG, Tong HL, Kocaballi AB, Chen J, Bashir R, et al. Conversational agents in healthcare: a systematic review. *J Am Med Inform Assoc* 2018 Sep 01;25(9):1248-1258 [FREE Full text] [doi: [10.1093/jamia/ocy072](https://doi.org/10.1093/jamia/ocy072)] [Medline: [30010941](https://pubmed.ncbi.nlm.nih.gov/30010941/)]
20. Xu L, Sanders L, Li K, Chow JC. Chatbot for health care and oncology applications using artificial intelligence and machine learning: systematic review. *JMIR Cancer* 2021 Nov 29;7(4):e27850 [FREE Full text] [doi: [10.2196/27850](https://doi.org/10.2196/27850)] [Medline: [34847056](https://pubmed.ncbi.nlm.nih.gov/34847056/)]
21. Boucher EM, Harake NR, Ward HE, Stoeckl SE, Vargas J, Minkel J, et al. Artificially intelligent chatbots in digital mental health interventions: a review. *Expert Rev Med Devices* 2021 Dec;18(sup1):37-49. [doi: [10.1080/17434440.2021.2013200](https://doi.org/10.1080/17434440.2021.2013200)] [Medline: [34872429](https://pubmed.ncbi.nlm.nih.gov/34872429/)]
22. Amiri P, Karahanna E. Chatbot use cases in the Covid-19 public health response. *J Am Med Inform Assoc* 2022 Apr 13;29(5):1000-1010 [FREE Full text] [doi: [10.1093/jamia/ocac014](https://doi.org/10.1093/jamia/ocac014)] [Medline: [35137107](https://pubmed.ncbi.nlm.nih.gov/35137107/)]
23. Chaix B, Bibault JE, Pienkowski A, Delamon G, Guillemassé A, Nectoux P, et al. When chatbots meet patients: one-year prospective study of conversations between patients with breast cancer and a chatbot. *JMIR Cancer* 2019 May 02;5(1):e12856 [FREE Full text] [doi: [10.2196/12856](https://doi.org/10.2196/12856)] [Medline: [31045505](https://pubmed.ncbi.nlm.nih.gov/31045505/)]
24. Kim AJ, Yang J, Jang Y, Baek JS. Acceptance of an informational antituberculosis chatbot among Korean adults: mixed methods research. *JMIR Mhealth Uhealth* 2021 Nov 09;9(11):e26424 [FREE Full text] [doi: [10.2196/26424](https://doi.org/10.2196/26424)] [Medline: [34751667](https://pubmed.ncbi.nlm.nih.gov/34751667/)]
25. Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. *Implement Sci* 2009 Aug 07;4:50 [FREE Full text] [doi: [10.1186/1748-5908-4-50](https://doi.org/10.1186/1748-5908-4-50)] [Medline: [19664226](https://pubmed.ncbi.nlm.nih.gov/19664226/)]
26. Davis FD, Bagozzi RP, Warshaw PR. User acceptance of computer technology: a comparison of two theoretical models. *Manag Sci* 1989 Aug;35(8):982-1003 [FREE Full text] [doi: [10.1287/mnsc.35.8.982](https://doi.org/10.1287/mnsc.35.8.982)]
27. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. *MIS Q* 2003 Sep;27(3):425-478 [FREE Full text] [doi: [10.2307/30036540](https://doi.org/10.2307/30036540)]
28. Eckhardt A, Laumer S, Weitzel T. Who influences whom? Analyzing workplace referents' social influence on it adoption and non-adoption. *J Inf Technol* 2009 Mar 01;24(1):11-24 [FREE Full text] [doi: [10.1057/jit.2008.31](https://doi.org/10.1057/jit.2008.31)]
29. Curtis L, Edwards C, Fraser KL, Gudelsky S, Holmquist J, Thornton K, et al. Adoption of social media for public relations by nonprofit organizations. *Public Relat Rev* 2010 Mar;36(1):90-92. [doi: [10.1016/j.pubrev.2009.10.003](https://doi.org/10.1016/j.pubrev.2009.10.003)]
30. Verhoeven JC, Heerwegh D, De Wit K. Information and communication technologies in the life of university freshmen: an analysis of change. *Comput Educ* 2010 Aug;55(1):53-66. [doi: [10.1016/j.compedu.2009.12.002](https://doi.org/10.1016/j.compedu.2009.12.002)]
31. Venkatesh V, Thong JY, Xu X. Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. *MIS Q* 2012 Mar;36(1):157-178. [doi: [10.2307/41410412](https://doi.org/10.2307/41410412)]
32. Udo G, Bagchi K, Maity M. Exploring factors affecting digital piracy using the norm activation and UTAUT models: the role of national culture. *J Bus Ethics* 2014 Dec 9;135(3):517-541. [doi: [10.1007/s10551-014-2484-1](https://doi.org/10.1007/s10551-014-2484-1)]
33. Gupta B, Dasgupta S, Gupta A. Adoption of ICT in a government organization in a developing country: an empirical study. *J Strategic Inf Syst* 2008 Jun;17(2):140-154. [doi: [10.1016/j.jsis.2007.12.004](https://doi.org/10.1016/j.jsis.2007.12.004)]
34. Nadri H, Rahimi B, Lotfnezhad Afshar H, Samadbeik M, Garavand A. Factors affecting acceptance of hospital information systems based on extended technology acceptance model: a case study in three paraclinical departments. *Appl Clin Inform* 2018 Apr;9(2):238-247 [FREE Full text] [doi: [10.1055/s-0038-1641595](https://doi.org/10.1055/s-0038-1641595)] [Medline: [29618139](https://pubmed.ncbi.nlm.nih.gov/29618139/)]
35. Kallio H, Pietilä AM, Johnson M, Kangasniemi M. Systematic methodological review: developing a framework for a qualitative semi-structured interview guide. *J Adv Nurs* 2016 Dec;72(12):2954-2965. [doi: [10.1111/jan.13031](https://doi.org/10.1111/jan.13031)] [Medline: [27221824](https://pubmed.ncbi.nlm.nih.gov/27221824/)]
36. Parviainen J, Rantala J. Chatbot breakthrough in the 2020s? An ethical reflection on the trend of automated consultations in health care. *Med Health Care Philos* 2022 Mar;25(1):61-71 [FREE Full text] [doi: [10.1007/s11019-021-10049-w](https://doi.org/10.1007/s11019-021-10049-w)] [Medline: [34480711](https://pubmed.ncbi.nlm.nih.gov/34480711/)]
37. Wong LP. Focus group discussion: a tool for health and medical research. *Singapore Med J* 2008 Mar;49(3):256-261 [FREE Full text] [Medline: [18363011](https://pubmed.ncbi.nlm.nih.gov/18363011/)]

38. Sewell J, Cambiano V, Speakman A, Lampe FC, Phillips A, Stuart D, et al. Changes in chemsex and sexual behaviour over time, among a cohort of MSM in London and Brighton: findings from the AURAH2 study. *Int J Drug Policy* 2019 Jun;68:54-61 [FREE Full text] [doi: [10.1016/j.drugpo.2019.03.021](https://doi.org/10.1016/j.drugpo.2019.03.021)] [Medline: [30999243](https://pubmed.ncbi.nlm.nih.gov/30999243/)]
39. Kijsanayotin B, Pannarunothai S, Speedie SM. Factors influencing health information technology adoption in Thailand's community health centers: applying the UTAUT model. *Int J Med Inform* 2009 Jun;78(6):404-416. [doi: [10.1016/j.ijmedinf.2008.12.005](https://doi.org/10.1016/j.ijmedinf.2008.12.005)] [Medline: [19196548](https://pubmed.ncbi.nlm.nih.gov/19196548/)]
40. Wills MJ, El-Gayar OF, Bennett D. Examining healthcare professionals' acceptance of electronic medical records using UTAUT. *Issues Inf Syst* 2008;IX(2):396-401. [doi: [10.48009/2\\_iis\\_2008\\_396-401](https://doi.org/10.48009/2_iis_2008_396-401)]
41. Mayer P, Volland D, Thiesse F, Fleisch E. User Acceptance of 'Smart Products': An Empirical Investigation. *Wirtschaftsinformatik Proceedings*. 2011. URL: [https://www.researchgate.net/publication/221200885\\_User\\_Acceptance\\_of\\_'Smart\\_Products'\\_An\\_Empirical\\_Investigation](https://www.researchgate.net/publication/221200885_User_Acceptance_of_'Smart_Products'_An_Empirical_Investigation) [accessed 2020-12-10]
42. Neufeld DJ, Dong L, Higgins C. Charismatic leadership and user acceptance of information technology. *Eur J Inf Syst* 2017 Dec 19;16(4):494-510. [doi: [10.1057/palgrave.ejis.3000682](https://doi.org/10.1057/palgrave.ejis.3000682)]
43. Nov O, Ye C. Resistance to change and the adoption of digital libraries: an integrative model. *J Am Soc Inf Sci* 2009 Aug;60(8):1702-1708. [doi: [10.1002/asi.21068](https://doi.org/10.1002/asi.21068)]
44. Rahmanti AR, Yang HC, Bintoro BS, Nursetyo AA, Muhtar MS, Syed-Abdul S, et al. SlimMe, a chatbot with artificial empathy for personal weight management: system design and finding. *Front Nutr* 2022 Jun 23;9:870775 [FREE Full text] [doi: [10.3389/fnut.2022.870775](https://doi.org/10.3389/fnut.2022.870775)] [Medline: [35811989](https://pubmed.ncbi.nlm.nih.gov/35811989/)]
45. Loos J, Manirankunda L, Platteau T, Albers L, Franssen K, Vermoesen T, et al. Acceptability of a community-based outreach HIV-testing intervention using oral fluid collection devices and web-based HIV test result collection among Sub-Saharan African migrants: a mixed-method study. *JMIR Public Health Surveill* 2016 Aug 04;2(2):e33 [FREE Full text] [doi: [10.2196/publichealth.5519](https://doi.org/10.2196/publichealth.5519)] [Medline: [27493067](https://pubmed.ncbi.nlm.nih.gov/27493067/)]
46. Proctor E, Silmere H, Raghavan R, Hovmand P, Aarons G, Bunger A, et al. Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. *Adm Policy Ment Health* 2011 Mar;38(2):65-76 [FREE Full text] [doi: [10.1007/s10488-010-0319-7](https://doi.org/10.1007/s10488-010-0319-7)] [Medline: [20957426](https://pubmed.ncbi.nlm.nih.gov/20957426/)]
47. Carminati L. Generalizability in qualitative research: a tale of two traditions. *Qual Health Res* 2018 Nov;28(13):2094-2101. [doi: [10.1177/1049732318788379](https://doi.org/10.1177/1049732318788379)] [Medline: [30043686](https://pubmed.ncbi.nlm.nih.gov/30043686/)]

## Abbreviations

**AI:** artificial intelligence

**mHealth:** mobile health

**MSM:** men who have sex with men

**NLP:** natural language processing

**TAM:** technology acceptance model

**UTAUT:** unified theory of acceptance and use of technology

*Edited by A Mavragani; submitted 19.08.22; peer-reviewed by J Chow, C Zhang; comments to author 15.09.22; revised version received 17.09.22; accepted 19.09.22; published 06.10.22.*

*Please cite as:*

*Peng ML, Wickersham JA, Altice FL, Shrestha R, Azwa I, Zhou X, Halim MAA, Ikhtiaruddin WM, Tee V, Kamarulzaman A, Ni Z Formative Evaluation of the Acceptance of HIV Prevention Artificial Intelligence Chatbots By Men Who Have Sex With Men in Malaysia: Focus Group Study*

*JMIR Form Res* 2022;6(10):e42055

URL: <https://formative.jmir.org/2022/10/e42055>

doi: [10.2196/42055](https://doi.org/10.2196/42055)

PMID: [36201390](https://pubmed.ncbi.nlm.nih.gov/36201390/)

©Mary L Peng, Jeffrey A Wickersham, Frederick L Altice, Roman Shrestha, Iskandar Azwa, Xin Zhou, Mohd Akbar Ab Halim, Wan Mohd Ikhtiaruddin, Vincent Tee, Adeeba Kamarulzaman, Zhao Ni. Originally published in *JMIR Formative Research* (<https://formative.jmir.org>), 06.10.2022. 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 Related to Smoking and Perceptions of a Behavioral Counseling and Messenger Service–Delivered Smoking Cessation Intervention for People With HIV in China: Qualitative Study

Shanyin Yang<sup>1</sup>, MPH; Jiegang Huang<sup>2</sup>, PhD; Li Ye<sup>2,3</sup>, PhD; Jianyan Lin<sup>4</sup>, PhD; Zhiman Xie<sup>4</sup>, MMED; Baodong Guo<sup>2</sup>, MMED; Yanjun Li<sup>4</sup>, MMED; Bingyu Liang<sup>2,3</sup>, PhD; Zhigang Zheng<sup>2</sup>, PhD; Karsten Lunze<sup>1,5</sup>, MD, MPH, DrPH; Abu S Abdullah<sup>5</sup>, MD, PhD; Hao Liang<sup>2,3,6</sup>, PhD; Lisa M Quintiliani<sup>1,5</sup>, PhD

<sup>1</sup>Section of General Internal Medicine, Boston Medical Center, Boston, MA, United States

<sup>2</sup>Guangxi Key Laboratory of AIDS Prevention and Treatment, School of Public Health, Guangxi Medical University, Nanning, Guangxi, China

<sup>3</sup>Collaborative Innovation Centre of Regenerative Medicine and Medical BioResource Development and Application Co-constructed by the Province and Ministry, Nanning, Guangxi, China

<sup>4</sup>The Fourth People's Hospital of Nanning, Nanning, Guangxi, China

<sup>5</sup>Department of Medicine, Boston University School of Medicine, Boston, MA, United States

<sup>6</sup>Life Science Institute, Guangxi Medical University, Nanning, Guangxi, China

**Corresponding Author:**

Lisa M Quintiliani, PhD

Department of Medicine

Boston University School of Medicine

72 E Concord St

Boston, MA, 02118

United States

Phone: 1 617 638 2777

Email: [lisa.quintiliani@bmc.org](mailto:lisa.quintiliani@bmc.org)

## Abstract

**Background:** China, where half of the adult male population smoke tobacco, has one of the highest global burdens of smoking. Smoking rates are even higher among people with HIV. People with HIV can be affected by smoking in multiple ways, including more severe HIV-related symptoms and worse antiretroviral therapy treatment outcomes. However, smoking cessation services targeted for people with HIV are not routinely integrated into HIV care in China. Given the widespread mobile phone ownership, an exploration of factors related to smoking among people with HIV in China who smoke could inform the design and implementation of mobile smoking cessation interventions that target the needs of this vulnerable population.

**Objective:** This study aims to explore the perspectives of smoking, barriers and facilitators to quitting, and perceptions related to a smoking cessation intervention delivered through behavioral counseling sessions and brief daily messenger service (WeChat)–delivered messages.

**Methods:** We recruited people with HIV from the People's 4th Hospital of Nanning, Guangxi, China, and conducted semistructured face-to-face interviews. All interviews were audio-recorded, transcribed verbatim in Chinese, and translated into English for data analysis. We conducted a thematic analysis using a codebook, which was guided by a team-based consensus approach to identify 5 main themes. We also explored themes according to the demographic groups.

**Results:** A total of 24 participants were enrolled in the study. The mean age was 37.2 (SD=13.5) years. The participants had lived with HIV for a mean of 2.4 years. The majority were male (18/24, 75%) and lived in urban or metropolitan settings (19/24, 79%). We identified five main themes: variable knowledge of the harms of smoking, both related and unrelated to HIV; willpower perceived as the primary quitting strategy; a duality of the effect of social factors on quitting; perceptions about optimal features of the smoking cessation intervention (eg, messages should be brief and most frequent during the first few weeks); and the largely negative impact of their HIV diagnosis on smoking behaviors. In addition, some themes differed according to participant demographic characteristics such as age, sex, and education level.

**Conclusions:** We identified barriers to and facilitators of smoking cessation among people with HIV in China by conducting semistructured qualitative interviews. Owing to the adverse impact of smoking on HIV outcomes, targeting cessation interventions

to the unique needs and preferences of people with HIV in China may be needed to increase the effectiveness of future interventions. A pilot clinical trial will be conducted in the future to evaluate this behavioral counseling and brief daily messenger service (WeChat)–delivered messages approach among people with HIV who smoke in China.

(*JMIR Form Res* 2022;6(10):e35923) doi:[10.2196/35923](https://doi.org/10.2196/35923)

## KEYWORDS

mobile health; mHealth; China; smoking; smoking cessation; HIV; qualitative research; SMS text messages; WeChat

## Introduction

### Background

Tobacco use is one of the most important preventable causes of premature death worldwide. In 2018, approximately 50.5% of adult males in China and 2.1% of adult females in China smoked cigarettes [1]. Smoking is more prevalent among people with HIV compared with the general population in China. For instance, a sample of males with HIV from Yunnan province in 2012 had a smoking prevalence of 92% [2]. In addition, the same sample showed that a majority (67.6%) of people with HIV who smoked in China were heavy smokers, smoking 20 cigarettes per day [2]. People with HIV are affected by tobacco in multiple ways: more severe HIV-related symptoms [2-4]; worse treatment outcomes, in part owing to lower antiretroviral therapy adherence [4]; and higher rates of tobacco-related comorbidities such as lung cancer, cervical cancer, and pulmonary disease [4,5]. In China, people with HIV who smoke experience high degrees of stigma, prejudice, and discrimination [6,7]. People with HIV who experience HIV stigma can have high psychosocial distress and low self-esteem and resilience [8]. To deal with HIV-related stigmatization and depression [9], both of which are highly prevalent among people with HIV and have adverse effects on HIV care engagement [10], they may choose to engage in smoking behavior as a coping strategy [8]. Owing to the harmful nature of smoking with negative implications on the course of HIV infection and success of treatment, targeted smoking cessation interventions are needed.

Although there is a large body of research on smoking cessation interventions for the general population [11,12], smoking cessation interventions specifically targeting people with HIV in China are lacking and have not been routinely integrated into clinical HIV services. This specific population has high motivation to engage in smoking cessation interventions to improve health outcomes [13,14]. A few studies have been published about smoking cessation efforts among people with HIV in China who smoke [13,14]; however, these studies are limited to cross-sectional surveys examining motivations for cessation, instead of evaluating cessation interventions [13,14]. An in-depth understanding of the factors related to smoking behaviors and preferences would present opportunities to design culturally targeted smoking cessation interventions for people with HIV who smoke in China.

Mobile phones are a widely used mode of delivery for behavioral interventions because they are readily accessible in many populations and have the potential for widespread use. In China, interventions using mobile technologies for smoking cessation such as SMS text messaging and mobile apps have demonstrated efficacy in smoking cessation [15,16]. In

particular, in China, the WeChat messenger service platform has demonstrated efficacy in changing behaviors across a range of behaviors and health conditions [17-21], including among people with HIV for decreasing depressive symptoms and improving medication adherence [19,20].

### Objectives

The goal of this study was to identify factors related to smoking and cessation among people with HIV in China who were current or former smokers and to explore their perspectives on how to modify a draft of *Quit for Life*, an 8-week intervention that combined 4 counseling sessions and brief daily messenger service (WeChat)–delivered messages. *Quit for Life* was based on a previous intervention, which consisted of cessation and nicotine replacement therapy (NRT) adherence counseling and did not include WeChat messages, tested among Chinese patients living in Hong Kong with erectile dysfunction in a randomized controlled trial conducted from 2004 to 2007. The previous intervention was shown to be effective; the intervention group participants were more likely to quit smoking compared with the control group [22]. On the basis of these efficacy findings, we selected this intervention and decided to boost its potential effectiveness and reach with mobile technology (ie, WeChat messaging). We focused on exploring participants' perspectives on the new component of the intervention (ie, WeChat messaging) and how to modify all intervention content for people with HIV.

## Methods

### Study Setting and Recruitment

The study was conducted at the People's 4th Hospital of Nanning, Guangxi, China. The Guangxi Zhuang Autonomous Region (Guangxi) is located in the southwest region of China. From 2013 to 2015, HIV prevalence in Guangxi increased from 6.6% to 11.2% [23]. In 2018, Guangxi reported more than 50,000 HIV cases in 1 year, ranking it third for the highest number of HIV cases among all provinces in China [24]. From 2003 onward, the Chinese government started providing HIV prevention services and free antiretroviral therapy to encourage individuals to participate in screening and treatment [24]. People's 4th Hospital of Nanning is the only grade A (the highest classification) tertiary hospital (comprehensive referral hospitals) in Nanning city.

To recruit participants, research assistants contacted individuals who received HIV treatment at the outpatient HIV clinic or who were hospitalized in 2021. The eligibility criteria were as follows: aged  $\geq 21$  years, HIV seropositive, current or former smokers, receiving antiretroviral therapy, receiving HIV care at the clinic at the 4th People's Hospital of Nanning, and

speaking Mandarin Chinese. Exclusion criteria were as follows: any serious health problems that preclude participation, only smoke tobacco products other than cigarettes (ie, cigars and electronic devices), or inability to give consent.

### Data Collection

A graduate student interviewer, trained by an investigator on the Guangxi Medical University team experienced with qualitative methods, conducted all interviews face-to-face in Mandarin Chinese. A health care provider from the HIV clinic (ie, physicians and nurse practitioners) was present during the interview and answered the participants' medical questions, if any. All interviews were audio-recorded. Before the interview, the interviewer administered a survey assessing demographic information such as age, sex, ethnicity, education level, current smoking behaviors, and HIV status (self-reported diagnosis, diagnosis and time since diagnosis were later verified in the health record). The interviewer then conducted the interview following a semistructured guide with open-ended questions ([Multimedia Appendix 1](#)) consisting of the following topics: smoking patterns and motivations; knowledge of general health impacts of smoking, second-hand and third-hand smoke, and e-cigarettes; knowledge of the impact of smoking on HIV and disease prognosis; and experiences with previous attempts to quit and perceived facilitators and barriers to quitting. The interviewer briefly described the draft *Quit for Life* intervention, including reading 4 sample WeChat messages, and then explored preferences for delivery modality, duration of contact, frequency of contact, and desired content. Finally, participants had the opportunity to provide additional comments. Participants received compensation (equivalent to US \$15) for completing the onetime interview.

### Data Analysis

The research team entered participants' demographic information using the Research Electronic Data Capture tools (Vanderbilt University) hosted at Boston University [25]. We computed the frequencies of demographic variables to characterize the study sample. Investigators from China transcribed the interview recordings verbatim in Mandarin, and

a bilingual second research assistant who is a native Mandarin speaker then translated the transcriptions into English. The research team in China reviewed the transcripts for validation purposes. All data analyses were performed using the final set of English language transcripts. Using a thematic coding process [26], the bilingual research assistant systematically reviewed all transcripts, generated and noted initial ideas, and developed a preliminary codebook. In consensus with another investigator with expertise in qualitative methods, we modified and refined the codes and subcodes. All transcripts were then coded using the final codebook, with NVivo version 12 (QSR International) as the data management platform. Two team members grouped the codes into themes. We then held 2 joint meetings with research team members from the Guangxi Medical University team in China (including the original interviewers) to collect their feedback about the themes and cultural context. During the joint meetings, we decided to explore demographic (age, sex, and education level) differences in our coding. As we did not address demographic differences directly in our interview guide, we categorized them as exploratory themes. At the end of this process, we developed a final version of the thematic categories.

### Ethical Considerations

The ethics approval for this study was obtained from the institutional review boards of the Boston Medical Center (Institutional Review Board number: H-40111) and the Guangxi Medical University.

## Results

### Overview

A total of 24 participants were enrolled; the majority of participants were male (18/24, 75%) and from outpatient HIV units (19/24, 79%; [Table 1](#)). Interviews were conducted either individually (18 interviews) or in pairs (3 interviews). The average duration of the interviews was approximately 30 minutes. We identified 5 thematic categories ([Table 2](#)) and explored themes among demographic categories ([Multimedia Appendix 2](#)).

**Table 1.** Sociodemographic characteristics of smokers and former smokers with HIV in China participating in qualitative interviews (N=24).

Characteristics	Value
Age (years), mean (SD)	37.3 (13.5)
<b>Sex, n (%)</b>	
Male	18 (75)
Female	6 (25)
<b>Education level<sup>a</sup>, n (%)</b>	
Primary or lower	2 (8)
Lower secondary	8 (33)
Upper secondary	8 (33)
Tertiary or higher	4 (17)
Missing	2 (8)
<b>Ethnicity, n (%)</b>	
Han	13 (54)
Zhuang	10 (42)
Other	1 (4)
<b>Marital status, n (%)</b>	
Divorced	4 (17)
Married	9 (37)
Unmarried	11 (46)
<b>Smoking status, n (%)</b>	
Current	18 (75)
Former	6 (25)
Smoking duration (years), mean (SD)	18.6 (13.6)
Smoking amount (cigarettes per day), mean (SD)	20.9 (15.4)
Duration of HIV diagnosis (years), mean (SD)	2.4 (2.7)
Duration of ART <sup>b</sup> (years), mean (SD)	2.1 (2.8)

<sup>a</sup>Primary or lower education level in China is equivalent to grades 6 or lower in the United States; lower secondary education level in China is equivalent to grades 7 to 9 in the United States; upper secondary education level in China is equivalent to high school grades 10 to 12 or vocational training in the United States; and tertiary education level in China includes junior college, undergraduate, and graduate school.

<sup>b</sup>ART: antiretroviral therapy.



**Table 2.** Frequently noted themes with supporting codes and example quotations of smokers and former smokers with HIV in China.

Theme and parent code	Subcode	Additional sub-code	Quotations
<b>Theme 1: Knowledge on the harms of smoking and personal risk perception</b>			
Knowledge	Source of knowledge	— <sup>a</sup>	<ul style="list-style-type: none"> <li>• “I watched TV ads saying that smoking is bad and harmful to your health.” [M<sup>b</sup>, 37, C<sup>c</sup>]</li> <li>• “[I learned smoking would impact my health from] Science, doctors, and people in the society.” [M, 33, C]</li> </ul>
Knowledge	Health effects	—	<ul style="list-style-type: none"> <li>• “Every year when I do physical examination, my lungs appear to be influenced by smoking to some extent...” [M, 24, Q<sup>d</sup>]</li> <li>• “...I know it’s bad for health. But I don’t know the details on how and what exactly does it impact.” [F<sup>e</sup>, 38, C]</li> </ul>
Knowledge	Effects on HIV	—	<ul style="list-style-type: none"> <li>• “I don’t think there’s a relationship between smoking and HIV infection. It may affect treatment.” [M, 48, Q]</li> <li>• “There’s certainly a damage to the body, but I don’t know much about it.” [M, 25, C]</li> </ul>
Knowledge	Second- and third-hand smoke	—	<ul style="list-style-type: none"> <li>• “I know second-hand smoke, but not third-hand smoke.” [M, 46, C]</li> </ul>
Barriers and motivators to quitting	Barriers to quit	Risk perception	<ul style="list-style-type: none"> <li>• “...people say it’s carcinogenic and affects my lungs. However, since I’m smoking, I shouldn’t think of those things too much.” [M, 21, C]</li> </ul>
Barriers and motivators to quitting	Daily stress	—	<ul style="list-style-type: none"> <li>• “Yes, it is possible [to not relapse if there were no recent stress from work].” [M, 24, C]</li> </ul>
Smoking behavior	Feelings associated with smoking	—	<ul style="list-style-type: none"> <li>• “When I talk about happy things, I smoke more cigarettes than usual.” [M, 25, C]</li> <li>• “I smoked more when I was moody or unhappy and smoked less when I had lighter moods. Having stress also made me smoke more.” [M, 71, Q]</li> </ul>
<b>Theme 2: Willpower as a primary quitting strategy</b>			
Strategies	Willpower	—	<ul style="list-style-type: none"> <li>• “...NRT<sup>f</sup> or other medications are only facilitators. One’s own willpower is the most important and the most crucial.” [M, 46, C]</li> </ul>
Strategies	Gradually decrease	—	<ul style="list-style-type: none"> <li>• “I used to smoke, but it decreased slowly after I got sick, and then I really stopped smoking. When others give me cigarettes, I refused by saying I don’t smoke anymore. I tend to avoid occasion involving drinking and socializing as much as possible.” [M, 43, Q]</li> </ul>
Strategies	Suddenly decrease	—	<ul style="list-style-type: none"> <li>• “I think it’s useless to slowly reduce the amount. If you want to resist the craving, give it up all at once, and you’ll have better chance of quitting successfully.” [M, 48, Q]</li> </ul>
Knowledge	Available cessation resources	—	<ul style="list-style-type: none"> <li>• “If I’m sick and the doctor says I shouldn’t smoke or drink, then I won’t smoke.” [M, 21, C]</li> </ul>
QFL <sup>g</sup> intervention	Useful	Initial interest in joining	<ul style="list-style-type: none"> <li>• “If I see it [smoking cessation counseling offered at outpatient clinic], I will go check it out and learn more about it.” [M, 29, C]</li> </ul>
Quit attempts	Perceptions of cessation resources	Positive	<ul style="list-style-type: none"> <li>• “It can be difficult to refrain from smoking without medications or other help.” [M, 37, C]</li> </ul>
Quit attempts	Perceptions of cessation resources	Negative	<ul style="list-style-type: none"> <li>• “No, there is no need [for other cessation resources]. I can rely on myself to quit.” [M, 30, C]</li> </ul>
Additional concepts	E-cigarettes	Not useful	<ul style="list-style-type: none"> <li>• “I smoked e-cigarettes and it was completely useless. So I shifted back to smoking cigarettes. After I tried smoking cessation, I smoked again and with greater quantity” [M, 48, Q]</li> </ul>

Theme and parent code	Subcode	Additional sub-code	Quotations
Additional concepts	E-cigarettes	Useful	<ul style="list-style-type: none"> <li>“...I’ve used it [e-cigarettes] before. It can replace real cigarettes for a period of time.” [M, 24, C]</li> </ul>
<b>Theme 3: Duality of social factors</b>			
Smoking behavior	Reasons to start smoking	Curiosity	<ul style="list-style-type: none"> <li>“I was bored...And I was curious about the feeling of smoking, so I bought cigarettes to try.” [M, 21, C]</li> </ul>
Smoking behavior	Reasons to start smoking	Social	<ul style="list-style-type: none"> <li>“I went out with friends, and they handed me the cigarettes.” [F, 22, C]</li> </ul>
Barriers and motivators to quitting	Barriers to quit	Social pressure	<ul style="list-style-type: none"> <li>“After I was diagnosed with HIV, I stopped smoking for half a year and then started again. It was during banquets...I was given cigarettes...and then became addicted to it again after a while.” [M, 25, C]</li> </ul>
Smoking behavior	Patterns of smoking	With whom	<ul style="list-style-type: none"> <li>“...when I’m hanging out with my friends, or when my friends give me cigarettes.” [F, 44, C]</li> </ul>
Smoking behavior	Patterns of smoking	Setting	<ul style="list-style-type: none"> <li>“There is no such feeling [influencing others’ health] because most people smoke.” [M, 50, C]</li> <li>“When there are other people together I won’t smoke in cars. I choose to park on the side of the road, smoke and then leave. If I’m alone driving, I smoke with my windows open.” [M, 37, C]</li> </ul>
Quit attempts	Details about attempts	—	<ul style="list-style-type: none"> <li>“...I was thinking too much, I couldn’t accept [the HIV diagnosis], so I started smoking again.” [M, 30, C]</li> </ul>
Barriers and motivators to quitting	Motivators to quit	Social support	<ul style="list-style-type: none"> <li>“Quitting together is definitely better than quitting alone, no matter if it’s having one person as the leader or keeping track of each other’s progress.” [M, 46, C]</li> <li>“The encouragement of my family [helped with smoking cessation].” [M, 43, Q]</li> </ul>
Barriers and motivators to quitting	Motivators to quit	Family pressure	<ul style="list-style-type: none"> <li>“My family didn’t allow me to smoke, doctors didn’t allow me to smoke either.” [M, 58, C]</li> <li>“When first I stopped smoking, I would go to places where I used to put cigarettes, but there was no way I could find them.” [M, 58, C]</li> </ul>
Barriers and motivators to quitting	Motivators to quit	Positive impacts on others	<ul style="list-style-type: none"> <li>“[An advice for other smokers:] For the health of people around you, try your best to reduce the frequency of smoking, and be mindful of your surrounding.” [M, 33, C]</li> <li>“Because you have the responsibility to take care of the family, things you do are closely linked to the future of the family.” [M, 24, Q]</li> </ul>
Barriers and motivators to quitting	Motivators to quit	Social influence and stigma	<ul style="list-style-type: none"> <li>“Smoking is not accepted by the society, it has no social status. Especially you are in a car with a woman. It is not good and you should not smoke continuously at least in terms of human nature...” [M, 46, C]</li> </ul>
Knowledge	Second- and third-hand smoke	—	<ul style="list-style-type: none"> <li>“My family used to blame me for inhaling my second-hand smoke, but I can’t control it.” [M, 48, Q]</li> </ul>

**Theme 4: Impact of HIV diagnosis**

Theme and parent code	Subcode	Additional sub-code	Quotations
HIV diagnosis	After HIV diagnosis	—	<ul style="list-style-type: none"> <li>“It probably does. I tried to control myself. After I was diagnosed with HIV, I did not touch a cigarette for at least two years. However, before Lunar New Year last year, I gradually started again. I smoked a little, then a little more, and then got used to it.” [F, 36, C]</li> <li>“Before, I smoked one pack in two days, after I learned it [HIV diagnosis], I decreased the amount.” [M, 29, C]</li> <li>“I don’t work anymore...Mainly because some jobs need a health certificate. Since I have HIV infection, I cannot apply for the health certificate.” [M, 25, C]</li> </ul>
Additional concepts	ART <sup>h</sup> use	—	<ul style="list-style-type: none"> <li>“After I was diagnosed with HIV, I stopped the quitting thoughts. I think smoking only has minor influence toward human body. I am taking medications everyday now, and that cause more problems.” [M, 46, C]</li> </ul>
Barriers and motivators to quitting	Motivators to quit	Health benefits and physical changes after quitting	<ul style="list-style-type: none"> <li>“If I’m pregnant, I will definitely quit. Reminding someone like me that they have something important to concern [pregnancy etc.]” [F, 44, C]</li> <li>“My HIV infection and severe lung infection caused me to be hospitalized. I quitted smoking when I was being hospitalized. After I returned home, I didn’t smoke anymore...family members didn’t allow me to smoke either. I almost died.” [M, 48, Q]</li> </ul>
Barriers and motivators to quitting	Barriers to quit	Daily stress	<ul style="list-style-type: none"> <li>“...I didn’t [smoke when hospitalized]. I started smoking three months after I was discharged...” [F, 25, C]</li> <li>“...I was thinking too much, I couldn’t accept [the HIV diagnosis], so I started smoking again.” [M, 30, C]</li> </ul>
QFL intervention	Length and time of intervention	—	<ul style="list-style-type: none"> <li>“Many people...are having a hard time accepting the fact that they’re diagnosed with HIV infection. If you talk to them about smoking cessation, they won’t be interested and won’t be willing to spend too much time on it.” [M, 24, Q]</li> </ul>
<b>Theme 5: QFL intervention modification</b>			
QFL intervention	Modality	—	<ul style="list-style-type: none"> <li>“I will read WeChat messages. I usually read WeChat but not text messages.” [M, 24, C]</li> <li>“[WeChat is effective for the intervention, but] don’t construct a WeChat group [vs. private messaging]. That involves personal privacy and no one wants others to know about their condition.” [M, 46, C]</li> <li>“A little bit. I am not very familiar with WeChat.” [M, 71, Q]</li> <li>“I might not answer all the calls...If the number is not from Nanning City, I won’t pick up.” [M, 58, C]</li> </ul>
QFL intervention	Content	—	<ul style="list-style-type: none"> <li>“...I think it’s good to not send any messages at all...you are reminding me that ‘I am a smoker’.” [F, 36, C]</li> <li>“...most mobile messages...are advertisements so I get annoyed and don’t even read them.” [M, 48, Q]</li> </ul>
QFL intervention	Length and time of intervention	—	<ul style="list-style-type: none"> <li>“When we have to wait in line to see a doctor, 40 minutes is a bit long, 20 minutes would be appropriate.” [M, 43, Q]</li> </ul>
QFL intervention	Setting	—	<ul style="list-style-type: none"> <li>“It’s fine if I talk to someone face-to-face when I’m picking up my medications.” [M, 29, C]</li> </ul>
QFL intervention	Frequency	—	<ul style="list-style-type: none"> <li>“The most important period for quitting in the first two weeks...If you can resist your cravings for the first two weeks, then you can stop smoking.” [M, 48, Q]</li> </ul>

<sup>a</sup>Dashes represent when there are no additional subcodes.

<sup>b</sup>M represents participants who report to be male.

<sup>c</sup>C represents participants who report to be current smokers.

<sup>d</sup>Q represents participants who report to be former smokers and have quit smoking.

<sup>e</sup>F represents participants who report to be female.

<sup>f</sup>NRT: nicotine replacement therapy.

<sup>g</sup>QFL: *Quit for Life*.

<sup>h</sup>ART: antiretroviral therapy.

### **Theme 1. Some Awareness of Smoking Harms in General, Yet Low Personal Risk Perceptions of Smoking and Impacts on HIV Treatment Care Outcomes**

Most participants understood that smoking harms their lungs and health, but to different extents. For example, some participants knew that smoking leads to a worsening of their HIV treatment outcomes. Other participants had a limited understanding of the harms from smoking on their own health; for example, a person noted the following: “at least at the moment smoking does not cause any harm for me so far.” A participant expressed that as they had not heard anything from authoritative sources, they would still “take a chance” and think smoking is not harmful. When asked about the harm of smoking and the harm of second- and third-hand smoke, most participants had a general awareness from their medical providers, from television, or through social media, but they lacked specific knowledge of these terms.

Participants saw a close connection among smoking and mood, emotions, and mental well-being. About half of the participants agreed that emotions were closely linked with cravings to smoke. Several smokers mentioned that they “smoked more when [they were] moody or unhappy and smoked less when [they] had lighter moods.” Workplaces were often mentioned as environments that induced stress and anxiety, leading to smoking to calm down and escape.

### **Theme 2: Willpower Was Identified as a Primary Quitting Strategy, With Less Awareness and Trust in Other Resources**

When asked about known strategies for quitting, participants most often mentioned individual effort, determination, and personal willpower; for example: “quitting smoking mainly depends on [one’s] self-control of the impulse to smoke.” A few participants had tried or thought about decreasing the number of cigarettes smoked gradually, whereas others tended to decrease suddenly. Participants had limited awareness of available smoking cessation resources other than they “can rely on [themselves] to quit.” However, participants were willing to trust their medical providers. Specifically, a participant pointed out the following: “If I see it [smoking cessation counseling offered at outpatient clinic], I will go check it out and learn more about it.” Among those who understood the potential of e-cigarettes as a smoking cessation aid, almost all considered them “completely useless” and felt that e-cigarettes would worsen their cravings for cigarettes. For the subset of participants who acknowledged the availability of cessation resources, most thought that these methods would not be useful. A few participants also mentioned that candies and snacks were the distraction methods they used to quit smoking.

### **Theme 3: The Duality of Social Factors: Both Discouraging and Encouraging for Quitting**

Most participants reported that they started smoking because of curiosity and social pressure. A participant noted that she started smoking when “[she] went out with friends, and they handed [her] the cigarettes.” Social occasions where they “were given cigarettes” prevented the participants from quitting.

Several participants mentioned that they often smoked with friends; relapse after a quit attempt tended to occur because they saw friends smoking. Smoking with friends was considered a necessary part of social interaction. Regarding societal perceptions, some participants expressed a few negative connotations about smoking in public places because most people do the same. Smoking in cars was also common when passengers were smokers, but participants refrained from smoking with nonsmoking passengers. Participants noted that social stigma about their HIV diagnosis induced stress, and smoking was perceived as needed to alleviate this stigma-induced stress.

By contrast, participants also mentioned social pressure and peer support as facilitators to quitting. This included support from family and friends, supervision, and the benefits of quitting together. A participant mentioned the following: “the encouragement of my family [helped with smoking cessation].” Other social facilitators to quitting included awareness that their smoking behavior harmed their loved ones (such as children) or had a negative influence on their family. In some instances, participants felt that family members blamed them for creating harm through second- and third-hand smoke, and family members would pressure them to quit. Social stigma is believed to play a role in facilitating quitting. There was a sense that “smoking is not accepted by the society,” especially when there are others (such as women) nearby.

### **Theme 4: HIV Diagnosis Raised Awareness of Participants’ Own Health but Also Induced Stress and Anxiety**

When asked about changes after receiving the HIV diagnosis, some participants decided to, or felt the necessity to, quit or reduce the number of cigarettes smoked immediately after the diagnosis or because of other health conditions. Participants expressed fear and sadness when they first learned about their HIV diagnoses. Some participants described the effect of hospitalization owing to severe lung infection as pivotal to quitting (“family members [prohibited him] from smoking”), whereas others described being forced to quit smoking while being hospitalized. Those people mostly “started smoking again after [they were] discharged from the hospital” owing to stress from life or work. Similarly, several participants “couldn’t accept [the HIV diagnosis], so [they] started smoking again.” Having an HIV diagnosis induced societal pressure and stigma toward patients, which was felt to prevent engagement in daily tasks. For instance, a few participants reported having trouble finding a job because some jobs required a health certificate that cannot be obtained with a HIV diagnosis. Therefore, many participants returned to smoking because of reasons such as dependency on nicotine, stress, and social pressure.

The participants were hesitant to use NRT, such as patches or gum, because they were not willing to take additional medications. Ultimately, for some participants, concerns about the side effects of antiretroviral treatment medications overshadowed the potential benefit of using NRT (eg, a participant thought the pain caused by antiretroviral treatment medications would cause more problems than smoking).

## Theme 5: Quit for Life Intervention Modification

Participants were asked about their perceptions of the existing draft of the *Quit for Life* intervention. Almost all participants preferred using WeChat as the primary messenger service platform to receive messages, as it is less likely to be ignored and considered as scam messages (compared with SMS text messaging, which are often scams). Although participants thought WeChat was an acceptable form of communication, one participant thought that messages would act as a constant reminder of the idea that “I am a smoker” and could potentially backfire. A participant who was less familiar with WeChat as a messenger service platform thought that they would still read all the WeChat messages they received. Participants were concerned with using WeChat group (vs direct one-on-one WeChat messaging) because “[WeChat groups] involve personal privacy and no one wants others to know about their condition.” Some participants suggested sending more messages during the first 2 weeks of the intervention, as that period is when cravings occur more often and with higher intensity. For counseling sessions, participants thought that calls originating from a reliable area code in Nanning city would have a higher chance of being picked up. Many participants agreed that scheduling counseling sessions to coincide with picking up HIV medications was desirable. Participants endorsed face-to-face counseling at the HIV clinic and suggested that the counseling sessions be brief (10-20 minutes).

## Exploratory Theme: Demographic Differences in Perceptions of Smoking and the Decision to Quit Smoking

Table 2 presents exploratory themes that were summarized from the interview results. When comparing younger (<30 years) and older (>30 years) participants, responses differed in that older participants tended to quit because of health concerns or social stigma; in contrast, younger participants expressed unwillingness to quit because of low-risk perception despite knowing the harms of smoking. Compared with older participants, younger participants were more likely to smoke when hanging out with friends. We also explored comparisons between male and female smokers. Some female participants considered smoking and NRT to be more harmful to women compared with men, and pregnancy would be an event that would motivate them to quit (a female participant noted HIV by itself would not stop her from smoking). Male participants mentioned that societal perceptions of masculine ideals, familial responsibilities, and being dissuaded from expressing negative emotions act as barriers to quitting. Regarding educational background, risk perception tended to be low among participants with both lower and higher levels of education. However, those with higher levels of education had heard of second-hand and third-hand smoke and could explain these terms more often, whereas those with lower educational levels had more limited knowledge of these terms.

## Discussion

### Principal Findings

This qualitative study is one of the first to explore the salient perspectives of smoking and quitting topics including smoking patterns, facilitators and barriers to cessation, perceptions of cessation resources, and HIV stigma in people with HIV who smoke in China. In general, participants were willing to quit if their health was affected by smoking or were told to quit by their health care providers. This finding is consistent with quantitative studies that found that people with HIV who smoke in China have a higher likelihood of reporting willingness to quit than the general population of smokers without HIV [13,14].

### Comparison With Prior Work

First, similar to previous studies examining smoking cessation strategies among smokers in China, participants identified willpower as a primary quitting strategy in this study [27-29]. A lack of awareness of available NRT combined with a lack of smoking cessation interventions designed specifically for people with HIV who smoke in China are barriers for this population to seek external help [13,28]. At the same time, limited knowledge prevented some participants from understanding and acknowledging the usefulness of the smoking cessation services. Future interventions could promote the availability and usefulness of NRT for people with HIV who smoke.

Second, perceived stress was also a factor across several themes (eg, themes 1, 3, and 4). This finding is similar to previous studies conducted among teenagers and adults in China and Hong Kong, where peer pressure induced stress and need to conform, so that one could feel like they belonged to a social group—factors that could be addressed when designing cessation intervention content [27,30,31]. This barrier is more pronounced when considering China’s collectivistic culture, which focuses on social acceptance and the perception of cigarettes as a necessary social connection tool [32]. Stress was also perceived to be closely linked to the workplace. People in China traditionally view work as an integral part of their life as they work to support their family, even more so than their Western counterparts [33]. Stressful events during work include intense competition between coworkers within the same vocation [34], along with long working hours for many occupations [35]. Higher levels of work stress have been shown to result in a lower quality of life and higher levels of somatic symptoms and burnout [30,35,36]. In addition, people with HIV endure stress from the risk of involuntary serostatus disclosure, health status confidentiality, risk of viral transmission to others, fear of stigma and discrimination, and impairment of physical functioning [37]. For people with HIV in China, the compounded effect of stress from workplace environments and learning of their HIV diagnosis may lead to symptoms of depression and anxiety and prompt risk behaviors such as tobacco use, substance use, and unhealthy alcohol use [8,37-40], with a high likelihood of engaging in multiple risk behaviors [8,38,40]. Thus, future work should address the mental health needs of people with HIV who smoke in China, help this population develop strategies to cope with stressors, and consider multiple risk factor interventions.

Third, regarding the exploratory demographic differences with respect to sex, we found several female participants who perceived NRT as more dangerous for them compared with men. This aligns with previous research among women who were pregnant or planning to become pregnant [41]. Especially for women during pregnancy, taking NRT was perceived to be a burden because of other medications they had to take [41]. Future qualitative studies should consider directly investigating this exploratory finding among a larger sample of women in addition to other demographic differences (age and education level).

In addition, the draft telephone- and messenger-delivered smoking cessation intervention, *Quit for Life*, will be modified based on these qualitative results. For example, doctors and nurses were found to be common, trusted, and motivating sources of information on the importance of cessation and available resources. Therefore, for the *Quit for Life* intervention, interactions between medical providers and participants can enhance this trust-based relationship and may improve effectiveness and adherence to the study protocol [19]. One way to build trust is to have participants attend face-to-face initial counseling sessions in the HIV clinic with a counselor. As another example, concerns about the health of others, social disapproval, and stress were mentioned as motivators to quit, which is aligned with motivators from prior research, where those who quit endorsed health concerns, family disapproval, and being an example for their children [32]. Therefore, WeChat messages and counseling sessions will incorporate content related to health impacts, social influences of smoking, and formulating treatment plans to help the participant resolve barriers. Intervention content will also focus on the harms of smoking among people with HIV, specifically to respond to variable knowledge on this topic. Counseling sessions will follow motivational interviewing techniques.

Similar to previous research on telephone-delivered interventions in China, participants preferred the WeChat messenger service platform rather than other mobile service platforms [12,19,21,28,42]. Messages from the *Quit for Life* intervention will use short informative phrasing, be sent through

a designated study account, and have a higher frequency at the start of the intervention period when cravings are thought to be the highest. Similar to our findings, previous cessation research in China has also reported generally low interest in cessation services and resources [27]. Given its efficacy for cessation [5], the *Quit for Life* intervention will provide widely available NRT gum at no cost for a time-limited duration. A pilot clinical trial will be conducted to test the preliminary efficacy of the *Quit for Life* intervention (WeChat messaging plus behavioral counseling) in addition to NRT plus educational booklets versus NRT plus educational booklets only on smoking cessation at 12 weeks.

### Limitations

Although we obtained in-depth information from our sample, there was an imbalance across demographic categories. For example, the number of female participants was low (reflecting the lower proportion of Chinese females who smoked in general). We did not conduct additional methods such as member checking; however, our multiteam collaboration provided regular input into the analysis, which lends credibility. In addition, more recent surveillance data on national smoking trends among people with HIV in China are needed to promote targeted cessation efforts.

### Conclusions

We examined the perspectives of people with HIV in China about smoking patterns, attitudes toward quitting, and HIV stigma to inform the development of a smoking cessation intervention to be delivered through counseling sessions and brief messenger service-delivered messages, specifically targeting this vulnerable population. In the future, the modified *Quit for Life* intervention will be evaluated in a pilot clinical trial among people with HIV who smoke in China recruited from the People's 4th Hospital of Nanning in Nanning, Guangxi, China. If effective, the *Quit for Life* intervention has the potential to be implemented on a wider scale for the people with HIV who smoke in China, in part using mobile messenger service platforms.

---

### Acknowledgments

The authors would like to acknowledge the funding support from the National Institutes of Health or the National Cancer Institute under grant number R21CA243835. The paper's contents are solely the responsibility of the authors and do not necessarily represent the official views of the National Institutes of Health or the National Cancer Institute. Study data were collected and managed using Research Electronic Data Capture tools hosted at the Boston University, Clinical & Translational Science Institute 1UL1TR001430.

---

### Authors' Contributions

As this is a United States-China collaborative project, joint authorship and joint corresponding authorship is necessary to ensure joint and equal responsibility. LMQ and HL are the joint corresponding authors. SY and JH are the joint first authors. LMQ and SY contributed to the planning, monitoring, and analyzing of the study. HL and JH worked on planning, implementing, and gathering data in Guangxi, China. All authors reviewed the final manuscript.

---

### Conflicts of Interest

None declared.

---

## Multimedia Appendix 1

Qualitative interview guide with open-ended questions.

[DOC File, 87 KB - [formative\\_v6i10e35923\\_app1.doc](#)]

## Multimedia Appendix 2

Frequently noted themes with supporting codes and example quotations of smokers and former smokers with HIV in China.

[DOCX File, 18 KB - [formative\\_v6i10e35923\\_app2.docx](#)]

## References

1. Global Tobacco Surveillance System Data (GTSSData). Centers for Disease Control and Prevention. URL: <https://nccd.cdc.gov/GTSSDataSurveyResources/Ancillary/DataReports.aspx?CAID=1> [accessed 2021-11-30]
2. Luo X, Duan S, Duan Q, Pu Y, Yang Y, Ding Y, et al. Tobacco use among HIV-infected individuals in a rural community in Yunnan Province, China. *Drug Alcohol Depend* 2014 Jan 01;134:144-150. [doi: [10.1016/j.drugalcdep.2013.09.023](#)] [Medline: [24144787](#)]
3. Chen W, Shiu C, Yang JP, Tun MM, Zhang L, Wang K, et al. Tobacco use and HIV symptom severity in Chinese people living with HIV. *AIDS Care* 2020 Feb 22;32(2):217-222 [FREE Full text] [doi: [10.1080/09540121.2019.1620169](#)] [Medline: [31116021](#)]
4. Reynolds NR. Cigarette smoking and HIV: more evidence for action. *AIDS Educ Prev* 2009 Jun;21(3 Suppl):106-121 [FREE Full text] [doi: [10.1521/aeap.2009.21.3\\_supp.106](#)] [Medline: [19537958](#)]
5. Ledgerwood DM, Yskes R. Smoking cessation for people living with HIV/AIDS: a literature review and synthesis. *Nicotine Tob Res* 2016 Dec 31;18(12):2177-2184 [FREE Full text] [doi: [10.1093/ntr/ntw126](#)] [Medline: [27245237](#)]
6. Rao D, Angell B, Lam C, Corrigan P. Stigma in the workplace: employer attitudes about people with HIV in Beijing, Hong Kong, and Chicago. *Soc Sci Med* 2008 Nov;67(10):1541-1549 [FREE Full text] [doi: [10.1016/j.socscimed.2008.07.024](#)] [Medline: [18760869](#)]
7. Zhang C, Li X, Liu Y, Qiao S, Zhang L, Zhou Y, et al. Emotional, physical and financial burdens of stigma against people living with HIV/AIDS in China. *AIDS Care* 2016 Feb 17;28 Suppl 1(sup1):124-131 [FREE Full text] [doi: [10.1080/09540121.2016.1146206](#)] [Medline: [26886415](#)]
8. Zhang C, Li X, Liu Y, Zhou Y, Shen Z, Chen Y. Impacts of HIV stigma on psychosocial well-being and substance use behaviors among people living with HIV/AIDS in China: across the life span. *AIDS Educ Prev* 2018 Apr;30(2):108-119. [doi: [10.1521/aeap.2018.30.2.108](#)]
9. Lasser KE, Lunze K, Cheng DM, Blokhina E, Walley AY, Tindle HA, et al. Depression and smoking characteristics among HIV-positive smokers in Russia: a cross-sectional study. *PLoS One* 2018 Feb 6;13(2):e0189207 [FREE Full text] [doi: [10.1371/journal.pone.0189207](#)] [Medline: [29408935](#)]
10. Vetrova MV, Cheng DM, Bendiks S, Gnatienco N, Lloyd-Travaglini C, Jiang W, et al. HIV and substance use stigma, intersectional stigma and healthcare among HIV-positive PWID in Russia. *AIDS Behav* 2021 Sep 28;25(9):2815-2826. [doi: [10.1007/s10461-021-03172-5](#)] [Medline: [33506305](#)]
11. Grau LE, Pham T, O'Leary T, Weiss J, Toll B, Bernstein SL. Smokers' perspectives on texting for tobacco dependence treatment: a qualitative analysis. *Nicotine Tob Res* 2017 Mar 01;19(3):307-313. [doi: [10.1093/ntr/ntw184](#)] [Medline: [27613935](#)]
12. Luk TT, Wong SW, Lee JJ, Chan SS, Lam TH, Wang MP. Exploring community smokers' perspectives for developing a chat-based smoking cessation intervention delivered through mobile instant messaging: qualitative study. *JMIR Mhealth Uhealth* 2019 Jan 31;7(1):e11954 [FREE Full text] [doi: [10.2196/11954](#)] [Medline: [30702431](#)]
13. Berg CJ, Nehl EJ, Wang X, Ding Y, He N, Wong FY. Utilization of cessation resources among HIV-positive and HIV-negative men who smoke and who have sex with men in Chengdu, China. *Nicotine Tob Res* 2014 Oct 14;16(10):1283-1288 [FREE Full text] [doi: [10.1093/ntr/ntu064](#)] [Medline: [24827789](#)]
14. Wang Y, Chen X, Li X, Wang Y, Shan Q, Zhou Y, et al. Cigarette smoking among Chinese PLWHA: an exploration of changes in smoking after being tested HIV positive. *AIDS Care* 2016 Oct 12;28(3):365-369 [FREE Full text] [doi: [10.1080/09540121.2015.1090536](#)] [Medline: [26457812](#)]
15. Liao Y, Wu Q, Kelly BC, Zhang F, Tang Y, Wang Q, et al. Effectiveness of a text-messaging-based smoking cessation intervention ("Happy Quit") for smoking cessation in China: a randomized controlled trial. *PLoS Med* 2018 Dec 18;15(12):e1002713 [FREE Full text] [doi: [10.1371/journal.pmed.1002713](#)] [Medline: [30562352](#)]
16. Augustson E, Engelgau MM, Zhang S, Cai Y, Cher W, Li R, et al. Text to quit China: an mHealth smoking cessation trial. *Am J Health Promot* 2016 Jan 05;31(3):217-225. [doi: [10.4278/ajhp.140812-quan-399](#)]
17. Lyu K, Zhao J, Wang B, Xiong G, Yang W, Liu Q, et al. Smartphone application WeChat for clinical follow-up of discharged patients with head and neck tumors: a randomized controlled trial. *Chin Med J (Engl)* 2016 Dec 05;129(23):2816-2823 [FREE Full text] [doi: [10.4103/0366-6999.194635](#)] [Medline: [27900995](#)]

18. Mao L, Lu J, Zhang Q, Zhao Y, Chen G, Sun M, et al. Family-based intervention for patients with type 2 diabetes via WeChat in China: protocol for a randomized controlled trial. *BMC Public Health* 2019 Apr 05;19(1):381 [FREE Full text] [doi: [10.1186/s12889-019-6702-8](https://doi.org/10.1186/s12889-019-6702-8)] [Medline: [30953483](https://pubmed.ncbi.nlm.nih.gov/30953483/)]
19. Guo Y, Xu Z, Qiao J, Hong YA, Zhang H, Zeng C, et al. Development and feasibility testing of an mHealth (text message and WeChat) intervention to improve the medication adherence and quality of life of people living with HIV in China: pilot randomized controlled trial. *JMIR Mhealth Uhealth* 2018 Sep 04;6(9):e10274. [doi: [10.2196/10274](https://doi.org/10.2196/10274)]
20. Guo Y, Hong YA, Cai W, Li L, Hao Y, Qiao J, et al. Effect of a WeChat-based intervention (Run4Love) on depressive symptoms among people living with HIV in China: a randomized controlled trial. *J Med Internet Res* 2020 Feb 09;22(2):e16715 [FREE Full text] [doi: [10.2196/16715](https://doi.org/10.2196/16715)] [Medline: [32044751](https://pubmed.ncbi.nlm.nih.gov/32044751/)]
21. Zhao Y, Zhu X, Pérez AE, Zhang W, Shi A, Zhang Z, et al. MHealth approach to promote Oral HIV self-testing among men who have sex with men in China: a qualitative description. *BMC Public Health* 2018 Sep 27;18(1):1146 [FREE Full text] [doi: [10.1186/s12889-018-6046-9](https://doi.org/10.1186/s12889-018-6046-9)] [Medline: [30261856](https://pubmed.ncbi.nlm.nih.gov/30261856/)]
22. Chan SS, Leung DY, Abdullah AS, Lo SS, Yip AW, Kok W, et al. Smoking-cessation and adherence intervention among Chinese patients with erectile dysfunction. *Am J Prev Med* 2010 Sep;39(3):251-258. [doi: [10.1016/j.amepre.2010.05.006](https://doi.org/10.1016/j.amepre.2010.05.006)] [Medline: [20709257](https://pubmed.ncbi.nlm.nih.gov/20709257/)]
23. Guanghua L, Yi C, Shuai T, Zhiyong S, Zhenzhu T, Yuhua R, et al. HIV, syphilis and behavioral risk factors among men who have sex with men in a drug-using area of southwestern China: results of 3 cross-sectional surveys from 2013 to 2015. *Medicine (Baltimore)* 2018 Apr;97(16):e0404 [FREE Full text] [doi: [10.1097/MD.00000000000010404](https://doi.org/10.1097/MD.00000000000010404)] [Medline: [29668597](https://pubmed.ncbi.nlm.nih.gov/29668597/)]
24. Wu Z, Chen J, Scott SR, McGoogan JM. History of the HIV epidemic in China. *Curr HIV/AIDS Rep* 2019 Dec 26;16(6):458-466 [FREE Full text] [doi: [10.1007/s11904-019-00471-4](https://doi.org/10.1007/s11904-019-00471-4)] [Medline: [31773405](https://pubmed.ncbi.nlm.nih.gov/31773405/)]
25. 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](https://doi.org/10.1016/j.jbi.2008.08.010)] [Medline: [18929686](https://pubmed.ncbi.nlm.nih.gov/18929686/)]
26. Bradley EH, Curry LA, Devers KJ. Qualitative data analysis for health services research: developing taxonomy, themes, and theory. *Health Serv Res* 2007 Aug;42(4):1758-1772 [FREE Full text] [doi: [10.1111/j.1475-6773.2006.00684.x](https://doi.org/10.1111/j.1475-6773.2006.00684.x)] [Medline: [17286625](https://pubmed.ncbi.nlm.nih.gov/17286625/)]
27. Abdullah AS, Ho WW. What Chinese adolescents think about quitting smoking: a qualitative study. *Subst Use Misuse* 2006;41(13):1735-1743. [doi: [10.1080/10826080601006433](https://doi.org/10.1080/10826080601006433)] [Medline: [17118813](https://pubmed.ncbi.nlm.nih.gov/17118813/)]
28. Jiang N, Zhang Y, Qian X, Thorpe L, Trinh-Shevrin C, Shelley D. Chinese immigrant smokers' access barriers to tobacco cessation services and experience using social media and text messaging. *Tob Prev Cessat* 2020 Sep 11;6(September):52 [FREE Full text] [doi: [10.18332/tpc/125942](https://doi.org/10.18332/tpc/125942)] [Medline: [33083680](https://pubmed.ncbi.nlm.nih.gov/33083680/)]
29. Yang T, Zhu Z, Barnett R, Zhang W, Jiang S. Tobacco advertising, anti-tobacco information exposure, environmental smoking restrictions, and unassisted smoking cessation among chinese male smokers: a population-based study. *Am J Mens Health* 2019 Jun 11;13(3):1557988319856152 [FREE Full text] [doi: [10.1177/1557988319856152](https://doi.org/10.1177/1557988319856152)] [Medline: [31185783](https://pubmed.ncbi.nlm.nih.gov/31185783/)]
30. Cui X, Rockett IR, Yang T, Cao R. Work stress, life stress, and smoking among rural-urban migrant workers in China. *BMC Public Health* 2012 Nov 14;12(1):979 [FREE Full text] [doi: [10.1186/1471-2458-12-979](https://doi.org/10.1186/1471-2458-12-979)] [Medline: [23151299](https://pubmed.ncbi.nlm.nih.gov/23151299/)]
31. Sheer VC, Mao CM, Chen YR. Focus group findings of smoking onset among male youth in China. *Subst Use Misuse* 2017 Jun 07;52(7):866-874. [doi: [10.1080/10826084.2016.1264972](https://doi.org/10.1080/10826084.2016.1264972)] [Medline: [28426356](https://pubmed.ncbi.nlm.nih.gov/28426356/)]
32. Im PK, McNeill A, Thompson ME, Fong GT, Xu S, Quah AC, et al. Individual and interpersonal triggers to quit smoking in China: a cross-sectional analysis. *Tob Control* 2015 Nov 17;24 Suppl 4(Supplement 4):iv40-iv47 [FREE Full text] [doi: [10.1136/tobaccocontrol-2014-052198](https://doi.org/10.1136/tobaccocontrol-2014-052198)] [Medline: [25888422](https://pubmed.ncbi.nlm.nih.gov/25888422/)]
33. Spector PE, Cooper CL, Poelmans S, Allen TD, O'Driscoll M, Sanchez JI, et al. A cross-national comparative study of work-family stressors, working hours, and well-being: China and Latin America versus the Anglo world. *Pers Psychol* 2004 Mar;57(1):119-142. [doi: [10.1111/j.1744-6570.2004.tb02486.x](https://doi.org/10.1111/j.1744-6570.2004.tb02486.x)]
34. Zhang L, Fu J, Yao B, Zhang Y. Correlations among work stressors, work stress responses, and subjective well-being of civil servants: empirical evidence from China. *Iran J Public Health* 2019 Jun;48(6):1059-1067 [FREE Full text] [Medline: [31341847](https://pubmed.ncbi.nlm.nih.gov/31341847/)]
35. Wang Y, Ramos A, Wu H, Liu L, Yang X, Wang J, et al. Relationship between occupational stress and burnout among Chinese teachers: a cross-sectional survey in Liaoning, China. *Int Arch Occup Environ Health* 2015 Jul 26;88(5):589-597. [doi: [10.1007/s00420-014-0987-9](https://doi.org/10.1007/s00420-014-0987-9)] [Medline: [25256806](https://pubmed.ncbi.nlm.nih.gov/25256806/)]
36. Li J, Ding H, Han W, Jin L, Kong L, Mao K, et al. The association of work stress with somatic symptoms in Chinese working women: a large cross-sectional survey. *J Psychosom Res* 2016 Oct;89:7-10. [doi: [10.1016/j.jpsychores.2016.08.001](https://doi.org/10.1016/j.jpsychores.2016.08.001)] [Medline: [27663103](https://pubmed.ncbi.nlm.nih.gov/27663103/)]
37. Huang Y, Luo D, Chen X, Zhang D, Huang Z, Xiao S. HIV-related stress experienced by newly diagnosed people living with HIV in China: a 1-year longitudinal study. *Int J Environ Res Public Health* 2020 Apr 14;17(8):2681 [FREE Full text] [doi: [10.3390/ijerph17082681](https://doi.org/10.3390/ijerph17082681)] [Medline: [32295107](https://pubmed.ncbi.nlm.nih.gov/32295107/)]
38. Berg CJ, Nehl EJ, Wong FY, He N, Huang ZJ, Ahluwalia JS, et al. Prevalence and correlates of tobacco use among a sample of MSM in Shanghai, China. *Nicotine Tob Res* 2011 Jan 08;13(1):22-28 [FREE Full text] [doi: [10.1093/ntr/ntq193](https://doi.org/10.1093/ntr/ntq193)] [Medline: [21059821](https://pubmed.ncbi.nlm.nih.gov/21059821/)]



39. Niu L, Luo D, Liu Y, Silenzio VM, Xiao S. The mental health of people living with HIV in China, 1998-2014: a systematic review. *PLoS One* 2016 Apr 15;11(4):e0153489 [FREE Full text] [doi: [10.1371/journal.pone.0153489](https://doi.org/10.1371/journal.pone.0153489)] [Medline: [27082749](https://pubmed.ncbi.nlm.nih.gov/27082749/)]
40. Liu Y, Ruan Y, Strauss SM, Yin L, Liu H, Amico KR, et al. Alcohol misuse, risky sexual behaviors, and HIV or syphilis infections among Chinese men who have sex with men. *Drug Alcohol Depend* 2016 Nov 01;168:239-246 [FREE Full text] [doi: [10.1016/j.drugalcdep.2016.09.020](https://doi.org/10.1016/j.drugalcdep.2016.09.020)] [Medline: [27723554](https://pubmed.ncbi.nlm.nih.gov/27723554/)]
41. England LJ, Tong VT, Koblitz A, Kish-Doto J, Lynch MM, Southwell BG. Perceptions of emerging tobacco products and nicotine replacement therapy among pregnant women and women planning a pregnancy. *Prev Med Rep* 2016 Dec;4:481-485 [FREE Full text] [doi: [10.1016/j.pmedr.2016.09.002](https://doi.org/10.1016/j.pmedr.2016.09.002)] [Medline: [27635381](https://pubmed.ncbi.nlm.nih.gov/27635381/)]
42. Wang J, Abdullah AS, Ma Z, Fu H, Huang K, Yu H, et al. Building capacity for information and communication technology use in global health research and training in China: a qualitative study among Chinese health sciences faculty members. *Health Res Policy Syst* 2017 Jun 28;15(1):59 [FREE Full text] [doi: [10.1186/s12961-017-0222-8](https://doi.org/10.1186/s12961-017-0222-8)] [Medline: [28659155](https://pubmed.ncbi.nlm.nih.gov/28659155/)]

## Abbreviations

**NRT:** nicotine replacement therapy

*Edited by A Mavragani; submitted 06.04.22; peer-reviewed by C Zhang, S Wen; comments to author 09.06.22; revised version received 14.07.22; accepted 15.08.22; published 12.10.22.*

*Please cite as:*

*Yang S, Huang J, Ye L, Lin J, Xie Z, Guo B, Li Y, Liang B, Zheng Z, Lunze K, Abdullah AS, Liang H, Quintiliani LM  
Factors Related to Smoking and Perceptions of a Behavioral Counseling and Messenger Service-Delivered Smoking Cessation  
Intervention for People With HIV in China: Qualitative Study  
JMIR Form Res 2022;6(10):e35923  
URL: <https://formative.jmir.org/2022/10/e35923>  
doi: [10.2196/35923](https://doi.org/10.2196/35923)  
PMID: [36222795](https://pubmed.ncbi.nlm.nih.gov/36222795/)*

©Shanyin Yang, Jiegang Huang, Li Ye, Jianyan Lin, Zhiman Xie, Baodong Guo, Yanjun Li, Bingyu Liang, Zhigang Zheng, Karsten Lunze, Abu S Abdullah, Hao Liang, Lisa M Quintiliani. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 12.10.2022. 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 Triage Capability of Laypersons: Retrospective Exploratory Analysis

Marvin Kopka<sup>1,2</sup>, MSc; Markus A Feufel<sup>3</sup>, Dipl-Ing (FH), MSc, PhD; Felix Balzer<sup>1</sup>, MSc, MD, PhD; Malte L Schmieding<sup>1</sup>, MD

<sup>1</sup>Institute of Medical Informatics, Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin, Germany

<sup>2</sup>Cognitive Psychology and Ergonomics, Department of Psychology and Ergonomics (IPA), Technische Universität Berlin, Berlin, Germany

<sup>3</sup>Division of Ergonomics, Department of Psychology and Ergonomics (IPA), Technische Universität Berlin, Berlin, Germany

**Corresponding Author:**

Marvin Kopka, MSc

Institute of Medical Informatics

Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin

Charitéplatz 1

Berlin, 10117

Germany

Phone: 49 30 450 581 052

Email: [marvin.kopka@charite.de](mailto:marvin.kopka@charite.de)

## Abstract

**Background:** Although medical decision-making may be thought of as a task involving health professionals, many decisions, including critical health-related decisions are made by laypersons alone. Specifically, as the first step to most care episodes, it is the patient who determines whether and where to seek health care (triage). Overcautious self-assessments (ie, overtriaging) may lead to overutilization of health care facilities and overcrowded emergency departments, whereas imprudent decisions (ie, undertriaging) constitute a risk to the patient's health. Recently, patient-facing decision support systems, commonly known as symptom checkers, have been developed to assist laypersons in these decisions.

**Objective:** The purpose of this study is to identify factors influencing laypersons' ability to self-triage and their risk averseness in self-triage decisions.

**Methods:** We analyzed publicly available data on 91 laypersons appraising 45 short fictitious patient descriptions (case vignettes; N=4095 appraisals). Using signal detection theory and descriptive and inferential statistics, we explored whether the type of medical decision laypersons face, their confidence in their decision, and sociodemographic factors influence their triage accuracy and the type of errors they make. We distinguished between 2 decisions: whether emergency care was required (decision 1) and whether self-care was sufficient (decision 2).

**Results:** The accuracy of detecting emergencies (decision 1) was higher (mean 82.2%, SD 5.9%) than that of deciding whether any type of medical care is required (decision 2, mean 75.9%, SD 5.25%;  $t_{>90}=8.4$ ;  $P<.001$ ; Cohen  $d=0.9$ ). Sensitivity for decision 1 was lower (mean 67.5%, SD 16.4%) than its specificity (mean 89.6%, SD 8.6%) whereas sensitivity for decision 2 was higher (mean 90.5%, SD 8.3%) than its specificity (mean 46.7%, SD 15.95%). Female participants were more risk averse and overtriaged more often than male participants, but age and level of education showed no association with participants' risk averseness. Participants' triage accuracy was higher when they were certain about their appraisal (2114/3381, 62.5%) than when being uncertain (378/714, 52.9%). However, most errors occurred when participants were certain of their decision (1267/1603, 79%). Participants were more commonly certain of their overtriage errors (mean 80.9%, SD 23.8%) than their undertriage errors (mean 72.5%, SD 30.9%;  $t_{>89}=3.7$ ;  $P<.001$ ;  $d=0.39$ ).

**Conclusions:** Our study suggests that laypersons are overcautious in deciding whether they require medical care at all, but they miss identifying a considerable portion of emergencies. Our results further indicate that women are more risk averse than men in both types of decisions. Layperson participants made most triage errors when they were certain of their own appraisal. Thus, they might not follow or even seek advice (eg, from symptom checkers) in most instances where advice would be useful.

(JMIR Form Res 2022;6(10):e38977) doi:[10.2196/38977](https://doi.org/10.2196/38977)

**KEYWORDS**

digital health; triage; self-triage; urgency assessment; patient-centered care; care navigation; decision support; symptom checker; care; support; medical; health professional; patient; self-assessment; decision; accuracy; error; sensitivity; emergency; female; male

## Introduction

### Background

Increased emergency department (ED) crowding and longer waiting times are associated with higher mortality [1,2], increased adverse events [3], and worse patient outcomes in general [4]. To mitigate these problems, each patient's urgency is assessed upon arrival in the ED—a process called triage. Usually, triage decisions are made by nurses and trained specialists, whose workload is generally high [5]. Increased workload and overcrowding pose huge risks as triage errors currently occur in roughly 16% of all cases [6] and nurses often rely on their intuition and speed up triage by not collecting further information, for example, medical history and physiological data [7,8]. These risks and the workload for health care workers could be reduced if patients were guided to appropriate health care facilities before visiting an ED to receive an initial assessment.

Getting patients to visit the health care facility most suitable for their symptoms can, for example, be achieved by offering phone triage, a telephone hotline patients can call to get a remote urgency assessment of their symptoms. For example, Roivainen et al [9] found that giving medical advice was sufficient for one-third of callers with nonemergent cases, and ED workload subsequently decreased by 36%. Another study by Midtbø et al [10] found that ED attendance decreased from 68.7% to 23.4% when advertising telephone triage hotlines. However, the accuracy of triage hotlines seems to be only around 71% with undertriaging occurring in 12% of all cases [11]—which could be potentially dangerous for some patients. Furthermore, a study based on commercial claims data from 2011 to 2013 concludes that direct-to-consumer telehealth may increase health care utilization and costs by making access to health care more convenient. Thus, such services may shift and increase, rather than reduce, the demand for health care services [12].

Another solution to disburden health care services lies in empowering patients to self-assess their medical complaints and thereby improving their ability to adequately decide which type of health care facility to visit (ie, self-triage), or where appropriate to stay at home and care for themselves. Since these decisions are made by laypersons instead of medical professionals, they come with various challenges. For example, previous studies indicate that laypersons tend to overtriage [13] and that women rate symptoms as more urgent than men [14] and are thus potentially even more inclined toward overtriage. Moreover, an Australian study showed that although laypersons are risk averse when making triage decisions, they cannot reliably detect emergencies either [15].

To address these deficits, decision support systems are designed to aid laypersons in their self-triage decision-making process, for example, kiosks in the ED [13] or symptom checkers [16]. When using them, around 25% of all patients seem to have

reduced perceived urgency of their complaints, and in an experimental study, most participants also followed the advice received by the symptom checkers [17,18]. However, the accuracy of these systems varies widely [19-22]. Implementing such decision aids raises several new questions: can laypersons adequately judge when to seek such decision support? Would all laypersons profit the same way from such decision support tools? Which are the most challenging decisions to laypersons, that is, where is advice from a decision support tool most needed?

### Objective

To provide a foundation for tackling these questions, in this study we investigate whether laypersons' triage accuracy and their risk averseness differ by sociodemographic variables and whether laypersons can potentially gauge whether they require advice in their decision-making, by exploring if laypersons' confidence in their triage appraisals functions as a reliable predictor of accuracy.

## Methods

### Ethical Considerations

No ethics approval was required for this study. Approval of the original study [23] was granted by the Ethics Committee of the Department of Psychology and Ergonomics (IPA) at Technische Universität Berlin (Tracking number: FEU\_03\_20180615). Participants volunteered to participate in the survey, and informed consent was obtained.

### Data Collection

This analysis builds on data collected in a previous study by Schmieding et al [23], which was made available in a public open data repository [24]. They compared laypersons' triage capabilities with symptom checker performance, which are tools developed to provide clinical decision support to laypersons. Their study found that laypersons' overall triage accuracy was mediocre (mean 60.9%, SD 6.8%) based on a set of 45 fictitious patient descriptions (case vignettes). These 45 case vignettes were originally compiled by Semigran et al [21], whose study reported a mediocre overall triage accuracy for a sample of 15 symptom checkers, which was similar to laypersons' accuracy in Schmieding et al's study [23].

The layperson sample consisted of 91 US residents without prior professional medical training. They were recruited from Amazon Mechanical Turk (MTurk) in March 2020. Participants were paid US \$4 for completing the web-based survey and assessing the 45 case vignettes unaided. As an incentive, a bonus of US \$3 was rewarded if they achieved an accuracy above 58%. Compared with the US general population, the layperson sample had a higher level of education (all participants had at least a high school degree) and included a higher proportion of men (55/91, 60.4%) than women (36/91, 39.5%).

A more detailed description of the participants' characteristics is provided by Schmieding et al [23]. Here, we use their data set to further explore individual differences influencing laypersons' triage assessment and decisions.

### Web-Based Survey

Schmieding et al [23] developed a web-based survey in which participants were asked to rate the urgency of presented case vignettes. They adapted 45 case vignettes from Semigran et al [21], comprising 15 cases for each of the 3 triage levels (self-care, nonemergency, or emergency care). The vignettes, as chosen by Semigran et al [21], included both common and uncommon chief complaints from a wide range of diagnoses and were collected from various clinical sources, including teaching materials for health care professionals.

After every triage assessment, participants were asked how certain they were in their assessment on a 4-point Likert-scale ("Very uncertain," "Rather uncertain," "Rather certain," and "Very certain"). Three sociodemographic variables were surveyed (gender, age, and level of education) and rated on a 5-tiered ordinal scale ("Non-high school graduate," "High school graduate," "Some college," "Bachelor's degree," and "Graduate degree").

### Data Analysis

We conducted analyses and generated the images using base R 4.0.5 (R Core Team) [25] and the packages ggplot2 [26], RColorBrewer [27], and tidyverse packages [28].

We dichotomized triage levels to explore whether laypersons could reliably distinguish whether emergency care was required or not (decision 1), and whether self-care was sufficient or not (decision 2). Whether and where health care should be sought are the 2 common questions symptom checkers are approached with [19,29]. For decision 1, we grouped self-care cases with nonemergency cases to assess whether participants were able to correctly detect when emergency care is necessary. For decision 2, we combined emergency and nonemergency cases to the category "health care" to verify whether participants were able to correctly assess when seeing a health care professional rather than staying at home is appropriate. For each of these binary decisions, we calculated means and standard deviations for common metrics of signal detection theory (accuracy, sensitivity, specificity, and negative and positive predictive values). When determining negative and positive predictive values (NPVs and PPVs), we used the occurrence rate within the sample of case vignettes as prevalence.

To explore which factors influence laypersons' risk averseness, triage accuracy, and confidence in their own triage appraisal,

we used linear models to quantify relationships between continuous variables (age and risk averseness) and compared proportions between subgroups for ordinal and categorical variables (gender, education, and certainty).

Overtriage errors were defined as appraising the case's urgency as more urgent than necessary (eg, the participant suggested emergency care when nonemergency care was appropriate) and undertriage errors as judging it as less urgent than required. Risk averseness was defined as the proportion of overtriage errors compared to all errors, whereas a participant's triage accuracy was calculated as the proportion of correctly solved cases related to all cases.

## Results

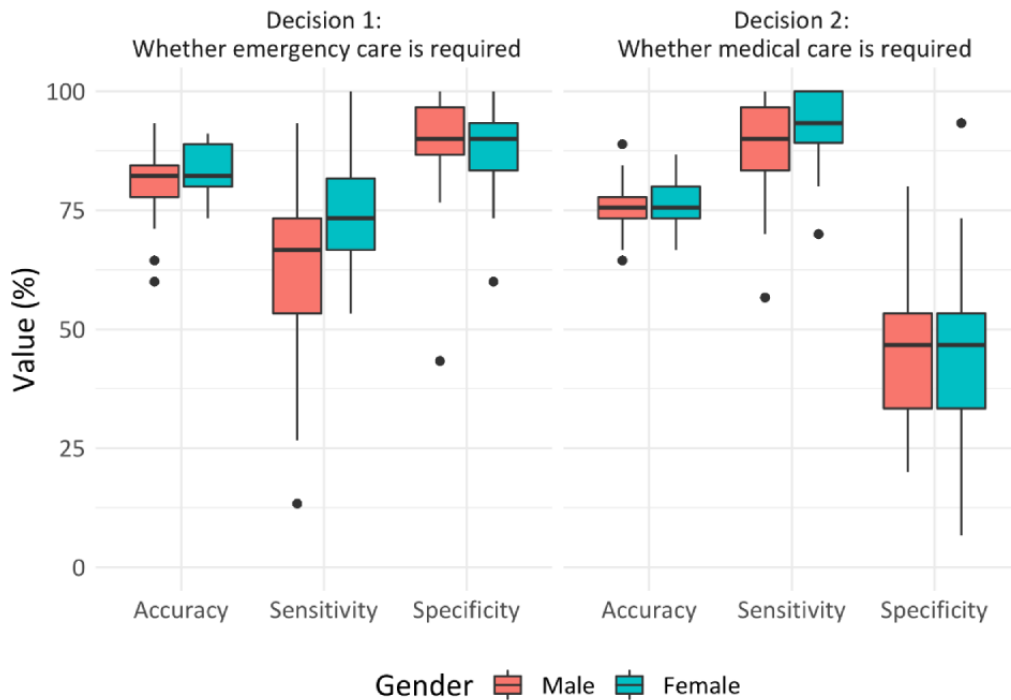
### Laypersons' Triage Capability in Binary Triage Levels

On average, the participants were able to correctly classify whether a fictitious patient required emergency care or not in about 36 out of 45 cases (mean 82.2%, SD 5.88%; see decision 1 in Figure 1). The majority of participants (85/91, 93.4%) achieved an accuracy of 75% or higher. Sensitivity for detecting emergencies (mean 67.5%, SD 16.4%) was lower than the corresponding specificity (mean 89.6%, SD 8.6%), that is, the rate of assessments where cases not requiring emergency care were correctly classified as such. The PPV for detecting emergencies (mean 78.5%, SD 11.8%) was lower than the NPV (mean 85.3%, SD 6.0%).

Concerning decision 2 (Figure 1) on whether or not professional medical care is required, the overall accuracy was lower on average (mean 75.9%, SD 5.25%). Two-thirds of participants (58/91, 63.7%) achieved an accuracy of 75% or greater concerning decision 2. Here the observed sensitivity (mean 90.5%, SD 8.3%) was higher than the corresponding specificity (mean 46.7%, SD 15.9%). The PPV for this decision (mean 77.6%, SD 4.7%) was similar to the NPV (mean 76.1%, SD 16.5%).

The difference in accuracy between decision 1 and decision 2 was found to be statistically significant in a post-hoc 2-sided  $t$  test ( $t_{>90}=8.44$ ;  $P<.001$ ) with a large effect size (Cohen  $d=0.88$ ). The described patterns are similar when values for accuracy, sensitivity, and specificity are broken down by gender (Figure 1). However, male participants missed a case that required medical care 1.4 times more often than female participants (false negative rates for decision 2 of 10.8%, 178/1650, for men and 7.6%, 82/180, for women), and 1.5 times more often that a case required emergency care (false negative rates for decision 1 of 37.8%, 312/825, for men and 24.4%, 132/540, for women).

**Figure 1.** Accuracy, sensitivity, and specificity for 2 binary triage decisions by participants' gender.



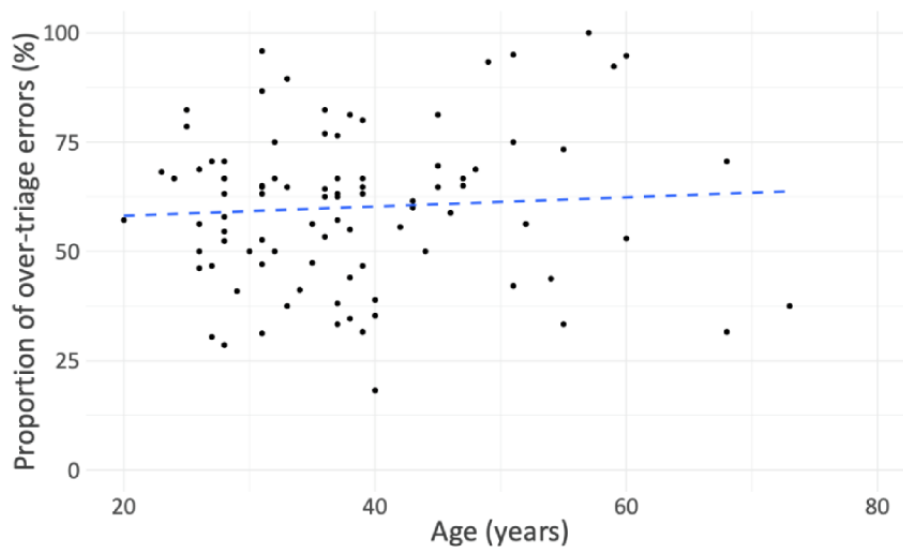
**Demographic Variables and Laypersons' Risk Averseness**

Participants were generally risk averse [23]. Most of their errors were overtriage errors. Only a small portion of participants (24/91, 27%) made more undertriage errors than overtriage errors.

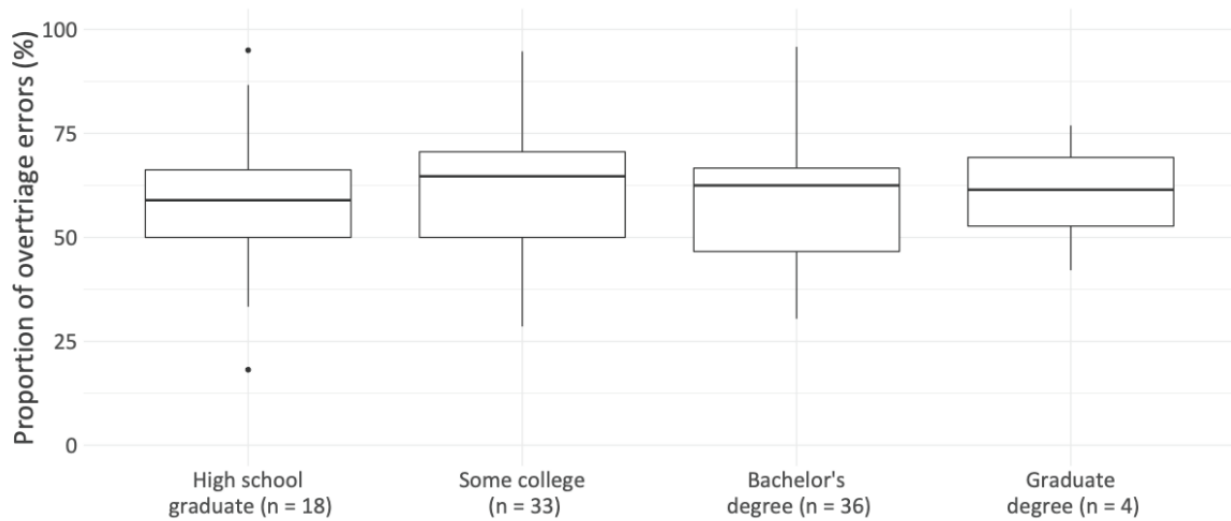
Age only explained little variance in the decisions made ( $R^2=0.004$ ; Figure 2), and median risk averseness was similar for each education level (Figure 3). Risk averseness varied with

gender; that is, female participants were more risk averse than their male counterparts (Figure 4): for male participants, the ratio of overtriage to undertriage error was 1.2:1 (549:452 vignette evaluations), in contrast to 2:1 for female participants (407:195 vignette evaluations). Subsequently, the average female participant's proportion of overtriage errors among all errors was higher (mean 65.9%, SD 17.3%) than the respective proportion for male participants (mean 55.6%, SD 16.0%). This difference was found to be statistically significant in a post-hoc 2-sided Welch  $t$  test ( $t_{>68.6}=2.85$ ;  $P=0.006$ ) with a medium effect size ( $d=0.62$ ).

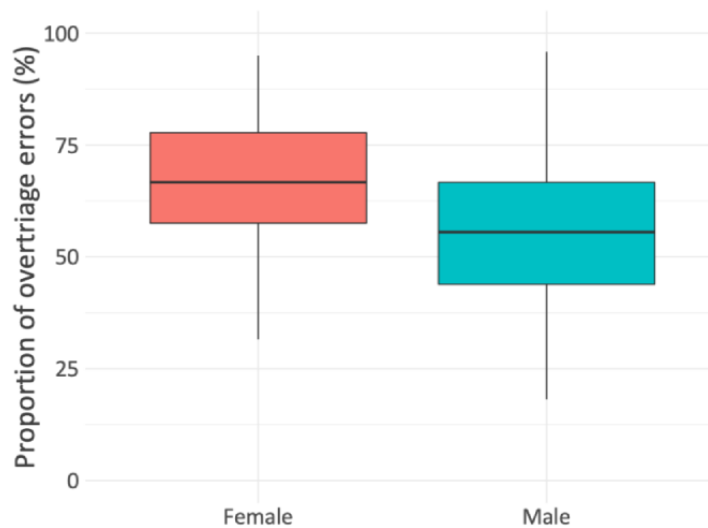
**Figure 2.** Risk averseness by age.



**Figure 3.** Risk averseness by education.



**Figure 4.** Risk averseness by gender.



**Decisional Certainty and Laypersons' Triage Behavior**

**Overall Decisional Certainty**

Participants were certain in most of their triage appraisals. They indicated being “very certain” or “rather certain” in about 4 out

of 5 (3381/4095, 82.6%) of the triage assessments, see [Table 1](#) for details. Only a small portion of participants (33/91, 36.2%) indicated having been “very uncertain” in one or more of their triage decisions.

**Table 1.** Triage assessment and accuracy by certainty.

Triage assessment	Degree of certainty				Total
	“Very uncertain”	“Rather uncertain”	“Rather certain”	“Very certain”	
Correct	24	354	1145	969	2492
Incorrect	34	302	886	381	1603
Total	58	656	2031	1350	4095
Accuracy (%)	41.4	53.9	56.3	71.8	60.8

**Decisional Certainty and Triage Accuracy**

Participants’ triage accuracy varied with their degree of certainty, see [Table 1](#): it was the highest when they indicated being “very certain” (969/1350, 71.8%) and the lowest when they indicated being “very uncertain” (24/58, 41.4%). For

decisions where participants were uncertain (pooling “rather uncertain” and “very uncertain”), the accuracy of their triage decision was lower (378/714, 52.9%) than when being rather or very certain (2114/3381, 62.5%; see [Table 2](#)).

When mistaken, participants were on average still commonly certain about their assessment, though less certain than when their appraisal was correct (see [Tables 1 and 2](#)). When being correct, they were certain in 84.9% (2114/2492) of the correct

assessments, whereas they were certain in 79.0% (1267/1603) of all incorrect assessments. This difference is statistically significant in a post-hoc paired-sample *t* test ( $t_{>90}=5.43$ ;  $P<.001$ ;  $d=0.57$ ).

**Table 2.** Triage assessment and accuracy by dichotomized certainty levels.

Degree of certainty	Triage assessment		Total	Accuracy (%)
	Correct	Incorrect		
Uncertain	378	336	714	52.9
Certain	2114	1267	3381	62.5
Total	3381	1603	4095	60.8
Proportion certain (%)	87	79	N/A <sup>a</sup>	N/A

<sup>a</sup>N/A: not applicable.

### Decisional Certainty and Type of Error

Regarding the 2 types of errors, that is, overtriaging and undertriaging, we observed that participants' decisional certainty was higher when overtriaging (proportion of overtriage errors where they were either rather or very certain, 783/956, 81.9%) than when undertriaging (484/647, 74.8%). The ratio of overtriage to undertriage errors increased with the level of

certainty, from 1.12:1 (18:16) for appraisals where participants were very uncertain to 2:1 (254:127) for appraisals where participants were very certain, see [Table 3](#). The average proportion of participants being certain while overtriaging was higher (mean 80.9%, SD 23.8%) than that while undertriaging (mean 72.5%, SD 30.9%). This difference was statistically significant in a post-hoc paired-sample *t* test ( $t_{>89}=3.70$ ;  $P<.001$ ) with a small effect size ( $d=0.39$ ).

**Table 3.** Triage errors by certainty.

Error type	Degree of certainty, n (%)				Total
	"Very uncertain"	"Rather uncertain"	"Rather certain"	"Very certain"	
Overtriage	18 (1.9)	155 (16.2)	529 (55.3)	254 (26.6)	956
Undertriage	16 (2.5)	147 (22.7)	357 (55.2)	127 (19.6)	647
Total	34 (2.1)	302 (18.8)	886 (55.3)	381 (23.8)	1603

## Discussion

### Principal Findings

As previously reported [15,23], laypersons' ability to triage is systematically biased toward overtriage, but they tend to miss emergency cases. However, when analyzing actionable metrics of triage ability, we see a majority of laypersons being fairly competent (accuracy  $\geq 75\%$ ) in deciding whether they should seek health care (decision 2) and whether emergency care is required (decision 1). The analysis of these binary decisions also revealed a more differentiated insight into laypersons' risk averseness: they seem to be risk averse regarding decision 2 and thus tend to seek care unnecessarily, but not regarding decision 1, which suggests that they have problems identifying emergencies.

Concerning decision 1, participants were more prone to undertriage (ie, not identifying emergencies) and made only a few overtriage errors. In decision 2, they mostly made overtriage errors (ie, not recognizing that self-care is sufficient) and only a few undertriage errors. This adds further evidence to previous findings that laypersons tend to overtriage [13], but it extends them and shows that this is true only when deciding between self-care and need for a health care professional. Indeed, laypersons were more likely to undertriage when looking at

emergency cases. This supports results from a study by Mills et al [15], who found that laypersons often do not recognize emergencies. It also suggests that even when assuming symptom checkers will become highly reliable and patients would use them to improve their decision-making, the benefit of these decision aids might help to disburden health care facilities of low-acuity care but not emergency care; the number of patients (rightly) presenting to the ED might increase with decision aids.

Our results indicate that women are more risk averse and overtriage more than men. This finding is in line with a previous study that found women to rate their symptoms as more urgent [14]. The two other demographic variables we examined (level of education and age) were not associated with accuracy and risk averseness.

As current symptom checkers used to assist laypersons in their triage decisions are rather risk averse [21,30], they would therefore be of greater benefit to men, who made more unsafe decisions. However, women appear to be more regular users of symptom checkers [29].

Participants' judgement of their decisional certainty predicted to some extent whether their stand-alone triage assessment was correct: when uncertain, participants were more likely to make incorrect assessments than when being certain. At first glance, this suggests that perceived uncertainty is a good prompt for

when to seek decision support. However, whether correct or incorrect, our study participants were certain about most of their judgements. Thus, perceived certainty is not a reliable predictor of triage errors, and based on their level of certainty, it seems likely that laypersons may not be able to correctly determine when they would benefit from a decision aid. Especially as participants' certainty was greater when overtriaging than when undertriaging, even perfect decision aids might not effectively reduce unnecessary doctor's visits and disburden health care facilities. This contrasts with the suggestions by Winn et al [17]. Their study found that users' perceived urgency to seek health care commonly decreased after encountering a symptom checker. Their study, however, does not consider whether the provided advice was correct. Therefore, it remains an open question, whether the use of symptom checkers (even when assuming perfect accuracy) can contribute to disburdening health care facilities.

### Limitations

Limitations regarding participants and the evaluation of symptom checker accuracy are reported in detail by Schmieding et al [23] and Semigran et al [21]. Here we report the most important limitations again. Among the main limitations are that the sample of case vignettes is neither exhaustive (eg, mental health issues were excluded) nor proportionate to the incidence of diseases or medical complaints in a real-world setting. Thus, in particular, the reported values for NPVs and PPVs are not to be taken at face value, because they only reflect the prevalence in the sample of vignettes.

Second, although it has been reported that case vignettes are a valid method to assess the health care decision-making of physicians [31-33], the external validity of case vignette-based approaches with layperson decision makers has not been explored yet; that is, laypersons may decide very differently when assessing clinical vignettes compared to when they assess their own or someone else's medical complaints in the real world [21,23].

Beyond that, there are further limitations specific to the analyses in this study. Our results may have limited external validity: first, the sample was not representative of the US general population; second, triage appraisals might be influenced by the context of the health care system and by recruiting

participants online; and third, population groups with no or low (information) technology affinity are not represented. Thus, future studies with more representative panels are required to determine whether our findings hold true for the broader population and whether factors other than gender and perceived certainty influence laypersons' health care decisions (eg, eHealth literacy [34], health anxiety or hypochondria [35], or propensity to trust [36]). In turn, this knowledge may help to specify for which decision and for whom decision aids would provide a benefit [34-36].

We applied statistical significance testing sparingly, since this study is a retrospective exploratory analysis and was primarily intended to generate hypotheses that can be tested (experimentally) in future studies to draw inferences.

Methodologically, we considered certainty as a measure of whether participants would consult decision aids and whether they would be open to incorporating the recommendations into their decision-making. This implies that perceived certainty in one's own appraisal correlates inversely with the openness to follow contradicting advice. However, this relationship is not based on any data. Future studies need to test this assumption.

### Conclusions

Our study suggests that laypersons are overcautious in deciding whether they require medical care (decision 2). At the same time, they miss identifying a considerable proportion of emergencies (decision 1). Our results also indicate that women are more risk averse than men in both these decisions. When providing correct advice, decision aids such as symptom checkers could be of benefit to users as they could help reduce the number of missed emergencies and unnecessary visits to low-acuity care facilities. Thus, from a health system's perspective, decision aids might disburden health care facilities more of low-acuity care than of emergency care. However, layperson participants made most triage errors, and especially overtriage errors, when being certain of their own appraisal. Thus, they might not follow or even seek such advice in most instances where advice would be useful. More studies are needed to better understand laypersons' ability to self-triage, how this ability could be improved, how decision aids may support laypersons' medical decision-making, and when laypersons are willing to take advice from decision aids.

### Conflicts of Interest

FB reports grants from German Federal Ministry of Education and Research, grants from German Federal Ministry of Health, grants from Berlin Institute of Health, personal fees from Elsevier Publishing, grants from Hans Böckler Foundation, other from Robert Koch Institute, grants from Einstein Foundation, grants from Berlin University Alliance, personal fees from Medtronic, and personal fees from GE Healthcare outside the submitted work.

### References

1. Lin S, Ge S, He W, Zeng M. Association of delayed time in the emergency department with the clinical outcomes for critically ill patients. *QJM* 2021 Aug 29;114(5):311-317. [doi: [10.1093/qjmed/hcaa192](https://doi.org/10.1093/qjmed/hcaa192)] [Medline: [32516375](https://pubmed.ncbi.nlm.nih.gov/32516375/)]
2. Woodworth L. Swamped: emergency department crowding and patient mortality. *J Health Econ* 2020 Mar;70:102279. [doi: [10.1016/j.jhealeco.2019.102279](https://doi.org/10.1016/j.jhealeco.2019.102279)] [Medline: [32062054](https://pubmed.ncbi.nlm.nih.gov/32062054/)]
3. Weissman JS, Rothschild JM, Bendavid E, Sprivulis P, Cook EF, Evans RS, et al. Hospital workload and adverse events. *Med Care* 2007 May;45(5):448-455. [doi: [10.1097/01.mlr.0000257231.86368.09](https://doi.org/10.1097/01.mlr.0000257231.86368.09)] [Medline: [17446831](https://pubmed.ncbi.nlm.nih.gov/17446831/)]



4. Sun BC, Hsia RY, Weiss RE, Zingmond D, Liang L, Han W, et al. Effect of emergency department crowding on outcomes of admitted patients. *Ann Emerg Med* 2013 Jun;61(6):605-611.e6 [FREE Full text] [doi: [10.1016/j.annemergmed.2012.10.026](https://doi.org/10.1016/j.annemergmed.2012.10.026)] [Medline: [23218508](https://pubmed.ncbi.nlm.nih.gov/23218508/)]
5. Innes K, Jackson D, Plummer V, Elliott D. A profile of the waiting room nurse in emergency departments: an online survey of Australian nurses exploring implementation and perceptions. *Int Emerg Nurs* 2019 Mar;43:67-73. [doi: [10.1016/j.ienj.2018.10.003](https://doi.org/10.1016/j.ienj.2018.10.003)] [Medline: [30381142](https://pubmed.ncbi.nlm.nih.gov/30381142/)]
6. Ausserhofer D, Zaboli A, Pfeifer N, Solazzo P, Magnarelli G, Marsoner T, et al. Errors in nurse-led triage: an observational study. *Int J Nurs Stud* 2021 Jan;113:103788. [doi: [10.1016/j.ijnurstu.2020.103788](https://doi.org/10.1016/j.ijnurstu.2020.103788)] [Medline: [33120136](https://pubmed.ncbi.nlm.nih.gov/33120136/)]
7. Wolf L. Does your staff really "get" initial patient assessment? Assessing competency in triage using simulated patient encounters. *J Emerg Nurs* 2010 Jul;36(4):370-374. [doi: [10.1016/j.jen.2010.04.016](https://doi.org/10.1016/j.jen.2010.04.016)] [Medline: [20624578](https://pubmed.ncbi.nlm.nih.gov/20624578/)]
8. Wolf LA, Delao AM, Perhats C, Moon MD, Zavotsky KE. Triageing the emergency department, not the patient: United States emergency nurses' experience of the triage process. *J Emerg Nurs* 2018 May;44(3):258-266. [doi: [10.1016/j.jen.2017.06.010](https://doi.org/10.1016/j.jen.2017.06.010)] [Medline: [28750891](https://pubmed.ncbi.nlm.nih.gov/28750891/)]
9. Roivainen P, Hoikka MJ, Raatinieniemi L, Silfvast T, Ala-Kokko T, Kääriäinen M. Telephone triage performed by nurses reduces non-urgent ambulance missions: a prospective observational pilot study in Finland. *Acta Anaesthesiol Scand* 2020 Apr 22;64(4):556-563. [doi: [10.1111/aas.13542](https://doi.org/10.1111/aas.13542)] [Medline: [31898315](https://pubmed.ncbi.nlm.nih.gov/31898315/)]
10. Midtbø V, Fotland SS, Johansen IH, Hunskaar S. From direct attendance to telephone triage in an emergency primary healthcare service: an observational study. *BMJ Open* 2022 May 02;12(5):e054046 [FREE Full text] [doi: [10.1136/bmjopen-2021-054046](https://doi.org/10.1136/bmjopen-2021-054046)] [Medline: [35501086](https://pubmed.ncbi.nlm.nih.gov/35501086/)]
11. Morreel S, Colliers A, Remmen R, Verhoeven V, Philips H. How accurate is telephone triage in out-of-hours care? An observational trial in real patients. *Acta Clin Belg* 2022 Apr 30;77(2):301-306. [doi: [10.1080/17843286.2020.1839719](https://doi.org/10.1080/17843286.2020.1839719)] [Medline: [33124524](https://pubmed.ncbi.nlm.nih.gov/33124524/)]
12. Ashwood JS, Mehrotra A, Cowling D, Uscher-Pines L. Direct-to-consumer telehealth may increase access to care but does not decrease spending. *Health Aff (Millwood)* 2017 Mar 01;36(3):485-491. [doi: [10.1377/hlthaff.2016.1130](https://doi.org/10.1377/hlthaff.2016.1130)] [Medline: [28264950](https://pubmed.ncbi.nlm.nih.gov/28264950/)]
13. Trivedi S, Littmann J, Stempien J, Kapur P, Bryce R, Betz M. A comparison between computer-assisted self-triage by patients and triage performed by nurses in the emergency department. *Cureus* 2021 Mar 19;13(3):e14002 [FREE Full text] [doi: [10.7759/cureus.14002](https://doi.org/10.7759/cureus.14002)] [Medline: [33884243](https://pubmed.ncbi.nlm.nih.gov/33884243/)]
14. Cooper AA, Humphreys KR. The uncertainty is killing me: self-triage decision making and information availability. *E-JAP* 2008 Oct 14;4(1):1-6. [doi: [10.7790/ejap.v4i1.124](https://doi.org/10.7790/ejap.v4i1.124)]
15. Mills B, Hill M, Buck J, Walter E, Howard K, Raisinger A, et al. What constitutes an emergency ambulance call? *Australas J Paramed* 2019 Mar 22;16 This is an online article and does not have a page number. Please delete the filler page number 0 in the XML. [doi: [10.33151/ajp.16.626](https://doi.org/10.33151/ajp.16.626)]
16. Coyle N, Kennedy A, Schull MJ, Kiss A, Hefferon D, Sinclair P, et al. The use of a self-check-in kiosk for early patient identification and queuing in the emergency department. *CJEM* 2019 Nov 06;21(6):789-792. [doi: [10.1017/cem.2019.349](https://doi.org/10.1017/cem.2019.349)] [Medline: [31057137](https://pubmed.ncbi.nlm.nih.gov/31057137/)]
17. Winn AN, Somai M, Fergestrom N, Crotty BH. Association of use of online symptom checkers with patients' plans for seeking care. *JAMA Netw Open* 2019 Dec 02;2(12):e1918561 [FREE Full text] [doi: [10.1001/jamanetworkopen.2019.18561](https://doi.org/10.1001/jamanetworkopen.2019.18561)] [Medline: [31880791](https://pubmed.ncbi.nlm.nih.gov/31880791/)]
18. Kopka M, Schmieding ML, Rieger T, Roesler E, Balzer F, Feufel MA. Determinants of laypersons' trust in medical decision aids: randomized controlled trial. *JMIR Hum Factors* 2022 May 03;9(2):e35219 [FREE Full text] [doi: [10.2196/35219](https://doi.org/10.2196/35219)] [Medline: [35503248](https://pubmed.ncbi.nlm.nih.gov/35503248/)]
19. Schmieding ML, Kopka M, Schmidt K, Schulz-Niethammer S, Balzer F, Feufel MA. Triage accuracy of symptom checker apps: 5-year follow-up evaluation. *J Med Internet Res* 2022 May 10;24(5):e31810 [FREE Full text] [doi: [10.2196/31810](https://doi.org/10.2196/31810)] [Medline: [35536633](https://pubmed.ncbi.nlm.nih.gov/35536633/)]
20. Ceney A, Tolond S, Glowinski A, Marks B, Swift S, Palser T. Accuracy of online symptom checkers and the potential impact on service utilisation. *PLoS One* 2021;16(7):e0254088 [FREE Full text] [doi: [10.1371/journal.pone.0254088](https://doi.org/10.1371/journal.pone.0254088)] [Medline: [34265845](https://pubmed.ncbi.nlm.nih.gov/34265845/)]
21. Semigran HL, Linder JA, Gidengil C, Mehrotra A. Evaluation of symptom checkers for self diagnosis and triage: audit study. *BMJ* 2015 Jul 08;351:h3480 [FREE Full text] [doi: [10.1136/bmj.h3480](https://doi.org/10.1136/bmj.h3480)] [Medline: [26157077](https://pubmed.ncbi.nlm.nih.gov/26157077/)]
22. Hill MG, Sim M, Mills B. The quality of diagnosis and triage advice provided by free online symptom checkers and apps in Australia. *Med J Aust* 2020 Jun;212(11):514-519. [doi: [10.5694/mja2.50600](https://doi.org/10.5694/mja2.50600)] [Medline: [32391611](https://pubmed.ncbi.nlm.nih.gov/32391611/)]
23. Schmieding ML, Mörgeli R, Schmieding MAL, Feufel MA, Balzer F. Benchmarking triage capability of symptom checkers against that of medical laypersons: survey study. *J Med Internet Res* 2021 Mar 10;23(3):e24475 [FREE Full text] [doi: [10.2196/24475](https://doi.org/10.2196/24475)] [Medline: [33688845](https://pubmed.ncbi.nlm.nih.gov/33688845/)]
24. Schmieding ML, Mörgeli R, Schmieding MAL, Feufel MA, Balzer F. Data set supplementing "Benchmarking triage capability of symptom checkers against that of medical laypersons: survey study". Zenodo. 2021. URL: <https://zenodo.org/record/4454538#.YyCACXZBzIU> [accessed 2022-09-13]

25. R Core Team. R: A Language and Environment for Statistical Computing. The R Project for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing; 2014. URL: <https://www.R-project.org/> [accessed 2022-09-12]
26. Wickham H. ggplot2.3.5. Vienna, Austria: R Foundation for Statistical Computing URL: <https://ggplot2.tidyverse.org/> [accessed 2022-09-12]
27. Neuwirth E. RColorBrewer: ColorBrewer Palettes. The Comprehensive R Archive Network. 2022. URL: <https://cran.r-project.org/web/packages/RColorBrewer/index.html> [accessed 2022-09-12]
28. Wickham H, Averick M, Bryan J, Chang W, McGowan L, François R, et al. Welcome to the tidyverse. JOSS 2019 Nov;4(43):1686. [doi: [10.21105/joss.01686](https://doi.org/10.21105/joss.01686)]
29. 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](https://doi.org/10.2196/14679)] [Medline: [32012052](https://pubmed.ncbi.nlm.nih.gov/32012052/)]
30. Gottlieb K, Petersson G. Limited evidence of benefits of patient operated intelligent primary care triage tools: findings of a literature review. BMJ Health Care Inform 2020 May 07;27(1):e100114 [FREE Full text] [doi: [10.1136/bmjhci-2019-100114](https://doi.org/10.1136/bmjhci-2019-100114)] [Medline: [32385041](https://pubmed.ncbi.nlm.nih.gov/32385041/)]
31. Meyer AND, Payne VL, Meeks DW, Rao R, Singh H. Physicians' diagnostic accuracy, confidence, and resource requests: a vignette study. JAMA Intern Med 2013 Nov 25;173(21):1952-1958. [doi: [10.1001/jamainternmed.2013.10081](https://doi.org/10.1001/jamainternmed.2013.10081)] [Medline: [23979070](https://pubmed.ncbi.nlm.nih.gov/23979070/)]
32. Peabody JW, Luck J, Glassman P, Dresselhaus TR, Lee M. Comparison of vignettes, standardized patients, and chart abstraction: a prospective validation study of 3 methods for measuring quality. JAMA 2000 Apr 05;283(13):1715-1722. [doi: [10.1001/jama.283.13.1715](https://doi.org/10.1001/jama.283.13.1715)] [Medline: [10755498](https://pubmed.ncbi.nlm.nih.gov/10755498/)]
33. Yager J, Linn LS, Leake B, Gastaldo G, Palkowski C. Initial clinical judgments by internists, family physicians, and psychiatrists in response to patient vignettes: I. Assessment of problems and diagnostic possibilities. Gen Hosp Psychiatry 1986 May;8(3):145-151. [doi: [10.1016/0163-8343\(86\)90072-1](https://doi.org/10.1016/0163-8343(86)90072-1)] [Medline: [3710146](https://pubmed.ncbi.nlm.nih.gov/3710146/)]
34. Bol N, Helberger N, Weert JCM. Differences in mobile health app use: a source of new digital inequalities? Inf Soc 2018 Apr 26;34(3):183-193. [doi: [10.1080/01972243.2018.1438550](https://doi.org/10.1080/01972243.2018.1438550)]
35. Loos A. Cyberchondria: too much information for the health anxious patient? J Consum Health Internet 2013 Oct;17(4):439-445. [doi: [10.1080/15398285.2013.833452](https://doi.org/10.1080/15398285.2013.833452)]
36. Jessup S, Schneider T, Alarcon G, Ryan T, Capiola A. The measurement of the propensity to trust automation. In: Chen JYC, Fragomeni G, editors. Virtual, Augmentation and Mixed Reality Applications and Case Studies. Cham: Springer International Publishing; 2019:476-489.

## Abbreviations

- ED:** emergency department  
**NPV:** negative predictive value  
**PPV:** positive predictive value

*Edited by A Mavragani; submitted 24.04.22; peer-reviewed by M Schmude, H Fraser, K Bin; comments to author 21.07.22; revised version received 08.08.22; accepted 16.08.22; published 12.10.22.*

*Please cite as:*

*Kopka M, Feufel MA, Balzer F, Schmieding ML  
The Triage Capability of Laypersons: Retrospective Exploratory Analysis  
JMIR Form Res 2022;6(10):e38977  
URL: <https://formative.jmir.org/2022/10/e38977>  
doi:[10.2196/38977](https://doi.org/10.2196/38977)  
PMID:[36222793](https://pubmed.ncbi.nlm.nih.gov/36222793/)*

©Marvin Kopka, Markus A Feufel, Felix Balzer, Malte L Schmieding. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 12.10.2022. 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

# Online Health Information Seeking and eHealth Literacy Among Spanish Language–Dominant Latino Adults Receiving Care in a Community Clinic: Secondary Analysis of Pilot Randomized Controlled Trial Data

Enmanuel A Chavarria<sup>1,2</sup>, PhD; Shannon M Christy<sup>3,4,5</sup>, PhD; Han Feng<sup>6</sup>, MS, PhD; Hongyu Miao<sup>7</sup>, PhD; Rania Abdulla<sup>8</sup>, MA; Liliana Gutierrez<sup>9</sup>, RN, BSN; Diana Lopez<sup>10</sup>, RN; Julian Sanchez<sup>3,5</sup>, MD; Clement K Gwede<sup>3,4,5</sup>, RN, MPH, PhD, FAAN; Cathy D Meade<sup>3,4</sup>, RN, PhD, FAAN

<sup>1</sup>Department of Behavioral, Social, and Health Education Sciences, Rollins School of Public Health, Emory University, Atlanta, GA, United States

<sup>2</sup>Cancer Prevention and Control Research Program, Winship Cancer Institute, Emory University, Atlanta, GA, United States

<sup>3</sup>Department of Health Outcomes and Behavior, Division of Population Science, H. Lee Moffitt Cancer Center and Research Institute, Tampa, FL, United States

<sup>4</sup>Department of Oncologic Sciences, Morsani College of Medicine, University of South Florida, Tampa, FL, United States

<sup>5</sup>Department of Gastrointestinal Oncology, H. Lee Moffitt Cancer Center and Research Institute, Tampa, FL, United States

<sup>6</sup>School of Medicine, Tulane University, New Orleans, LA, United States

<sup>7</sup>College of Nursing, Florida State University, Tallahassee, FL, United States

<sup>8</sup>Non-Therapeutic Research Office, H. Lee Moffitt Cancer Center and Research Institute, Tampa, FL, United States

<sup>9</sup>Evergreen Health, Kirkland, WA, United States

<sup>10</sup>Suncoast Community Health Centers, Brandon, FL, United States

**Corresponding Author:**

Enmanuel A Chavarria, PhD

Department of Behavioral, Social, and Health Education Sciences

Rollins School of Public Health

Emory University

1518 Clifton Road, NE

Atlanta, GA, 30322

United States

Phone: 1 4047278741

Email: [enmanuel.chavarria@emory.edu](mailto:enmanuel.chavarria@emory.edu)

## Abstract

**Background:** eHealth literacy is the ability to seek, obtain, and decipher online health information (OHI) for health and disease management. Rapid developments in eHealth (eg, health care services and online information) place increased demands on patients to have high eHealth literacy levels. Yet, greater emphasis on eHealth may disproportionately affect groups with limited eHealth literacy. Cultural background, language, and eHealth literacy are influential considerations affecting health care and information access, health care use, and successful eHealth resource use, and they may influence OHI seeking for behavioral change toward cancer prevention.

**Objective:** This study aimed to characterize the extent of OHI seeking and eHealth literacy among Spanish-dominant (SD) Latino adults aged 50 to 75 years. Further, we aimed to examine potential associations between sociodemographic characteristics, Preventive Health Model (PHM) constructs, OHI-seeking behaviors, and eHealth literacy, separately.

**Methods:** Participants (N=76) self-identified as Latino, were enrolled in a colorectal cancer (CRC) screening intervention, were aged 50 to 75 years, were at average risk for CRC, were not up to date with CRC screening, and preferred receiving health information in Spanish. We describe participants' sociodemographic characteristics, PHM constructs, OHI-seeking behaviors, and eHealth literacy—among those seeking OHI—assessed at enrollment. Descriptive analyses were first performed for all variables. Next, primary univariate logistic analyses explored possible associations with OHI seeking. Finally, using data from those seeking OHI, exploratory univariate analyses sought possible associations with eHealth literacy.

**Results:** A majority (51/76, 67%) of the participants were female, 62% (47/76) reported not having graduated high school, and 41% (31/76) reported being unemployed or having an annual income of less than US \$10,000. Additionally, 75% (57/76) of the participants reported not having health insurance. In total, 71% (54/76) of the participants reported not having sought OHI for themselves or others. Univariate logistic regression suggested that higher educational attainment was significantly associated with an increased likelihood of having sought OHI (odds ratio 17.4, 95% CI 2.0-150.7;  $P=.009$ ). Among those seeking OHI (22/76, 29%), 27% (6/22) were at risk of having low eHealth literacy based on an eHealth Literacy Scale score of less than 26. Among OHI seekers (22/76, 29%), an examination of associations found that higher eHealth literacy was associated with greater self-efficacy for screening with the fecal immunochemical test ( $\beta=1.20$ , 95% CI 0.14-2.26;  $P=.02$ ).

**Conclusions:** Most SD Latino participants had not sought OHI for themselves or others (eg, family or friends), thus potentially limiting access to beneficial online resources. Preliminary findings convey that higher eHealth literacy occurs among those with higher self-efficacy for CRC screening. Findings inform areas of focus for future larger-scale investigations, including further exploration of reasons for not seeking OHI among SD Latino adults and an in-depth look at eHealth literacy and cancer screening behaviors.

**Trial Registration:** ClinicalTrials.gov NCT03078361; <https://clinicaltrials.gov/ct2/show/NCT03078361>

(*JMIR Form Res* 2022;6(10):e37687) doi:[10.2196/37687](https://doi.org/10.2196/37687)

## KEYWORDS

eHealth literacy; online health information seeking; medically underserved; Hispanic; Latino; Spanish language–dominant; health communications; colorectal cancer screening

## Introduction

Today, an ever-increasing quantity of health information can be accessed online [1,2]. Accordingly, online health information (OHI) is now becoming commonplace in health care interactions and health education for patients and caregivers [3-5]. Such developments place an increased demand on patients to be able to seek, obtain, and decipher OHI for health and disease management [6-9]. eHealth tools provide little value if the intended users require added training and skills to effectively engage these resources. These skills are termed “eHealth literacy” and comprise a multifaceted dynamic construct, including previous and current technology use, demographic and cognitive status, and health-related quality of life (HRQOL) [9,10].

A growing evidence base now attests to the promise of eHealth for promoting positive health behavior change, self-efficacy, and knowledge acquisition [11-13].

In recognition of the importance of health information access, the US Department of Health and Human Services set objectives in 2010 for health communication (HC) and health information technology (HIT) as part of the Healthy People initiative. Today, Healthy People 2030 offers HC and HIT objectives aimed at enhancing the use of OHI in public health, including the following: (1) “Increase the proportion of people who can view, download, and send their electronic health information” [14] and (2) “Increase the proportion of people who say their online medical record is easy to understand” [15].

Access to OHI is not uniformly distributed throughout the population, which may exacerbate disparities in health and health care [16,17].

Among Latino adults, the fastest growing demographic group in the United States, 72% overall are using the internet [18]. Yet, among Spanish-dominant (SD) Latino adults aged 50 to 64 years, internet use drops to 67%, and only 42% of those aged

65 years and older use the internet [19]. Since disparities exist, it is necessary to understand the current OHI and eHealth literacy levels of SD Latino adults aged 50 to 75 years, yet there is insufficient knowledge in this area.

Further, among Latino adults, cancer continues to be the leading cause of death, accounting for 21% of overall deaths [8,20,21]. Specifically, colorectal cancer (CRC) is the third leading cause of cancer deaths among Latino adults [21].

The Preventive Health Model (PHM) is a conceptual framework that aims to explain how health beliefs are related to CRC screening [22]. Specifically, among the CRC screening literature, 26 items measure seven PHM constructs. The PHM constructs include salience and coherence, perceived susceptibility, self-efficacy, response efficacy, cancer worry, social influence, and religious beliefs [23]. These are assessed using a 5-point Likert scale, ranging from 1 (strongly disagree) to 5 (strongly agree) [22,24-26]. Reliability and validity for these subscales have been demonstrated previously [22,24-26]. The salience and coherence subscale measures one’s belief that CRC screening is important and makes sense in one’s life. The perceived susceptibility subscale assesses one’s perceived risk of being diagnosed with CRC. The self-efficacy subscale measures the belief that one could complete the steps necessary for fecal immunochemical test (FIT) collection. The response efficacy subscale measures the belief that CRC screening is beneficial for early detection and prevention of CRC. The cancer worry subscale measures the degree to which one is worried about having an abnormal CRC screening result. The social influence subscale assesses the perception that important others (eg, family members, friends, and one’s health care provider) would want the individual to complete CRC screening. The religious beliefs subscale assesses the degree to which one relies on one’s religious beliefs to make health decisions. Thus, in relating a need to understand current OHI and eHealth literacy levels of SD Latino adults aged 50 to 75 years, it is also of interest to explore how these may relate to PHM constructs in this understudied population. Prior studies suggest that

sociodemographic and PHM constructs, such as higher social influence and religious beliefs, were associated with lower health literacy among English-preferring individuals [27,28]. By contrast, lower cancer worry and lower religious beliefs [29] were associated with adequate health literacy to understand written health information, and higher educational attainment was significantly associated with adequate health literacy in completing health forms among SD Latino adults aged 50 to 75 years [29]. However, less is known about how sociodemographic variables and PHM constructs might influence OHI seeking or eHealth literacy among SD Latino adults.

Thus, the aims of this study were to (1) describe the prevalence of OHI seeking and eHealth literacy and (2) examine preliminary associations between sociodemographic characteristics, PHM constructs, OHI seeking, and eHealth literacy, separately, among SD Latino adults.

To achieve our study aims, we posed the following four research questions:

1. Research question 1 (RQ1): To what extent do SD Latino adults use the internet to locate OHI?
2. Research question 2 (RQ2): How are self-reported levels of eHealth literacy described among SD Latino adults who seek OHI?
3. Research question 3 (RQ3): What sociodemographic characteristics and PHM constructs are associated with OHI seeking by SD Latino adults?
4. Research question 4 (RQ4): What sociodemographic characteristics and PHM constructs are preliminarily associated with eHealth literacy among SD Latino adults who seek OHI?

Taken together, this study examines OHI and eHealth literacy among a diverse understudied population (ie, SD Latino adults). Further, the study provides a preliminary look at possible associations among sociodemographic characteristics, PHM constructs, OHI seeking, and eHealth literacy through an innovative examination within the literature.

## Methods

### Overview

Data for this report were collected as part of a larger pilot randomized controlled trial—Latino Colorectal Cancer Awareness, Research, Education, and Screening (Latino CARES)—that promoted CRC screening by providing education and a FIT. The pilot study aimed to evaluate the feasibility and impact of a Spanish-language, low-literacy, culturally targeted intervention (ie, a photonovella and DVD) plus FIT compared with a standard Spanish-language booklet developed by the Centers for Disease Control and Prevention plus FIT. Of note, recruitment took place at two participating clinic sites that are part of a Federally Qualified Health Center (FQHC) in Southwest Florida. The FQHC sites are centrally located in agricultural farmworker communities and annually serve a large number (ie, approximately 5000) of medically underserved patients aged 50 to 75 years, a majority of whom are of Latino origin from diverse nationalities and include farmworker

populations [30]. Detailed methods are provided in prior publications that communicate the results of the main outcomes of the Latino CARES study [29-31].

Participant consent was obtained prior to baseline interview and randomization. Eligible participants (1) were receiving care at two participating FQHC clinic locations; (2) were between the ages of 50 and 75 years; (3) self-identified as Latino; (4) were able to read, speak, and understand Spanish; (5) preferred to receive health information in Spanish; (6) were currently not up to date per CRC screening guidelines (ie, had never screened or previously screened but were now overdue); and (7) were at average risk for CRC.

### Ethical Considerations

The University of South Florida Institutional Review Board approved the study (approval No. MCC-17665) prior to participant enrollment. Study procedures were conducted in accordance with the Declaration of Helsinki. During enrollment, participants provided informed consent to be included in the study.

### Measures

Study assessments consisted of validated measures that were translated into Spanish and refined for cultural relevance using the following established procedures. Our study team included three bilingual (ie, fluent in English and Spanish) researchers and a bilingual (ie, fluent in English and Spanish) community advisory board. Applying the Brislin method [32], measures were first translated into Spanish by a bilingual study coordinator and then back-translated by a second bilingual study team member. Any discrepancies were arbitrated by a third bilingual study team member. Items were pretested among community members using learner verification methodology [33,34]; they were further refined in consultation with the community advisory board for cultural relevance and comprehension. Baseline items were administered by bilingual (ie, English and Spanish) study coordinators at the time of interview. To minimize literacy issues, all questions were read aloud for all participants.

### Online Health Information Seeking

A single item, gathered from the Pew Research Center [8] and used in previous OHI behavioral studies among Latino adults [4,35], assessed OHI-seeking usage: “Have you previously personally searched for health information on the Internet/Online for yourself or for others? For example, [have you] sought on Google/Yahoo information on high blood pressure, healthy recipes, or efficient exercises?” Response options were yes or no.

### eHealth Literacy

eHealth literacy was assessed only among participants who reported engaging in OHI seeking. Thus, the number of participants for whom eHealth literacy was assessed was lower than that for those responding to the other measures. The Spanish-translated eHealth Literacy Scale (eHEALS) [9,36] was administered to assess eHealth literacy, including knowledge, comfort, and perceived skills at finding, evaluating, and applying OHI to health problems. The eHEALS [9]

comprises eight items scored on a 5-point Likert scale, ranging from 1 (strongly disagree) to 5 (strongly agree), and aims to reflect the individuals' own perceptions of their knowledge and skills at using eHealth information [9,37]. The final result is the sum of all items and ranges from 8 to 40, with higher scores reflecting a higher level of eHealth literacy. The validity and reliability of the eHEALS has been demonstrated across various health conditions [38,39], ages [40-42], and languages [43], including Spanish [36]. Following other studies with similar target populations, the cutoff for high eHealth literacy was set at 26 [10,38,44-48]. Thus, in maintaining consistency in terminology with the literature, this study defined high self-perceived eHealth literacy as an eHEALS score equal to or greater than 26 out of 40, and low self-perceived eHealth literacy was defined as an eHEALS score of less than 26 [10,38,44-48].

### Sociodemographic Characteristics

Sociodemographic variables assessed included age, gender, race and ethnicity, marital status, education, health insurance status, employment status, and income. Additionally, as aggregated data in prior literature masked substantial heterogeneity within the Latino population, we assessed parental foreign-born status and participant foreign-born status; if foreign-born, the country of origin and years lived in the United States were also assessed [30].

### Preventive Health Model Variables

Seven constructs of the PHM were assessed in this study using 26 total items referenced from prior CRC studies [24,25,27,28,30,49-52]. For each item assessing PHM constructs, response options were based on a 5-point Likert scale, ranging from 1 (strongly disagree) to 5 (strongly agree). Four items assessed salience and coherence, or the perception that performing a health behavior is consistent with their beliefs about how to protect and maintain health [24-26,29,30,49,50]. Three items assessed perceived susceptibility, or one's perceived personal risk for developing CRC or colon polyps [24,25,30,49]. Two items assessed cancer worry, or one's concern that completing CRC screening will reveal a health concern [24,26,29,30,49,50,53].

Two items measured response efficacy, or the belief that adopting a behavior will be effective in reducing disease threat [24,29,30,49,50]. Four items measured social influence, or the influence of family members and doctors or health professionals on an individual's willingness to comply with CRC screening [24,26,29,30,49,50]. Five items assessed religiosity, or the extent to which religious beliefs might influence medical decision-making, such as CRC screening [29,30,51,54]. Six items measured self-efficacy for screening using FIT, or attitudes and confidence toward completing FIT testing [24-26,29,30,49,55]. For each respective construct, the corresponding items' scores reported by participants were added together; construct total scores were used for analyses.

### Statistical Analysis

Sociodemographic characteristics and PHM constructs were summarized using descriptive statistics. The assessment details

for each research question are discussed in the following four sections.

#### Research Question 1

RQ1 is as follows: To what extent do SD Latino adults use the internet to locate OHI? OHI seeking was characterized using descriptive statistics.

#### Research Question 2

RQ2 is as follows: How are self-reported levels of eHealth literacy described among SD Latino adults who seek OHI? eHealth literacy was characterized using descriptive statistics.

#### Research Question 3

RQ3 is as follows: What sociodemographic characteristics and PHM constructs are associated with OHI seeking by SD Latino adults? Univariate logistic regression analyses were conducted to examine potential sociodemographic and PHM constructs associated with OHI seeking. Data from all study participants (N=76) were available for this exploration. For these analyses, we coded responses to the item "Personally looked online for health information for self or others" as a binary outcome. Gender, insurance status, employment status, marriage status, age, annual income, educational attainment, perceived salience, perceived susceptibility, response efficacy, cancer worry, social influence, religious beliefs, and self-efficacy for screening with FIT were treated as independent variables, and whether participants sought OHI for themselves or others—yes or no response—was treated as the dependent variable.

#### Research Question 4

RQ4 is as follows: What sociodemographic characteristics and PHM constructs are preliminarily associated with eHealth literacy among SD Latino adults who seek OHI? Univariate analyses were conducted leveraging linear regressions to examine potential sociodemographic characteristics and PHM constructs associated with eHealth literacy (ie, the eHEALS score). The eHealth literacy outcome score was treated as a continuous outcome, ranging from 8 to 40. For this exploration, data were available only from those seeking OHI (22/76, 29%). Gender, insurance status, employment status, marriage status, age, annual income, educational attainment, perceived salience, perceived susceptibility, response efficacy, cancer worry, social influence, religious beliefs, and self-efficacy for screening with FIT were treated as independent variables, and the eHEALS score was treated as the dependent variable. Due to the sample size limitation, this exploration was underpowered; hence, the goal was to provide reasonably reliable estimates to guide the design of a future, larger, appropriately powered study.

### General Analysis

Analyses were conducted using SAS (version 9.4 [TS1M6]; SAS Institute Inc). A *P* value of less than .05 was considered statistically significant. Of note, analyses are exploratory and not for definitive inferential interpretations.

## Results

---

### Sample Characteristics

Sociodemographic characteristics are described in [Table 1](#) and in the main reported outcomes of Gwede et al [[30](#)]. In total, 67% (51/76) of participants were female. The mean age of the participants was 57.2 (SD 6.0) years (range 50-74). In total, 62% (47/76) of participants reported not completing high school.

In addition, 41% (31/76) of participants reported being unemployed and having an annual income of less than US \$10,000. Further, 75% (57/76) of participants lacked health insurance. In total, 93% (71/76) of participants reported being born outside of the United States. Among those born outside of the United States, a majority (49/71, 69%) reported Mexico as their country of birth. Further, among those born outside of the United States, the mean length of time reported living in the United States was 23.4 (SD 10.9) years (range 2-56).

**Table 1.** Summary of descriptive statistics.

Variables	Full sample (N=76)	OHI <sup>a</sup> seekers (n=22)
Age (years), mean (SD), range	57.2 (6.0), 50-74	56.6 (5.8), 50-69
Years in the United States, mean (SD), range	23.4 (10.9), 2-56	19.4 (11.2), 2-56
<b>Gender, n (%)</b>		
Male	25 (33)	3 (4)
Female	51 (67)	19 (25)
<b>Race, n (%)</b>		
White	23 (30)	8 (11)
Black	1 (1)	1 (1)
Other or more than one race	52 (68)	13 (17)
<b>Marital status, n (%)</b>		
Married or partnered	53 (70)	14 (18)
Divorced or separated	10 (13)	5 (7)
Widowed	6 (8)	3 (4)
Never married or single	7 (9)	0 (0)
<b>Employment status, n (%)</b>		
Employed or self-employed	40 (53)	16 (21)
Unemployed	31 (41)	6 (8)
Student	1 (1)	0 (0)
Retired	3 (4)	0 (0)
<b>Education, n (%)</b>		
Less than high school	47 (62)	3 (4)
High school graduate	13 (17)	6 (8)
Some college or technical school	7 (9)	5 (7)
College graduate	7 (9)	6 (8)
Graduate or professional (postcollege)	2 (3)	2 (3)
<b>Annual income (US \$), n (%)</b>		
<10,000	31 (41)	8 (11)
10,000-25,000	29 (38)	10 (13)
25,001-35,000	8 (11)	2 (3)
35,001-75,000	2 (3)	1 (1)
Don't know or prefer not to answer	6 (8)	1 (1)
<b>Insurance status, n (%)</b>		
No insurance	57 (75)	14 (18)
Medicaid or Medicare	4 (5)	0 (0)
County health insurance	8 (11)	3 (4)
Private health insurance	7 (9)	5 (7)
<b>Country of birth, n (%)</b>		
United States	5 (7)	2 (3)
Other	71 (94)	20 (26)
<b>Self-reported country of birth (if that reported was other than the United States), n (%)</b>		
Mexico	49 (64)	7 (9)
Colombia	6 (8)	4 (5)



Variables	Full sample (N=76)	OHI <sup>a</sup> seekers (n=22)
Puerto Rico	5 (7)	2 (3)
Costa Rica	3 (4)	2 (3)
Dominican Republic	2 (3)	0 (0)
Peru	2 (3)	2 (3)
Chile	1 (1)	1 (1)
Cuba	1 (1)	1 (1)
El Salvador	1 (1)	0 (0)
Venezuela	1 (1)	1 (1)
<b>Parents born outside the United States, including Puerto Rico, n (%)</b>		
Yes	73 (96)	21 (28)
No	3 (4)	1 (1)
<b>Personally sought OHI, n (%)</b>		
Yes	22 (29)	22 (29)
No	54 (71)	0 (0)
eHEALS <sup>b</sup> score, mean (SD), range	N/A <sup>c</sup>	29.73 (6.6), 15-40

<sup>a</sup>OHI: online health information.

<sup>b</sup>eHEALS: eHealth Literacy Scale; the total score ranged from 8 to 40; 22 out of 76 participants completed the scale.

<sup>c</sup>N/A: not applicable; this was not calculated for the full sample, since not all participants completed the scale.

## Main Findings

### *Extent to Which SD Latino Adults Use the Internet to Locate OHI (RQ1)*

All participants (N=76) responded to the item assessing personal OHI seeking for themselves or others. In total, 71% (54/76) of participants reported not having personally sought OHI for themselves or others (Table 1), whereas the remaining 29% (22/76) of participants reported having personally looked for OHI information for themselves or others.

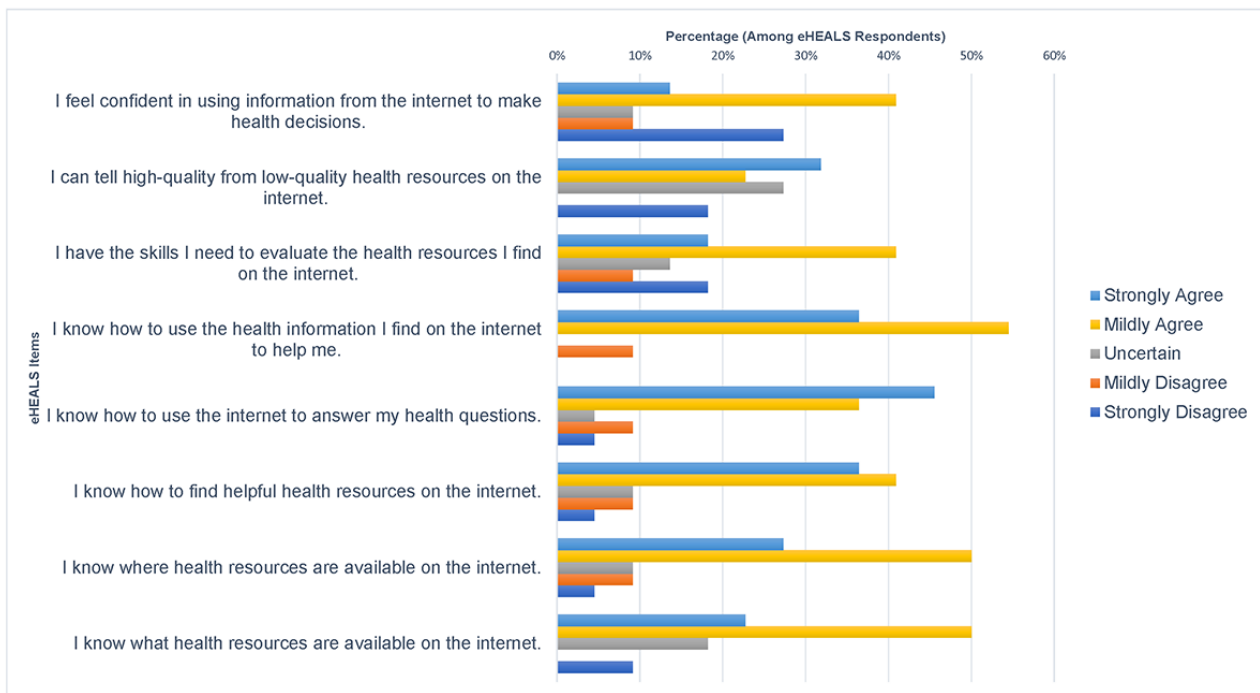
### *Variability in Self-reported Levels of eHealth Literacy Among SD Latino Adults (RQ2)*

Assessment of eHealth literacy using the eHEALS was only completed for those who answered yes to having sought OHI for themselves or others (22/76, 29%). Of those individuals, the

mean eHEALS score was 29.7 (SD 6.6; range 15-40). In total, 27% (6/22) of those seeking OHI had an eHEALS score of less than 26, indicating that these participants were in the low-eHealth literacy category.

Frequency of responses to the eight-item eHEALS—note that only 22 participants completed this scale, as they sought OHI—is reported in Figure 1. The item with the greatest degree of agreement was “I know how to use the health resources I find on the internet to help me,” with 91% (20/22) of eHEALS respondents self-reporting either mildly agree or strongly agree. The item with the least amount of agreement was “I feel confident in using information from the internet to make health decisions,” with 45% (10/22) of eHEALS respondents self-reporting being uncertain, mildly disagree, or strongly disagree.

**Figure 1.** Responses to the eight-item eHEALS; 22 out of 76 respondents completed the scale, and percentages are out of 22. eHEALS: eHealth Literacy Scale.



**Sociodemographics and PHM Constructs Associated With OHI Seeking by SD Latino Adults (RQ3)**

Univariate logistic regression analyses on data from 76 participants were performed to examine preliminary associations between sociodemographic variables, PHM constructs, and the likelihood of seeking OHI. Only educational attainment was found to be significantly associated with OHI seeking (Table 2). Higher educational attainment was significantly associated

with an increased likelihood of having sought OHI (odds ratio 17.4, 95% CI 2.0-150.7;  $P=.009$ ). Employment status approached significance ( $P=.07$ ). None of the PHM health beliefs were significantly associated with OHI. The logistic regression model was statistically significant ( $\chi^2_{14}=60.3$ ,  $P<.001$ ). The model explained 81.9% (Nagelkerke  $R^2$ ) of the variance in OHI and correctly classified 90.0% of cases. Sensitivity was 85.7%, specificity was 91.8%, positive predictive value was 81.8%, and negative predictive value was 93.8%.

**Table 2.** Factors associated with seeking of online health information using univariate logistic regression analyses.

Covariate	Respondents (N=76), n (%)	Odds ratio (95% CI)	P value <sup>a</sup>
<b>Gender</b>			
Female	51 (67)	170.05 (0.37 to >999.99)	.10
Male	25 (33)	Reference	N/A <sup>b</sup>
<b>Insured</b>			
Yes	19 (25)	59.19 (0.42 to >999.99)	.11
No	57 (75)	Reference	N/A
<b>Employed</b>			
Yes	40 (53)	>999.99 (0.57 to >999.99)	.07
No	35 (46)	Reference	N/A
<b>Married</b>			
Yes	53 (70)	62.01 (0.20 to >999.99)	.16
No	23 (30)	Reference	N/A
Age	76 (100)	0.93 (0.64 to 1.35)	.68
Annual income	70 (92)	2.03 (0.50 to 8.29)	.32
Educational attainment	76 (100)	17.40 (2.01 to 150.72)	.009
Perceived salience	76 (100)	1.84 (0.45 to 7.44)	.39
Perceived susceptibility	76 (100)	0.99 (0.53 to 1.85)	.98
Response efficacy	76 (100)	2.33 (0.32 to 17.10)	.41
Cancer worry	76 (100)	1.18 (0.70 to 2.02)	.53
Social influence	76 (100)	1.07 (0.60 to 1.90)	.83
Religious beliefs	76 (100)	0.95 (0.74 to 1.22)	.68
Self-efficacy for screening with FIT <sup>c</sup>	76 (100)	0.96 (0.43 to 2.14)	.91

<sup>a</sup> $P < .05$  indicates statistical significance.

<sup>b</sup>N/A: not applicable.

<sup>c</sup>FIT: fecal immunochemical test.

### ***Sociodemographics and PHM Constructs Associated With the eHealth Literacy of SD Latino Adults (RQ4)***

Univariate analyses using linear regressions were completed to examine preliminary associations between eHEALS scores, sociodemographic variables, and PHM constructs. Higher

self-efficacy for screening with FIT was significantly associated with higher eHEALS scores (ie, eHealth literacy;  $\beta=1.20$ , 95% CI 0.14-2.26;  $P=.02$ ). Table 3 reports the regression coefficients and standard errors. The best-fit model was significantly associated with eHEALS score ( $F_{1,20}=5.53$ ,  $P=.02$ ; adjusted  $R^2=0.18$ ).

**Table 3.** Results from univariate analyses using linear regressions for two models.

Factors associated with the eHEALS <sup>a</sup>	$\beta$ coefficient (SE)	Standardized $\beta$	<i>t</i> test ( <i>df</i> =1)	<i>P</i> value <sup>b</sup>
<b>Model 1</b>				
Constant	4.22 (56.50)	0	0.07	.94
Gender	-3.19 (10.53)	-0.18	-0.30	.77
Insured	1.12 (7.49)	0.09	0.15	.88
Employed	-1.71 (9.88)	-0.12	-0.17	.87
Married	3.51 (7.86)	0.26	0.45	.67
Age	0.17 (0.45)	0.16	0.37	.72
Annual income	-0.60 (5.04)	-0.08	-0.12	.91
Educational attainment	-0.19 (2.99)	-0.04	-0.06	.95
Perceived salience	-0.57 (2.88)	-0.11	-0.20	.84
Perceived susceptibility	-0.82 (1.18)	-0.32	-0.69	.51
Response efficacy	-0.46 (2.36)	-0.08	-0.19	.85
Cancer worry	0.09 (1.66)	0.04	0.06	.96
Social influence	-0.10 (1.30)	-0.05	-0.08	.94
Religious beliefs	-0.33 (1.00)	-0.25	-0.34	.75
Self-efficacy for screening with FIT <sup>c</sup>	1.58 (1.17)	0.64	1.35	.22
<b>Model 2</b>				
Constant	-6.41 (14.51)	0	-0.44	.66
Self-efficacy for screening with FIT	1.20 (0.51)	0.47	2.35	.02

<sup>a</sup>eHEALS: eHealth Literacy Score.

<sup>b</sup>*P*<.05 indicates statistical significance.

<sup>c</sup>FIT: fecal immunochemical test.

## Discussion

### Principal Findings

OHI has become more routine in health care interactions and health education for patients and caregivers [3-5]. Yet, greater emphasis on eHealth may disproportionately affect groups with limited eHealth literacy. Healthy People 2030 continues the national effort initiated in 2010 in recognition of the importance of HC and HIT to support better communication, care, and outcomes toward achieving health equity [56]. There is a dearth of information available in the literature that communicates specifically on factors (eg, sociodemographic characteristics and health beliefs) associated with OHI seeking and eHealth literacy among SD Latino adults.

In this study, nearly three-quarters (54/76, 71%) of participants reported not having personally sought OHI for themselves or others. Among those for whom eHealth literacy was assessed (22/76, 29%), nearly one-third (6/22, 27%) had an eHEALS score of less than 26, indicating self-perceived low eHealth literacy. In this study, preliminary findings suggest that higher educational attainment was associated with an increased likelihood of having sought OHI. Although further research to confirm directionality is necessary, this association between higher educational attainment and seeking OHI among SD Latino adults is consistent with prior findings on OHI behaviors

conducted among English-speaking populations [57,58]. Specifically, Jacobs et al [58] highlight that a heavy reliance on e-technologies for disseminating health information may increase the likelihood of further perpetuating health disparities; they suggest a need for interventions and efforts focused on developing training and services to boost internet self-efficacy tailored to patients' learning styles and their cultural and demographic characteristics to reduce this digital disparity.

This study was conducted among SD Latino adults of whom the majority were born outside of the United States, which is acutely different from studies conducted among participants preferring the English language [57,58]. Our efforts are preliminary and, thus, additional research is needed to explore the relationships between educational attainment and OHI seeking among various groups who prefer non-English languages and groups newly arrived in the United States. Of interest, in our exploration of factors associated with the act of OHI seeking, there was a high proportion of individuals who had not sought OHI (54/76, 71%). This lack of OHI seeking may disproportionately affect SD Latino adults, as reliance on eHealth resources continues to gain emphasis in the United States [3-5] and globally [59]. Indeed, our efforts are cautiously interpreted, yet our data signal a timely opportunity to examine the reasons preventing OHI-seeking behaviors among SD Latino adults aged 50 to 75 years, as well as the potential for intervention research. Thus, in recognizing this distinction,

considerations for culture and language are a fertile area of future research examining reasons for the lack of OHI seeking (eg, learning styles and preferences, availability and access to Spanish-language computer resources, knowledge, and technological skills, such as search strategies, among others). In considering culture, previous studies exploring OHI behaviors suggested that younger, English-dominant Latino and Hispanic generations may be OHI brokers for older generations of SD Latino adults [4,35,60]. Therefore, whether OHI seeking could be facilitated by intergenerational co-learning approaches is an avenue for future exploration. In a study examining digital health disparities among ethnically diverse older adults who prefer the English language [61], researchers suggested important steps to close this digital divide, including cultural adaptation based on preferences for receiving health information. Research has found that personal instruction through the process of internet use and assistance with using new digital devices helps older adults to adopt a daily habit of internet use [62]. However, whether these outcomes may be replicated among SD Latino adults aged 50 to 75 years remains to be examined. Further still, the literature communicates that smartphone ownership offers no statistically significant difference relevant to race or ethnicity [63] and is the primary source of internet use among all Latino adults [63]. Thus, considerations for examining the potential of leveraging smartphones for interventions on promoting OHI seeking among SD Latino adults is worth noting. Moreover, in addition to reaching underserved populations in health centers [64], community libraries offer another potential avenue for engaging with diverse community members [65].

In our study, exploratory findings suggest that higher eHealth literacy was preliminarily associated with higher self-efficacy for FIT screening ( $\beta=1.20$ , 95% CI 0.14-2.26;  $P=.02$ ) among SD Latino adults. This preliminary finding suggests that further investigation is warranted. Future in-depth examinations are necessary to confirm directionality of association. Prior to this study, the literature did not communicate specifically regarding findings that have examined eHealth literacy and self-efficacy for CRC screening. Efforts by Park et al [66] conveyed an association between higher levels of eHealth literacy and greater confidence in seeking online cancer information. Yet, the study was completed among English-language participants; importantly, the “confidence” examined by Park et al [66] differs from our use of “self-efficacy” that we occupy via the PHM.

This preliminary study communicates promising results suggesting that increased eHealth literacy may be associated with increased self-efficacy for CRC screening with FIT. While this study’s results must be replicated in a larger sample, these findings are encouraging and set an important small guiding step in suggesting that eHealth literacy skills development might benefit self-efficacy of CRC screening. Future appropriately powered research is necessary to examine whether eHealth literacy training and building self-efficacy for CRC screening may ultimately impact CRC screening uptake.

## Limitations and Strengths

This study has several limitations and strengths to acknowledge. First, the study was conducted in the context of a small pilot trial, underpowered for inferential analyses. Thus, our findings should be interpreted cautiously and call for further study in larger samples of SD Latino individuals. Second, this study did not assess access to the internet nor barriers to OHI seeking, limiting potential findings. Third, Latino ethnicity is shared among many Americans from various backgrounds and, most certainly, does not comprise a homogenous group. Characteristics of Latino adults vary by region. In this study, participants communicated their Latino ethnicity, and most were of Mexican heritage, thus potentially limiting generalizability. Yet, the sample representing 10 different Latin American countries offers preliminary evidence of feasibility for recruiting Latino adults from diverse backgrounds into future studies.

## Implications

There are several implications to consider and acknowledge. First, to reach the Healthy People 2030 HC and HIT objectives, it is necessary to expand research efforts to highlight and address barriers and leverage facilitators to OHI behaviors; it is also necessary to increase access to the internet. Second, future intervention research to promote eHealth literacy should examine and consider how cultural background and language affect access, health care use, successful use of eHealth resources, and behavior change. Finally, this study provided preliminary findings that suggested a need for pursuing further research on promoting OHI-seeking behaviors and eHealth literacy training that may impact CRC self-efficacy and, ultimately, CRC screening behaviors. Nevertheless, larger research studies are needed to corroborate these findings before clear implications for practice can be reached.

## Conclusions

This preliminary study adds to an extremely small evidence base and is the first to communicate findings on the assessment and analyses of OHI seeking and eHealth literacy in the context of the PHM among SD Latino adults. A high proportion of SD Latino participants in our study have not sought OHI for themselves or others, thus limiting their access to beneficial resources. In light of the growing use and reliance on technologies in health care, factors preventing OHI seeking remain to be further examined. In addition, there is a need for resources to improve eHealth literacy among SD Latino adults. Our study communicates preliminary evidence that higher eHealth literacy is occurring among those with higher self-efficacy for CRC screening. Appropriately powered research in the future is warranted to further examine this preliminary finding. Additionally, the next logical step that future research should examine is whether eHealth literacy training and building self-efficacy for CRC screening could increase the uptake of CRC screening.

## Acknowledgments

Funding for the project was provided by the Florida Department of Health–Bankhead Coley Cancer Research Program (4BB09; principal investigators [PIs]: CKG and CDM). The data collection efforts of EAC and SMC were supported by a grant from the National Cancer Institute (NCI) while they were postdoctoral fellows (grant R25CA090314; PIs: PB Jacobsen and TH Brandon). Support was provided by the Center for Biostatistics Collaboration and Data Service at the University of Texas Health Science Center at Houston. The publication of the research was supported, in part, by the Winship Cancer Institute of Emory University and the National Institutes of Health, NCI (award P30CA138292). The funders had no role in the study design; the collection, analysis, and interpretation of the data; the writing of the manuscript; or the decision to submit the manuscript for publication.

The authors communicate a special thank-you to Dr Colleen M McBride (Professor, Department of Behavioral, Social, and Health Education Sciences, Rollins School of Public Health, Emory University) for the valuable and much appreciated review of earlier drafts of this article.

## Conflicts of Interest

HM is the cofounder of SmartCT Studio LLC, which is not related to or involved in this published work.

## References

1. Prestin A, Vieux SN, Chou WS. Is online health activity alive and well or flatlining? Findings from 10 years of the health information national trends survey. *J Health Commun* 2015;20(7):790-798. [doi: [10.1080/10810730.2015.1018590](https://doi.org/10.1080/10810730.2015.1018590)] [Medline: [26042588](https://pubmed.ncbi.nlm.nih.gov/26042588/)]
2. Percheski C, Hargittai E. Health information-seeking in the digital age. *J Am Coll Health* 2011;59(5):379-386. [doi: [10.1080/07448481.2010.513406](https://doi.org/10.1080/07448481.2010.513406)] [Medline: [21500056](https://pubmed.ncbi.nlm.nih.gov/21500056/)]
3. Cutrona SL, Mazor KM, Vieux SN, Luger TM, Volkman JE, Finney Rutten LJ. Health information-seeking on behalf of others: Characteristics of "surrogate seekers". *J Cancer Educ* 2015 Mar;30(1):12-19 [FREE Full text] [doi: [10.1007/s13187-014-0701-3](https://doi.org/10.1007/s13187-014-0701-3)] [Medline: [24989816](https://pubmed.ncbi.nlm.nih.gov/24989816/)]
4. Chavarria EA, Chaney EH, Stollefson ML, Chaney JD, Chavarria N, Dodd VJ. Types and factors associated with online health information seeking among college men in Latino fraternities: A qualitative study. *Am J Mens Health* 2017 Nov;11(6):1692-1702 [FREE Full text] [doi: [10.1177/1557988315626510](https://doi.org/10.1177/1557988315626510)] [Medline: [26783157](https://pubmed.ncbi.nlm.nih.gov/26783157/)]
5. Bundorf MK, Wagner TH, Singer SJ, Baker LC. Who searches the internet for health information? *Health Serv Res* 2006 Jun;41(3 Pt 1):819-836 [FREE Full text] [doi: [10.1111/j.1475-6773.2006.00510.x](https://doi.org/10.1111/j.1475-6773.2006.00510.x)] [Medline: [16704514](https://pubmed.ncbi.nlm.nih.gov/16704514/)]
6. Rice RE. Influences, usage, and outcomes of internet health information searching: Multivariate results from the Pew surveys. *Int J Med Inform* 2006 Jan;75(1):8-28. [doi: [10.1016/j.ijmedinf.2005.07.032](https://doi.org/10.1016/j.ijmedinf.2005.07.032)] [Medline: [16125453](https://pubmed.ncbi.nlm.nih.gov/16125453/)]
7. Arora NK, Hesse BW, Rimer BK, Viswanath K, Clayman ML, Croyle RT. Frustrated and confused: The American public rates its cancer-related information-seeking experiences. *J Gen Intern Med* 2008 Mar;23(3):223-228 [FREE Full text] [doi: [10.1007/s11606-007-0406-y](https://doi.org/10.1007/s11606-007-0406-y)] [Medline: [17922166](https://pubmed.ncbi.nlm.nih.gov/17922166/)]
8. Fox S, Duggan M. Health Online 2013. Washington, DC: Pew Research Center's Internet & American Life Project; 2013 Jan 15. URL: <https://www.pewresearch.org/internet/2013/01/15/health-online-2013/> [accessed 2022-08-23]
9. Norman CD, Skinner HA. eHEALS: The eHealth Literacy Scale. *J Med Internet Res* 2006 Nov 14;8(4):e27 [FREE Full text] [doi: [10.2196/jmir.8.4.e27](https://doi.org/10.2196/jmir.8.4.e27)] [Medline: [17213046](https://pubmed.ncbi.nlm.nih.gov/17213046/)]
10. Neter E, Brainin E. eHealth literacy: Extending the digital divide to the realm of health information. *J Med Internet Res* 2012 Jan 27;14(1):e19 [FREE Full text] [doi: [10.2196/jmir.1619](https://doi.org/10.2196/jmir.1619)] [Medline: [22357448](https://pubmed.ncbi.nlm.nih.gov/22357448/)]
11. Sanchez MA, Rabin BA, Gaglio B, Henton M, Elzarrad MK, Purcell P, et al. A systematic review of eHealth cancer prevention and control interventions: New technology, same methods and designs? *Transl Behav Med* 2013 Dec;3(4):392-401 [FREE Full text] [doi: [10.1007/s13142-013-0224-1](https://doi.org/10.1007/s13142-013-0224-1)] [Medline: [24294327](https://pubmed.ncbi.nlm.nih.gov/24294327/)]
12. Slev VN, Mistiaen P, Pasmán HRW, Verdonck-de Leeuw IM, van Uden-Kraan CF, Francke AL. Effects of eHealth for patients and informal caregivers confronted with cancer: A meta-review. *Int J Med Inform* 2016 Mar;87:54-67. [doi: [10.1016/j.ijmedinf.2015.12.013](https://doi.org/10.1016/j.ijmedinf.2015.12.013)] [Medline: [26806712](https://pubmed.ncbi.nlm.nih.gov/26806712/)]
13. Horvath KJ, Bauermeister JA. eHealth literacy and intervention tailoring impacts the acceptability of a HIV/STI testing intervention and sexual decision making among young gay and bisexual men. *AIDS Educ Prev* 2017 Feb;29(1):14-23 [FREE Full text] [doi: [10.1521/aeap.2017.29.1.14](https://doi.org/10.1521/aeap.2017.29.1.14)] [Medline: [28195779](https://pubmed.ncbi.nlm.nih.gov/28195779/)]
14. Increase the proportion of people who can view, download, and send their electronic health information — HC/HIT-D09. US Department of Health and Human Services, Healthy People 2030. 2020. URL: <https://tinyurl.com/3t9e6az2> [accessed 2022-08-23]
15. Increase the proportion of people who say their online medical record is easy to understand — HC/HIT-D10. US Department of Health and Human Services, Healthy People 2030. 2020. URL: <https://health.gov/healthypeople/objectives-and-data/browse-objectives/health-it/increase-proportion-people-who-say-their-online-medical-record-easy-understand-hchit-d10> [accessed 2022-08-23]

16. McCloud RF, Okechukwu CA, Sorensen G, Viswanath K. Beyond access: Barriers to internet health information seeking among the urban poor. *J Am Med Inform Assoc* 2016 Nov;23(6):1053-1059 [FREE Full text] [doi: [10.1093/jamia/ocv204](https://doi.org/10.1093/jamia/ocv204)] [Medline: [27206459](https://pubmed.ncbi.nlm.nih.gov/27206459/)]
17. Nguyen A, Mosadeghi S, Almario CV. Persistent digital divide in access to and use of the internet as a resource for health information: Results from a California population-based study. *Int J Med Inform* 2017 Jul;103:49-54. [doi: [10.1016/j.ijmedinf.2017.04.008](https://doi.org/10.1016/j.ijmedinf.2017.04.008)] [Medline: [28551001](https://pubmed.ncbi.nlm.nih.gov/28551001/)]
18. Goldberg R, Robinson A, Carlson E. Digital divide is shrinking for America's Hispanic population, NTIA data show. National Telecommunications and Information Administration. 2019 Oct 28. URL: <https://www.ntia.gov/blog/2019/digital-divide-shrinking-america-s-hispanic-population-ntia-data-show> [accessed 2022-08-23]
19. Anna A, López G, Lopez MH. Digital Divide Narrows for Latinos as More Spanish Speakers and Immigrants Go Online. Washington, DC: Pew Research Center; 2016 Jul. URL: [https://www.pewresearch.org/hispanic/wp-content/uploads/sites/5/2016/07/PH\\_2016.07.21\\_Broadbank\\_Final.pdf](https://www.pewresearch.org/hispanic/wp-content/uploads/sites/5/2016/07/PH_2016.07.21_Broadbank_Final.pdf) [accessed 2022-08-23]
20. Lee YJ, Boden-Albala B, Larson E, Wilcox A, Bakken S. Online health information seeking behaviors of Hispanics in New York City: A community-based cross-sectional study. *J Med Internet Res* 2014 Jul 22;16(7):e176 [FREE Full text] [doi: [10.2196/jmir.3499](https://doi.org/10.2196/jmir.3499)] [Medline: [25092120](https://pubmed.ncbi.nlm.nih.gov/25092120/)]
21. Cancer Facts & Figures for Hispanics/Latinos 2018-2020. Atlanta, GA: American Cancer Society; 2018. URL: <https://tinyurl.com/5949prpv> [accessed 2022-08-23]
22. Myers RE, Ross E, Jepson C, Wolf T, Balshem A, Millner L, et al. Modeling adherence to colorectal cancer screening. *Prev Med* 1994 Mar;23(2):142-151. [doi: [10.1006/pmed.1994.1020](https://doi.org/10.1006/pmed.1994.1020)] [Medline: [8047519](https://pubmed.ncbi.nlm.nih.gov/8047519/)]
23. Christy SM, Sutton SK, Gwede CK, Chavarria EA, Davis SN, Abdulla R, et al. Examining the durability of colorectal cancer screening awareness and health beliefs among medically underserved patients: Baseline to 12 months post-intervention. *J Cancer Educ* 2017 Nov 25;34(2):297-303. [doi: [10.1007/s13187-017-1301-9](https://doi.org/10.1007/s13187-017-1301-9)]
24. Tiro JA. Factorial validity and invariance of a survey measuring psychosocial correlates of colorectal cancer screening among African Americans and Caucasians. *Cancer Epidemiol Biomarkers Prev* 2005 Dec 01;14(12):2855-2861. [doi: [10.1158/1055-9965.epi-05-0217](https://doi.org/10.1158/1055-9965.epi-05-0217)]
25. Vernon SW, Myers RE, Tilley BC, Li S. Factors associated with perceived risk in automotive employees at increased risk of colorectal cancer. *Cancer Epidemiol Biomarkers Prev* 2001 Jan;10(1):35-43. [Medline: [11205487](https://pubmed.ncbi.nlm.nih.gov/11205487/)]
26. Myers RE, Sifri R, Hyslop T, Rosenthal M, Vernon SW, Cocroft J, et al. A randomized controlled trial of the impact of targeted and tailored interventions on colorectal cancer screening. *Cancer* 2007 Nov 01;110(9):2083-2091 [FREE Full text] [doi: [10.1002/ncr.23022](https://doi.org/10.1002/ncr.23022)] [Medline: [17893869](https://pubmed.ncbi.nlm.nih.gov/17893869/)]
27. Davis SN, Wischhusen JW, Sutton SK, Christy SM, Chavarria EA, Sutter ME, et al. Demographic and psychosocial factors associated with limited health literacy in a community-based sample of older Black Americans. *Patient Educ Couns* 2020 Feb;103(2):385-391 [FREE Full text] [doi: [10.1016/j.pec.2019.08.026](https://doi.org/10.1016/j.pec.2019.08.026)] [Medline: [31466881](https://pubmed.ncbi.nlm.nih.gov/31466881/)]
28. Christy SM, Gwede CK, Sutton SK, Chavarria E, Davis SN, Abdulla R, et al. Health literacy among medically underserved: The role of demographic factors, social influence, and religious beliefs. *J Health Commun* 2017 Nov;22(11):923-931 [FREE Full text] [doi: [10.1080/10810730.2017.1377322](https://doi.org/10.1080/10810730.2017.1377322)] [Medline: [29125435](https://pubmed.ncbi.nlm.nih.gov/29125435/)]
29. Christy S, Cousin L, Sutton S. Characterizing health literacy among Spanish language-preferring Latinos ages 50-75. *Nurs Res* 2021;70(5):344-353. [doi: [10.1097/nnr.0000000000000519](https://doi.org/10.1097/nnr.0000000000000519)]
30. Gwede C, Sutton S, Chavarria E, Gutierrez L, Abdulla R, Christy SM, et al. A culturally and linguistically salient pilot intervention to promote colorectal cancer screening among Latinos receiving care in a Federally Qualified Health Center. *Health Educ Res* 2019 Jun 01;34(3):310-320 [FREE Full text] [doi: [10.1093/her/cyz010](https://doi.org/10.1093/her/cyz010)] [Medline: [30929015](https://pubmed.ncbi.nlm.nih.gov/30929015/)]
31. Aguado Loi CX, Martinez Tyson D, Chavarria EA, Gutierrez L, Klasko L, Davis S, et al. 'Simple and easy:' Providers' and Latinos' perceptions of the fecal immunochemical test (FIT) for colorectal cancer screening. *Ethn Health* 2020 Feb;25(2):206-221 [FREE Full text] [doi: [10.1080/13557858.2017.1418298](https://doi.org/10.1080/13557858.2017.1418298)] [Medline: [29319331](https://pubmed.ncbi.nlm.nih.gov/29319331/)]
32. Brislin RW. Back-translation for cross-cultural research. *J Cross Cult Psychol* 2016 Jul 27;1(3):185-216. [doi: [10.1177/135910457000100301](https://doi.org/10.1177/135910457000100301)]
33. Doak CC, Doak LG, Root JH. Learner verification and revision of materials. In: *Teaching Patients With Low Literacy Skills*. 2nd edition. Philadelphia, PA: J. B. Lippincott Company; 1996:167-188.
34. Chavarria EA, Christy SM, Simmons VN, Vadaparampil ST, Gwede CK, Meade CD. Learner verification: A methodology to create suitable education materials. *Health Lit Res Pract* 2021 Jan 11;5(1):e49-e59 [FREE Full text] [doi: [10.3928/24748307-20210201-02](https://doi.org/10.3928/24748307-20210201-02)] [Medline: [34251862](https://pubmed.ncbi.nlm.nih.gov/34251862/)]
35. Chavarria EA, Chaney BH, Stollefson ML, Chaney JD, Chavarria NE, Dodd VJ. Predictors of online health information-seeking behaviors among college Latino men. *Health Behav Policy Rev* 2017 Jul 01;4(4):367-379. [doi: [10.14485/hbpr.4.4.6](https://doi.org/10.14485/hbpr.4.4.6)]
36. Paramio Pérez G, Almagro BJ, Hernando Gómez Á, Aguaded Gómez JI. [Validation of the eHealth Literacy Scale (eHEALS) in Spanish university students] [Article in Spanish]. *Rev Esp Salud Publica* 2015;89(3):329-338 [FREE Full text] [doi: [10.4321/S1135-57272015000300010](https://doi.org/10.4321/S1135-57272015000300010)] [Medline: [26388346](https://pubmed.ncbi.nlm.nih.gov/26388346/)]

37. Collins SA, Currie LM, Bakken S, Vawdrey DK, Stone PW. Health literacy screening instruments for eHealth applications: A systematic review. *J Biomed Inform* 2012 Jun;45(3):598-607 [[FREE Full text](#)] [doi: [10.1016/j.jbi.2012.04.001](https://doi.org/10.1016/j.jbi.2012.04.001)] [Medline: [2521719](https://pubmed.ncbi.nlm.nih.gov/2521719/)]
38. Milne RA, Puts MTE, Papadakos J, Le LW, Milne VC, Hope AJ, et al. Predictors of high eHealth literacy in primary lung cancer survivors. *J Cancer Educ* 2015 Dec;30(4):685-692. [doi: [10.1007/s13187-014-0744-5](https://doi.org/10.1007/s13187-014-0744-5)] [Medline: [25355524](https://pubmed.ncbi.nlm.nih.gov/25355524/)]
39. Robinson C, Graham J. Perceived internet health literacy of HIV-positive people through the provision of a computer and internet health education intervention. *Health Info Libr J* 2010 Dec;27(4):295-303 [[FREE Full text](#)] [doi: [10.1111/j.1471-1842.2010.00898.x](https://doi.org/10.1111/j.1471-1842.2010.00898.x)] [Medline: [21050372](https://pubmed.ncbi.nlm.nih.gov/21050372/)]
40. Xie B. Experimenting on the impact of learning methods and information presentation channels on older adults' e-health literacy. *J Am Soc Inf Sci Technol* 2011 Jun 06;62(9):1797-1807. [doi: [10.1002/asi.21575](https://doi.org/10.1002/asi.21575)]
41. Chung S, Nahm E. Testing reliability and validity of the eHealth Literacy Scale (eHEALS) for older adults recruited online. *Comput Inform Nurs* 2015 Apr;33(4):150-156 [[FREE Full text](#)] [doi: [10.1097/CIN.000000000000146](https://doi.org/10.1097/CIN.000000000000146)] [Medline: [25783223](https://pubmed.ncbi.nlm.nih.gov/25783223/)]
42. Nguyen J, Moorhouse M, Curbow B, Christie J, Walsh-Childers K, Islam S. Construct validity of the eHealth Literacy Scale (eHEALS) among two adult populations: A Rasch analysis. *JMIR Public Health Surveill* 2016 May 20;2(1):e24 [[FREE Full text](#)] [doi: [10.2196/publichealth.4967](https://doi.org/10.2196/publichealth.4967)] [Medline: [27244771](https://pubmed.ncbi.nlm.nih.gov/27244771/)]
43. De Caro W, Corvo E, Marucci A, Mitello L, Lancia L, Sansoni J. eHealth Literacy Scale: A nursing analysis and Italian validation. *Stud Health Technol Inform* 2016;225:949. [Medline: [27332425](https://pubmed.ncbi.nlm.nih.gov/27332425/)]
44. Richtering SS, Hyun K, Neubeck L, Coorey G, Chalmers J, Usherwood T, et al. eHealth literacy: Predictors in a population with moderate-to-high cardiovascular risk. *JMIR Hum Factors* 2017 Jan 27;4(1):e4 [[FREE Full text](#)] [doi: [10.2196/humanfactors.6217](https://doi.org/10.2196/humanfactors.6217)] [Medline: [28130203](https://pubmed.ncbi.nlm.nih.gov/28130203/)]
45. Tennant B, Stellefson M, Dodd V, Chaney B, Chaney D, Paige S, et al. eHealth literacy and Web 2.0 health information seeking behaviors among baby boomers and older adults. *J Med Internet Res* 2015 Mar 17;17(3):e70 [[FREE Full text](#)] [doi: [10.2196/jmir.3992](https://doi.org/10.2196/jmir.3992)] [Medline: [25783036](https://pubmed.ncbi.nlm.nih.gov/25783036/)]
46. Choi NG, Dinitto DM. The digital divide among low-income homebound older adults: Internet use patterns, eHealth literacy, and attitudes toward computer/internet use. *J Med Internet Res* 2013 May 02;15(5):e93 [[FREE Full text](#)] [doi: [10.2196/jmir.2645](https://doi.org/10.2196/jmir.2645)] [Medline: [23639979](https://pubmed.ncbi.nlm.nih.gov/23639979/)]
47. van der Vaart R, van Deursen AJ, Drossaert CH, Taal E, van Dijk JA, van de Laar MA. Does the eHealth Literacy Scale (eHEALS) measure what it intends to measure? Validation of a Dutch version of the eHEALS in two adult populations. *J Med Internet Res* 2011 Nov 09;13(4):e86 [[FREE Full text](#)] [doi: [10.2196/jmir.1840](https://doi.org/10.2196/jmir.1840)] [Medline: [22071338](https://pubmed.ncbi.nlm.nih.gov/22071338/)]
48. Mitsutake S, Shibata A, Ishii K, Okazaki K, Oka K. [Developing Japanese version of the eHealth Literacy Scale (eHEALS)] [Article in Japanese]. *Nihon Kosshu Eisei Zasshi* 2011 May;58(5):361-371. [Medline: [21905612](https://pubmed.ncbi.nlm.nih.gov/21905612/)]
49. McQueen A, Tiro JA, Vernon SW. Construct validity and invariance of four factors associated with colorectal cancer screening across gender, race, and prior screening. *Cancer Epidemiol Biomarkers Prev* 2008 Sep 02;17(9):2231-2237. [doi: [10.1158/1055-9965.epi-08-0176](https://doi.org/10.1158/1055-9965.epi-08-0176)]
50. Vernon SW, Myers RE, Tilley BC. Development and validation of an instrument to measure factors related to colorectal cancer screening adherence. *Cancer Epidemiol Biomarkers Prev* 1997 Oct;6(10):825-832. [Medline: [9332766](https://pubmed.ncbi.nlm.nih.gov/9332766/)]
51. Davis SN, Christy SM, Chavarria EA, Abdulla R, Sutton SK, Schmidt AR, et al. A randomized controlled trial of a multicomponent, targeted, low-literacy educational intervention compared with a nontargeted intervention to boost colorectal cancer screening with fecal immunochemical testing in community clinics. *Cancer* 2017 Apr 15;123(8):1390-1400 [[FREE Full text](#)] [doi: [10.1002/cncr.30481](https://doi.org/10.1002/cncr.30481)] [Medline: [27906448](https://pubmed.ncbi.nlm.nih.gov/27906448/)]
52. Gwede CK, William CM, Thomas KB, Tarver WL, Quinn GP, Vadapampil ST, et al. Exploring disparities and variability in perceptions and self-reported colorectal cancer screening among three ethnic subgroups of US Blacks. *Oncol Nurs Forum* 2010 Aug 26;37(5):581-591. [doi: [10.1188/10.onf.581-591](https://doi.org/10.1188/10.onf.581-591)]
53. Christy SM, Schmidt A, Wang HL. Understanding cancer worry among patients in a community clinic-based colorectal cancer screening intervention study. *Nurs Res* 2018;67(4):275-285. [doi: [10.1097/nnr.0000000000000275](https://doi.org/10.1097/nnr.0000000000000275)]
54. Christy SM, Davis SN, Williams KR, Zhao X, Govindaraju SK, Quinn GP, et al. A community-based trial of educational interventions with fecal immunochemical tests for colorectal cancer screening uptake among Blacks in community settings. *Cancer* 2016 Nov 15;122(21):3288-3296 [[FREE Full text](#)] [doi: [10.1002/cncr.30207](https://doi.org/10.1002/cncr.30207)] [Medline: [27420119](https://pubmed.ncbi.nlm.nih.gov/27420119/)]
55. Flight IH, Wilson CJ, McGillivray J, Myers RE. Cross-cultural validation of the preventive health model for colorectal cancer screening: An Australian study. *Health Educ Behav* 2010 Oct;37(5):724-736. [doi: [10.1177/1090198110364107](https://doi.org/10.1177/1090198110364107)] [Medline: [20930134](https://pubmed.ncbi.nlm.nih.gov/20930134/)]
56. Patient Engagement Playbook. Washington, DC: US Department of Health and Human Services, Office of the National Coordinator for Health Information Technology URL: <https://www.healthit.gov/playbook/pe/> [accessed 2022-08-23]
57. Lee HY, Jin SW, Henning-Smith C, Lee J, Lee J. Role of health literacy in health-related information-seeking behavior online: Cross-sectional study. *J Med Internet Res* 2021 Jan 27;23(1):e14088 [[FREE Full text](#)] [doi: [10.2196/14088](https://doi.org/10.2196/14088)] [Medline: [33502332](https://pubmed.ncbi.nlm.nih.gov/33502332/)]
58. Jacobs W, Amuta AO, Jeon KC. Health information seeking in the digital age: An analysis of health information seeking behavior among US adults. *Cogent Soc Sci* 2017 Mar 13;3(1):1302785 [[FREE Full text](#)] [doi: [10.1080/23311886.2017.1302785](https://doi.org/10.1080/23311886.2017.1302785)]



59. Global Diffusion of eHealth: Making Universal Health Coverage Achievable. Report of the Third Global Survey on eHealth. Geneva, Switzerland: World Health Organization; 2016. URL: <http://apps.who.int/iris/bitstream/handle/10665/252529/9789241511780-eng.pdf> [accessed 2022-08-23]
60. Gonzalez C, Bollinger B, Yip J, Pina L, Roldan W, Nieto Ruiz C. Intergenerational online health information searching and brokering: Framing health literacy as a family asset. *Health Commun* 2022 Apr;37(4):438-449. [doi: [10.1080/10410236.2020.1847445](https://doi.org/10.1080/10410236.2020.1847445)] [Medline: [33205663](https://pubmed.ncbi.nlm.nih.gov/33205663/)]
61. Tappen RM, Cooley ME, Luckmann R, Panday S. Digital health information disparities in older adults: A mixed methods study. *J Racial Ethn Health Disparities* 2022 Feb;9(1):82-92 [FREE Full text] [doi: [10.1007/s40615-020-00931-3](https://doi.org/10.1007/s40615-020-00931-3)] [Medline: [33415705](https://pubmed.ncbi.nlm.nih.gov/33415705/)]
62. Smith A. Older Adults and Technology Use. Washington, DC: Pew Research Center; 2014 Apr 03. URL: <https://www.pewresearch.org/internet/2014/04/03/older-adults-and-technology-use/> [accessed 2022-08-23]
63. Atske S, Perrin A. Home broadband adoption, computer ownership vary by race, ethnicity in the US. Pew Research Center. 2021 Jul 16. URL: <https://www.pewresearch.org/fact-tank/2021/07/16/home-broadband-adoption-computer-ownership-vary-by-race-ethnicity-in-the-u-s/> [accessed 2022-08-23]
64. Davis S, Govindaraju S, Jackson B. Recruitment techniques and strategies in a community-based colorectal cancer screening study of men and women of African ancestry. *Nurs Res* 2018;67(3):212-221. [doi: [10.1097/nnr.0000000000000274](https://doi.org/10.1097/nnr.0000000000000274)]
65. Whitney W, Keselman A, Humphreys B. Libraries and librarians: Key partners for progress in health literacy research and practice. *Inf Serv Use* 2017 Mar 07;37(1):85-100. [doi: [10.3233/isu-170821](https://doi.org/10.3233/isu-170821)]
66. Park H, Moon M, Baeg JH. Association of eHealth literacy with cancer information seeking and prior experience with cancer screening. *Comput Inform Nurs* 2014 Sep;32(9):458-463. [doi: [10.1097/CIN.0000000000000077](https://doi.org/10.1097/CIN.0000000000000077)] [Medline: [25105588](https://pubmed.ncbi.nlm.nih.gov/25105588/)]

## Abbreviations

**CRC:** colorectal cancer

**eHEALS:** eHealth Literacy Scale

**FIT:** fecal immunochemical test

**FQHC:** Federally Qualified Health Center

**HC:** health communication

**HIT:** health information technology

**HRQOL:** health-related quality of life

**Latino CARES:** Latino Colorectal Cancer Awareness, Research, Education, and Screening

**NCI:** National Cancer Institute

**OHI:** online health information

**PHM:** Preventive Health Model

**PI:** principal investigator

**RQ1:** research question 1

**RQ2:** research question 2

**RQ3:** research question 3

**RQ4:** research question 4

**SD:** Spanish-dominant

*Edited by A Mavragani; submitted 03.03.22; peer-reviewed by S Song, H Liu; comments to author 25.04.22; revised version received 12.07.22; accepted 09.08.22; published 13.10.22.*

*Please cite as:*

Chavarria EA, Christy SM, Feng H, Miao H, Abdulla R, Gutierrez L, Lopez D, Sanchez J, Gwede CK, Meade CD

Online Health Information Seeking and eHealth Literacy Among Spanish Language-Dominant Latino Adults Receiving Care in a Community Clinic: Secondary Analysis of Pilot Randomized Controlled Trial Data

*JMIR Form Res* 2022;6(10):e37687

URL: <https://formative.jmir.org/2022/10/e37687>

doi: [10.2196/37687](https://doi.org/10.2196/37687)

PMID: [35238785](https://pubmed.ncbi.nlm.nih.gov/35238785/)

©Enmanuel A Chavarria, Shannon M Christy, Han Feng, Hongyu Miao, Rania Abdulla, Liliana Gutierrez, Diana Lopez, Julian Sanchez, Clement K Gwede, Cathy D Meade. Originally published in *JMIR Formative Research* (<https://formative.jmir.org/>), 13.10.2022. 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

# Evaluating the Feasibility of a Digital Therapeutic Program for Patients With Cancer During Active Treatment: Pre-Post Interventional Study

G Haukur Gudmundsson<sup>1</sup>, MSc; Judit Mészáros<sup>2</sup>, PhD; Ágústa E Björnsdóttir<sup>1</sup>, MA; María L Ámundadóttir<sup>1</sup>, PhD; Gudrun E Thorvardardóttir<sup>3</sup>, BSc; Erna Magnúsdóttir<sup>4</sup>, MSc; Halla Helgadóttir<sup>1</sup>, MSc; Saemundur Oddsson<sup>1</sup>, MD

<sup>1</sup>Medical and Research Department, Sidekick Health, Kopavogur, Iceland

<sup>2</sup>Medical and Research Department, Sidekick Health, Berlin, Germany

<sup>3</sup>Physical rehabilitation, Ljosid Cancer Rehabilitation Centre, Reykjavík, Iceland

<sup>4</sup>Management, Ljosid Cancer Rehabilitation Centre, Reykjavík, Iceland

**Corresponding Author:**

Judit Mészáros, PhD

Medical and Research Department

Sidekick Health

Oberwallstraße 6

Berlin, 10117

Germany

Phone: 49 1744789708

Email: [judit@sidekickhealth.com](mailto:judit@sidekickhealth.com)

## Abstract

**Background:** Increasing evidence shows that lifestyle interventions can improve the symptoms, quality of life (QoL), and even overall survival of patients with cancer. Digital therapeutics (DTx) can help implement behavioral modifications and empower patients through education, lifestyle support, and remote symptom monitoring.

**Objective:** We aimed to test the feasibility of a DTx program for patients with cancer, as measured by engagement, retention, and acceptability. In addition, we explored the effects of the program on cancer-related QoL.

**Methods:** We conducted a 4-week single-arm trial in Iceland, where DTx was delivered through a smartphone app. The intervention consisted of patient education about mindfulness, sleep, stress, and nutrition; lifestyle coaching; and the completion of daily missions for tracking physical activity and exercise, reporting patient-reported outcomes (PROs), practicing mindfulness, and logging healthy food intake. Information on program engagement and retention, step goal attainment, as well as PROs were collected throughout the study. QoL was measured using the European Organization for Research and Treatment of Cancer Quality of Life Questionnaire C30 at baseline and follow-up.

**Results:** In total, 30 patients with cancer undergoing active therapy were enrolled, and 29 registered in the app (23 female, 18 with breast cancer; mean age 52.6, SD 11.5 years). Overall, 97% (28/29) of participants were active in 3 of the 4 weeks and completed the pre- and postprogram questionnaires. The weekly active days (median) were 6.8 (IQR 5.8-6.8), and 72% (21/29) of participants were active at least 5 days a week. Users interacted with the app on average 7.7 (SD 1.9) times per day. On week 1, all 29 participants used the step counter and logged an average of 20,306 steps; 21 (72%) participants reached their step goals of at least 3000 steps per day. On week 4, of the 28 active users, 27 (96%) were still logging their steps, with 19 (68%) reaching their step goals. Of the 28 participants who completed the satisfaction questionnaire, 25 (89%) were likely to recommend the program, 23 (82%) said the program helped them deal with the disease, and 24 (86%) said it helped them remember their medication. QoL assessment showed that the average global health status, functioning, and symptom burden remained stable from baseline to follow-up. In all, 50% (14/28) of participants reported less pain, and the average pain score decreased from 31 (SD 20.1) to 22.6 (SD 23.2;  $P=.16$ ). There was no significant change in PROs on the quality of sleep, energy, and stress levels from the first to the last week.

**Conclusions:** The high retention, engagement, and acceptability found in this study demonstrate that multidisciplinary DTx is feasible for patients with cancer. A longer, full-scale randomized controlled trial is currently being planned to evaluate the efficacy of the intervention.

**KEYWORDS**

cancer; lifestyle; quality of life; mobile app; digital therapeutics; self-management; physical activity; mobile phone

## Introduction

According to the latest statistics, the global prevalence of all types of cancers is projected to increase by nearly 50% in the next 20 years, with female breast cancer being the most prevalent in 2020 [1]. In Iceland, approximately 1700 people received a new cancer diagnosis in 2018, but due to improved awareness, early detection, and treatment options, survival rates have been increasing, and by 2018, there were >15,000 people living with a previously diagnosed cancer [2,3]. Thus, there is now an increased population of people living a long, productive life with a history of cancer. Nevertheless, the side effects of current treatments, as well as the stress associated with the diagnosis and fear of disease recurrence, pose a serious burden on patients' mental and physical well-being [4,5].

Research over the last 2 decades has shown that lifestyle modifications can effectively improve the quality of life (QoL) of patients with cancer. Mindfulness exercises, muscle relaxation, and cognitive behavioral therapy can help patients cope with stress [6]. Traditionally, patients were advised to rest while undergoing cancer treatment, but guidelines now recommend avoiding sedentary behaviors and doing regular aerobic and resistance training [7]. Increasing evidence suggests that regular physical activity can help combat disease-related physical and psychological symptoms and improve QoL [7]. Research has shown that more walking during recovery can reduce the chance of readmission after cancer surgery and engaging in home-based exercise programs can aid people in increasing their physical fitness [8,9]. In addition, healthy dietary habits are equally important to maintain for patients with cancer and cancer survivors [10]. However, lifestyle and related behavioral modifications are often hard to achieve, and there is a need for a structured implementation of lifestyle support for patients.

The advent of digital technology and the wide reach of smartphones provide a potential avenue for motivating and delivering structured lifestyle programs for patients. Several digital intervention programs have been developed for patients with cancer and cancer survivors to provide psychological support and help manage symptoms [11]. Studies integrating electronic patient-reported outcomes (PROs) and advice on symptom management in care found reduced fatigue and increased QoL and overall survival [12-14]. In addition, collecting electronic PROs can foster more efficient patient-health care provider interaction and thus reduce the need for consultations [15]. Studies on digital therapeutics (DTx) for patients with cancer have been largely focused on providing and evaluating psychosocial interventions. These studies found that electronically delivered cognitive behavioral therapy increases the ability to relax and decreases anxiety in men with prostate cancer [16], combining it with exercise and patient education can increase QoL, reduce fatigue, and improve dietary habits [17,18]. Furthermore, a study found that an internet-based

exercise program in itself can improve cognitive abilities in patients with breast cancer for 6 months after the intervention [19]. Thus, existing evidence points toward the benefits of a holistic digital therapeutic that combines multiple aspects of patient support during cancer treatment (psychological support, education, exercise, and nutrition).

The primary aim of this study was to assess the feasibility of a holistic DTx program to improve the lifestyle and health-related QoL of patients in active anticancer therapy. This was measured in terms of user engagement, retention, acceptability, and step goal attainment. An additional objective of the study was to gather preliminary indications of the program's efficacy through secondary endpoint measures. The results of this feasibility trial will be used to inform a future definitive randomized controlled trial (RCT).

## Methods

### Study Design and Recruitment

We conducted a 4-week single-arm trial from August to November 2021 at the Ljósið cancer rehabilitation clinic in Iceland. Patients were invited to participate in the study, which was promoted as a support program aimed at improving QoL for patients with cancer, via emails and educational lectures at the clinic, and they were recruited after voluntarily reaching out. Inclusion criteria were (1) diagnosed with cancer and receiving anticancer treatment at the National University Hospital of Iceland (chemotherapy, radiation therapy, or other nonhormonal cancer medication) at the start of participation; (2) aged  $\geq 18$  years; (3) speaks Icelandic; (4) has the capacity to give informed consent; and (5) owns and knows how to operate a smartphone.

### Ethical Considerations

All participants provided informed consent before enrolling in the study. The protocol was approved by the National Bioethics Committee (institutional review board registration number VSN-21-102). This study was conducted in accordance with the ethical principles of the Declaration of Helsinki 2008.

### Procedures

After signing informed consent forms, participants completed a preintervention QoL questionnaire on the web, and exercise physiologists or physical therapists at the cancer clinic collected information about baseline physical measurements, physical fitness and body composition. Participants were also instructed to download the Sidekick smartphone app and received an access code to the program. During the intervention, data on participants' retention, engagement, and self-reported in-app activity were collected through the app. QoL, fitness, and body composition measurements were repeated after the intervention, and participants' feedback on the program was collected through a web-based satisfaction questionnaire.

### Intervention

The program was delivered through the Sidekick app, which was created by a group of data scientists, designers, gamification experts, behavioral scientists, psychologists, medical doctors, and other health care professionals and uses the principles of behavioral economics combined with gamification elements to achieve behavioral modifications for the primary and secondary prevention of lifestyle-related chronic diseases [20]. Gamification elements in the app are intended to motivate and engage users and examples include the collection of digital water drops—which are converted into donations for charity—when completing any task in the app; a leveling system and progress bar based on the water drops collected; and a checklist for completion of daily tasks. In addition, the visual appearance, animations, and sounds within the app were designed to be enjoyable, pleasing, and mimic games.

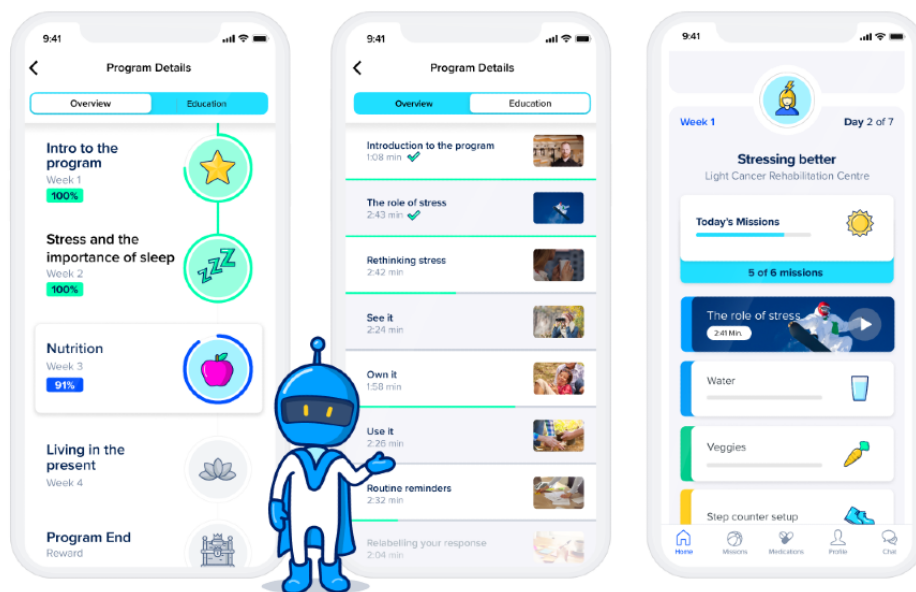
The intervention was designed to provide patients with cancer with tools to better deal with side effects during cancer treatment

and to improve their overall QoL, with the main focus on stress management and improving sleeping habits. A general program overview is provided in Table 1 and Figure 1, and overview of the educational content is provided in the Multimedia Appendix 1. The intervention was designed with inputs from oncologists and was composed of assigned daily missions for educational materials relating to each weekly theme (stress, sleep, nutrition, and mindfulness from week 1 to 4), step counter, food logging, meditation, and PRO surveys. Missions were defined as any assignment available in the app, and users received on average 5 of these a day. Users could also proactively explore the app for other content that was not assigned (eg, exercises and medication reminders). An additional feature was remote patient monitoring in which an exercise physiologist with experience in cancer rehabilitation provided lifestyle coaching through weekly motivational messages and general support with the program.

**Table 1.** Presentation of main program missions and expected outcomes.

Mission type	Content or action	Outcome
Education	Videos and other content every day	Increased knowledge of tools for building healthy habits
Food logger	Log vegetables and water consumed every day	Healthier eating habits and better nutrition
Step counter	Log number of steps every day	Increased awareness and motivation for physical activity
Guided meditation	Meditate 3 times per week	Improved mental health and stress management
Patient-reported outcomes	Indicate energy levels, stress levels, and quality of sleep 3 times per week	Increased self-awareness

**Figure 1.** The Sidekick app interface. Program overview (left): the program was composed of 4 weekly modules with introduction on week 1, followed by modules on stress and sleep, nutrition, and mindfulness in the following weeks. Users could access educational videos in each topic (middle) as part of their daily missions. The missions included reaching step goals and logging food and water intake (right).



### Primary Outcomes

Retention and engagement were assessed using data on user-reported interactions by completing missions on the Sidekick app. Treatment completion was defined as being active

in-app at least 3 of 4 weeks (or 75% of the program duration) and completing the prestudy and poststudy QoL questionnaires. Although previous studies found that attrition rates can be as high as 50% [21-23], we based this more ambitious target on

previous engagement and completion rates found on our platform. Program retention was defined as the number of users who returned to the app during the last week of the study, and highly engaged users were those who were active in-app for at least five days a week. We focused on 2 measures of engagement: the number of days the participant was active in the app and the average number of mission interactions during active days.

## Secondary Outcomes

Participants' health-related QoL was assessed using the European Organization for Research and Treatment of Cancer Quality of Life Questionnaire C30 (QLQ-C30). QLQ-C30 is a clinically validated and well-established questionnaire composed of 30 questions that measure QoL across 3 domains: functional scales (physical, role, emotional, cognitive, and social), symptom scales (fatigue, nausea or vomiting, pain, dyspnea, insomnia, appetite loss, constipation, diarrhea, and financial difficulties), and global health status [24].

Aerobic fitness was measured using the Åstrand submaximal test on a cycle ergometer (Monark) [25]. In this test, the participants were instructed to cycle for 6 minutes with a pedal frequency of 50 to 70 rpm and reach approximately 85% of the maximum heart rate based on age. Their maximum oxygen uptake ( $VO_{2max}$ ; mL/kg/min) was estimated based on the heart rate, workload (in watt), age, sex, body weight, and heart rate at the end of the test.

Body composition was measured using the InBody 770 bioelectrical impedance analyzer [26,27]. The participants were instructed to stand on the device platform with bare feet, stand upright, and hold on to the handles. Built-in scales and electrodes in the device measure body composition values, such as weight, fat mass, and fat-free mass (lean mass). Participants were instructed not to consume food at least two hours before their appointment, refrain from physical exercise earlier that day, and avoid putting creams or lotions on their bodies.

Step counts were automatically measured by the app based on data from the built-in accelerometer of the smartphone. However, at the end of each day, the steps had to be manually registered by the user by clicking *claim steps* in the app. The minimum step goal was set at 15,000 per week or 3000 at least 5 days per week.

PROs on quality of sleep, stress, and energy levels were measured using the app 3 times a week. Users received these as part of their daily missions on 3 random days of the week and were composed of a prompt ("Please indicate last night's quality of your sleep/today's energy level/today's stress level") and a 10-point visual analog sliding scale, where 10 represents the highest sleep quality, energy levels, or stress levels. Users could rate these at any time of the day but received no additional reminders for them.

## Data Analysis

Characteristics are presented as mean and SD with the corresponding number and percentage of participants. User engagement information is presented as medians and IQR for weekly active days and total active days (active days out of 28

days), as these were nonparametric variables. Active days were defined as days when the user logged at least one mission. The average number of daily mission interactions is the number of events a user completes per mission; it was calculated as total mission interactions divided by total active days and is presented as mean and SD. Step goal attainment is shown as the number and percentage of users who used a step counter in the first and last weeks, along with weekly step counts as mean and SD, and step goal attainment is shown as the percentage of users who used the step counter.

Scores from QLQ-C30 were calculated according to instructions in the scoring manual (open source) [24]. To calculate the score for each subscale (global health status, physical functioning, role functioning, fatigue, pain, etc), we averaged the raw scores given to each of the questions contributing to that subscale. Most questions were scored on a scale of 1 (not at all) to 4 (very much), except for the questions contributing to global health status, which could be scored from 1 to 7. This raw score was then linearly transformed to 0 to 100 so that all subscales had the same range of values. For functional scales and global health status, a higher number represents a higher level of functioning and higher QoL; conversely, a higher number means a higher symptom burden for symptom scales. The number and percentage of participants whose scores increased, decreased, or did not change after the study was calculated along with the scores.

PROs on sleep, stress, and energy levels were compared at the beginning and end of the study using 2-tailed paired *t* tests. The normalized change was calculated as the change in the weekly average score divided by the maximum possible gain to be able to compare rating scales in different directions (evaluating increased energy levels and quality of sleep but decreased stress levels). The correlation between preprogram QLQ-C30 scores and the in-app PROs during the first week was calculated using Spearman's rank correlations.

## Results

### Baseline Characteristics

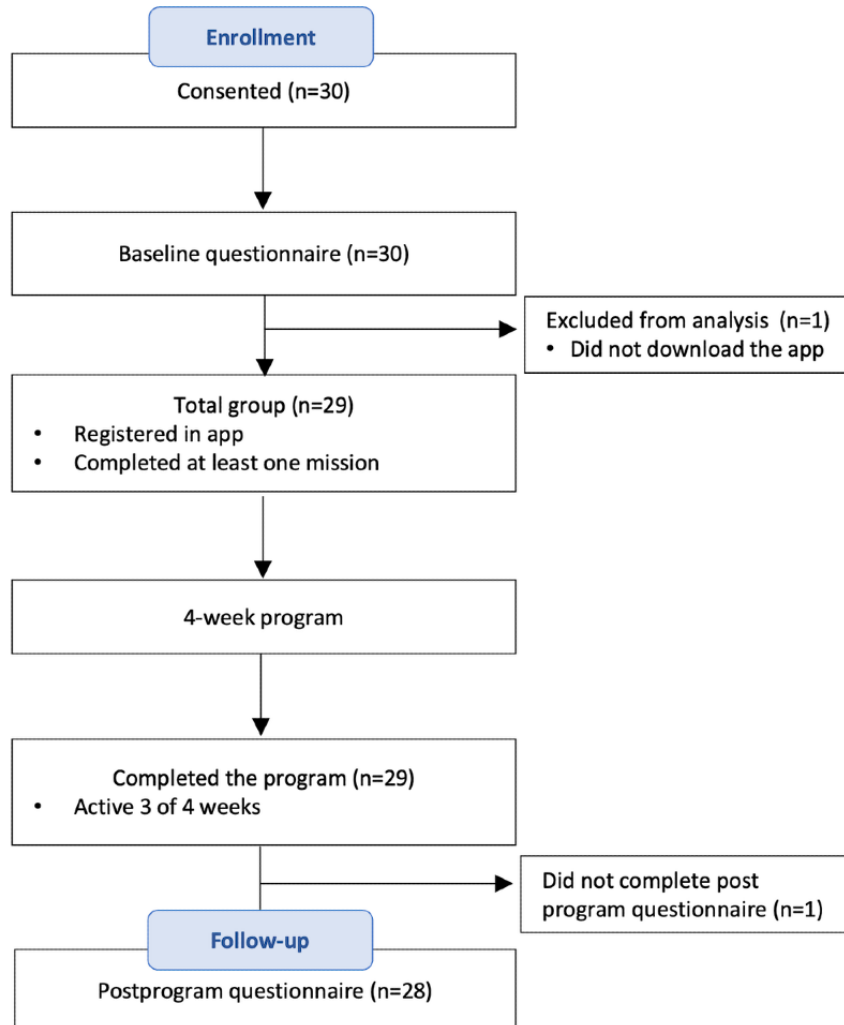
In total, 30 patients with cancer were initially enrolled in the program; 1 (3%) did not download the app and was therefore excluded from further analysis, and 29 (97%) completed the program with in-app activity in 3 of 4 weeks (Figure 2). Of the 29 participants, 1 (3%) was inactive in the last week but completed the follow-up questionnaire, and 1 (3%) participant who engaged with the program every week did not complete the postprogram questionnaire.

The baseline characteristics and measurements are shown in Table 2. Most participants were female (23/29, 79%); the average age was 52.6 (SD 11.5) years, and 62% (18/29) had breast cancer, while 38% (11/29) had other types of cancer. As there were 2 apparent patient groups with respect to cancer type (breast cancer and other cancer), we compared the baseline measurements of the 2 groups and found no statistically significant differences; thus, we present our results for the total patient population. A total of 28% (8/29) of participants had stage IV metastatic cancer, and 72% (21/29) had stage I to III

cancer; all participants were receiving cancer treatment, either chemotherapy or radiation therapy, at the start of the study. The sample population was slightly obese, with an average BMI of 30 (SD 5.8) kg/m<sup>2</sup>, and most participants (23/29, 79%) were in the overweight or obese BMI ranges. As expected, the mean percentage of body fat was higher in females (41.2%) than in

males (26.5%), and VO<sub>2max</sub> was 27.7 mL/kg/min and 31.2 mL/kg/min in females and males, respectively. These measurements were repeated at the end of the study, but only small nonsignificant changes were observed (data not shown). Of the 29 participants, 28 (97%) completed the full program and all postprogram questionnaires.

**Figure 2.** Participant flow through the study.



**Table 2.** Participants' baseline characteristics (N=29).

Characteristics	Value
Age (years), mean (SD)	52.6 (11.5)
Female, n (%)	23 (79)
<b>Cancer type and stage, n (%)</b>	
Breast cancer	18 (62)
Other cancer	11 (38)
Stage I-III	21 (72)
Stage IV (metastasis)	8 (28)
<b>Current therapy, n (%)</b>	
Chemotherapy	26 (90)
Radiation therapy	2 (7)
Both	1 (3)
<b>BMI categories (kg/m<sup>2</sup>), n (%)</b>	
Normal weight (18.5-24.9)	5 (17)
Overweight (25.0-29.9)	9 (31)
Obese (≥30.0)	14 (52)
<b>Body composition, mean (SD)</b>	
Height (cm)	168.3 (7.5)
Weight (kg)	85.0 (17.6)
BMI (kg/m <sup>2</sup> )	30.0 (5.8)
<b>Percentage body fat (%)</b>	
Female <sup>a</sup>	41.2 (7.6)
Male <sup>b</sup>	26.5 (8.6)
<b>Fat mass (kg)</b>	
Female <sup>a</sup>	32.6 (11.8)
Male <sup>b</sup>	25.4 (17.5)
<b>Lean mass (kg)</b>	
Female <sup>a</sup>	48.1 (5.5)
Male <sup>b</sup>	63.1 (9.3)
<b>Maximal aerobic capacity (VO<sub>2max</sub><sup>c</sup>, mL/kg/min), mean (SD)</b>	
Female <sup>d</sup>	27.7 (8.1)
Male <sup>e</sup>	31.2 (11.6)

<sup>a</sup>Data available for 23 participants.

<sup>b</sup>Data available for 5 participants.

<sup>c</sup>VO<sub>2max</sub>: maximum oxygen uptake.

<sup>d</sup>Data available for 15 participants.

<sup>e</sup>Data available for 6 participants.

## Primary Outcomes

### Program Feasibility

Engagement metrics are presented in Table 3. The median number of active days per week was 6.8 (IQR 5.8-6.8), while

the total active days out of the maximum of 28 days of the program was 27 (IQR 23-27). Users interacted with missions on average 7.7 times a day, and the number of highly engaged users was 72% (21/29), of whom 57% (12/21) were patients with breast cancer and 43% (9/21) had other types of cancer. A



total of 97% (28/29) participants continued using the app until week 4 (4-week retention), with only 1 with breast cancer being inactive in the last week.

Data on step counter use show that all 29 participants used this function in the first week, with 72% (21/29) achieving the set

target. The number of users logging their steps remained high in the last week (27/28, 96% of active users), with 68% (19/28) reaching the target goal (Table 4). The average weekly step count increased by over 3000, from 21,307 in the first week to 24,449 in the last week.

**Table 3.** Engagement metrics (N=29).

Metric	Value
Completion rate <sup>a</sup> , n (%)	28 (97)
Weekly active days, median (IQR)	6.8 (5.8-6.8)
Total active days, median (IQR)	27.0 (23.0-27.0)
Daily mission interactions, mean (SD)	7.7 (1.9)
Highly engaged users <sup>b</sup> , n (%)	21 (72)
4-week retention, n (%)	28 (97)

<sup>a</sup>Completed 75% of the program and all preprogram and postprogram questionnaires.

<sup>b</sup>Users who were active in the app at least five days a week.

**Table 4.** In-app measured step counts and goal attainment (N=29).

Metric	Week 1	Week 4
Used step counter, n (%)	29 (100)	27 (96)
Attained step goal, n (%)	21 (72)	19 (68)
Weekly step counts, mean (SD)	21,306 (11,411)	24,449 (17,445)

### Program Acceptability

Overall, 28 participants completed the postintervention satisfaction survey. These results showed that program acceptability was high, with 89% (25/28) of participants likely to recommend the program to others and 93% (26/28) who found the Sidekick app user friendly. Regarding program content, of the 28 participants, 26 (93%) found the educational content helpful, 23 (82%) said they felt better equipped to deal with their illness after participating in the program, and 24 (86%) said the app helped them remember to take their medication. With regard to the lifestyle coaching feature, 93% (26/28) of the participants said they found the weekly messages from the coach useful, but only 54% (15/28) somewhat agreed with the statement that they would have liked more feedback from the coach. Overall, 86% (24/28) of participants agreed that the program had positive effects on their lives and well-being.

### Secondary Outcomes

#### Quality of Life Questionnaire C30

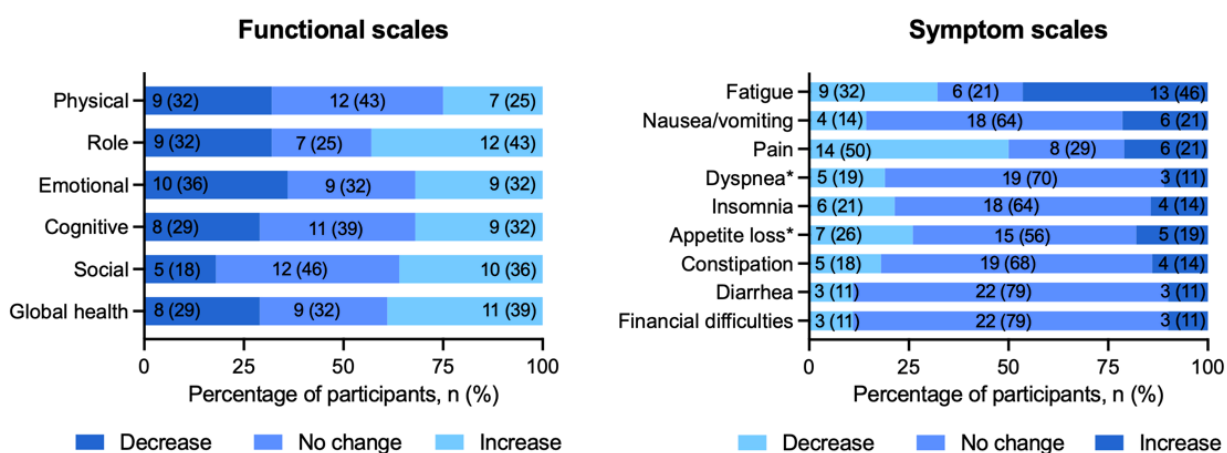
The results of the self-reported QoL questionnaire are shown in Table 5. One individual did not complete the follow-up questionnaire; thus, only 28 completed questionnaires were analyzed. We found no significant change in the mean scores from inclusion to follow-up for any of the functional or symptom items or for global health (Table 4). The largest change was observed in pain scores, which decreased by 8.4 points (27%), and social functioning, which increased by 6 points (10%) at follow-up.

Functioning in all subcategories remained the same or increased in approximately two-thirds of the participants, with most participants seeing an improvement in role functioning (Figure 3). Most participants (15-22 of 28, 56%-79%) saw no change in symptoms, except for pain, which decreased in 50% (12/28) of participants, while fatigue increased in 46% (13/28) of participants by the end of 4 weeks.

**Table 5.** Quality of Life Questionnaire C30 scores for each item scale at inclusion and follow-up.

Domain	Inclusion (n=28)	Follow-up (n=28)	P value
<b>Functional scales, mean (SD)</b>			
Physical	85.5 (10.7)	83.8 (14.1)	.39
Role	60.7 (20.9)	62.5 (26.7)	.54
Emotional	72.3 (16.5)	73.5 (18.0)	.63
Cognitive	69.0 (21.6)	70.8 (20.6)	.65
Social	58.3 (23.4)	64.3 (21.6)	.21
<b>Symptom scales, mean (SD)</b>			
Fatigue	35.7 (14.3)	40.1 (22.1)	.55
Nausea or vomiting	8.3 (12.4)	8.9 (14.0)	.67
Pain	31.0 (20.1)	22.6 (23.2)	.16
Dyspnea <sup>a</sup>	22.2 (22.6)	23.5 (27.4)	.94
Insomnia	31.0 (25.5)	29.8 (27.7)	.87
Appetite loss <sup>a</sup>	19.8 (19.1)	19.0 (26.3)	.48
Constipation	11.9 (22.6)	11.9 (20.7)	.85
Diarrhea	11.9 (20.7)	10.7 (18.3)	.74
Financial difficulties	19.0 (30.7)	15.5 (21.2)	.52
Global health status, mean (SD)	61.6 (17.0)	61.9 (15.8)	.78

<sup>a</sup>1 missing value; n=27 answers were analyzed.

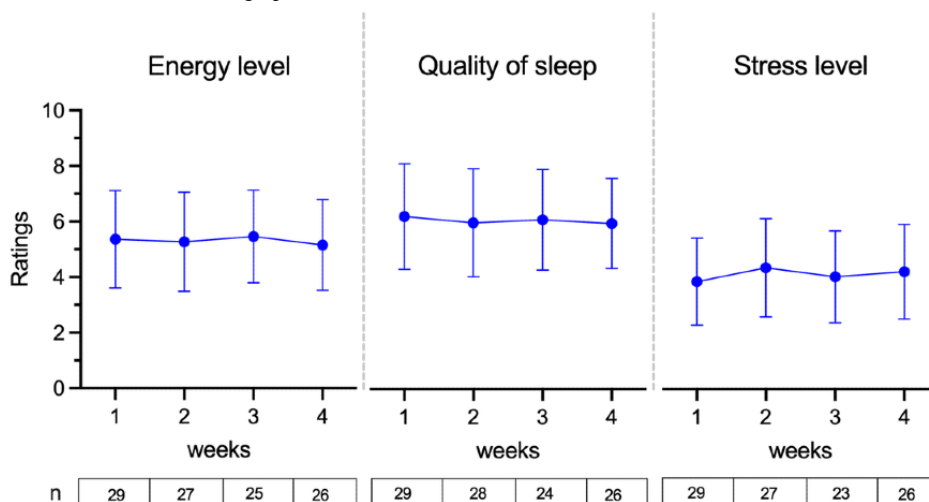
**Figure 3.** Percentage of individuals with decreased, increased, or unchanged Quality of Life Questionnaire C30 scores on the functional and symptom scales. \*1 missing value; n=27 answers were analyzed.

### In-App PROs

PROs on energy levels, quality of sleep, and stress levels collected within the app showed that they remained stable over

time, with no significant changes (Figure 4). In week 1, all 29 users reported that these PROs and engagement remained high, as 26 users still engaged with them in week 4.

**Figure 4.** Average weekly changes in users' ratings of energy levels, quality of sleep, and stress levels (error bars show SD), with the number of users reporting these (n) each week shown under the graphs.

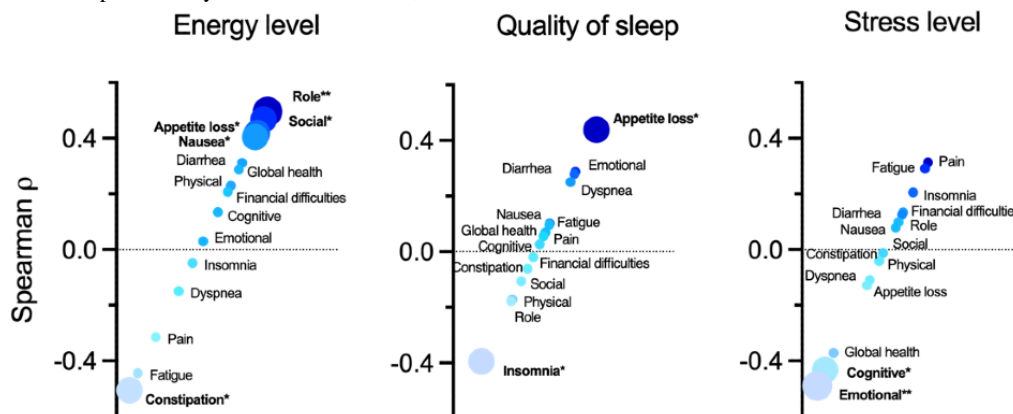


### Exploratory Analysis

We carried out an exploratory correlation analysis between the QLQ-C30 and in-app-reported QoL measures to gain insight into the clinical validity of the QoL outcomes measured in the app (Figure 5). We found that energy levels correlated most strongly with role ( $\rho=0.49$ ;  $P=.007$ ) and social functioning ( $\rho=0.47$ ;  $P=.01$ ) as well as with nausea, vomiting and appetite

loss ( $\rho=0.41$ ;  $P=.03$  each), while it inversely correlated with fatigue ( $\rho=-0.44$ ;  $P=.02$ ) and constipation ( $\rho=-0.5$ ;  $P=.006$ ). Quality of sleep significantly correlated with appetite loss ( $\rho=0.44$ ;  $P=.02$ ) and negatively correlated with insomnia ( $\rho=-0.39$ ;  $P=.04$ ), while stress levels significantly inversely correlated with cognitive ( $\rho=-0.43$ ;  $P=.02$ ) and emotional ( $\rho=-0.49$ ;  $P<.01$ ) functioning.

**Figure 5.** Spearman rank correlation analysis of in-app quality of life (QoL) and Quality of Life Questionnaire C30 (QLQ-C30) item scores. The listed QLQ-C30 items below 0 are inversely correlated, while above 0, they are positively correlated with energy level, quality of sleep, and stress levels. Items that significantly correlate with either in-app QoL measure are shown in bold and are represented by large circles, while those items with no significant correlations are represented by small circles. \* $P\leq.05$ ; \*\* $P\leq.01$ .



## Discussion

### Principal Findings

This trial tested the feasibility of a DTx intervention targeted at patients with cancer in active treatment and gathered preliminary information on its effectiveness. We obtained encouraging results regarding retention and engagement; the program had a very high completion rate (97%) and high acceptability (>80%), and thus, the feasibility criteria were met. Engagement metrics painted a similar picture, with users staying active 96% of the time, or 27 out of 28 days, and completing on average 7.7 (SD 1.9) missions a day.

These engagement metrics are somewhat higher than those reported in other studies. A recent systematic review of 6 studies

found that the average retention rate of digital behavioral interventions was 90.7% among cancer survivors [28]. However, 1 trial with a patient population similar to ours reported a 50% retention rate during a 6-week intervention and 80% questionnaire completion at follow-up, with 51% to 76% engagement with their app (the amount of content viewed) [21]. Another trial reported that only 41% to 65% of the participants logged in more than twice during the 10-week intervention period when they were not given personalized messaging; however, all those receiving personalized messaging used the app at least twice [29]. Similarly, a 12-week pilot research of breast cancer survivors revealed that 70% of the participants were continuously using the app, with 7.26 log-ins on average in the first month, which later declined [30]. Given that most participants in this study interacted with the app nearly daily,

we can conclude that it succeeded in motivating patients with cancer to continually engage with its content. Multimedia content and tailoring, as well as reminders and personalized messaging, were found to increase user engagement with web-based tools [29,31]. These elements were all used by the Sidekick app, which may help explain the relatively low attrition. In addition, the app was designed to associate the completion of missions with appealing sounds or visuals (ie, users receive a supporting animation when they complete a mission) and charitable donations, which in itself might motivate users and increase their engagement. However, cancer survivors are often internally motivated to seek help and use digital interventions to improve their health and fight the disease [32].

In terms of goal attainment, >70% of users were able to achieve their step goals in this trial, and the average step count (approximately 3000-3500 per day) was comparable with that found in other studies with patients in active cancer treatment [9,33-35]. Although these results might underestimate participants' real step counts (as these had to be manually claimed each day), they suggest that users are interested in tracking their physical activity, particularly their step counts.

The patients in this study were on average more obese than similar patient populations in other trials [9,36,37], and their average BMI was higher than that recommended for cancer survivors [38]. Regarding body composition, normative data are available for healthy adults in Sweden, Switzerland, and wider Europe [39,40]. According to these data, normal body fat percentage falls between 19% and 25% for males and 26% and 36% for females—the results found in our study (26.5% body fat in males and 41% in females) were somewhat higher than these normative values. Regarding cardiorespiratory fitness, reference values for a normal  $VO_{2max}$  range (30-43 mL/kg/min for males and 28-34 mL/kg/min for females) have been published for healthy Swedish and Norwegian adults [41,42]. The  $VO_{2max}$  values of patients with cancer in this study (28 mL/kg/min for females and 31 mL/kg/min for males) fell on the lower end of these ranges, suggesting lower cardiorespiratory fitness. Although we did not find significant changes in physical fitness or body composition, other studies promoting exercise intervention for patients in active treatment found significant improvements after 6 to 12 weeks [9,43]. Improving physical fitness and reducing sedentary behaviors can improve health-related QoL and reduce the risk of hospital readmissions [8,36,38]; thus, these will be important outcomes in future longer RCTs.

### QoL Outcomes

The global QoL of patients with cancer found in this study agrees with scores reported from patients with cancer in previous studies [9,37,44]. Compared with reference values from a large sample of patients aged 50 to 59 years with breast cancer or cancer in general, our participants scored slightly lower on role, cognitive, and social functioning and had higher symptom burden in most items [45]. An important question, however, is what scores represent clinically significant problems or symptom burden for patients. A previous study sought to answer this question and established cutoff values for 4 subitem scales [46]. According to that study, scores <83 for physical and <70 for

emotional functioning and >39 for fatigue and >25 for pain (the most commonly reported symptoms among patients with advanced cancers [47-49]) likely mean significant problems for patients. Compared with these cutoff values, our participants reported less pain by the end of the program, suggesting that pain caused less clinically significant burden to them. In addition, they reported better physical and emotional functioning both before and after the program. Overall, these results showed a generally stable health-related QoL during these 4 weeks, with a trend for improved pain scores. Future RCTs should further evaluate the effectiveness of the program in improving or maintaining QoL. In the long term, even maintaining stable health can be important for patients with cancer, as they usually experience a decline during prolonged treatments or as the disease progresses [50,51].

We found that the in-app PROs positively or negatively correlated with certain QLQ-C30 items as expected. Higher energy levels indicated higher role and social functioning and lower fatigue and constipation, while higher stress levels indicated lower cognitive and emotional functioning. Surprisingly however, higher energy levels were also associated with increased loss of appetite and nausea or vomiting, and better sleep was associated with reduced appetite. It is important to further assess these associations in larger sample trials to better validate PROs used by the Sidekick app.

### Strengths and Limitations

The strengths of this study were excellent retention, engagement, and questionnaire completion, which eliminated the need to correct for missing data. Feedback from the patients suggested that the supportive and familiar environment at the rehabilitation clinic could have played a key role in this finding. An additional strength is the multidisciplinary nature of the intervention, which has been shown to benefit the rehabilitation of patients with breast cancer [37].

A limitation of this study was the small sample size, which was composed of self-selected and likely self-motivated individuals. This restricts the generalizability of the results and reduces the study power for testing preliminary program efficiency. Another limitation arising from the study design was the short program duration of 4 weeks, which is likely too short a time frame to detect significant changes in physical and mental health parameters. Finally, the known limitations of the app are the lack of automatic step counting and the fact that step count missions could not be completed retrospectively for previous days; thus, if users did not claim their steps, the records showed 0 steps for the given day. Therefore, this feature likely underestimated the real physical activity that participants completed and hence should be further optimized in future programs and trials.

### Conclusions

On the basis of evidence gathered, digital support delivered through the Sidekick app is feasible for patients with cancer, and a large-scale RCT can be initiated. Preliminary results suggest that participants' health-related QoL remained stable for 4 weeks, but a longer, controlled trial will be required to gauge the efficacy of the digital intervention for improving QoL.

Changes in the most burdensome side effects, fatigue and pain, should also be the primary focus and assessed using specific measures in future trials. In addition, the digital program could be further tailored to the cancer experience by including education about treatments and specific side effects, providing

symptom tracking and medication reminders, and adapting the assigned daily tasks and workload to the individual's stage on the cancer treatment journey and actual energy and motivation levels.

## Acknowledgments

This study was sponsored by Sidekick Health. The authors thank Kolbrún Halla Gudjónsdóttir for providing support with recruitment and coaching at Ljósíð as well as all the staff members at Ljósíð who helped with the study procedures.

## Conflicts of Interest

GHG, JM, ÁB, MLÁ, and HH are employees of Sidekick Health; in addition, SO is the cofounder of Sidekick Health.

## Multimedia Appendix 1

Details of the weekly education content.

[[DOCX File , 15 KB - formative\\_v6i10e39764\\_app1.docx](#) ]

## References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2021 May;71(3):209-249 [FREE Full text] [doi: [10.3322/caac.21660](https://doi.org/10.3322/caac.21660)] [Medline: [33538338](https://pubmed.ncbi.nlm.nih.gov/33538338/)]
2. Aðalvalmynd. Krabbameinsfélagið. URL: <https://www.krabb.is/rannsoknasetur/upplysingar-um-krabbamein/yfirlitstolfraedi/> [accessed 2022-02-08]
3. Lu WL, Jansen L, Post WJ, Bonnema J, Van de Velde JC, De Bock GH. Impact on survival of early detection of isolated breast recurrences after the primary treatment for breast cancer: a meta-analysis. *Breast Cancer Res Treat* 2009 Apr;114(3):403-412. [doi: [10.1007/s10549-008-0023-4](https://doi.org/10.1007/s10549-008-0023-4)] [Medline: [18421576](https://pubmed.ncbi.nlm.nih.gov/18421576/)]
4. Ochoa Arnedo C, Sánchez N, Sumalla EC, Casellas-Grau A. Stress and growth in cancer: mechanisms and psychotherapeutic interventions to facilitate a constructive balance. *Front Psychol* 2019 Feb 4;10:177 [FREE Full text] [doi: [10.3389/fpsyg.2019.00177](https://doi.org/10.3389/fpsyg.2019.00177)] [Medline: [30778323](https://pubmed.ncbi.nlm.nih.gov/30778323/)]
5. Zoberi K, Tucker J. Primary care of breast cancer survivors. *Am Fam Physician* 2019 Mar 15;99(6):370-375 [FREE Full text] [Medline: [30874405](https://pubmed.ncbi.nlm.nih.gov/30874405/)]
6. Greenlee H, DuPont-Reyes MJ, Balneaves LG, Carlson LE, Cohen MR, Deng G, et al. Clinical practice guidelines on the evidence-based use of integrative therapies during and after breast cancer treatment. *CA Cancer J Clin* 2017 May 06;67(3):194-232 [FREE Full text] [doi: [10.3322/caac.21397](https://doi.org/10.3322/caac.21397)] [Medline: [28436999](https://pubmed.ncbi.nlm.nih.gov/28436999/)]
7. Campbell KL, Winters-Stone KM, Wiskemann J, May AM, Schwartz AL, Courneya KS, et al. Exercise guidelines for cancer survivors: consensus statement from international multidisciplinary roundtable. *Med Sci Sports Exerc* 2019 Nov;51(11):2375-2390 [FREE Full text] [doi: [10.1249/MSS.0000000000002116](https://doi.org/10.1249/MSS.0000000000002116)] [Medline: [31626055](https://pubmed.ncbi.nlm.nih.gov/31626055/)]
8. Low CA, Bovbjerg DH, Ahrendt S, Choudry MH, Holtzman M, Jones HL, et al. Fitbit step counts during inpatient recovery from cancer surgery as a predictor of readmission. *Ann Behav Med* 2018 Jan 05;52(1):88-92 [FREE Full text] [doi: [10.1093/abm/kax022](https://doi.org/10.1093/abm/kax022)] [Medline: [29538623](https://pubmed.ncbi.nlm.nih.gov/29538623/)]
9. Delrieu L, Pialoux V, Pérol O, Morelle M, Martin A, Friedenreich C, et al. Feasibility and health benefits of an individualized physical activity intervention in women with metastatic breast cancer: intervention study. *JMIR Mhealth Uhealth* 2020 Jan 28;8(1):e12306 [FREE Full text] [doi: [10.2196/12306](https://doi.org/10.2196/12306)] [Medline: [32012082](https://pubmed.ncbi.nlm.nih.gov/32012082/)]
10. Arends J, Bachmann P, Baracos V, Barthelemy N, Bertz H, Bozzetti F, et al. ESPEN guidelines on nutrition in cancer patients. *Clin Nutr* 2017 Feb;36(1):11-48. [doi: [10.1016/j.clnu.2016.07.015](https://doi.org/10.1016/j.clnu.2016.07.015)] [Medline: [27637832](https://pubmed.ncbi.nlm.nih.gov/27637832/)]
11. Gussoni G, Ravot E, Zecchina M, Recchia G, Santoro E, Ascione R, et al. Digital therapeutics in oncology: findings, barriers and prospects. A narrative review. *Ann Res Oncol* 2022 Mar;2(1):55-69 [FREE Full text] [doi: [10.48286/aro.2022.39](https://doi.org/10.48286/aro.2022.39)]
12. Basch E, Deal AM, Kris MG, Scher HI, Hudis CA, Sabbatini P, et al. Symptom monitoring with patient-reported outcomes during routine cancer treatment: a randomized controlled trial. *J Clin Oncol* 2016 Feb 20;34(6):557-565 [FREE Full text] [doi: [10.1200/JCO.2015.63.0830](https://doi.org/10.1200/JCO.2015.63.0830)] [Medline: [26644527](https://pubmed.ncbi.nlm.nih.gov/26644527/)]
13. Denis F, Basch E, Septans AL, Bennouna J, Urban T, Dueck AC, et al. Two-year survival comparing Web-based symptom monitoring vs routine surveillance following treatment for lung cancer. *JAMA* 2019 Jan 22;321(3):306-307 [FREE Full text] [doi: [10.1001/jama.2018.18085](https://doi.org/10.1001/jama.2018.18085)] [Medline: [30667494](https://pubmed.ncbi.nlm.nih.gov/30667494/)]
14. Kearney N, McCann L, Norrie J, Taylor L, Gray P, McGee-Lennon M, et al. Evaluation of a mobile phone-based, advanced symptom management system (ASyMS) in the management of chemotherapy-related toxicity. *Support Care Cancer* 2009 Apr;17(4):437-444. [doi: [10.1007/s00520-008-0515-0](https://doi.org/10.1007/s00520-008-0515-0)] [Medline: [18953579](https://pubmed.ncbi.nlm.nih.gov/18953579/)]

15. Schmalz O, Jacob C, Ammann J, Liss B, Iivanainen S, Kammermann M, et al. Digital monitoring and management of patients with advanced or metastatic non-small cell lung cancer treated with cancer immunotherapy and its impact on quality of clinical care: interview and survey study among health care professionals and patients. *J Med Internet Res* 2020 Dec 21;22(12):e18655 [[FREE Full text](#)] [doi: [10.2196/18655](https://doi.org/10.2196/18655)] [Medline: [33346738](https://pubmed.ncbi.nlm.nih.gov/33346738/)]
16. Penedo FJ, Fox RS, Oswald LB, Moreno PI, Boland CL, Estabrook R, et al. Technology-based psychosocial intervention to improve quality of life and reduce symptom burden in men with advanced prostate cancer: results from a randomized controlled trial. *Int J Behav Med* 2020 Oct;27(5):490-505 [[FREE Full text](#)] [doi: [10.1007/s12529-019-09839-7](https://doi.org/10.1007/s12529-019-09839-7)] [Medline: [31898309](https://pubmed.ncbi.nlm.nih.gov/31898309/)]
17. Spahrkäs SS, Looijmans A, Sanderman R, Hagedoorn M. Beating cancer-related fatigue with the Untire mobile app: results from a waiting-list randomized controlled trial. *Psychooncology* 2020 Nov;29(11):1823-1834 [[FREE Full text](#)] [doi: [10.1002/pon.5492](https://doi.org/10.1002/pon.5492)] [Medline: [33393199](https://pubmed.ncbi.nlm.nih.gov/33393199/)]
18. Holdirak F, Mehnert A, Weiss M, Mayer J, Meyer B, Bröde P, et al. Results of the Optimune trial: a randomized controlled trial evaluating a novel Internet intervention for breast cancer survivors. *PLoS One* 2021 May 7;16(5):e0251276 [[FREE Full text](#)] [doi: [10.1371/journal.pone.0251276](https://doi.org/10.1371/journal.pone.0251276)] [Medline: [33961667](https://pubmed.ncbi.nlm.nih.gov/33961667/)]
19. Galiano-Castillo N, Arroyo-Morales M, Lozano-Lozano M, Fernández-Lao C, Martín-Martín L, Del-Moral-Ávila R, et al. Effect of an Internet-based telehealth system on functional capacity and cognition in breast cancer survivors: a secondary analysis of a randomized controlled trial. *Support Care Cancer* 2017 Nov;25(11):3551-3559. [doi: [10.1007/s00520-017-3782-9](https://doi.org/10.1007/s00520-017-3782-9)] [Medline: [28639097](https://pubmed.ncbi.nlm.nih.gov/28639097/)]
20. Thorgeirsson T, Torfadottir JE, Egilsson E, Oddsson S, Gunnarsdottir T, Aspelund T, et al. Randomized trial for weight loss using a digital therapeutic application. *J Diabetes Sci Technol* 2022 Sep;16(5):1150-1158. [doi: [10.1177/19322968211000815](https://doi.org/10.1177/19322968211000815)] [Medline: [33736484](https://pubmed.ncbi.nlm.nih.gov/33736484/)]
21. Wright H, Martin F, Clyne W, Clark CC, Matouskova G, McGillion M, et al. A digital self-management program (help to overcome problems effectively) for people living with cancer: feasibility randomized controlled trial. *J Med Internet Res* 2021 Nov 05;23(11):e28322 [[FREE Full text](#)] [doi: [10.2196/28322](https://doi.org/10.2196/28322)] [Medline: [34738912](https://pubmed.ncbi.nlm.nih.gov/34738912/)]
22. Enrique A, Palacios JE, Ryan H, Richards D. Exploring the relationship between usage and outcomes of an Internet-based intervention for individuals with depressive symptoms: secondary analysis of data from a randomized controlled trial. *J Med Internet Res* 2019 Aug 01;21(8):e12775 [[FREE Full text](#)] [doi: [10.2196/12775](https://doi.org/10.2196/12775)] [Medline: [31373272](https://pubmed.ncbi.nlm.nih.gov/31373272/)]
23. Waller R, Gilbody S. Barriers to the uptake of computerized cognitive behavioural therapy: a systematic review of the quantitative and qualitative evidence. *Psychol Med* 2009 May;39(5):705-712. [doi: [10.1017/S0033291708004224](https://doi.org/10.1017/S0033291708004224)] [Medline: [18812006](https://pubmed.ncbi.nlm.nih.gov/18812006/)]
24. Fayers PM, Aaronson NK, Bjordal K, Groenvold M, Curran D, Bottomley A, EORTC Quality of Life Group. The EORTC QLQ-C30 Scoring Manual. 3rd edition. European Organisation for Research and Treatment of Cancer. 2001. URL: <https://www.eortc.org/app/uploads/sites/2/2018/02/SCmanual.pdf> [accessed 2022-08-02]
25. Astrand I. Aerobic work capacity in men and women with special reference to age. *Acta Physiol Scand Suppl* 1960;49(169):1-92. [Medline: [13794892](https://pubmed.ncbi.nlm.nih.gov/13794892/)]
26. Antonio J, Kenyon M, Ellerbroek A, Carson C, Burgess V, Tyler-Palmer D, et al. Comparison of dual-energy X-ray absorptiometry (DXA) versus a multi-frequency bioelectrical impedance (InBody 770) device for body composition assessment after a 4-week hypoenergetic diet. *J Funct Morphol Kinesiol* 2019 Apr 25;4(2):23 [[FREE Full text](#)] [doi: [10.3390/jfmk4020023](https://doi.org/10.3390/jfmk4020023)] [Medline: [33467338](https://pubmed.ncbi.nlm.nih.gov/33467338/)]
27. Lin CL, Yu NC, Wu HC, Lee YY, Lin WC, Chiu IY, et al. Association of body composition with type 2 diabetes: a retrospective chart review study. *Int J Environ Res Public Health* 2021 Apr 21;18(9):4421 [[FREE Full text](#)] [doi: [10.3390/ijerph18094421](https://doi.org/10.3390/ijerph18094421)] [Medline: [33919339](https://pubmed.ncbi.nlm.nih.gov/33919339/)]
28. Gan T, Cheng HL, Tse MM. Feasibility, acceptability, and effects of behavior change interventions for improving multiple dietary behaviors among cancer survivors: a systematic review. *Support Care Cancer* 2022 Mar;30(3):2877-2889. [doi: [10.1007/s00520-021-06582-2](https://doi.org/10.1007/s00520-021-06582-2)] [Medline: [34581862](https://pubmed.ncbi.nlm.nih.gov/34581862/)]
29. Marthick M, Dhillon HM, Alison JA, Cheema BS, Shaw T. An interactive Web portal for tracking oncology patient physical activity and symptoms: prospective cohort study. *JMIR Cancer* 2018 Dec 21;4(2):e11978 [[FREE Full text](#)] [doi: [10.2196/11978](https://doi.org/10.2196/11978)] [Medline: [30578217](https://pubmed.ncbi.nlm.nih.gov/30578217/)]
30. Lally RM, Kupzyk K, Gallo S, Berry D. Use of an unguided, Web-based distress self-management program after breast cancer diagnosis: sub-analysis of CaringGuidance pilot study. *J Med Internet Res* 2020 Jul 06;22(7):e19734 [[FREE Full text](#)] [doi: [10.2196/19734](https://doi.org/10.2196/19734)] [Medline: [32628117](https://pubmed.ncbi.nlm.nih.gov/32628117/)]
31. Milward J, Drummond C, Fincham-Campbell S, Deluca P. What makes online substance-use interventions engaging? A systematic review and narrative synthesis. *Digit Health* 2018 Feb 1;4:2055207617743354 [[FREE Full text](#)] [doi: [10.1177/2055207617743354](https://doi.org/10.1177/2055207617743354)] [Medline: [29942622](https://pubmed.ncbi.nlm.nih.gov/29942622/)]
32. Phillips SM, Conroy DE, Keadle SK, Pellegrini CA, Lloyd GR, Penedo FJ, et al. Breast cancer survivors' preferences for technology-supported exercise interventions. *Support Care Cancer* 2017 Oct;25(10):3243-3252 [[FREE Full text](#)] [doi: [10.1007/s00520-017-3735-3](https://doi.org/10.1007/s00520-017-3735-3)] [Medline: [28470368](https://pubmed.ncbi.nlm.nih.gov/28470368/)]

33. Nyrop KA, Deal AM, Choi SK, Wagoner CW, Lee JT, Wood WA, et al. Measuring and understanding adherence in a home-based exercise intervention during chemotherapy for early breast cancer. *Breast Cancer Res Treat* 2018 Feb;168(1):43-55. [doi: [10.1007/s10549-017-4565-1](https://doi.org/10.1007/s10549-017-4565-1)] [Medline: [29124455](https://pubmed.ncbi.nlm.nih.gov/29124455/)]
34. Hall CC, Skipworth RJ, Blackwood H, Brown D, Cook J, Diernberger K, et al. A randomized, feasibility trial of an exercise and nutrition-based rehabilitation programme (ENeRgy) in people with cancer. *J Cachexia Sarcopenia Muscle* 2021 Dec;12(6):2034-2044 [FREE Full text] [doi: [10.1002/jcsm.12806](https://doi.org/10.1002/jcsm.12806)] [Medline: [34612012](https://pubmed.ncbi.nlm.nih.gov/34612012/)]
35. Cheville AL, Kollasch J, Vandenberg J, Shen T, Grothey A, Gamble G, et al. A home-based exercise program to improve function, fatigue, and sleep quality in patients with Stage IV lung and colorectal cancer: a randomized controlled trial. *J Pain Symptom Manage* 2013 May;45(5):811-821 [FREE Full text] [doi: [10.1016/j.jpainsymman.2012.05.006](https://doi.org/10.1016/j.jpainsymman.2012.05.006)] [Medline: [23017624](https://pubmed.ncbi.nlm.nih.gov/23017624/)]
36. Low CA, Danko M, Durica KC, Kunta AR, Mulukutla R, Ren Y, et al. A real-time mobile intervention to reduce sedentary behavior before and after cancer surgery: usability and feasibility study. *JMIR Perioper Med* 2020 Mar 23;3(1):e17292 [FREE Full text] [doi: [10.2196/17292](https://doi.org/10.2196/17292)] [Medline: [33393915](https://pubmed.ncbi.nlm.nih.gov/33393915/)]
37. Leclerc AF, Foidart-Dessalle M, Tomasella M, Coucke P, Devos M, Bruyère O, et al. Multidisciplinary rehabilitation program after breast cancer: benefits on physical function, anthropometry and quality of life. *Eur J Phys Rehabil Med* 2017 Oct;53(5):633-642 [FREE Full text] [doi: [10.23736/S1973-9087.17.04551-8](https://doi.org/10.23736/S1973-9087.17.04551-8)] [Medline: [28322035](https://pubmed.ncbi.nlm.nih.gov/28322035/)]
38. Rock CL, Doyle C, Demark-Wahnefried W, Meyerhardt J, Courneya KS, Schwartz AL, et al. Nutrition and physical activity guidelines for cancer survivors. *CA Cancer J Clin* 2012;62(4):243-274 [FREE Full text] [doi: [10.3322/caac.21142](https://doi.org/10.3322/caac.21142)] [Medline: [22539238](https://pubmed.ncbi.nlm.nih.gov/22539238/)]
39. Pichard C, Kyle UG, Bracco D, Slosman DO, Morabia A, Schutz Y. Reference values of fat-free and fat masses by bioelectrical impedance analysis in 3393 healthy subjects. *Nutrition* 2000 Apr;16(4):245-254. [doi: [10.1016/s0899-9007\(00\)00256-2](https://doi.org/10.1016/s0899-9007(00)00256-2)] [Medline: [10758358](https://pubmed.ncbi.nlm.nih.gov/10758358/)]
40. Larsson I, Lissner L, Samuelson G, Fors H, Lantz H, Näslund I, et al. Body composition through adult life: Swedish reference data on body composition. *Eur J Clin Nutr* 2015 Jul;69(7):837-842. [doi: [10.1038/ejcn.2014.268](https://doi.org/10.1038/ejcn.2014.268)] [Medline: [25514897](https://pubmed.ncbi.nlm.nih.gov/25514897/)]
41. Loe H, Rognmo Ø, Saltin B, Wisløff U. Aerobic capacity reference data in 3816 healthy men and women 20-90 years. *PLoS One* 2013 May 15;8(5):e64319 [FREE Full text] [doi: [10.1371/journal.pone.0064319](https://doi.org/10.1371/journal.pone.0064319)] [Medline: [23691196](https://pubmed.ncbi.nlm.nih.gov/23691196/)]
42. Ekblom B, Engström LM, Ekblom O. Secular trends of physical fitness in Swedish adults. *Scand J Med Sci Sports* 2007 Jun;17(3):267-273. [doi: [10.1111/j.1600-0838.2006.00531.x](https://doi.org/10.1111/j.1600-0838.2006.00531.x)] [Medline: [17501867](https://pubmed.ncbi.nlm.nih.gov/17501867/)]
43. Dittus KL, Gramling RE, Ades PA. Exercise interventions for individuals with advanced cancer: a systematic review. *Prev Med* 2017 Nov;104:124-132. [doi: [10.1016/j.ypmed.2017.07.015](https://doi.org/10.1016/j.ypmed.2017.07.015)] [Medline: [28716654](https://pubmed.ncbi.nlm.nih.gov/28716654/)]
44. Yee J, Davis GM, Beith JM, Wilcken N, Currow D, Emery J, et al. Physical activity and fitness in women with metastatic breast cancer. *J Cancer Surviv* 2014 Dec;8(4):647-656. [doi: [10.1007/s11764-014-0378-y](https://doi.org/10.1007/s11764-014-0378-y)] [Medline: [24986228](https://pubmed.ncbi.nlm.nih.gov/24986228/)]
45. Scott NW, Fayers PM, Aaronson NK, Bottomley A, de Graeff A, Groenvold M, EORTC Quality of Life Group. EORTC QLQ-C30 Reference Values. EORTC Quality of Life Group. 2008 Jul. URL: [https://www.eortc.org/app/uploads/sites/2/2018/02/reference\\_values\\_manual2008.pdf](https://www.eortc.org/app/uploads/sites/2/2018/02/reference_values_manual2008.pdf) [accessed 2022-08-02]
46. Giesinger JM, Kuijpers W, Young T, Tomaszewski KA, Friend E, Zabernigg A, et al. Thresholds for clinical importance for four key domains of the EORTC QLQ-C30: physical functioning, emotional functioning, fatigue and pain. *Health Qual Life Outcomes* 2016 Jun 07;14:87 [FREE Full text] [doi: [10.1186/s12955-016-0489-4](https://doi.org/10.1186/s12955-016-0489-4)] [Medline: [27267486](https://pubmed.ncbi.nlm.nih.gov/27267486/)]
47. Yennurajalingam S, Palmer JL, Zhang T, Poulter V, Bruera E. Association between fatigue and other cancer-related symptoms in patients with advanced cancer. *Support Care Cancer* 2008 Oct;16(10):1125-1130. [doi: [10.1007/s00520-008-0466-5](https://doi.org/10.1007/s00520-008-0466-5)] [Medline: [18504622](https://pubmed.ncbi.nlm.nih.gov/18504622/)]
48. Kokkonen K, Saarto T, Mäkinen T, Pohjola L, Kautio H, Järvenpää S, et al. The functional capacity and quality of life of women with advanced breast cancer. *Breast Cancer* 2017 Jan;24(1):128-136. [doi: [10.1007/s12282-016-0687-2](https://doi.org/10.1007/s12282-016-0687-2)] [Medline: [27002988](https://pubmed.ncbi.nlm.nih.gov/27002988/)]
49. Teunissen SC, Wesker W, Kruitwagen C, de Haes HC, Voest EE, de Graeff A. Symptom prevalence in patients with incurable cancer: a systematic review. *J Pain Symptom Manage* 2007 Jul;34(1):94-104 [FREE Full text] [doi: [10.1016/j.jpainsymman.2006.10.015](https://doi.org/10.1016/j.jpainsymman.2006.10.015)] [Medline: [17509812](https://pubmed.ncbi.nlm.nih.gov/17509812/)]
50. Cardoso F, Spence D, Mertz S, Corneliussen-James D, Sabelko K, Gralow J, et al. Global analysis of advanced/metastatic breast cancer: decade report (2005-2015). *Breast* 2018 Jun;39:131-138. [doi: [10.1016/j.breast.2018.03.002](https://doi.org/10.1016/j.breast.2018.03.002)] [Medline: [29679849](https://pubmed.ncbi.nlm.nih.gov/29679849/)]
51. Foucaut AM, Berthouze SE, Touillaud M, Morelle M, Bourne-Branchu V, Kempf-Lépine AS, et al. Deterioration of physical activity level and metabolic risk factors after early-stage breast cancer diagnosis. *Cancer Nurs* 2015;38(4):E1-E9. [doi: [10.1097/NCC.000000000000187](https://doi.org/10.1097/NCC.000000000000187)] [Medline: [25207592](https://pubmed.ncbi.nlm.nih.gov/25207592/)]

## Abbreviations

**DTx:** digital therapeutics

**PRO:** patient-reported outcome

**QLQ-C30:** Quality of Life Questionnaire C30

**QoL:** quality of life

**RCT:** randomized controlled trial

**VO<sub>2max</sub>:** maximum oxygen uptake

*Edited by A Mavragani; submitted 23.05.22; peer-reviewed by C Drummond, J Johnson, C Varsi; comments to author 29.06.22; revised version received 17.08.22; accepted 29.08.22; published 13.10.22.*

*Please cite as:*

*Gudmundsson GH, Mészáros J, Björnsdóttir ÁE, Ámundadóttir ML, Thorvardardóttir GE, Magnúsdóttir E, Helgadóttir H, Oddsson S*

*Evaluating the Feasibility of a Digital Therapeutic Program for Patients With Cancer During Active Treatment: Pre-Post Interventional Study*

*JMIR Form Res 2022;6(10):e39764*

*URL: <https://formative.jmir.org/2022/10/e39764>*

*doi: [10.2196/39764](https://doi.org/10.2196/39764)*

*PMID: [36227639](https://pubmed.ncbi.nlm.nih.gov/36227639/)*

©G Haukur Gudmundsson, Judit Mészáros, Ágústa E Björnsdóttir, María L Ámundadóttir, Gudrun E Thorvardardóttir, Erna Magnúsdóttir, Halla Helgadóttir, Saemundur Oddsson. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 13.10.2022. 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 Technologies for Managing Problematic Pornography Use: Content Analysis

Nathan Henry<sup>1</sup>, BTech, MSc; Liesje Donkin<sup>1</sup>, BSc, MSc, PhD; Matt Williams<sup>2</sup>, BBS, GDipArts, MA, PhD; Mangor Pedersen<sup>1</sup>, BA, MSc, PhD

<sup>1</sup>Auckland University of Technology, Auckland, New Zealand

<sup>2</sup>Massey University, Auckland, New Zealand

**Corresponding Author:**

Nathan Henry, BTech, MSc

Auckland University of Technology

55 Wellesley Street East, Auckland CBD

Auckland, 1010

New Zealand

Phone: 64 9 921 9666 ext 7304

Email: [nathan.henry@aut.ac.nz](mailto:nathan.henry@aut.ac.nz)

## Abstract

**Background:** Several mobile apps are currently available that purportedly help with managing *pornography addiction*. However, the utility of these apps is unclear, given the lack of literature on the effectiveness of mobile health solutions for problematic pornography use. Little is also known about the content, structure, and features of these apps.

**Objective:** This study aims to characterize the purpose, content, and popularity of mobile apps that claim to manage pornography addiction.

**Methods:** The phrase “pornography addiction” was entered as a search term in the app stores of the two major mobile phone platforms (Android and iOS). App features were categorized according to a coding scheme that contained 16 categories. Apps were included in the analysis if they were described as helpful for reducing pornography use, and data were extracted from the store descriptions of the apps. Metrics such as number of user ratings, mean rating score, and number of installations were analyzed on a per-feature basis.

**Results:** In total, 170 apps from both app stores met the inclusion criteria. The five most common and popular features, both in terms of number of apps with each feature and minimum possible number of installations, were the ability to track the time since last relapse (apps with feature=72/170, 42.4%; minimum possible number of installations=6,388,000), tutorials and coaching (apps with feature=63/170, 37.1%; minimum possible number of installations=9,286,505), access to accountability partners or communities (apps with feature=51/170, 30%; minimum possible number of installations=5,544,500), content blocking or content monitoring (apps with feature=46/170, 27.1%; minimum possible number of installations=17,883,000), and a reward system for progress (apps with feature=34/170, 20%; minimum possible number of installations=4,425,300). Of these features, content-blocking apps had the highest minimum possible number of installations. Content blocking was also the most detected feature combination in a combinatorial analysis (with 28 apps having only this feature), but it also had the lowest mean consumer satisfaction rating (4.04) and second-lowest median rating (4.00) out of 5 stars. None of the apps reviewed contained references to literature that provided direct evidence for the app’s efficacy or safety.

**Conclusions:** There are several apps with the potential to provide low- or zero-cost real-time interventions for people struggling to manage problematic pornography use. Popular app features include blockers of pornographic content, behavior monitoring, and tutorials that instruct users how to eliminate pornography use. However, there is currently no empirical evidence to support the effectiveness and safety of these apps. Further research is required to be able to provide recommendations about which apps (and app features) are safe for public consumption.

(JMIR Form Res 2022;6(10):e39869) doi:[10.2196/39869](https://doi.org/10.2196/39869)

## KEYWORDS

pornography; compulsive sexual behavior disorder; CSBD; mobile health; mHealth; problematic pornography use; PPU; mobile intervention; just-in-time adaptive intervention; smartphone-based therapy; addiction; internet addiction; behavioral addiction; mobile phone

## Introduction

### Background

Internet pornography is becoming increasingly normalized worldwide because the rapid proliferation of the internet and personal computing devices, coupled with widespread cultural acceptance of pornography, has accelerated its spread [1]. An analysis of website traffic in Poland found a 310% increase in the number of people who used pornography on the web between October 2004 (2.76 million) and October 2016 (8.54 million) [2]. In 2019 alone, there were >42 billion visits to Pornhub, the world's most popular pornography website [3]. The use of the internet for sexual pleasure is likely to become more prevalent as people's time spent on the internet increases [4].

Problematic pornography use (PPU)—defined here as compulsive, dysregulated, or excessive pornography use—is considered a subcategory of compulsive sexual behavior disorder in the International Classification of Diseases, Eleventh Revision. However, there remains controversy over whether *pornography addiction* in the classical sense of an addiction exists, with opponents of this construct suggesting that compulsive pornography use is a coping mechanism for issues such as depression and anxiety. Still, there remains a high prevalence of anecdotal evidence for self-perceived pornography addiction in popular culture and web-based discourse, with many believing that pornography use can cause both depression and anxiety, among other disorders [5].

### Technological Interventions for PPU

The introduction of smartphones into modern society provides an unprecedented opportunity to deliver mobile therapies to people on a global scale. A mobile app that can anonymously guide the user through the steps required to manage PPU or regulate internet content could be valuable to people who are uncomfortable disclosing their struggle with PPU to another person. Such mobile health (mHealth) apps already exist, and unlike traditional therapy, these apps are often free and immediately available, thus supporting user autonomy and expanding access. However, it should be noted that such interventions are often hindered by low treatment commitment and adherence, limited flexibility and tailoring, and higher dropout rates than in-person therapies [6,7].

In a qualitative review of abstinence journals on a web-based *rebooting* forum for PPU, Fernandez et al [8] found that many of the forum's members struggled with “the seeming inescapability of cues” that triggered pornography cravings, particularly on electronic media such as the internet or television. Furthermore, these cues often appear unpredictably, making casual internet browsing a risky behavior for those attempting to reduce PPU [4,8-11]. Before internet pornography was introduced, one simply had to distance oneself from

physical sources of pornography, such as magazines [12]. Today, distancing is harder to achieve, given the amount of time we are connected to the internet and the ready accessibility of web-based pornography through our devices.

### Currently Available mHealth Therapies

Despite the popularity and potential of mHealth apps for managing PPU [13], there remains practically no literature assessing their effectiveness. This is consistent with the lack of literature on many other categories of mHealth apps, such as those for nonmedical cannabis use, mindfulness, and posttraumatic stress disorder, all of which have a large variety of apps available in both Google's and Apple's app ecosystems [14-16]. This is concerning because it makes it difficult to determine whether people are receiving evidence-based or effective interventions. Colbert et al [17] reviewed studies of 19 mobile apps designed to manage alcohol consumption; of these apps, only 8 were available in public app stores, and only 4 of these 8 apps had been demonstrated to help reduce alcohol consumption in the literature. A similarly low conversion rate has been observed in related fields [18].

One example of a therapeutic technique for PPU management is an app that can block pornographic websites, which hypothetically increases the number of barriers to accessing pornography, giving the user time to resist their urges better. However, these apps can often be turned off or bypassed and may not block the entire spectrum of sexual content, such as audio or text. Progress has been made on the use of machine learning models to detect pornographic imagery automatically, and these models may be incorporated into this kind of software in future [19]. However, these systems need to be used with care because the potential for false positives is high, leading to censorship of harmless content such as tutorials on sexual health [20], which could ultimately be harmful. Such a filter must be sensitive enough to capture a substantively high percentage of pornographic material but not so sensitive that it makes internet browsing a frustrating experience because of too many false positives, which can cause users to uninstall the program; for instance, biasing the software toward a skin-based detection method can lead to false positives in many domains (such as swimming, wrestling, or underwear modeling) and false negatives for sexual content of a nonnude nature [19]. This classification problem will also require constant updating as new sexual content and imagery forms are produced. In addition, these classifiers often do not recognize alternative media such as erotic literature or audio.

Browser extensions that provide broad-spectrum content blocking, such as uBlock Origin [21] and BlockSite [22], can also be configured to block pornography. Many of these digital self-control tools (DSCTs) are free (or have a free tier) and are thus available and effective for a broad consumer base [23]. These DSCTs are also examples of a just-in-time adaptive intervention (JITAI), where the intervention is delivered when

the user needs it most—as they are beginning down the path toward relapse by browsing a website that they know to be tempting [24]. An example of a scientifically grounded JITAI is the Addiction–Comprehensive Health Enhancement Support System, a smartphone app designed to improve continuing care for alcohol use disorders [25]. However, there seems to be no evidence-based version of this for PPU.

### Interventional Studies

Content-blocking apps are just 1 example of a wide variety of tools available to help people with PPU. However, very little research has examined the effectiveness of any of these tools. Although some studies have examined the efficacy of short interventions to reduce PPU, these have generally been small scale and focused on in-person cognitive behavioral therapy (CBT) and acceptance and commitment therapy, rather than mHealth [26–29]. At the time of writing, only 1 large-scale study has investigated the effectiveness of a web-based intervention for people struggling with PPU [26]. The *Hands-off* trial, currently ongoing, is a 2-armed randomized controlled trial examining the effectiveness of a 6-week web-based PPU intervention [26,30]. The intervention arm draws from techniques such as motivational therapy, CBT, and mindfulness to provide 6 modules to the participant over 6 weeks, with baseline and follow-up surveys used to measure self-reported scores on the Problematic Pornography Consumption Scale, along with other measurements such as frequency and duration of pornography use as well as mood tracking [26,30]. Preliminary analysis shows that participants in the intervention group reported lower levels of PPU use, including frequency of use, whereas the control group showed no change in use [26]. However, the *Hands-off* intervention only assesses the effectiveness of a subset of potential features for PPU interventions and thus can only enhance the credibility of a few

apps with comparable features. Furthermore, no research exists that can provide an overview of the techniques that are currently being used by consumers for PPU management.

Thus, this research aimed to categorize the features of currently available smartphone apps for managing PPU to obtain an overview of the techniques that are most prevalent and in demand. This will help to direct future research in this area and create potential therapies that can use the combined strengths of in-demand features that have been scientifically validated.

## Methods

### Feature Analysis

We performed a restricted systematic feature analysis of mobile apps available on the two leading mobile app stores: Google Play Store (Android) and Apple App Store (iOS). Apps available on either mobile phones or tablets were included in the review. The review methodology was based loosely on that used in the study by Shen et al [31], but this is not an exhaustive systematic review of all software; rather, it is an assessment of what is currently available in the mobile app space only.

On February 3, 2022, the key phrase “pornography addiction” was entered into the search bar of the 2 main mobile app stores. The results were categorized by a member of the authorship team (NH) according to a coding scheme that was iteratively updated throughout the categorization process (Textbox 1). This process was performed under the supervision of the other members of the authorship team (LD, MW, and MP). Modifications to the procedure were made when unique, nonclassifiable features were observed, in which case an adjacent category was expanded to include the feature, or a new category was created. The final coding scheme contained 16 app feature categories (Textbox 1).

**Textbox 1.** Final coding scheme used to categorize app features.

#### Feature name and type of feature

- Track: variable tracking
- Stats: statistical insights derived from variable trackers or streak trackers
- Tutor: tutorials or coaching sessions, often delivered in written, audio, or video format
- Exercise: exercises or meditations, often in the form of cognitive behavioral therapy, motivational therapy, or hypnosis
- Block: content blocking or monitoring
- Streak: tracking of streak length (the time since the last relapse), often called a day counter, streak timer, or progress tracker
- Account: accountability partner or access to a community forum
- Diary: diary or journal with the ability to take notes or set reminders
- Badge: badges, often as rewards for reaching a new streak length
- Distract: distractions for users with high urge levels, often in the form of games, soothing music, or relaxing scenery
- Quote: motivational quotes, often from famous and historical figures
- Finance: financial tracker providing an indication of money saved via abstinence from pornography use
- Locate: location tracker
- Panic: panic button, often sending the user to a site with motivational quotes, encouraging videos, or blog posts
- Religion: an explicitly religious element that is conservatively opposed to problematic pornography use
- Test: a survey that screens for pornography addiction

## Inclusion and Exclusion Criteria

Apps were screened and included in the study based on the app's description, title, and screenshots. Only apps described as helpful for reducing pornography use were included in the review, and this included apps that had a broader focus but still mentioned pornography; for example, an app designed to target *internet addiction*, *gaming addiction*, and *pornography addiction* would have been included. Apps were excluded if they did not provide sufficient information to classify the app's functions, offered pornographic content or did not claim to help participants to reduce or manage pornography use. Apps that

were apparent duplicates of higher ranked apps as determined by their near-identical descriptions or user interfaces were recorded and excluded from further analysis.

The following information was extracted from the store descriptions of the apps: app name, creator, number of user ratings, mean rating score, and number of installations (Android only). There were several notable differences between the available information in Apple's App Store and the Google Play Store ([Textbox 2](#)), making it challenging to aggregate statistics between the two stores.

**Textbox 2.** The differences between the available information in Apple's App Store and the Google Play Store.

### Notable differences between the App Store and Google Play Store with regard to app information

- The number of app installations was only provided in the Google Play Store.
- The number and value of ratings were only recorded for the most recent version of iOS apps, whereas they were recorded for all historical versions of Android apps. This reduced the number of ratings available for iOS apps but also introduced a potential bias toward previous versions of Android apps that did not contain newer features.
- Although some apps were identical between the stores, their creator was listed under a different name for each store. Hence, best judgment had to be used to categorize identical apps between the stores.
- The Apple App Store provides specific details about in-app purchases, whereas the Google Play Store offers a range of potential in-app costs. However, the variety of pricing models for each app made the cost of apps a difficult metric to assess objectively; hence, this was excluded from the analysis.

## Results

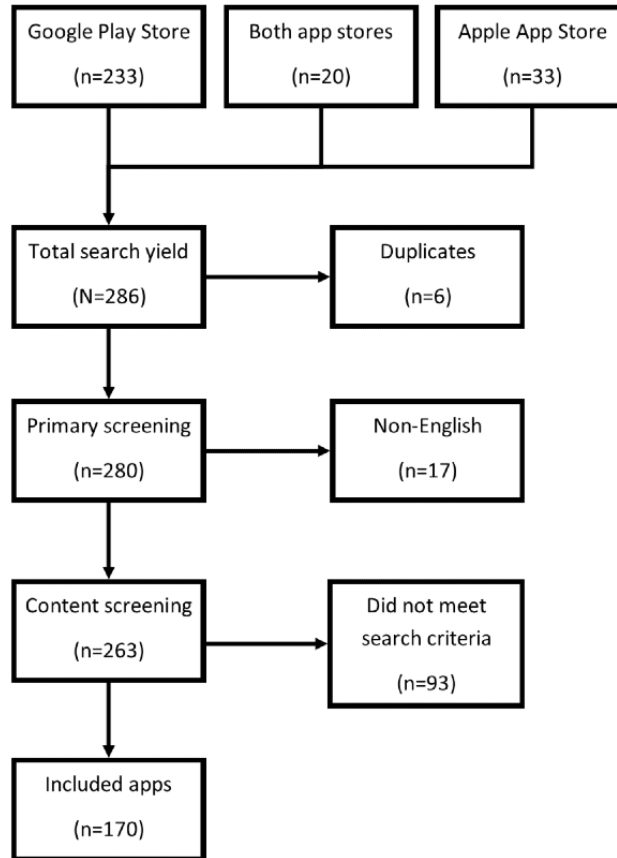
### General Characteristics

The data set for the content analysis can be found in [Multimedia Appendix 1](#). The search yielded 286 apps across both platforms, of which 17 (5.9%) apps had descriptions in languages other than English and were excluded from further analysis. Of the remaining 269 apps, 93 (34.6%) were excluded because they did not aim to manage PPU, whereas 6 (2.2%) were excluded because they were duplicating an app higher in the search

rankings. This left 170 relevant English-language apps to analyze, of which 121 (71.2%) were for Android devices only, and 29 (17.1%) were for iOS devices only, whereas 20 (11.8%) apps were found on both platforms ([Figure 1](#)).

Only the Google Play Store (Android) reported the number of app installations and only in predefined ranges. The most frequent range of installations for Android apps was 10,000 to 50,000, achieved by 25.6% (31/121) of the apps, whereas 2.5% (3/121) of the apps were installed in the range of 5 million to 10 million, the highest recorded range.

**Figure 1.** Flow diagram illustrating exclusion of apps at various stages of the study.

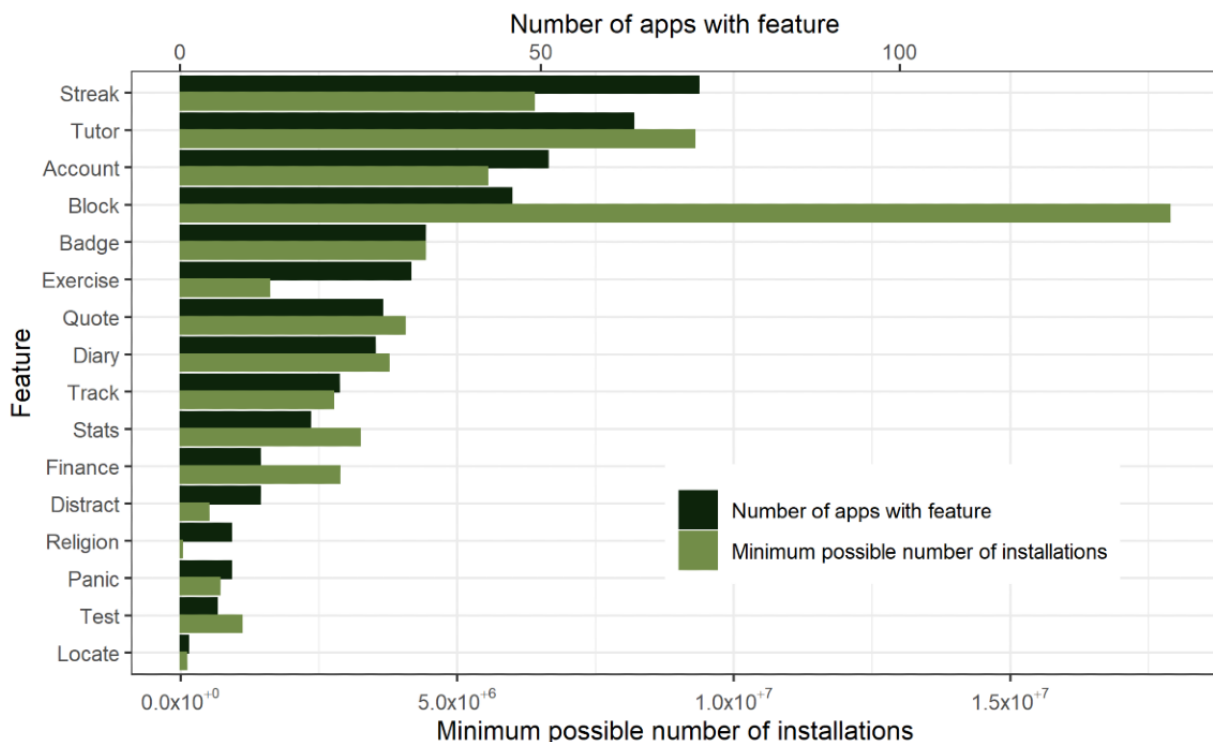


**Breakdown by Feature**

The number of apps with each feature from both stores was tallied (Figure 2). The *streak* feature, which tracks the time since the last relapse, was the most common in terms of app numbers,

with 42.4% (72/170) of the apps providing this feature. The *tutor* feature, providing tutorials and coaching for users, was the second most common feature, offered by 37.1% (63/170) of the apps.

**Figure 2.** Number of apps containing each feature for both Android and iOS, and minimum possible number of installations for each feature (Android only).

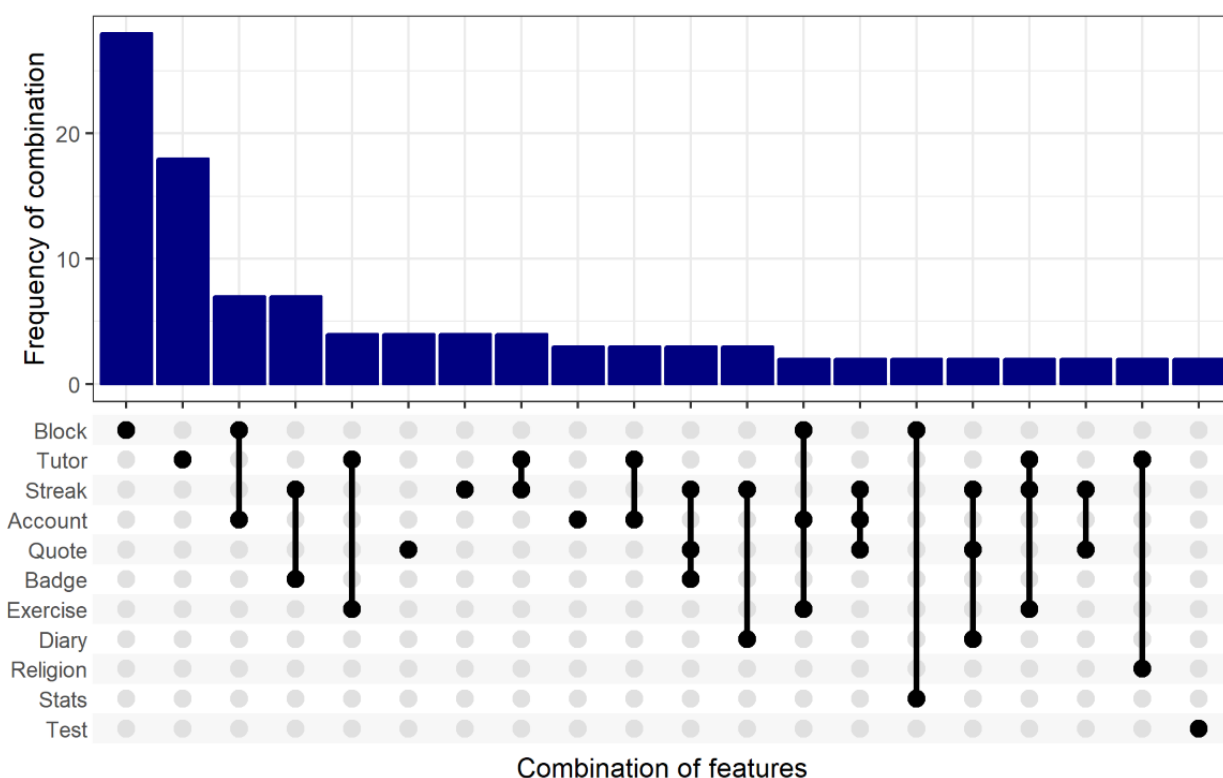


In addition, it was possible to calculate the minimum possible number of installations for the Android apps with each feature by summing the minimum value in the range of installations for each app containing the feature (Figure 2). The four most common features in terms of app number (*streak*, *tutor*, *account*, and *block*) were also the four most popular features in terms of downloads, although their rank changed slightly. Notably, the *block* feature, which provided content blocking and monitoring, was only included as a feature on 27.1% (46/170) of the apps but had the highest minimum number of installations at 17,883,000, almost double that of the next highest-ranked feature (*tutor* with 9,286,505 installations).

A combinatorial analysis was also performed. The frequency of different combinations of features was plotted to determine

the most common combinations, both in terms of app frequency and the minimum possible number of installations. The results can be seen in Figure 3 for the 20 most frequent feature combinations. Of note, the two top feature combinations were single features—*block* and *tutor*—both in terms of the number of apps with only these features (28/170, 16.5%, and 18/170, 10.6%, respectively) and in terms of the minimum possible number of installations (16,002,000 and 7,244,205, respectively). The following 3 most popular combinations of features (based on minimum possible number of installations) were all individual apps: BlockerX used the *block*, *account*, *diary*, and *test* features; Quizzilla used the *stats*, *streak*, *diary*, *badge*, *quote*, and *finance* features; and I Am Sober used the *track*, *streak*, *account*, *quote*, and *finance* features.

Figure 3. Combinatorial plot showing the most common combinations of features. Only the top 20 combinations are shown.

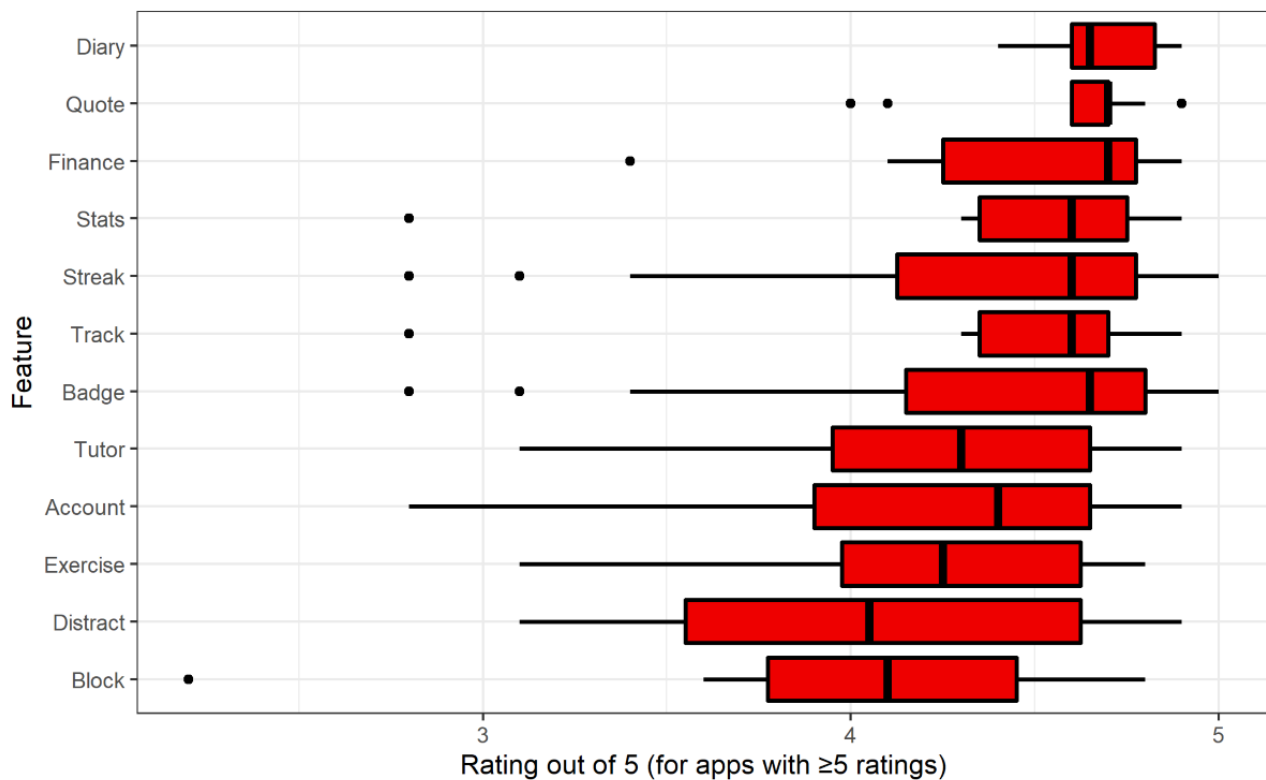


App ratings out of 5 were aggregated between both stores and grouped by feature. The distributions of these ratings are summarized in Figure 4, ordered by mean rating. Only apps that had received at least five ratings were recorded for each feature, and only features with at least five eligible app ratings were plotted. None of the 4 most common features remained in the top 4 highest-rated apps by mean or median rating; *streak* came fifth on both counts, with a mean rating of 4.34 and median rating of 4.60. The *quote* and *finance* features, providing motivational quotes and financial tracking, respectively, had the equal highest median rating of 4.70, with *diary*, the journaling feature, having the highest mean rating of 4.69. *Distract*, a feature consisting of distracting activities, had the lowest median rating (4.00). *Block* had the lowest mean rating

(4.04) and second-lowest median rating (4.10), despite having the highest minimum possible number of installations. *Tutor* and *account* (a feature providing an accountability partner), both of which scored highly in the minimum possible number of installations and commonality, were either eighth or ninth in the mean and median rating scores (means of 4.24 and 4.21, respectively, and medians of 4.30 and 4.40, respectively). This may indicate that features in higher demand are held to a high standard.

It should be noted that the ratings shown in Figure 4 only consist of apps that received enough ratings to be included in the results. Hence, there may be some bias toward superior-quality apps, as recognized by the market population.

**Figure 4.** Distribution of ratings for apps with each feature, for Android and iOS, ordered by mean rating. Only apps that had received at least five ratings were recorded for each feature, and only features with at least five eligible app ratings were plotted.



### Review of Features Based on Installations

Of the 170 included apps, 46 (27.1%) provided the *block* feature, which provided content blocking or monitoring. Going by the minimum-possible-number-of-installations metric, this feature was the most popular by a large margin (with a score of 16,002,000). Most of this score was due to apps such as BlockSite, which has >5,000,000 downloads on the Google Play Store alone. Of these 46 apps, 15 (32.6%) provided content blocking associated with the *account* feature, usually providing an accountability partner. An example of an app with only these 2 features was Triple, which has been downloaded >100,000 times on the Google Play Store alone. This combination of features was the third most common, with 4.1% (7/170) of the apps having this combination. Still, in the ranking of the minimum possible number of installations, this combination only came 12th with a score of 270,000. However, when combined with other features, the minimum possible number of installations increased to 1,780,000. Most of this score came from the BlockerX app, which has >1,000,000 downloads on the Google Play Store alone.

Of the 170 included apps, 63 (37.1%) provided the *tutor* feature, which offered tutorials or coaching in written, audio, or video format; for example, the Life After Pornography Coach app offers >80 video lessons to aid users in their journey. However, for many apps, tutorials were not necessarily the app's main feature but were offered as part of a complete feature set.

Of the 170 included apps, 32 (18.8%) provided the *exercise* feature, with exercises such as guided meditations or modules for CBT and motivational therapy. This feature was often combined with the *tutor* feature; for instance, an audio recording

often accompanied guided meditations to direct the user through the meditation. Furthermore, 6.5% (11/170) of the apps provided the *distract* feature, often in the form of games or soothing music for users who were feeling urges.

Of the 170 included apps, 51 (30%) provided access to the *account* feature, either in the form of a web-based forum where users could discuss their PPU with others or the ability to communicate with an accountability partner; for example, users were often redirected to subreddit forums such as r/NoFap (which encourages abstinence from pornography, masturbation, and orgasm for 90 days; the name comes from the slang term "fap," referring to masturbation) or r/pornfree (which encourages abstinence from pornography-induced masturbation but not masturbation itself).

Of the 170 included apps, 72 (42.4%) provided the *streak* feature, allowing users to track their *streak length*: the time (usually in days) since their most recent lapse. This feature was often paired with the *badge* feature, giving the user extra motivation to reach their next target or badge, and also with the *stats* feature, which provided statistical feedback on the user's streak history and associated metrics.

Of the 170 included apps, 28 (16.5%) provided the *quote* feature in the form of motivational quotes, usually as a secondary feature of a more extensive program. Associated with this, 4% (7/170) of the apps had the *panic* feature, generally in the form of a *panic button*, which the user can press when their urge levels escalate. This typically directed a user to a website with randomized links, often with motivational quotes or a suggestion to perform a quick workout, such as 20 push-ups. An example

of this is the panic button created by the r/NoFap subreddit community.

Of the 170 included apps, 27 (15.9%) apps provided the *diary* feature, allowing users to record self-reflective notes on their recovery process. This helps users to identify triggers, relapse pathways, and successful strategies for overcoming PPU.

Of the 170 included apps, 22 (12.9%) provided the *track* feature, allowing users to record their variables on a daily or instantaneous basis. Of these 22 apps, 10 (45.5%) also provided the *stats* feature, generally as a summary of insights derived from these tracked variables, often in association with the *streak* feature. Of the 170 included apps, 11 (6.5%) also included the *finance* feature, revealing how much money the user was potentially saving by reducing their pornography use.

The three least popular features were *test*, *locate*, and *religion*. Only 2.9% (5/170) of the apps provided any form of survey or test for PPU, whereas 0.6% (1/170) of the apps provided *locate* as a feature. This feature represents a type of relapse prediction service, in that the aim is to learn and highlight locations where the user may be more prone to relapse. Only 4.1% (7/170) of the apps had the *religion* feature, which had the lowest minimum possible number of installations of all features listed, with a score of 21,110. Of these 7 apps, 6 (86%) also provided the *tutor* feature, generally providing helpful scriptures, devotional passages, or prayers.

## Discussion

### Principal Findings

It seems that the academic literature on PPU is divorced from the technological interventions currently available. None of the apps reviewed contained references to literature that provided direct evidence for the app's efficacy or safety. However, several apps reported anecdotal evidence that supported their real-world effectiveness. Furthermore, a lack of scientific studies examining a feature's effectiveness does not necessarily mean that there is no rationale for specific in-app features. The perceived utility of a feature can be partially inferred from its demand in the real world, even if the implementation is imperfect. The relationship between feature utility and real-world demand may be mediated through the number of positive reviews left on apps with the feature, which is indicative of the feature's effectiveness and likely to encourage a higher number of installations for these apps. Still, further research is required to consolidate the evidence base into a white list of safe apps for public consumption.

### Utility of Features

On the basis of the minimum possible number of installations, content-blocking apps are the most popular on the Google Play Store but have a lower mean rating than apps that do not offer content blocking. The high popularity score suggests that many users attempt to distance themselves from sources of pornography while trying to remain connected to smartphones and the internet. There is much anecdotal evidence on the internet supporting these apps, which are also popular in related fields such as internet addiction. Notably, Lyngs et al [32] performed a systematic review of DSCTs and found that the

most prevalent feature (found in 74% of the tools) involved blocking or removing distractions. Still, despite the robust demand for these apps and their apparent utility, there are no studies investigating the effectiveness of content-blocking apps with regard to PPU reduction. However, the comparatively low mean rating suggests that these apps are not as effective as their users would hope, compared with their expectations for other features.

It is possible that these content-blocking apps may be advertised more than apps with other features, and customers may resent paying for these apps, which are generally more expensive than other apps because of technical requirements such as virtual private network services. By contrast, as discussed earlier, it is challenging to create and maintain a content filter with sufficient sensitivity and specificity to satisfy the market mainly because web-based content keeps evolving. The mismatch between the difficulty of this task and the expected quality of these apps may be more significant than the difficulty-quality mismatch for other features; for instance, an app that provides CBT tutorials may only need to provide minimal benefit for its users to rate it highly because the user will not blame the app for causing a relapse. However, if a content-blocking app fails to block a new porn site, which leads to a user relapsing, this will likely lead to a negative review. This may explain why the *block* feature scored very highly in terms of the minimum possible number of installations but had a relatively low median rating compared with other features.

An interesting alternative to counteract this weakness is Triple, an app that takes screenshots of all browsing behavior, analyses them, and flags potential sexual content to an accountability partner [33]. In theory, this reduces the temptation to uninstall or bypass the app, which is a problem for some apps with the *block* feature. However, as is the case with using a therapist, not all users are willing to share their private browsing information with even trusted friends, which may reduce use of this feature. In addition, the app does not prevent users from viewing the websites, which indicates that it may be fortified by the addition of the *block* and *quote* features. Motivational quotes have some potential to provide positive boosts to self-esteem and self-efficacy for patients using mHealth interventions to recover from PPU [34]. The addition of motivational quotes may provide timely reminders to stay away from distracting material that has the potential to devolve into PPU.

Badges seem to have potential as a means of gamifying the recovery process and generating positive reinforcement [35] and have been shown to improve motivation in therapeutically administered tasks [36]. However, the use of progress trackers is controversial. Despite being popular, it is plausible they can be detrimental to users who relapse because they may induce additional feelings of guilt and anxiety upon breaking a streak, perpetuating the feelings that led to a relapse in the first place [31]. Still, these feelings may be a necessary part of the recovery process. Using the prospect theory framework developed by Kahneman and Tversky [37], we can define the risk of experiencing postrelapse depression as a potential loss, which the user must compare with the momentary pleasure of relapse as a gain. In this framework, healthy use of the *streak* feature



creates a positive feedback loop, whereby the user experiences increasing loss aversion as their streak length increases, and memories of past failures inspire greater motivation to reduce future losses. However, an alternative positive feedback loop is also possible, where the feelings of guilt and depression brought about by relapse cause the user to turn to pornography to avoid these feelings [38]. These questions invite further research to determine whether the utility of this feature varies within a population of participants struggling with PPU.

In many ways, apps combining the *track* and *stats* features act as a self-monitored ecological momentary assessment, where the user both records and analyses their data [39]. Such an app's effectiveness will be inherently limited by the ability of its users to interpret such data accurately, particularly those users without a sufficient level of scientific training. Hence, app creators must balance the complexity of the statistical analysis they provide against the ability of their users to interpret these data meaningfully.

The Fortify app (as part of Impact Suite) allows users to track their variables over time, including urges and variables related to well-being. It then performs analytics on these variables over time, allowing users to self-examine their weaknesses. The app also includes a personal journal, training and meditation sessions, community forums, and accountability partners. This is the only mobile app we found with a comparable alternative in the literature, although that trial (the *Hands-off* study) is not yet complete and focuses more on providing coaching modules and exercises for users [30]. However, it has shown promising results so far, with indications that the combination of variable tracking with modules and community messaging are an effective way to manage PPU [26].

Although several apps provide the ability to record one's urges at the end of each day, we found no app that allows the user to record their urges instantaneously, with associated follow-up recommendations such as taking a break from the internet or performing exercise. If a user performs urge tracking at regular intervals, they can monitor responses to various stimuli in real time and adjust their behaviors when urge levels increase; for example, a user may find that their urges increase rapidly while browsing a particular website and choose to regulate their exposure to this website. In addition, once they have been tracking their urges for a while, the user may become more attuned to scenarios that increase their urges and formulate strategies to escape these scenarios. This could take the form of a self-regulated JITAI, where the user records their urge levels and responds to the app's subsequent prompt to change their environment, perform exercise, or meditate.

### Motivational Readiness

Despite their lack of grounding in the literature, many of the apps in this analysis seem to have features that align with addiction-management frameworks such as self-regulation theory [40] and self-efficacy theory [41]. However, one theory allows us to quantify the game-theoretic effectiveness of the features seen in this review: the dual-threshold model of motivational readiness developed by Kruglanski [42,43]. In this theory, the degree to which a person pursues a goal state (such as gratification of pornographic urges) is a function of both their

motivational readiness to achieve the state and their belief that the state can be achieved (expectancy) [42]. To put this in the context of PPU, if the user believes that they can access pornography (expectancy), and their desire for pornography is stronger than their will to reduce PPU (motivational readiness), then they will take steps to use pornography in proportion to their desire to use it. However, if they do not believe that they can access pornography, then they will not attempt to use pornography, irrespective of their desire to do so.

On the basis of this model, we propose two potential strategies to reduce PPU: either raise the person's motivation to reduce PPU (strategy 1) or increase barriers to pornography access (strategy 2). From a game theory perspective, strategy 2 is superior to strategy 1 because if the user does not meet the requisite threshold for expectancy, they cannot take action to use pornography, regardless of their motivation to do so.

If we try to categorize the features presented in [Textbox 1](#) based on these strategies, only two features, *block* and *locate*, clearly fall under strategy 1, whereas the other features all fall under strategy 2 (excluding *test*, which is for assessment only). The *locate* feature has a fundamental flaw, in that users regularly spend time in what the feature would consider a high-risk location for pornography use: their own bedroom. Hence, *block* is perhaps the only feature with the potential to eliminate PPU entirely by simply removing the ability to view pornography. Of course, it is virtually impossible to cut off all sources of pornography in the postinternet age. However, the hypothetical superiority of strategy 2 may explain the popularity of content-blocking apps in this paper's analysis.

### Defense in Depth, or the Swiss Cheese Model

Given that both these strategies have their weaknesses, a potential user may consider combining the strategies to reduce PPU; for example, they may use a *block* app in conjunction with a *track* app to simultaneously monitor their urge levels, place barriers between themselves and exposure, and predict times and environments where they will be at higher risk of relapse. If the *exercise* and *tutor* features with training modules and CBT elements are added, the user would have tools available to both increase their motivation and reduce their expectancy.

This combining of techniques is an example of the *defense in depth* (DiD), or *Swiss cheese* model of risk management, applied to PPU. DiD is a concept created by the National Security Agency to harden their computer security. Multiple layers of defense are placed within a DiD system to provide redundancy of protection if a vulnerability is uncovered [44]. Applying this model to PPU, one could view pornography as a *virus* targeted at the brain. If the user has not put up enough layers of *security*, then the virus will *infiltrate* the mind and cause a relapse. Using one layer of security such as the *block* feature will provide some protection but leave multiple vulnerabilities open for the virus to penetrate, much like the holes in a slice of Swiss cheese. However, combining multiple layers of security (for instance, adding regular CBT to the user's routines) will reduce the probability of relapse even further by removing some of the *holes* left behind by prior layers of security. Having a unique feature combination may also help create a novel market niche, which seems to have been a successful strategy for several apps.

The DiD framework may also benefit apps seeking to differentiate themselves in a competitive app market. From the combinatorics analysis performed in this review, it can be deduced that there are two main strategies implemented by the most successful apps in this space: (1) either implement one desirable feature (such as *block*) exceptionally well, or (2) create a niche market by combining a unique set of features. The second strategy, which aligns with the DiD framework, has the potential upside of avoiding competition from numerous other apps, which is the challenge facing apps that only use either the *block* feature or the *tutor* feature.

Even when combined, these technologies are likely not entirely sufficient. People who struggle with compulsive behaviors tend to search for new pathways to indulging in their behavior and may find ways to bypass multiple protective layers, mainly when triggered or under stress [23]. Hence, the defense set up by the user should comprise both static layers that are always present and dynamic layers that are updated regularly to adjust for any blind spots that may appear [44]. Using this framework, an example of a dynamic layer would be using the *track* feature to identify new triggers, whereas an example of a static layer would be using pornography blocker software on a web browser. Most users trying to overcome PPU will find that complete abstinence is difficult, but by using the DiD model, they can treat each relapse as a weakness to be fixed in their quest to eliminate PPU.

In addition, the *track*, *streak*, and *stats* features can guide the user to adjust their strategy when new weaknesses are uncovered through self-analysis of their data. However, not all users will be equipped with the ability to interpret their data accurately. Furthermore, a software-generated interpretation of user data may provide some insights. Still, it will be limited in the amount of nuance it can collect, particularly when the user's data do not match regular trends. This is where a trained therapist would prove to be a significant addition to the DiD strategy, in that they could provide insights that both the user and their software tool are unable to uncover. The therapist will also be less vulnerable to self-evaluation bias.

### Weaknesses of mHealth Solutions

The use of apps to counter apps is counterintuitive [13] but may be necessary for certain users to form new habits. One of the difficulties of providing a technological intervention for PPU is that pornography is delivered on the same technologies; for example, if one uses a mobile app to track one's relapse statistics, one will likely be using the same device one uses to access pornography. This can have both positive and negative consequences. Some app users may be able to use the intervention to develop alternative habits. However, it is possible that for other users, the reverse will happen, where they link the intervention to continued pornography use, making it a double-edged sword. A separate device (such as a pager) is always an option, but these devices generally have limited functionality compared with a smartphone. To increase the reach of an app, highly accessible technologies should be used, which in the case of pornography users is most likely their smartphone [45].

No apps in this review had any screening tools for PPU grounded in the literature. It could be argued that a scientifically validated assessment should be mandatory for any app claiming to help with PPU management. Although there remains controversy around the assessment criteria for PPU [46], the lack of a test implies that the user has a problem by default. It could be argued that a false positive assessment of PPU can induce moral incongruence and subsequent anxiety or depression, potentially leading to a positive feedback loop that increases PPU in some users.

This also highlights that many web-based tools are most effective when guided and recommended by a qualified therapist [47]. Much work needs to be done to bring the apps reviewed in this study to the therapeutic space, especially considering the lack of therapists trained to treat PPU [48,49]. A therapist may also be able to warn against unintended consequences of certain apps; for example, several apps in this review provided distractions in the form of games or calming music. This feature could be viewed as an alternative to guided meditation, in that it offers a way to direct attention away from pornography. However, many users also struggle with internet or gaming addictions, meaning that there is a risk that the user could justify replacing PPU with addictive gaming behaviors. It is possible that the most effective apps for reducing PPU are designed so that the user no longer requires the app once they have achieved their goals and even uses their smartphone less, reducing temptation risk in general. This may be more difficult for apps with the *distract* feature if the app has an addictive design.

### Limitations

The aforementioned quantitative feature analysis focused on mobile apps and excluded other viable software; for example, several browser extensions provide a *block* feature, such as uBlock Origin [21] or BlockSite [22], that may not specifically be designed for reducing PPU but can be used to the same effect. One such app, Plucky, blocks all images and videos on the internet (besides a white list of allowed websites) and makes itself *inconvenient* both to uninstall and to edit the white list [50]. The user can customize the app but only after a predefined delay period. In theory, this provides the user with time to overcome their urges and resume baseline behavior.

There are several apps available on the internet that provide variations on the accountability feature, including X3Watch [51] and Qustodio [52], that were excluded from this review. Many of these apps are designed to broadly reduce internet use, of which pornography is a use case. Even the r/NoFap subreddit and the associated NoFap website could be considered a technological *intervention* of sorts, providing the *streak*, *account*, *tutor*, *exercise*, and *panic* features [53]. In summary, there are many self-help resources and associated software-based interventions available on the web, many of which could be used to complement the mobile apps included in this review.

Any statistical analysis of mobile app stores, including descriptive statistics, is hampered by a lack of information. Both app stores use various methods, including hiding in-app payment and advertising information and using large categories for aggregating and blurring popularity metrics. Variables are not identical between the stores and reflect the preferences of

different customer bases; for instance, on average, Apple customers have greater financial resources than Android customers because Apple mobile phones tend to be more expensive than equivalent Android mobile phones [54]. In addition, it is impossible to deduce how much promotional work has been done behind the scenes to advertise these apps. Hence, the popularity metrics provided here are a poor representation of the actual demand for, and quality of, features and should only be interpreted as indicative of potential trends in this space. Non-English-language apps were also excluded, providing another potential source of bias to these results.

Furthermore, any measure of popularity is likely to be biased in cases where apps are used for purposes other than managing PPU; for example, BlockSite, which has at least 5,000,000 downloads on the Google Play Store and contains the *block* feature, is also provided to the user for *blocking distracting websites*, not just pornographic ones [22]. Removing this app and others with multiple purposes may affect the ranking of features in the popularity metrics used for this study.

## Conclusions

The mobile app space is replete with apps that seem to be beneficial to users attempting to manage PPU based on their generally high app store ratings, positive reviews, and high number of installations. In particular, content-blocking apps have clear potential for reducing PPU by removing access to pornography on the user's device. However, there remains a substantive lack of evidence in the scientific literature to quantify the effectiveness of such apps. Although each of these apps has its purported benefits, the most effective method for reducing PPU may lie in combining the most robust features in this space, using a DiD strategy of risk management. This could be achieved by using multiple apps at once, although this should probably be performed under the guided hand of a trained therapist, if possible. However, such a strategy is inefficient and cumbersome. Further work is required in this space, both to research the effectiveness of these app features and to consolidate the evidence base into a white list of safe apps for public consumption.

## Authors' Contributions

NH performed the content analysis and was the primary author of this paper. LD, MW, and MP supervised and critiqued the analysis and contributed to the writing and editing of the paper.

## Conflicts of Interest

None declared.

## Multimedia Appendix 1

This Microsoft Excel file contains the data set generated during the app review. It is separated into two sheets: the first contains the data set, and the second contains a data dictionary describing each variable in the data set.

[[XLSX File \(Microsoft Excel File\), 122 KB - formative\\_v6i10e39869\\_app1.xlsx](#)]

## References

1. Perry SL. Pornography use and depressive symptoms: examining the role of moral incongruence. *Soc Ment Health* 2017 Sep 01;8(3):195-213. [doi: [10.1177/2156869317728373](https://doi.org/10.1177/2156869317728373)]
2. Lewczuk K, Wójcik A, Gola M. Increase in the prevalence of online pornography use: objective data analysis from the period between 2004 and 2016 in Poland. *Arch Sex Behav* 2022 Feb;51(2):1157-1171 [FREE Full text] [doi: [10.1007/s10508-021-02090-w](https://doi.org/10.1007/s10508-021-02090-w)] [Medline: [34750777](https://pubmed.ncbi.nlm.nih.gov/34750777/)]
3. The 2019 Year in Review - Pornhub Insights. Pornhub. 2019 Dec 11. URL: <https://www.pornhub.com/insights/2019-year-in-review> [accessed 2022-03-05]
4. Cooper A, Putnam DE, Planchon LA, Boies SC. Online sexual compulsivity: getting tangled in the net. *Sex Addict Compulsivity* 1999 Apr;6(2):79-104. [doi: [10.1080/10720169908400182](https://doi.org/10.1080/10720169908400182)]
5. Chasioti D, Binnie J. Exploring the etiological pathways of problematic pornography use in NoFap/PornFree rebooting communities: a critical narrative analysis of internet forum data. *Arch Sex Behav* 2021 Jul;50(5):2227-2243 [FREE Full text] [doi: [10.1007/s10508-021-01930-z](https://doi.org/10.1007/s10508-021-01930-z)] [Medline: [34143364](https://pubmed.ncbi.nlm.nih.gov/34143364/)]
6. Schaub MP, Blankers M, Lehr D, Boss L, Riper H, Dekker J, et al. Efficacy of an Internet-based self-help intervention to reduce co-occurring alcohol misuse and depression symptoms in adults: study protocol of a three-arm randomised controlled trial. *BMJ Open* 2016 May 25;6(5):e011457 [FREE Full text] [doi: [10.1136/bmjopen-2016-011457](https://doi.org/10.1136/bmjopen-2016-011457)] [Medline: [27225652](https://pubmed.ncbi.nlm.nih.gov/27225652/)]
7. Haug S, Paz Castro R, Wenger A, Schaub MP. Efficacy of a mobile phone-based life-skills training program for substance use prevention among adolescents: study protocol of a cluster-randomised controlled trial. *BMC Public Health* 2018 Sep 10;18(1):1102 [FREE Full text] [doi: [10.1186/s12889-018-5969-5](https://doi.org/10.1186/s12889-018-5969-5)] [Medline: [30200928](https://pubmed.ncbi.nlm.nih.gov/30200928/)]
8. Fernandez DP, Kuss DJ, Griffiths MD. The pornography "rebooting" experience: a qualitative analysis of abstinence journals on an online pornography abstinence forum. *Arch Sex Behav* 2021 Feb;50(2):711-728 [FREE Full text] [doi: [10.1007/s10508-020-01858-w](https://doi.org/10.1007/s10508-020-01858-w)] [Medline: [33403533](https://pubmed.ncbi.nlm.nih.gov/33403533/)]
9. Wolak J, Mitchell K, Finkelhor D. Unwanted and wanted exposure to online pornography in a national sample of youth Internet users. *Pediatrics* 2007 Feb;119(2):247-257. [doi: [10.1542/peds.2006-1891](https://doi.org/10.1542/peds.2006-1891)] [Medline: [17272613](https://pubmed.ncbi.nlm.nih.gov/17272613/)]

10. Cooper A. Sexuality and the Internet: surfing into the new millennium. *CyberPsychol Behav* 1998 Jan;1(2):187-193. [doi: [10.1089/cpb.1998.1.187](https://doi.org/10.1089/cpb.1998.1.187)]
11. Mihajlov M, Vejmelka L. Internet addiction: a review of the first twenty years. *Psychiatr Danub* 2017 Sep;29(3):260-272 [FREE Full text] [doi: [10.24869/psyd.2017.260](https://doi.org/10.24869/psyd.2017.260)] [Medline: [28949307](https://pubmed.ncbi.nlm.nih.gov/28949307/)]
12. Regan H. Playboy and pornification: 65 years of the playboy centerfold. *Sex Cult* 2021 Jan 08;25(3):1058-1075. [doi: [10.1007/s12119-020-09809-2](https://doi.org/10.1007/s12119-020-09809-2)]
13. Aboujaoude E. Problematic Internet use two decades later: apps to wean us off apps. *CNS Spectr* 2019 Aug;24(4):371-373. [doi: [10.1017/S109285291800127X](https://doi.org/10.1017/S109285291800127X)] [Medline: [30507371](https://pubmed.ncbi.nlm.nih.gov/30507371/)]
14. Sedrati H, Belrhiti Z, Nejari C, Ghazal H. Evaluation of mobile health apps for non-medical cannabis use: a scoping review. *Procedia Comput Sci* 2022;196(C):581-589. [doi: [10.1016/j.procs.2021.12.052](https://doi.org/10.1016/j.procs.2021.12.052)]
15. Plaza I, Demarzo MM, Herrera-Mercadal P, García-Campayo J. Mindfulness-based mobile applications: literature review and analysis of current features. *JMIR Mhealth Uhealth* 2013 Nov 01;1(2):e24 [FREE Full text] [doi: [10.2196/mhealth.2733](https://doi.org/10.2196/mhealth.2733)] [Medline: [25099314](https://pubmed.ncbi.nlm.nih.gov/25099314/)]
16. Rodriguez-Paras C, Tippey K, Brown E, Sasangohar F, Creech S, Kum HC, et al. Posttraumatic stress disorder and mobile health: app investigation and scoping literature review. *JMIR Mhealth Uhealth* 2017 Oct 26;5(10):e156 [FREE Full text] [doi: [10.2196/mhealth.7318](https://doi.org/10.2196/mhealth.7318)] [Medline: [29074470](https://pubmed.ncbi.nlm.nih.gov/29074470/)]
17. Colbert S, Thornton L, Richmond R. Smartphone apps for managing alcohol consumption: a literature review. *Addict Sci Clin Pract* 2020 May 07;15(1):17 [FREE Full text] [doi: [10.1186/s13722-020-00190-x](https://doi.org/10.1186/s13722-020-00190-x)] [Medline: [32381062](https://pubmed.ncbi.nlm.nih.gov/32381062/)]
18. Staiger PK, O'Donnell R, Liknaitzky P, Bush R, Milward J. Mobile apps to reduce tobacco, alcohol, and illicit drug use: systematic review of the first decade. *J Med Internet Res* 2020 Nov 24;22(11):e17156 [FREE Full text] [doi: [10.2196/17156](https://doi.org/10.2196/17156)] [Medline: [33231555](https://pubmed.ncbi.nlm.nih.gov/33231555/)]
19. Perez M, Avila S, Moreira D, Moraes D, Testoni V, Valle E, et al. Video pornography detection through deep learning techniques and motion information. *Neurocomputing* 2017 Mar;230:279-293. [doi: [10.1016/j.neucom.2016.12.017](https://doi.org/10.1016/j.neucom.2016.12.017)]
20. Richardson CR, Resnick PJ, Hansen DL, Derry HA, Rideout VJ. Does pornography-blocking software block access to health information on the Internet? *JAMA* 2002 Dec 11;288(22):2887-2894. [doi: [10.1001/jama.288.22.2887](https://doi.org/10.1001/jama.288.22.2887)] [Medline: [12472332](https://pubmed.ncbi.nlm.nih.gov/12472332/)]
21. uBlock Origin - Free, open-source ad content blocker. uBlock Origin. URL: <https://ublockorigin.com/> [accessed 2022-05-21]
22. BlockSite: Easily block distracting websites and apps. BlockSite. URL: <https://blocksite.co/> [accessed 2022-05-21]
23. Gardner EL. Addiction and brain reward and antireward pathways. *Adv Psychosom Med* 2011;30:22-60 [FREE Full text] [doi: [10.1159/000324065](https://doi.org/10.1159/000324065)] [Medline: [21508625](https://pubmed.ncbi.nlm.nih.gov/21508625/)]
24. Nahum-Shani I, Smith SN, Spring BJ, Collins LM, Witkiewitz K, Tewari A, et al. Just-in-Time Adaptive Interventions (JITAs) in mobile health: key components and design principles for ongoing health behavior support. *Ann Behav Med* 2018 May 18;52(6):446-462 [FREE Full text] [doi: [10.1007/s12160-016-9830-8](https://doi.org/10.1007/s12160-016-9830-8)] [Medline: [27663578](https://pubmed.ncbi.nlm.nih.gov/27663578/)]
25. Gustafson DH, McTavish FM, Chih MY, Atwood AK, Johnson RA, Boyle MG, et al. A smartphone application to support recovery from alcoholism: a randomized clinical trial. *JAMA Psychiatry* 2014 May;71(5):566-572 [FREE Full text] [doi: [10.1001/jamapsychiatry.2013.4642](https://doi.org/10.1001/jamapsychiatry.2013.4642)] [Medline: [24671165](https://pubmed.ncbi.nlm.nih.gov/24671165/)]
26. Bóthe B, Baumgartner C, Schaub MP, Demetrovics Z, Orosz G. Hands-off: feasibility and preliminary results of a two-armed randomized controlled trial of a web-based self-help tool to reduce problematic pornography use. *J Behav Addict* 2021 Oct 29;10(4):1015-1035 [FREE Full text] [doi: [10.1556/2006.2021.00070](https://doi.org/10.1556/2006.2021.00070)] [Medline: [34727088](https://pubmed.ncbi.nlm.nih.gov/34727088/)]
27. Crosby JM, Twohig MP. Acceptance and commitment therapy for problematic internet pornography use: a randomized trial. *Behav Ther* 2016 May;47(3):355-366. [doi: [10.1016/j.beth.2016.02.001](https://doi.org/10.1016/j.beth.2016.02.001)] [Medline: [27157029](https://pubmed.ncbi.nlm.nih.gov/27157029/)]
28. Twohig MP, Crosby JM. Acceptance and commitment therapy as a treatment for problematic Internet pornography viewing. *Behav Ther* 2010 Sep;41(3):285-295. [doi: [10.1016/j.beth.2009.06.002](https://doi.org/10.1016/j.beth.2009.06.002)] [Medline: [20569778](https://pubmed.ncbi.nlm.nih.gov/20569778/)]
29. Minarcik J. Proposed Treatment of Problematic Pornography Use: A Cognitive-Behavioral Approach. University of Arkansas. 2016 Aug 1. URL: <https://scholarworks.uark.edu/etd/1453> [accessed 2022-10-03]
30. Bóthe B, Baumgartner C, Schaub MP, Demetrovics Z, Orosz G. Hands-off: study protocol of a two-armed randomized controlled trial of a Web-based self-help tool to reduce problematic pornography use. *J Behav Addict* 2020 Jun;9(2):433-445 [FREE Full text] [doi: [10.1556/2006.2020.00037](https://doi.org/10.1556/2006.2020.00037)] [Medline: [32663382](https://pubmed.ncbi.nlm.nih.gov/32663382/)]
31. Shen N, Levitan MJ, Johnson A, Bender JL, Hamilton-Page M, Jadad AA, et al. Finding a depression app: a review and content analysis of the depression app marketplace. *JMIR Mhealth Uhealth* 2015 Feb 16;3(1):e16 [FREE Full text] [doi: [10.2196/mhealth.3713](https://doi.org/10.2196/mhealth.3713)] [Medline: [25689790](https://pubmed.ncbi.nlm.nih.gov/25689790/)]
32. Lyngs U, Lukoff K, Slovak P, Binns R, Slack A, Inzlicht M, et al. Self-control in cyberspace: applying dual systems theory to a review of digital self-control tools. In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 2019 Presented at: CHI '19; May 4-9, 2019; Glasgow, UK p. 1-18. [doi: [10.31235/osf.io/685tm](https://doi.org/10.31235/osf.io/685tm)]
33. Defeat porn and online filth with Screenshot Accountability. Triple. URL: <https://triple.io/> [accessed 2022-05-21]
34. Bedrov A, Bulaj G. Improving self-esteem with motivational quotes: opportunities for digital health technologies for people with chronic disorders. *Front Psychol* 2018 Nov 2;9:2126 [FREE Full text] [doi: [10.3389/fpsyg.2018.02126](https://doi.org/10.3389/fpsyg.2018.02126)] [Medline: [30450071](https://pubmed.ncbi.nlm.nih.gov/30450071/)]

35. Rajani NB, Mastellos N, Filippidis FT. Impact of gamification on the self-efficacy and motivation to quit of smokers: observational study of two gamified smoking cessation mobile apps. *JMIR Serious Games* 2021 Apr 27;9(2):e27290 [FREE Full text] [doi: [10.2196/27290](https://doi.org/10.2196/27290)] [Medline: [33904824](https://pubmed.ncbi.nlm.nih.gov/33904824/)]
36. Cheng VW, Davenport T, Johnson D, Vella K, Hickie IB. Gamification in apps and technologies for improving mental health and well-being: systematic review. *JMIR Ment Health* 2019 Jun 26;6(6):e13717 [FREE Full text] [doi: [10.2196/13717](https://doi.org/10.2196/13717)] [Medline: [31244479](https://pubmed.ncbi.nlm.nih.gov/31244479/)]
37. Kahneman D, Tversky A. Prospect theory: an analysis of decision under risk. *Econometrica* 1979 Mar;47(2):263-292. [doi: [10.2307/1914185](https://doi.org/10.2307/1914185)]
38. Grubbs JB, Stauner N, Exline JJ, Pargament KI, Lindberg MJ. Perceived addiction to Internet pornography and psychological distress: examining relationships concurrently and over time. *Psychol Addict Behav* 2015 Dec;29(4):1056-1067. [doi: [10.1037/adb0000114](https://doi.org/10.1037/adb0000114)] [Medline: [26372200](https://pubmed.ncbi.nlm.nih.gov/26372200/)]
39. Shiffman S, Stone AA, Hufford MR. Ecological momentary assessment. *Annu Rev Clin Psychol* 2008;4:1-32. [doi: [10.1146/annurev.clinpsy.3.022806.091415](https://doi.org/10.1146/annurev.clinpsy.3.022806.091415)] [Medline: [18509902](https://pubmed.ncbi.nlm.nih.gov/18509902/)]
40. Mithaug DE. *Self-Regulation Theory: How Optimal Adjustment Maximizes Gain*. Westport, CT, USA: Praeger Publishers/Greenwood Publishing Group; 1993.
41. Maddux JE. Self-efficacy theory: an introduction. In: Maddux JE, editor. *Self-Efficacy, Adaptation, and Adjustment: Theory, Research, and Application*. Boston, MA, USA: Springer; 1995:3-33.
42. Kruglanski AW, Chernikova M, Rosenzweig E, Kopetz C. On motivational readiness. *Psychol Rev* 2014 Jul;121(3):367-388. [doi: [10.1037/a0037013](https://doi.org/10.1037/a0037013)] [Medline: [25090424](https://pubmed.ncbi.nlm.nih.gov/25090424/)]
43. Kruglanski AW. Motivational phases on the road to action. *Motiv Sci* 2017 Sep;3(3):196-207. [doi: [10.1037/mot0000072](https://doi.org/10.1037/mot0000072)]
44. Groat S, Tront J, Marchany R. Advancing the defense in depth model. In: *Proceedings of the 7th International Conference on System of Systems Engineering*. 2012 Presented at: SoSE '12; July 16-19,2012; Genova, Italy p. 285-290. [doi: [10.1109/sysose.2012.6384127](https://doi.org/10.1109/sysose.2012.6384127)]
45. 2021 Year in Review - Pornhub Insights. Pornhub. 2021 Dec 14. URL: <https://www.pornhub.com/insights/yir-2021> [accessed 2022-03-05]
46. Bóthe B, Tóth-Király I, Zsila Á, Griffiths MD, Demetrovics Z, Orosz G. The development of the problematic pornography consumption scale (PPCS). *J Sex Res* 2018;55(3):395-406. [doi: [10.1080/00224499.2017.1291798](https://doi.org/10.1080/00224499.2017.1291798)] [Medline: [28276929](https://pubmed.ncbi.nlm.nih.gov/28276929/)]
47. Wiecezorek L, Klingemann JI. Mobile applications used to limit alcohol consumption – a literature review. *Alcohol Drug Addict* 2020;33(1):43-64. [doi: [10.5114/ain.2020.95979](https://doi.org/10.5114/ain.2020.95979)]
48. Short MB, Wetterneck CT, Bistricky SL, Shutter T, Chase TE. Clinicians' beliefs, observations, and treatment effectiveness regarding clients' sexual addiction and internet pornography use. *Community Ment Health J* 2016 Nov;52(8):1070-1081. [doi: [10.1007/s10597-016-0034-2](https://doi.org/10.1007/s10597-016-0034-2)] [Medline: [27345497](https://pubmed.ncbi.nlm.nih.gov/27345497/)]
49. Ayres MM, Haddock SA. Therapists' approaches in working with heterosexual couples struggling with male partners' online sexual behavior. *Sex Addict Compulsivity* 2009 Mar 04;16(1):55-78. [doi: [10.1080/10720160802711952](https://doi.org/10.1080/10720160802711952)]
50. Plucky - a self-control filter. Plucky. URL: <https://www.pluckyfilter.com> [accessed 2022-05-21]
51. X3watch - Internet Porn Accountability Software. X3watch. URL: <https://x3watch.com/> [accessed 2022-05-21]
52. Parental control and digital wellbeing software. Qustodio. URL: <https://www.qustodio.com/en/> [accessed 2022-05-21]
53. NoFap® Porn Addiction Recovery. NoFap. URL: <https://nofap.com/> [accessed 2022-05-21]
54. Sheikh AA, Ganai PT, Malik NA, Dar KA. Smartphone: Android Vs IOS. *SIJ Trans Comput Sci Eng Its Appl (CSEA)* 2013 Oct 09;1(4):31-38. [doi: [10.9756/sijcsea/v1i4/0104600401](https://doi.org/10.9756/sijcsea/v1i4/0104600401)]

## Abbreviations

- CBT:** cognitive behavioral therapy
- DiD:** defense in depth
- DSCT:** digital self-control tool
- JITAI:** just-in-time adaptive intervention
- mHealth:** mobile health
- PPU:** problematic pornography use

*Edited by A Mavragani; submitted 26.05.22; peer-reviewed by B Bóthe, C Fakh El Khoury; comments to author 22.08.22; revised version received 01.09.22; accepted 14.09.22; published 13.10.22.*

*Please cite as:*

*Henry N, Donkin L, Williams M, Pedersen M*

*mHealth Technologies for Managing Problematic Pornography Use: Content Analysis*

*JMIR Form Res 2022;6(10):e39869*

*URL: <https://formative.jmir.org/2022/10/e39869>*

*doi: [10.2196/39869](https://doi.org/10.2196/39869)*

*PMID: [36227634](https://pubmed.ncbi.nlm.nih.gov/36227634/)*

©Nathan Henry, Liesje Donkin, Matt Williams, Mangor Pedersen. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 13.10.2022. 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

# An Evaluation of a Mobile App for Chronic Low Back Pain Management: Prospective Pilot Study

Jonathan D Browne<sup>1</sup>, BS; Michael Vaninetti<sup>2</sup>, MD; David Giard<sup>3</sup>, MS, MA; Konstantinos Kostas<sup>4</sup>, PhD; Ankur Dave<sup>4</sup>, MD

<sup>1</sup>School of Medicine, California University of Science and Medicine, Colton, CA, United States

<sup>2</sup>Center for Pain Medicine, University of California San Diego, La Jolla, CA, United States

<sup>3</sup>College of Psychology, California Northstate University, Rancho Cordova, CA, United States

<sup>4</sup>Ascension Illinois, Alexian Brothers Medical Center, Elk Grove Village, IL, United States

**Corresponding Author:**

Ankur Dave, MD

Ascension Illinois

Alexian Brothers Medical Center

Eberle Building, Suite 610

800 Biesterfield Rd

Elk Grove Village, IL, 60007

United States

Phone: 1 630 379 9815

Email: [ankur.dave@ascension.org](mailto:ankur.dave@ascension.org)

## Abstract

**Background:** Chronic low back pain is challenging to manage due to multidisciplinary considerations. It has substantial socioeconomic impacts and cannot be simply treated with pharmacotherapy, nonsurgical intervention, or spine surgery. Medical consensus recommends optimizing conservative self-management therapies (eg, home exercise, wellness strategies, yoga, etc) as first-line treatment options for chronic low back pain. However, access to these modalities is often limited and secondary to cost, convenience, and ease of use. Mobile health apps have emerged as a cost-effective and accessible option for chronic low back pain self-management. Established in-person pain programs can provide the structure for an optimal mobile app adaptation. PainNavigator (PainNavigator, Inc) is an example of a mobile app that is based on an Ascension-Illinois group-based pain program—Pain Rehabilitation Outpatient-Camp.

**Objective:** This was a prospective pilot clinical trial that evaluated the PainNavigator platform's utility in low back pain management to inform future trial development.

**Methods:** A total of 75 participants who used PainNavigator were studied. Pain, Enjoyment, and General Activity (PEG-3) scale scores and scores from a brief anxiety and depression scale based on the Patient Health Questionnaire-4 (PHQ-4) were obtained at baseline and following program completion. The PEG-3 total score was used, in addition to individual items—*Average Pain*, *Pain Effect on Enjoyment*, and *Pain Effect on Activity*. The PHQ-4 total score was also used, in addition to other individual items, including *Felt Depressed*, *Loss of Interest*, *Felt Anxious*, and *Difficult to Control Worry*. Paired sample *t* tests (2-tailed) compared mean differences in scores from before and after participants received the intervention.

**Results:** The analysis found that PEG-3 (n=27) and PHQ-4 (n=27) total scores were significantly lower upon the completion of PainNavigator ( $P<.001$  and  $P=.001$ , respectively). The findings showed a 36% reduction in PEG-3 total scores, a 40% reduction in pain intensity, and a 40% reduction in PHQ-4 total scores. Scores for individual PEG-3 scale and PHQ-4 items also significantly decreased. All PEG-3 measures had large effect sizes. The PHQ-4 total score and *Difficult to Control Worry* item had large effect sizes, while the other three measures had medium effect sizes.

**Conclusions:** These findings show that PainNavigator has clinical significance in managing chronic low back pain and can be easily utilized to improve patient care. All PEG-3 scale and PHQ-4 measures significantly improved following the use of the platform, supporting the multidimensional, biopsychosocial approach to low back pain management. Differences in effect sizes may inform quality improvement investigations, such as optimizing features that impact measures with only medium effect sizes. This feasibility study demonstrates an effective protocol, and it will inform future, more extensive randomized controlled trials.

(JMIR Form Res 2022;6(10):e40869) doi:[10.2196/40869](https://doi.org/10.2196/40869)

**KEYWORDS**

back pain; chronic pain; mobile; app; multidisciplinary care; biopsychosocial; self-management; mHealth; mobile health; mobile app

**Introduction**

Chronic low back pain (LBP) is a multidisciplinary condition with significant socioeconomic implications. The Centers for Disease Control and Prevention has reported that approximately 20% of US adults experience chronic pain [1]. The economic impact of pain, when factoring in health care costs and productivity, has been estimated to be greater than that of heart disease, cancer, and diabetes [2]. LBP is a subset pain condition that also is the leading cause of years lost to disability [3]. LBP prevalence may also increase due to an aging population and can disproportionately burden low-middle-income people [4].

In addition to its profound impact on health care, LBP poses further challenges in management. LBP therapy begins with self-management, focusing on physical improvement and lifestyle modification; this can be scaled up to nonsurgical (including pharmacotherapy) and surgical interventions, as necessary [5]. In the past, medical management has emphasized these nonsurgical and surgical interventions. Although beneficial for the right individuals, this has potentially contributed to overusing imaging, opioids, and surgery, often complicating efficient patient care [6]. Additionally, external factors, such as insurance limitations, care provider reimbursements, and shifts in opioid prescribing practices, may affect a patient's treatment plan. More recently, there has been a focus on utilizing interdisciplinary therapies to treat LBP. This emphasizes a more holistic approach that focuses on the physical, psychological, and social impacts on an individual's life. This is known as the biopsychosocial management of chronic pain. Evidence of the benefits of this approach has been demonstrated across disciplines, with improvements seen in outcomes and significant cost savings [7,8]. However, most physicians lack sufficient training in biopsychosocial-based chronic pain treatment plans. Additionally, health care professionals' attitudes and beliefs about LBP have been linked to patients' attitudes and treatment adherence, showing that effective treatment encompasses all involved [9]. Patients are often unable to manage concurrent visits and the costs of physical therapy, health psychology, health care providers, and therapists while also maintaining daily personal and professional responsibilities. Given the multidisciplinary nature of LBP and the socioeconomic implications of effective management, recent attention has been drawn to the importance of comparative effectiveness and randomized controlled trials for emerging LBP management strategies [5].

Mobile health (mHealth) apps may address this gap in LBP management. In addition to research demonstrating app efficacy, care providers and patients have gradually adopted such technology, which is likely a result of the accessibility of mHealth apps. Back pain apps have shown utility in delivering therapeutic pain interventions while bridging the gap between patients and care providers [10]. Apps have been developed to assess chronic LBP patients' thoracolumbar range of motion [11] and postural re-education [12]. Pain-centered mobile apps

can also address postoperative pain assessment [13] and cancer pain management [14], demonstrating patient acceptance in ambulatory care. Meta-analyses that investigated the role of self-management and eHealth in LBP found sufficient evidence supporting their roles in pain and disability [15-17]. Du et al [16] further described that mobile platforms have advantages over web platforms with regard to their effect on pain and disability. Although many apps have been developed, differences among user interfaces may impact practicality, which may be addressed by focusing on user-centered designs [18]. LBP management has been evolving, and mobile apps appear to play an essential role in achieving lasting results, which is the goal of effective patient care.

An Ascension-Illinois group-based pain program—Pain Rehabilitation Outpatient-Camp—reported successful longitudinal results, including a 52% reduction in pain within 3 months, a 47% reduction in the risk of opioid misuse, a 40% reduction in pain disability, and a 60% reduction in depression. However, this in-person, 50-hour program (conducted over 6 weeks) was limited by accessibility and was correspondingly adapted into a mobile app interface—PainNavigator (PainNavigator, Inc). This self-guided program provides educational and movement-based modules to improve pain and function. PainNavigator's interdisciplinary approach and digital accessibility may exponentially increase cost savings for medical, insurance, and wellness groups. However, the platform's feasibility has not yet been assessed.

This research was a prospective pilot clinical trial that studied pain management questionnaire data from patients with chronic LBP before and after they used the PainNavigator software app. This study aimed to evaluate PainNavigator's utility by assessing pain scores and functional outcomes. We hypothesized that participants would experience positive outcomes, as depicted by improved subjective survey scores, after using the PainNavigator mobile app.

**Methods****Study Design**

This was a prospective pilot clinical trial that investigated the utility of the PainNavigator mobile app. This study assessed pain scores and functional outcome data from patients with chronic LBP before and after they used the platform. As this study was exploratory, all participants received the intervention, and there was no control group.

**Ethics Approval**

Study participants were provided with a digital informed consent form through the mobile app, and they provided a signature on acceptance of the terms of consent. Participant data were deidentified for privacy and confidentiality protection. Compensation in the form of cash gift cards was provided for participation. This research was approved by the Ascension-Illinois Institutional Review Board (RIL20210036)



and was carried out in accordance with the Helsinki Declaration's ethical standards.

## Participants

A total of 75 participants were recruited via health care providers within the Ascension-Illinois system. Participants either were given PainNavigator enrollment information, as deemed appropriate by their primary care provider during a clinic visit, or obtained the Ascension-Illinois flyer through Ascension-Illinois Marketing outlets. Participants were incrementally compensated up to US \$80 in the form of cash gift cards based on the completion of modules. The inclusion criteria were men and women aged 18 years or older and those experiencing LBP for greater than 4 weeks that was nonsurgical and nonmalignant in etiology. Given the limited access among many people to multidisciplinary pain management clinics, this

study aimed to target patients before they considered pain management clinics or those without access to such resources. The exclusion criteria were as follows: an age of under 18 years; the inability to complete the pain questionnaire in written English; the inability to utilize the software app; recent back surgery in the past 6 months; back pain due to malignancy; a diagnosis requiring surgery; uncontrolled depression, anxiety, or a severe mental disorder; severe medical conditions (including heart disease, lung disease, a history of strokes, and neurological disorders such as paralysis or uncontrolled seizures); patients who are advised against physical exercise or mental health self-therapy by a health care provider; patients undergoing interventional pain management techniques; patients undergoing any form of outpatient psychotherapy; pregnant patients; and adults who are unable to consent. Participant age and baseline pain duration demographics are shown in [Table 1](#).

**Table 1.** Participant age and baseline pain duration demographics.

Variables	Participants who signed up for the program (n=75)		Participants who completed the program (n=30)	
	n	%	n	%
<b>Age range (years)</b>				
18-29	7	9	0	0
30-39	11	15	4	13
40-49	13	17	7	23
50-59	16	21	7	23
≥60	28	37	12	40
<b>Pain duration</b>				
1 to 3 months	6	8	1	3
3 to 6 months	3	4	1	3
6 months to 1 year	11	15	4	13
1 to 3 years	12	16	5	17
>3 years	43	57	19	63

## Protocol

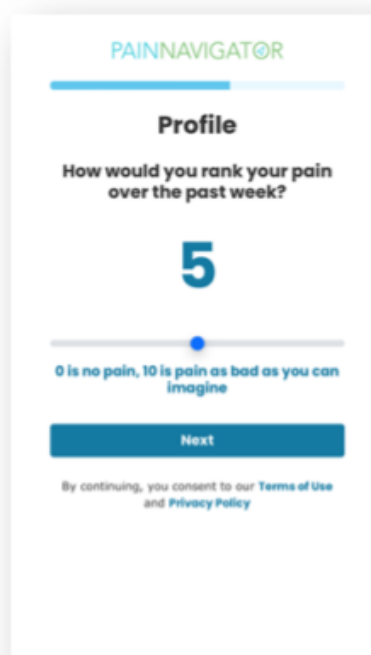
PainNavigator is a web-based mobile app that users can access for pain self-management ([Figure 1](#)). At baseline, the app gave participants pain and function questionnaires ([Multimedia Appendix 1](#)). These included demographics (ie, age, gender, and activity level); symptomatology (ie, the duration of LBP and the effect of LBP on behavior); treatments (ie, the names of opioids, the doses being taken [if applicable], and the likelihood of pursuing opioid or surgical intervention); the Pain, Enjoyment, and General Activity (PEG-3) scale; and the 4-item anxiety and depression questionnaire that was based on the Patient Health Questionnaire-4 (PHQ-4). This foundation promoted the development of a personal functional goal for users to achieve during the program. The app then guided users through prerecorded medical education and wellness strategy content, including evidence-based cognitive behavioral therapy techniques, yoga, mindfulness, and exercise therapy. The

prerecorded content was led by a pain management physician, a health psychologist, a certified yoga instructor, and a doctor of physical therapy. The educational videos taught common causes of and solutions for pain and wellness strategies that users could learn for managing pain better. The movement modules taught exercises and stretches for back pain and mindfulness to emphasize a mind-body connection. Users set goals for activities that pain was holding them back from and, via weekly phone calls, worked one-on-one with a live wellness coach who supported goal achievement and program consistency. To solidify their understanding of the content, users completed postvideo actions, including the use of pain, mood, and food journals. Throughout the program, users leveraged these journals to identify triggers and drive behavior change. Upon finishing the program, users answered the completion survey, which included a pain and function questionnaire ([Multimedia Appendix 2](#)).

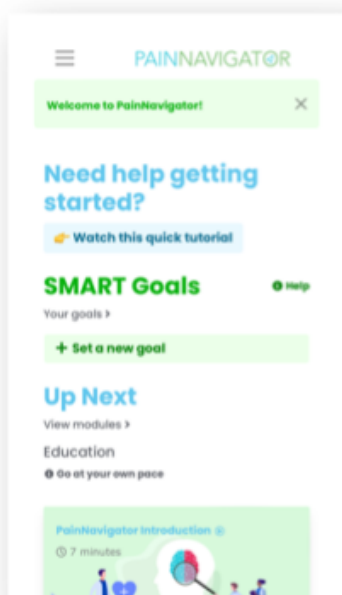
**Figure 1.** Screenshots of the PainNavigator app. SMART: Specific, Measurable, Attainable, Realistic, and Time-limited.



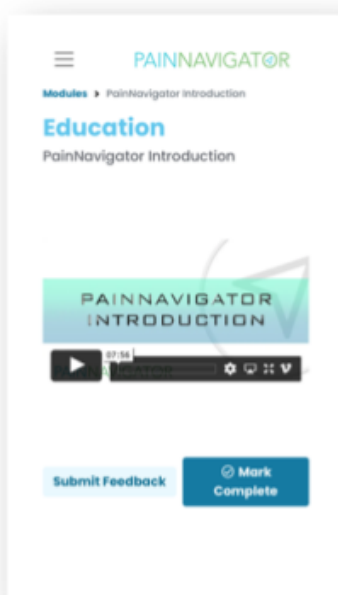
**Sign Up**



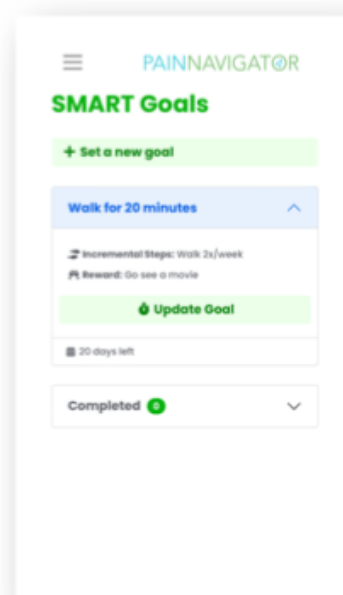
**User Profile Completion**



**Home Dashboard**



**Introduction Video**



**Setting a SMART Goal**

**Statistical Analysis**

The data analysis included a comparison of deidentified, participant-inputted responses to questionnaires. Responses that were given before, during, and upon the completion of PainNavigator were analyzed for statistically significant changes in pain scores, as defined by standard statistical methods. Analyses were completed only for participants who answered each question, and composite scores, including the PHQ-4 and

PEG-3 total scores, were only calculated for participants who answered each item that contributed to the scores. Data were analyzed in SPSS version 26 (IBM Corporation). The significance level was set to an  $\alpha$  of .05. Data were first assessed for skewness. Paired sample *t* tests were used to compare mean differences in scores from before and after participants received the intervention. The Cohen *d* was calculated from the raw means that yielded effect sizes. Power was also calculated with G\*Power version 3.1.9.7 (Heinrich-Heine-Universität

Düsseldorf) [19] post hoc, using the sample size,  $\alpha$ , and the effect size.

## Results

A total of 30 participants completed the program; however, 3 were excluded from analysis due to incomplete baseline PHQ-4 answers. Paired sample *t* tests were conducted for participants who completed all survey items before and after the intervention ( $n=27$ ) to determine the effect of PainNavigator on pain, as measured by the PEG-3 scale, and its effect on anxiety and depression symptoms, as measured by the PHQ-4. The results indicated a significant difference in PEG-3 total scores from before (mean 16.703, SD 6.550) and after (mean 10.629, SD 5.204) participants completed PainNavigator ( $t_{26}=7.639$ ;

$P<.001$ ). The 95% CI of the difference in means ranged from 4.439 to 7.708. Effect sizes ( $d$ ) were also calculated by using the Cohen  $d$ . Cohen [20] suggested that a  $d$  of 0.2 is considered a small effect size, 0.5 represents a medium effect size, and 0.8 represents a large effect size. For the PEG-3 total score, a large effect size ( $d=1.027$ ) was observed. Each of the three individual items—*Average Pain*, *Pain Effect on Enjoyment*, and *Pain Effect on Activity*—comprising the PEG-3 were also analyzed, suggesting significant differences and large effect sizes, as detailed in Table 2. The magnitude of change for each PEG-3 scale measure, which was standardized based on the baseline SD, is shown in Figure 2. Overall, on average, participants experienced pain decreases between baseline and the completion of PainNavigator.

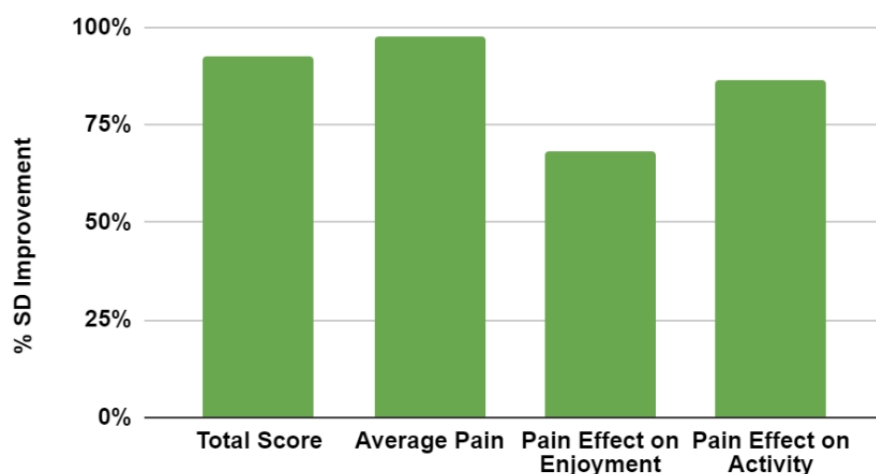
**Table 2.** Pre- and postintervention measures for Pain, Enjoyment, and General Activity (PEG-3;  $n=27$ ) and Patient Health Questionnaire-4 (PHQ-4;  $n=27$ ) total scores and subscores. Paired *t* tests evaluated mean differences.

Variables	Baseline, score		End, score		Analysis	
	Mean	SD	Mean	SD	<i>P</i> value	Effect size, <i>d</i>
<b>PEG-3 items</b>						
Total score	16.703	6.550	10.629	5.204	<.001 <sup>a</sup>	1.027
Average Pain	5.482	2.230	3.310	1.605	<.001 <sup>a</sup>	1.118
Pain Effect on Enjoyment	5.629	2.662	3.814	2.253	<.001 <sup>a</sup>	0.736
Pain Effect on Activity	5.583	2.357	3.541	1.793	<.001 <sup>a</sup>	0.975
<b>PHQ-4 items</b>						
Total score	5.814	3.101	3.481	3.309	.001 <sup>a</sup>	0.728
Felt Depressed	0.963	0.979	0.555	0.751	.031 <sup>b</sup>	0.468
Loss of Interest	1.333	1.000	0.814	1.001	.013 <sup>b</sup>	0.519
Felt Anxious	1.814	0.962	1.148	1.099	.003 <sup>a</sup>	0.645
Difficult to Control Worry	1.703	0.912	0.963	1.055	.002 <sup>a</sup>	0.750

<sup>a</sup>Significant at  $P<.01$ .

<sup>b</sup>Significant at  $P<.05$ .

**Figure 2.** The percent SD changes from baseline to after the completion of PainNavigator for the Pain, Enjoyment, and General Activity scale ( $n=27$ ). Effect size is denoted by color; medium effect sizes are yellow, and large effect sizes are green.



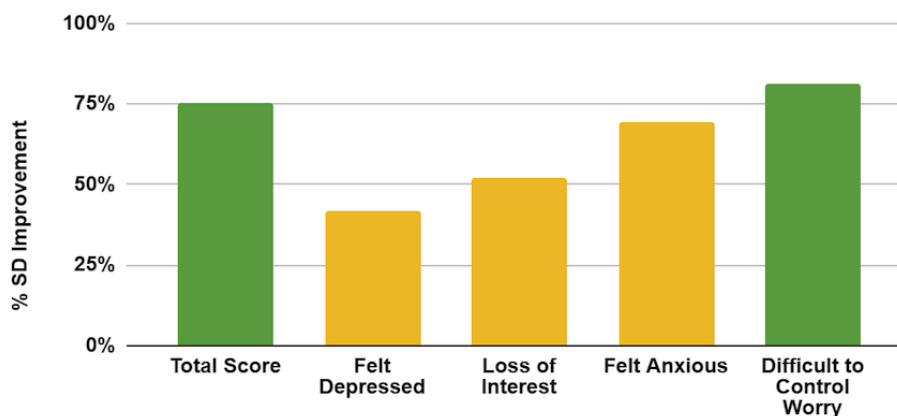
The results also indicated a significant difference in PHQ-4 total scores from before (mean 5.814, SD 3.101) and after (mean 3.481, SD 3.309) participants completed PainNavigator ( $t_{26}=3.805$ ;  $P=.001$ ). The 95% CI of the difference in means ranged from 1.072 to 3.593. An effect size approaching large ( $d=0.728$ ) was observed. Each of the four individual items from the PHQ-4—*Felt Depressed*, *Loss of Interest*, *Felt Anxious*, and *Difficult to Control Worry*—were also analyzed, suggesting significant mean differences with medium to large effect sizes between pre- and postintervention scores, as reported in Table 2. The magnitude of change for each PHQ-4 measure, which was standardized based on the baseline SD, is shown in Figure 3. Overall, participants observed decreases in anxiety and depression between baseline and the completion of PainNavigator. All metrics demonstrated significant decreases.

Moderately skewed results from the PHQ-4 at the end of the program (0.708) and the PEG-3 scale at the start of the program (0.627) were addressed by performing nonparametric tests to

decrease the chances of making a type 1 error. Findings from related-samples Wilcoxon signed rank tests indicated that the null hypothesis would still be rejected for the PEG-3 scale ( $P<.001$ ) and PHQ-4 ( $P=.002$ ). Other data were also skewed and were analyzed the same way, and all data indicated that the decision to reject the null hypothesis appeared correct.

A post hoc power analysis was completed to determine achieved power. This was done with G\*Power version 3.1.9.7 [19] for matched pair mean comparisons, using the sample size ( $n=27$ ),  $\alpha$  (.05), and the effect sizes of the PEG-3 total score ( $d=1.027$ ) and PHQ-4 total score ( $d=0.728$ ). Additionally, the highest ( $d=1.118$ ) and lowest ( $d=0.468$ ) calculated effect sizes were used to provide a range. For the PEG-3 total score  $t$  test, the power was calculated to be 0.99. For the PHQ-4 total score, the power was calculated to be 0.95. In the case of the largest effect size, the power was estimated to be 0.99. In the case of the smallest effect size, the power was estimated to be 0.64.

**Figure 3.** The percent SD changes from baseline to after the completion of PainNavigator for the Patient Health Questionnaire-4 ( $n=27$ ). Effect size is denoted by color; medium effect sizes are yellow, and large effect sizes are green.



## Discussion

### Principal Findings

The preliminary results from this pilot trial support PainNavigator's initial acceptability and can help inform a larger randomized controlled clinical trial. Using the app significantly impacted all PEG-3 scale and PHQ-4 measures, indicating that the platform addressed essential components for the self-management of LBP. The findings show a 36% reduction in the PEG-3 total score, a 40% reduction in pain intensity, and a 40% reduction in the PHQ-4 total score. The magnitude of change for each measure illustrated these effects, with a positive percent change indicating score improvement (Figure 2 and Figure 3). The sample size was limited but acceptable for an initial pilot study. Given this sample size, the large effect sizes play an important role in study power. The strong power in this study indicates that the observed significant changes are likely clinically relevant and warrant a more rigorous study.

### Comparison With Prior Work

Published literature on pain-focused mHealth apps shows efficacy data that can be compared with data from PainNavigator. A meta-analysis of pain eHealth applications

found that program durations of  $\leq 8$  weeks (“less intensive duration”) had an advantageous effect on reducing pain intensity [16]. The 8-week program described in this study similarly found significant improvements in pain, as assessed by the PEG-3 scale item *Average Pain*. These findings may suggest that concise programming facilitates user engagement and retention. Furthermore, the meta-analysis effect sizes of immediate and short-term pain intensity were found to be  $-0.16$  (95% CI  $-0.30$  to  $-0.02$ ) and  $-0.27$  (95% CI  $-0.43$  to  $-0.11$ ), respectively [16]. The absolute effect size for average pain following the PainNavigator intervention was 1.118, which is greatly higher in comparison. This may indicate that intrinsic app factors, such as design and functionality, assist with pain reduction.

Beyond pain measures, psychology may also play an important role in LBP management. An investigation of musculoskeletal pain rehabilitation outcomes found psychological and pain indices to be significantly intercorrelated [21]. This supports the biopsychosocial approach, as it appears to incorporate multidimensional patient care. PainNavigator demonstrated significant improvements in the psychological measures assessed by the PHQ-4. These findings may indicate the importance of psychological components in pain management programs.

As an emerging field, mHealth apps also have challenges. These include a general lack of quality mHealth app assessments [22] and the number of chronic pain apps that only offer 1 self-management strategy [23]. Many health apps are also mostly used until initial milestones, with significantly decreased usage afterward [24]. PainNavigator's effect may be the result of the mobile app addressing some of the aforementioned challenges. PainNavigator's interdisciplinary approach introduces users to multiple strategies (eg, home exercise, yoga, wellness, nutrition education, etc) instead of focusing on only 1 modality. As such, users can utilize the modalities that provide the most benefits. Additionally, PainNavigator emphasizes function as an equal or greater part of the self-management of chronic pain when compared to pain scores. The wellness coach and personal goals hold users accountable for completing the program and provide a framework for continuing usage once initial milestones are met.

### Limitations

Due to the single-arm design and the lack of a control group, our results should be interpreted cautiously. Nevertheless, our results were quite promising for informing future rigorous tests of the PainNavigator intervention. Another limitation may be the usability of mHealth technology in older populations. Despite this, over half of the participants who completed this study were aged over 50 years (19/30, 63%). Other back pain mobile app studies additionally reported disability data, which this study did not explicitly assess. There were also differences between the number of enrolled participants and the number of those who completed all surveys, indicating a drop-off in responses. Lastly, financial compensations were given to

participants for their participation in this study. Although this is a well-accepted practice in research studies, there is some debate about how this impacts compliance.

### Future Research

This study's findings will guide future research into pain-centered mHealth app efficacy and scalability. For example, the analysis found that the *Felt Anxious*, *Felt Depressed*, and *Loss of Interest* PHQ-4 submeasures had medium effect sizes, whereas large effect sizes were found in all other measures; a future focus may be directed on these targets. Since user engagement with mHealth apps is essential to their efficacy, many studies have prioritized users' needs during quality improvement and described their methodology [25,26]. Clinician feedback also may help direct LBP mobile app quality improvement [27]. Accounting for both types of input is essential, given the need for dynamic communication between patients and physicians. This feasibility study demonstrates that our protocol works, and more extensive randomized controlled trials will be conducted. mHealth app development is a dynamic process centered around providing patients with the best care.

### Conclusion

The PainNavigator mHealth app showed LBP management utility in this initial pilot trial. The significant improvements in all PEG-3 scale and PHQ-4 measures illustrate potential multidimensional, biopsychosocial management that is easily accessible to patients. This platform demonstrates clinical significance and can be easily utilized to improve patient care. Further randomized controlled trials are needed to expand upon these initial findings and explore the functional role of the PainNavigator platform in clinical settings.

---

### Acknowledgments

We appreciate the participants' willingness to give their time generously to support this work. PainNavigator, Inc, supported this study.

---

### Authors' Contributions

AD and KK designed the study and supervised participant recruitment. DG analyzed the data. JDB and MV drafted the manuscript. All authors made revisions to the manuscript. JDB had final responsibility for the submission of the manuscript.

---

### Conflicts of Interest

AD and KK have an equity interest in PainNavigator. The authors report no other conflicts of interest.

---

#### Multimedia Appendix 1

Initial user questionnaire.

[PNG File, 665 KB - [formative\\_v6i10e40869\\_app1.png](#)]

---

#### Multimedia Appendix 2

Completion questionnaire.

[PNG File, 604 KB - [formative\\_v6i10e40869\\_app2.png](#)]

---

### References

1. 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/)]

2. Gaskin DJ, Richard P. Appendix C: The economic costs of pain in the United States. In: *Relieving Pain in America: A Blueprint for Transforming Prevention, Care, Education, and Research*. Washington (DC): National Academies Press; 2011.
3. GBD 2015 Disease and Injury Incidence and Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet* 2016 Oct 08;388(10053):1545-1602 [[FREE Full text](#)] [doi: [10.1016/S0140-6736\(16\)31678-6](https://doi.org/10.1016/S0140-6736(16)31678-6)] [Medline: [27733282](https://pubmed.ncbi.nlm.nih.gov/27733282/)]
4. Buchbinder R, van Tulder M, Öberg B, Costa LM, Woolf A, Schoene M, Lancet Low Back Pain Series Working Group. Low back pain: a call for action. *Lancet* 2018 Jun 09;391(10137):2384-2388. [doi: [10.1016/S0140-6736\(18\)30488-4](https://doi.org/10.1016/S0140-6736(18)30488-4)] [Medline: [29573871](https://pubmed.ncbi.nlm.nih.gov/29573871/)]
5. Knezevic NN, Candido KD, Vlaeyen JWS, Van Zundert J, Cohen SP. Low back pain. *Lancet* 2021 Jul 03;398(10294):78-92. [doi: [10.1016/S0140-6736\(21\)00733-9](https://doi.org/10.1016/S0140-6736(21)00733-9)] [Medline: [34115979](https://pubmed.ncbi.nlm.nih.gov/34115979/)]
6. Maher C, Underwood M, Buchbinder R. Non-specific low back pain. *Lancet* 2017 Feb 18;389(10070):736-747. [doi: [10.1016/S0140-6736\(16\)30970-9](https://doi.org/10.1016/S0140-6736(16)30970-9)] [Medline: [27745712](https://pubmed.ncbi.nlm.nih.gov/27745712/)]
7. Friedman R, Sobel D, Myers P, Caudill M, Benson H. Behavioral medicine, clinical health psychology, and cost offset. *Health Psychol* 1995 Nov;14(6):509-518. [doi: [10.1037//0278-6133.14.6.509](https://doi.org/10.1037//0278-6133.14.6.509)] [Medline: [8565925](https://pubmed.ncbi.nlm.nih.gov/8565925/)]
8. Blount A, Schoenbaum M, Kathol R, Rollman BL, Thomas M, O'Donohue W, et al. The economics of behavioral health services in medical settings: A summary of the evidence. *Prof Psychol Res Pr* 2007;38(3):290-297. [doi: [10.1037/0735-7028.38.3.290](https://doi.org/10.1037/0735-7028.38.3.290)]
9. Darlow B, Fullen BM, Dean S, Hurley DA, Baxter GD, Dowell A. The association between health care professional attitudes and beliefs and the attitudes and beliefs, clinical management, and outcomes of patients with low back pain: a systematic review. *Eur J Pain* 2012 Jan;16(1):3-17 [[FREE Full text](#)] [doi: [10.1016/j.ejpain.2011.06.006](https://doi.org/10.1016/j.ejpain.2011.06.006)] [Medline: [21719329](https://pubmed.ncbi.nlm.nih.gov/21719329/)]
10. Wilson MA, Fouts BL, Brown KN. Development of a mobile application for acute pain management in U.S. military healthcare. *Appl Nurs Res* 2021 Apr;58:151393. [doi: [10.1016/j.apnr.2020.151393](https://doi.org/10.1016/j.apnr.2020.151393)] [Medline: [33745549](https://pubmed.ncbi.nlm.nih.gov/33745549/)]
11. de Brito Macedo L, Borges DT, Melo SA, da Costa KSA, de Oliveira Sousa C, Brasileiro JS. Reliability and concurrent validity of a mobile application to measure thoracolumbar range of motion in low back pain patients. *J Back Musculoskeletal Rehabil* 2020;33(1):145-151. [doi: [10.3233/BMR-181396](https://doi.org/10.3233/BMR-181396)] [Medline: [31127756](https://pubmed.ncbi.nlm.nih.gov/31127756/)]
12. Abadiyan F, Hadadnezhad M, Khosrokiani Z, Letafatkar A, Akhshik H. Adding a smartphone app to global postural re-education to improve neck pain, posture, quality of life, and endurance in people with nonspecific neck pain: a randomized controlled trial. *Trials* 2021 Apr 12;22(1):274 [[FREE Full text](#)] [doi: [10.1186/s13063-021-05214-8](https://doi.org/10.1186/s13063-021-05214-8)] [Medline: [33845880](https://pubmed.ncbi.nlm.nih.gov/33845880/)]
13. Carlier J, Robin F, Pages N, Quinart A, Roy M, Pauchard JC, et al. Pain evaluation after day-surgery using a mobile phone application. *Anaesth Crit Care Pain Med* 2021 Jun;40(3):100879. [doi: [10.1016/j.accpm.2021.100879](https://doi.org/10.1016/j.accpm.2021.100879)] [Medline: [33965646](https://pubmed.ncbi.nlm.nih.gov/33965646/)]
14. Yang J, Weng L, Chen Z, Cai H, Lin X, Hu Z, et al. Development and testing of a mobile app for pain management among cancer patients discharged from hospital treatment: Randomized controlled trial. *JMIR Mhealth Uhealth* 2019 May 29;7(5):e12542 [[FREE Full text](#)] [doi: [10.2196/12542](https://doi.org/10.2196/12542)] [Medline: [31144672](https://pubmed.ncbi.nlm.nih.gov/31144672/)]
15. Du S, Hu L, Dong J, Xu G, Chen X, Jin S, et al. Self-management program for chronic low back pain: A systematic review and meta-analysis. *Patient Educ Couns* 2017 Jan;100(1):37-49. [doi: [10.1016/j.pec.2016.07.029](https://doi.org/10.1016/j.pec.2016.07.029)] [Medline: [27554077](https://pubmed.ncbi.nlm.nih.gov/27554077/)]
16. Du S, Liu W, Cai S, Hu Y, Dong J. The efficacy of e-health in the self-management of chronic low back pain: A meta analysis. *Int J Nurs Stud* 2020 Jun;106:103507. [doi: [10.1016/j.ijnurstu.2019.103507](https://doi.org/10.1016/j.ijnurstu.2019.103507)] [Medline: [32320936](https://pubmed.ncbi.nlm.nih.gov/32320936/)]
17. Pfeifer AC, Uddin R, Schröder-Pfeifer P, Holl F, Swoboda W, Schiltenswolf M. Mobile application-based interventions for chronic pain patients: A systematic review and meta-analysis of effectiveness. *J Clin Med* 2020 Nov 05;9(11):3557 [[FREE Full text](#)] [doi: [10.3390/jcm9113557](https://doi.org/10.3390/jcm9113557)] [Medline: [33167300](https://pubmed.ncbi.nlm.nih.gov/33167300/)]
18. Smaradottir BF, Fagerlund AJ, Bellika JG. User-centred design of a mobile application for chronic pain management. *Stud Health Technol Inform* 2020 Jun 26;272:272-275. [doi: [10.3233/SHTI200547](https://doi.org/10.3233/SHTI200547)] [Medline: [32604654](https://pubmed.ncbi.nlm.nih.gov/32604654/)]
19. Faul F, Erdfelder E, Lang AG, Buchner A. G\*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res Methods* 2007 May;39(2):175-191. [doi: [10.3758/bf03193146](https://doi.org/10.3758/bf03193146)] [Medline: [17695343](https://pubmed.ncbi.nlm.nih.gov/17695343/)]
20. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*, Second Edition. New York, NY: Routledge; 1988.
21. Woznowski-Vu A, Aternali A, Gervais A, Pavilanis ADS, Nijs J, Sullivan MJL, et al. The prospective prognostic value of biopsychosocial indices of sensitivity to physical activity among people with back pain. *Clin J Pain* 2021 Oct 01;37(10):719-729. [doi: [10.1097/AJP.0000000000000965](https://doi.org/10.1097/AJP.0000000000000965)] [Medline: [34419972](https://pubmed.ncbi.nlm.nih.gov/34419972/)]
22. McKay FH, Cheng C, Wright A, Shill J, Stephens H, Uccellini M. Evaluating mobile phone applications for health behaviour change: A systematic review. *J Telemed Telecare* 2018 Jan;24(1):22-30. [doi: [10.1177/1357633X16673538](https://doi.org/10.1177/1357633X16673538)] [Medline: [27760883](https://pubmed.ncbi.nlm.nih.gov/27760883/)]
23. Piguët V, Mazouri S, Rochat J, Ehrler F, Cedraschi C, Luthy C. [Mobile applications and self-management of chronic pain]. *Rev Med Suisse* 2020 Jul 15;16(700):1367-1371. [Medline: [32672016](https://pubmed.ncbi.nlm.nih.gov/32672016/)]
24. Anderson K, Burford O, Emmerton L. Mobile health apps to facilitate self-care: A qualitative study of user experiences. *PLoS One* 2016 May 23;11(5):e0156164 [[FREE Full text](#)] [doi: [10.1371/journal.pone.0156164](https://doi.org/10.1371/journal.pone.0156164)] [Medline: [27214203](https://pubmed.ncbi.nlm.nih.gov/27214203/)]

25. Kavanagh ME, Chiavaroli L, Glenn AJ, Heijmans G, Grant SM, Chow CM, et al. A web-based health application to translate nutrition therapy for cardiovascular risk reduction in primary care (PortfolioDiet.app): Quality improvement and usability testing study. *JMIR Hum Factors* 2022 Apr 21;9(2):e34704 [FREE Full text] [doi: [10.2196/34704](https://doi.org/10.2196/34704)] [Medline: [35451981](https://pubmed.ncbi.nlm.nih.gov/35451981/)]
26. Grolier M, Arefyev A, Pereira B, Figueiredo IT, Gerbaud L, Coudeyre E. Refining the design of a smartphone application for people with chronic low back pain using mixed quantitative and qualitative approaches. *Disabil Rehabil Assist Technol* 2020 Nov 05:1-6. [doi: [10.1080/17483107.2020.1839575](https://doi.org/10.1080/17483107.2020.1839575)] [Medline: [33151760](https://pubmed.ncbi.nlm.nih.gov/33151760/)]
27. Richardson J, Letts L, Sinclair S, Chan D, Miller J, Donnelly C, et al. Using a web-based app to deliver rehabilitation strategies to persons with chronic conditions: Development and usability study. *JMIR Rehabil Assist Technol* 2021 Mar 18;8(1):e19519 [FREE Full text] [doi: [10.2196/19519](https://doi.org/10.2196/19519)] [Medline: [33734090](https://pubmed.ncbi.nlm.nih.gov/33734090/)]

## Abbreviations

**LBP:** low back pain

**mHealth:** mobile health

**PEG-3:** Pain, Enjoyment, and General Activity

**PHQ-4:** Patient Health Questionnaire-4

*Edited by A Mavragani; submitted 08.07.22; peer-reviewed by C Yazdi, R Chow; comments to author 15.08.22; revised version received 01.09.22; accepted 02.09.22; published 13.10.22.*

*Please cite as:*

*Browne JD, Vaninetti M, Giard D, Kostas K, Dave A*

*An Evaluation of a Mobile App for Chronic Low Back Pain Management: Prospective Pilot Study*

*JMIR Form Res* 2022;6(10):e40869

URL: <https://formative.jmir.org/2022/10/e40869>

doi: [10.2196/40869](https://doi.org/10.2196/40869)

PMID: [36227637](https://pubmed.ncbi.nlm.nih.gov/36227637/)

©Jonathan D Browne, Michael Vaninetti, David Giard, Konstantinos Kostas, Ankur Dave. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 13.10.2022. 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

# Developing Population Health Surveillance Using mHealth in Low-Resource Settings: Qualitative Assessment and Pilot Evaluation

Natalie C Benda<sup>1</sup>, PhD; Sakie Zawtha<sup>2</sup>, MBBS; Kyrie Anderson<sup>1</sup>, BS; Mohit Manoj Sharma<sup>1</sup>, MPH; Phoe Be Lin<sup>2</sup>, BSN; Beichoza Zawtha<sup>2</sup>, MBBS; Ruth Masterson Creber<sup>1</sup>, RN, MSc, PhD

<sup>1</sup>Department of Population Health Sciences, Weill Cornell Medicine, New York, NY, United States

<sup>2</sup>Health and Hope Myanmar, Yangon, Myanmar

**Corresponding Author:**

Natalie C Benda, PhD

Department of Population Health Sciences

Weill Cornell Medicine

425 E 61st Street

Suite 301

New York, NY, 10065

United States

Phone: 1 5635432222

Email: [ncb4001@med.cornell.edu](mailto:ncb4001@med.cornell.edu)

## Abstract

**Background:** Population surveillance data are essential for understanding population needs and evaluating health programs. Governmental and nongovernmental organizations in western Myanmar did not previously have means for conducting robust, electronic population health surveillance.

**Objective:** This study involved developing mobile health (mHealth)-based population health surveillance in a rural, low-resource setting with minimal cellular infrastructure in western Myanmar. This was an early formative study in which our goal was to establish the initial feasibility of conducting mHealth population health surveillance, optimizing procedures, and building capacity for future work.

**Methods:** We used an iterative design process to develop mHealth-based population health surveillance focused on general demographics (eg, total census, age category, sex, births, and deaths). Interviews were conducted with international consultants (nurse midwives) and local clinicians (nurses and physicians) in Myanmar. Our analytic approach was informed by the Systems Engineering Initiative for Patient Safety work systems model to capture the multilevel user needs for developing health interventions, which was used to create a prototype data collection tool. The prototype was then pilot-tested in 33 villages to establish an initial proof of concept.

**Results:** We conducted 7 interviews with 5 participants who provided feedback regarding the domains of the work system, including environmental, organizational, sociocultural, technological, informational, and task- and people-based considerations, for adapting an mHealth tool. Environmental considerations included managing limited electricity and internet service. Organizational needs involved developing agreements to work within existing government infrastructure as well as leveraging the communal nature of societies to describe the importance of surveillance data collection and gain buy-in. Linguistic diversity and lack of experience with technology were both cited as people- and technology-based aspects to inform prototype design. The use of mobile tools was also viewed as a means to improve the quality of the data collected and as a feasible option for working in settings with limited internet access. Following the prototype design based on the findings of initial interviews, the mHealth tool was piloted in 33 villages, allowing our team to collect census data from 11,945 people for an initial proof of concept. We also detected areas of potentially missing data, which will need to be further investigated and mitigated in future studies.

**Conclusions:** Previous studies have not focused heavily on the early stages of developing population health surveillance capacity in low- and middle-income countries. Findings related to key design considerations using a work systems lens may be informative to others developing technology-based solutions in extremely low-resource settings. Future work will involve collecting additional health-related data and further evaluating the quality of the data collected. Our team established an initial proof of concept for using an mHealth tool to collect census-related information in a low-resource, extremely rural, and low-literacy environment.



**KEYWORDS**

mobile health; mHealth; low- and middle-income countries; population health surveillance; user-centered design; mobile phone

## *Introduction*

Reliable and valid population surveillance data are essential for understanding population needs and evaluating health programs [1,2]. Electronic data capture is widely integrated for population health surveillance in high-income countries. Low- and middle-income countries have increasingly demonstrated the feasibility of implementing electronic population health surveillance systems [2-7]. Extremely low-resource settings still face multilevel challenges in establishing sustainable, robust electronic population health surveillance, including sparse or unreliable internet, competing demands of daily life, lack of access to technology devices, low literacy, diverse value structures, and heterogeneous cultural and linguistic design needs [4,7,8].

Mobile health (mHealth) has been presented by the World Health Organization as a means to mitigate some of the aforementioned challenges owing to its relatively low cost, expanding ubiquity throughout the world, and the ability to use cellular networks, which are much more widespread than broadband internet [5,9,10]. Many publications report the evaluation of mHealth interventions in low-resource settings [11,12], but few have described the formative development of population health surveillance systems that support program evaluation [4,5]. In this study, we described the formative development of an mHealth-based tool for conducting population health surveillance in an extremely rural, low-resource, low-literacy environment in western Myanmar. The long-term goal is to support the development of robust infrastructure for population health surveillance in the Chin State of Western Myanmar.

Myanmar presents a unique case related to development of population health surveillance tools, as its borders fully opened to foreign technological influence in the early 2010s; thus, the technological infrastructure lags behind many other low-income countries [9]. As of 2020, Myanmar ranked 147 of 189 on the United Nations Human Development Index [13], and the rural Chin State of Western Myanmar, where this work took place, had the second lowest life expectancy in Myanmar [14]. Residents of this ethnic minority region face barriers to accessing care, including overt racism and an average of 3 days travel to reach a health care facility.

We partnered with a local nongovernmental organization (NGO) that served as the primary provider of health services in the region. The NGO provides health care through local clinics by training community health workers and traditional birth attendants. One of our collaborative goals was to develop a system for mHealth-based population health surveillance to facilitate ongoing needs assessments in villages and evaluation of health programs. Our NGO partner previously collected population surveillance information via paper forms, which presented multiple challenges owing to the time lag from data

collection to data entry, as well as the harsh environmental conditions of the remote geography and rainy climate that impede the collection of the data forms in the rainy season. Population health surveillance information helped the NGO track general trends in population (eg, births and deaths), health needs that drive the programs they implement, and evaluate the impact of their programs on population health. Owing to the limitations related to paper-based surveillance, the NGO had never aggregated paper-based data collection in a systematic manner to understand the demographics and needs of their population. mHealth-based electronic data capture of population health surveillance information was proposed given the lower cost of mobile devices, their growing availability in the region, and their flexibility for use in settings with limited internet connectivity [5]. Here, we described formative studies that included interviews and field testing of an mHealth-based population health surveillance data collection tool. The aims of this evaluation were to establish an initial proof of concept and refine processes with the long-term goal-building capacity to use mHealth to provide real-time information regarding census-related information (births and deaths), common health issues addressed by community health workers, and outcomes of local programs (eg, women and infant health education). Our work can be used to inform the needs and best practices to establish a population health surveillance infrastructure in extremely low-resource settings.

## *Methods*

### **Phase One: Semistructured Interviews and Iterative Design**

#### *Setting*

All data were collected via teleconference with stakeholders residing in the United Kingdom or Myanmar by a research team residing in the United States. Interviews with stakeholders in the United Kingdom were conducted via Zoom (Zoom Video Communications, Inc; audio and video), and interviews with stakeholders in Myanmar were conducted via WhatsApp (Meta) or Slack (Slack Technologies; audio only) owing to limitations on internet bandwidth in Myanmar.

#### *Description of the Original (Paper-Based) Data Collection Instrument*

This study aimed to translate population health surveillance forms previously collected via paper into an mHealth-based tool. Our NGO partners operated in the United Kingdom and Myanmar with clinical experts from the United Kingdom traveling to Myanmar to advise with the design and evaluation of the NGO's health programs. Paper-based population health surveillance forms were developed by the clinical staff in Myanmar with guidance from United Kingdom-based midwife consultants.

The population health surveillance forms collected census data stratified by age, sex, births per year, and deaths per year. The census data included the total number of men and women in each of the following age categories: 0 to 5 years, 6 to 15 years, 16 to 59 years, and >60 years. Owing to the limitations of the collected paper data, this information had not been previously aggregated; therefore, there was no opportunity for high-level summary and comparison of information over time. The immediate target end users of the tool were the clinical staff from the NGO during outreach trips to villages. In the longer term, the clinical staff will train community health workers living in the villages to use the tool.

**Study Design and Sample**

We first conducted iterative phases of semistructured interviews to select a technology platform and create a prototype of the mHealth-based census form. Our goal was not to achieve theoretical saturation of qualitative data but to iteratively complete formative data collection sufficient for developing a prototype that could be used in a pilot evaluation [15,16]. Stakeholders interviewed included (1) 2 midwife consultants from the United Kingdom and (2) 3 clinical staff from the NGO in Myanmar. Midwives developed paper-based data collection forms used by the NGO. Clinical staff from the NGO included a nurse and 2 physicians with responsibilities such as development and execution of training programs for community health workers, assessment of community health worker competencies, oversight of a local clinic, clinical care at the local clinic, and NGO program evaluation. All participants worked closely with community health workers, who were the eventual end users of the tool. The local NGO staff also had the opportunity to introduce data collection tools to community health workers between rounds of interviews so that they could relay their feedback to the US-based team. Unfortunately, owing to language barriers and internet bandwidth limitations, it was not possible for the US-based team to interview community health workers directly. All interviews were conducted between February and September 2020. The participants were not directly compensated for their time.

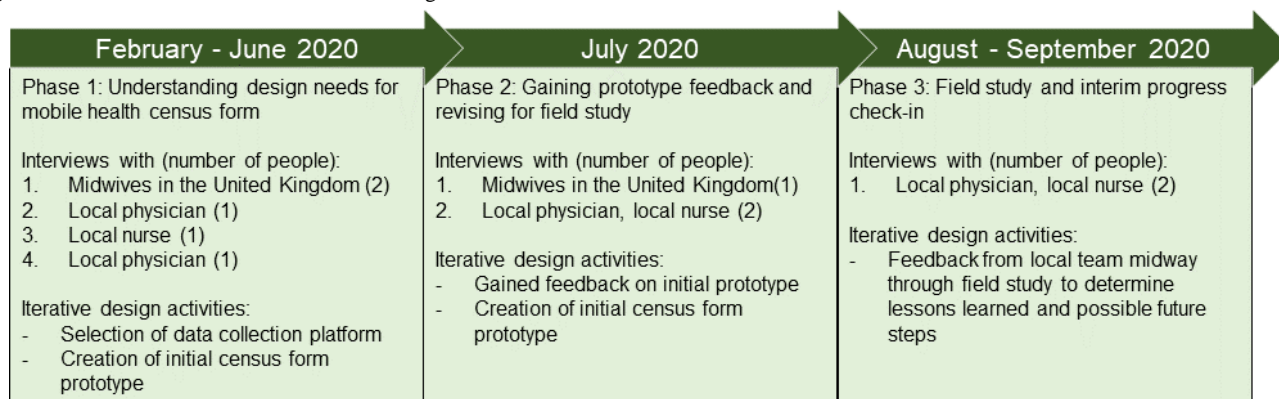
**Data Collection**

We completed 3 rounds of interviews with 5 key stakeholders as part of an iterative, formative design process, where smaller numbers of participants (ie, 4-6) were appropriate for needs assessment and design generation [16]. In round 1, we elicited information from all participants about the current processes for paper-based population health surveillance data collection and determined design needs for mHealth-based data collection. On the basis of the findings from the first round of interviews, a design platform (Survey123) was selected (by the NGO staff in the United Kingdom and Myanmar), and initial designs for the mHealth form were created by a member of the US-based team (NCB). Designs were created directly in the survey design, and the data collection platform was selected using a desktop-based program from the vendor.

In round 2, participants viewed the prototype mHealth form and provided their feedback. Round 2 included 2 separate interviews: 1 with a nurse midwife from the United Kingdom and 1 with a nurse and physician in Myanmar. For the staff in Myanmar, forms for second-round interviews were shared via the Survey123 platform before the interview, as live screen sharing was not possible. A final design based on follow-up interviews was created before the field study. An interview was then conducted in the middle of the field studies with the same local physician and nurse from round 2 to understand how the prototype form was suited to the needs of the NGO. Figure 1 presents a timeline of the interviews, prototype design, and feasibility assessment activities.

All interviews focused on current practices, perceived barriers, and facilitators of mHealth data collection. Textbox 1 provides sample questions for both the initial needs assessment interviews and follow-up interviews in which design feedback was collected. Additional probing questions were included in the full interview guide based on the work system components of the Systems Engineering Initiative for Patient Safety (SEIPS), for example, physical environment, organization, people, and task. In line with a semistructured approach, these probes were asked directly in instances when interviewees did not initially discuss each system component.

**Figure 1.** Timeline of interview and iterative design activities.



**Textbox 1.** Sample interview questions.**Needs assessment (round 1) sample questions:**

- Can you tell us about how community health workers fill out the current data collection forms?
- What are the positives of using the paper data collection forms?
- What are some of the challenges for using the paper data collection forms?
- Do you think the mobile phone app could make data collection better? If so, how?
- Do you think the mobile phone could make data collection harder? If so, how?
- Can you tell us more about the community health workers? For example, do most of them read or speak Burmese?
- Do you think the community health workers have used a mobile phone before? A smartphone? Do you think it would be possible for them to learn if we helped train them?
- Is there any more information you think is important for us to know?

**Prototype feedback (rounds 2 and 3) sample questions:**

- We have made a census form. What do you think of it?
  - What can we improve?
- What kinds of problems do you think your team may have using the data collection form?
- What new information did you hope to learn from the population surveys?
- How is the information you hope to learn going to help improve (nongovernmental organization name) programs?

All interviews were conducted by a single interviewer (NCB), with other research team members attending and participating as feasible. All interviews were audio recorded and transcribed, excluding an instance where internet connectivity issues prevented the interview from taking place synchronously, so written answers to the interview questions were provided by the interviewee asynchronously.

**Data Analysis**

Our team conducted qualitative interviews using thematic analysis and a constant comparative process [17]. First, a team of 3 coders (NCB, KSA, and MMS) read the transcripts to inductively elicit key themes. The codes were then organized into higher level categories based on work system components highlighted in the SEIPS model, given the complex, intertwined, multilevel constraints involved in a low-resource setting [18-20]. These categories and subcategories were used to develop a codebook consisting of themes, definitions, and exemplars [21]. Once the codebook had been developed, 2 researchers (KSA and MMS) independently coded each transcript, meeting with a third reviewer (NCB) to resolve discrepancies through consensus. After all the transcripts were analyzed, the research team reviewed the results code-by-code to ensure consistency and develop a list of emerging themes. All data were coded using Dedoose (version 8.0.35).

**Phase Two: Pilot Field Study****Setting, Study Design, and Sample**

The interviews were followed by a field study to establish an initial proof of concept for the use of the mHealth population health surveillance tool in this low-resource setting. Real-time population health surveillance using the census forms developed allowed for the detection of large population changes that required further investigation, as well as allocation of resources

to villages based on population. The goal of this pilot study was to establish a proof of concept for using mHealth to collect census-related information and determine potential data quality issues (eg, missingness) that could be optimized through the improvement of data collection tools or training.

The mHealth forms were piloted in villages in the Chin State during outreach visits. The region is characterized by remote, mountainous terrain, so the 33 villages included were a convenience sample located within 1 to 2 days motorbike ride from the NGO's central headquarters.

**Data Collection**

In the selected villages visited, NGO staff members filled out the population health surveillance forms on mobile devices using a commercially available mapping and survey platform (Survey123). This platform was selected because of its ability to integrate survey-related information with detailed maps necessary to identify the villages served in this remote region. Survey123 is available via app (on Android and Apple) and via web browser. Our partners collected data using tablets and a web browser-based form. Staff members involved in data collection included trained nurses and "area coordinators" (former community health workers who had been promoted into an oversight role). Nurses and area coordinators completed the population health surveillance forms by reviewing paper-based census documentation in the villages (procured from the government and local churches). Paper-based documents often had different age categories, which required the team to add categories from the paper-based forms to suit the categories in the mHealth forms. Nurses and area coordinators then went door-to-door to confirm the census in each sex or age category, number of births that year, and number of deaths. All data were collected between August and September 2020.

**Data Analysis**

Data were downloaded from the Survey123 and summarized descriptively using Microsoft Excel. The proof of concept was assessed by investigating the completeness of the fields included in the form. Summarization included the total number of villages from which data were captured, total persons accounted for, total births, total deaths, and the number of potential missing data fields (Phase Two: Pilot Field Study in the Results section).

**Ethics Approval**

This study was approved (20-02021442) by the Weill Cornell Medicine Institutional Review Board and conducted in 2 phases.

The first phase involved a series of semistructured interviews and iterative design of an mHealth-based population health surveillance tool with key stakeholders. The second phase involved piloting the tool into a field study.

**Results**

**Overview**

A total of 7 interviews were conducted with 5 unique participants: 2 midwife consultants and 3 staff members from the NGO. Further description of the interviewees is provided in Table 1. A midwife consultant and 2 staff members were interviewed multiple times.

**Table 1.** Description of participants.

Identifier	Participant description
LocalNurse1	Nurse working for the NGO <sup>a</sup> in Myanmar, in charge of CHW <sup>b</sup> training programs
LocalPhysician1	Physician working in Myanmar also serving as the NGO chief operating officer at the time of the interviews
LocalPhysician2	Physician working in Myanmar overseeing the local clinic and medical outreach programs
MidwifeConsultant1	United Kingdom-based midwife that had developed training and data collection for traditional birth attendants affiliated with the NGO
MidwifeConsultant2	United Kingdom-based midwife that had developed training and data collection for traditional birth attendants affiliated with the NGO

<sup>a</sup>NGO: nongovernmental organization.

<sup>b</sup>CHW: community health worker.

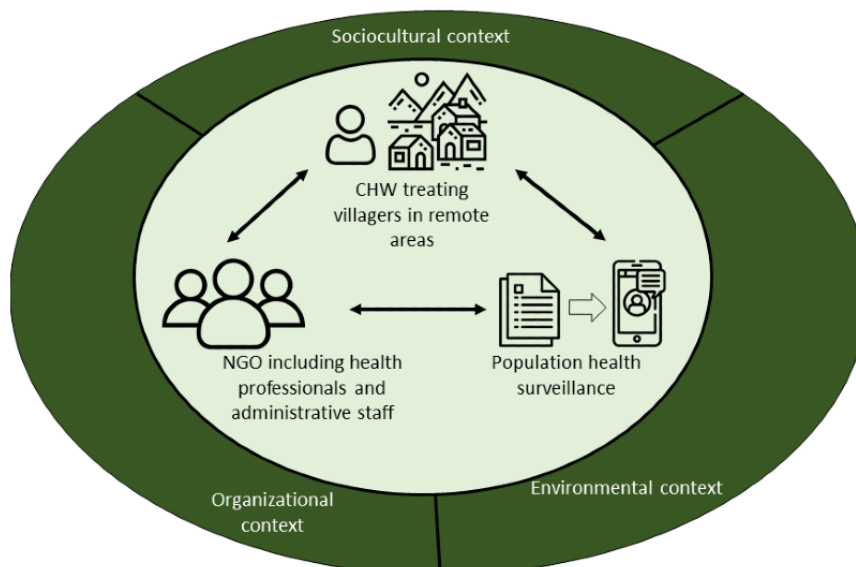
**Phase One: Semistructured Interviews**

**Description of Themes**

Key themes pertained to various components of the work system, including the physical environment, organizational concerns, tasks, technology or physical tools, people (user skills and motivation), and information. In line with the traditional SEIPS model, many themes intersected across work system components [20,22]. Figure 2 presents the SEIPS-informed work system for the given app, where community health workers

treating patients in villages (people) must communicate information (census and symptoms) to the NGO (people) for the purpose of population health surveillance (task). As this communication previously occurred via paper, the goal of our work was to translate paper data collection tools into an mHealth-based population health surveillance tool to develop capacity for real-time, easy-to-collate data collection. This work was situated in a particular organizational, sociocultural, and physical context that must be considered in the adaptation of new technological tools.

**Figure 2.** Work system depiction for the current app. CHW: community health worker; NGO: nongovernmental organization.



The themes described herein focus on important design considerations for population health surveillance. We have presented the quotes to bolster the description of themes, followed by an identifier indicating the source of the quote. Information included in brackets within the quotes was not explicitly spoken by the participant but provides a clarifying context. Table 1 provides a legend of how identifiers were provided at the end of each quote map for each participant.

### Physical Environmental Themes

Discussion of environmental themes highlighted the need to consider access to electricity, availability of devices, and internet in low-resource settings when surveillance involved the use of mHealth:

*90% [of the villages] they don't get electricity and they depend on [a] solar system...so whenever they want to charge the phone they would plug in the solar. And not even every household gets solar...only 10 in 30 households get solar. [LocalNurse1]*

It was also important to consider weather and topography, and how they may impact whether resources can be transported to certain areas during these times of the year:

*But in monsoon season, you know we have almost five months rainy season in our country...They rarely...travel. But when they have a better road, a better access to villages to go from here to there by walking or by motorbike, then we ask them to do that. But again, it's quite difficult. [LocalPhysician1]*

### Organizational or Sociocultural Themes

The organizational themes were multilevel, involving the government, cultural norms, and training support from the local organization (ie, our NGO partners). From a governmental perspective, planning was required to effectively manage collaborations with government-funded health entities. Our participants explained that successful integration would involve being aware of the formal and informal structures that exist and using both positively, for example, interfacing with government-funded midwives for maternal and childcare:

*But the only problem there is that because they're [non-NGO affiliated midwives] government-paid...so you'd have to get permission from the government...to use them somehow. [MidwifeConsultant1]*

Consideration of cultural norms was viewed as a necessity. The region contains predominantly agrarian societies where community health workers provide care to other villagers during their free time when they are not farming. Therefore, it was important to make data collection as simple and efficient as possible to avoid interfering with the competing demands of their daily lives:

*They have CHWs, but they are farmers, and they are just volunteer people...so they have to work for their living as well. At the same time, they have to serve the community, so it is really challenging for them. [LocalPhysician2]*

The cohesive nature of the villages was seen as a potential motivator for buy-in with data collection if the NGO trainers

could illustrate how collecting population health surveillance information impacted the health of the larger community (People-based considerations). Finally, providing in-person training and additional ongoing support, possibly remotely, was also a necessary consideration.

### Task-Related Themes

Participants described the need to support key tasks related to population health surveillance, including detection of problematic public health patterns, measurement of program progress or effectiveness, and promotion of safe health practices. The detection of public health problems would allow for better allocation of resources (eg, medicines and first aid) and additional training to address emerging problems, such as the COVID-19 pandemic:

*Data were very helpful especially during Covid-19 outreach. It was easy for us to choose the villages based on the data form and was easy for preparation of materials based on the population they provided. As for me data is most useful in preparation of medicines and materials to distribute to the active villages. I can prepare medicines and materials based on the number of active CHWs and population of the villages. [LocalNurse1]*

### Technology Themes

Key desired specifications related to the technological tools included: low cost, inclusion of data security protection (eg, password protection and remote wipe capabilities), ability to transmit information to cloud-based databases, ease of use, flexibility (to edit forms as needed), and capabilities to work in the given environment (eg, protection during the rainy season and store-and-forward capabilities for areas with limited cell service):

*Yes, I think using [a] mobile phone app for data collection will be a perfect choice for the...CHWs...It is easy to carry and once it is filled there is no need to worry about the loss of information. It will be done at once so time saving...The outcome of data will be more accurate. It is more simple to understand on mobile. [LocalNurse1]*

### Person-Related Themes

Person-based design needs pertained to user skills and motivations. Different user needs or skills to accommodate multiple languages, have limited literacy, and little to no experience with mobile technologies:

*One of our huge obstacles has been the translation...they really genuinely don't read each other's dialects, don't always follow each other's dialects, and so on. So, again, it's not just illiteracy, it's overcoming language barriers that would make some sort of pictorial and touchscreen thing easier. [MidwifeConsultant2]*

*[The] mobile phone thing is very new. Last year (2019) we have got this telephone tower and people start[ed] using phones. [LocalPhysician1]*

Another person-based theme involved understanding and leveraging potential motivators for community health work and data collection. Avenues for motivation overlapped with other work system components such as reimbursement or payment provided by the local NGO, respect for other villagers and desire to help, excitement related to new technologies (ie, mobile phones), and personal altruism. It was also viewed that connecting the data collection activities to the larger purpose of the different organizations (both the NGO and village) would serve as a motivating factor:

*We also have to explain why we make them use this app...because we would like [to know about] their village and how the patients are being taken care of by them, and then it'll be...a strong connection between the community and us [the NGO].* [LocalPhysician2]

**Information-Related Themes**

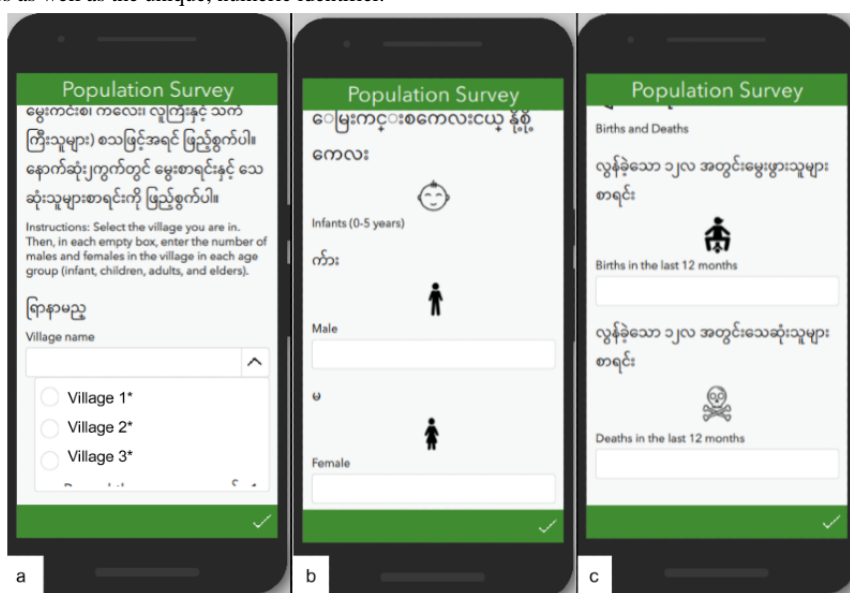
Information-related themes touched on data quality and data structure. In general, data quality was perceived to be higher when using the mHealth tool over paper forms as it eliminated issues with reading handwriting. Structured data entry was also viewed as helpful in reducing issues related to limited literacy, although local team members wanted unstructured data entry to account for atypical cases:

*Some of the CHWs are not so educated so they do not know the exact words to fill sometimes.* [LocalNurse1]  
*Actually, here “others” is something that involves nothing that we have here...If it's out of this topic and you see something other than this or that, please describe it. That's how we taught them.* [LocalPhysician2]

**Iterative Design**

Figure 3 shows key design features of the population health surveillance form developed for census tracking based on the iterative interviews with key stakeholders previously described. Figure 3A shows instructions provided in Burmese (and English for comparison purposes). Users were first able to search the village for which they are completing the survey. The form entry contained both Burmese text and icon-based representation of the information to be entered (Figures 3B and 3C). In addition to the visual design features, Survey123 allowed for local device data storage when there was insufficient internet service but would then automatically upload the information to a secure cloud server once the device had service again, thus mitigating previously described concerns related to internet connectivity. Table 2 provides a complete list of updates made to the prototype in advance of the pilot study on the basis of the interviews.

**Figure 3.** Population health surveillance form for collecting village census information. Village names have been redacted but included the village name in Burmese characters as well as the unique, numeric identifier.



**Table 2.** Approaches implemented for the design considerations derived from interview themes. Work system domain.

Particulars and design considerations	Approach implemented
<b>Environmental</b>	
Lack of electricity for charging devices	<ul style="list-style-type: none"> <li>Included solar charged batteries in budget to supplement existing solar-based systems</li> </ul>
Mountainous topography, transportation challenges owing to weather	<ul style="list-style-type: none"> <li>Choose a mapping software to find villages in need</li> </ul>
<b>Organizational</b>	
Planning of government vs NGO <sup>a</sup> ownership of materials	<ul style="list-style-type: none"> <li>Developed a plan where NGO would purchase and own devices, developing memorandum of understanding with the government</li> </ul>
Need to balance data collection with demands of daily life	<ul style="list-style-type: none"> <li>Created simple-form design so data collection is as efficient as possible</li> </ul>
<b>Task to support</b>	
Detection of problematic public health patterns	<ul style="list-style-type: none"> <li>Chose a mapping platform, which allows geographic tracking and data summarization</li> </ul>
Program measurement	<ul style="list-style-type: none"> <li>Chose a mapping platform, which allows geographic tracking and data summarization</li> </ul>
<b>Technology</b>	
Low cost	<ul style="list-style-type: none"> <li>Found a program usable on Android (ie, for mapping and survey) offering nonprofit pricing</li> </ul>
Flexibility to edit forms	<ul style="list-style-type: none"> <li>Used a web platform for NGO staff members to edit forms easily</li> </ul>
Work in the given environment	<ul style="list-style-type: none"> <li>Purchased waterproof cases for devices</li> </ul>
<b>People</b>	
Easy to learn and use	<ul style="list-style-type: none"> <li>Leveraged an iterative participatory design process to create data collection forms</li> </ul>
Multiple languages	<ul style="list-style-type: none"> <li>Included Burmese text in forms</li> </ul>
Limited literacy	<ul style="list-style-type: none"> <li>Included icons in forms</li> </ul>
Data structure (unstructured vs structured)	<ul style="list-style-type: none"> <li>Made most fields structured</li> <li>Included open-ended text boxes for additional notes not conforming to structured data collection elements</li> </ul>

<sup>a</sup>NGO: nongovernmental organization.

## Phase Two: Pilot Field Study

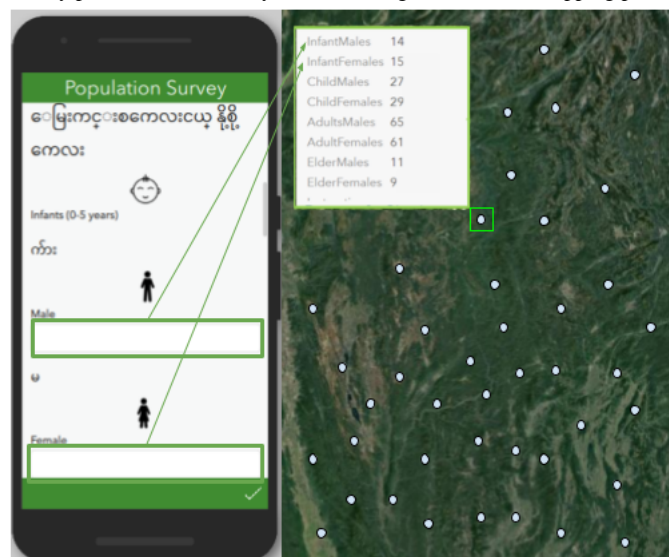
The prototype displayed in [Figure 3](#) was used to capture population information in 33 villages by the NGO staff, including nurses and area coordinators who oversaw the community health worker training program. The prototype was created using Survey123 Connect for ArcGIS (version 3.9.120; Esri) [Figure 4](#) illustrates how the census information collected using the survey app is integrated with the mapping app to allow users to view specific population statistics and detect patterns in population growth or decline. Data collection in 33 villages allowed our NGO partners to account for 11,945 people in prespecified demographic groups based on age and sex. In addition, they captured 205 births and 27 deaths across the villages. There were potential missing data from 8 and 17 villages regarding birth and death counts, respectively, but there were no missing data indicated for the general census

information (ie, census of villagers in age and sex groups). The missing data may indicate that birth- and death-related information is more challenging to determine or that the users simply did not know that they should enter the number “0” to indicate none instead of leaving the field blank.

Feedback from the study in the field (third round interviews) provided a proof of concept for electronic (mHealth-based) population health surveillance data collection and underscored the need for store-and-forward capabilities for villages that may not have sufficient cellular service for real-time data upload. For example, in one of the follow-up interviews, an NGO-employed nurse described the following:

*So far, it's very good and it is very easy to use...I think compared to the paper form...I prefer these Survey123 survey data forms. [LocalNurse1]*

**Figure 4.** Demonstration of how the survey platform from Survey123 (left) integrates with the mapping platform (right).



## Discussion

### Principal Findings

Our formative, iterative design approach allowed us to develop an mHealth tool for population health surveillance in an extremely rural, low-resource, limited-literacy environment with minimal cellular infrastructure in western Myanmar. We first summarized our results and illustrated how they were translated into concrete next steps, as the implementation of our findings may be informative for others conducting similar work. We then described our findings in the context of the work system model. Finally, we discussed the overlap of our findings with those of other studies and highlighted how our results extend currently published knowledge.

Key design needs were organized based on environmental, organizational, sociocultural, technological, and task-, people-, and information-based constructs. Many of the design considerations derived from the interviews were addressed before the field study (Table 2). Our field study provided an initial proof of concept by piloting the mHealth tool for in situ data collection in 33 villages. However, additional considerations were not necessary or feasible to implement at the time of the pilot study. For example, training considerations for community health workers were not needed for the pilot study, as it was conducted by the NGO staff. Table 3 presents further design considerations from our interviews, which were operationalized in later ongoing studies.

As shown in Tables 2 and 3, many of the considerations did not involve the design of the mHealth tool but instead pertained to the design of processes, training, and selection of technologies. In addition, some considerations were addressed through multiple solutions, and others addressed multiple considerations. Although example design considerations are listed under the primary work system domain, the examples also highlight the overlap of the domains. For example, building logic checks in the data collection form (information) may also mitigate issues related to limited literacy and ease of use (people). However, error messages may need to be conveyed in both written and

audio formats to those with limited literacy. Additionally, built-in logic checks must be facilitated by the chosen technology.

Table 3 describes the steps implemented in subsequent ongoing work. We further explored the reasons for data missingness. In addition, our ongoing work extended the described mHealth tool to allow for collection of noncommunicable diseases and maternal or child health information to better understand patterns in health needs within the villages that may also explain changes in census. Our approach for evaluating data plausibility described in Table 3 will also provide quantitative metrics to evaluate data quality and make further improvements. The ongoing refinement of population health surveillance tools is underway in preparation for future clinical trials to evaluate maternal health interventions.

Previous studies regarding the integration of electronic health surveillance and mHealth integration into low-resource settings are similar to ours, but others may not yet be applicable in particularly underdeveloped settings. For example, McIntosh et al [4] also noted the importance of store-and-forward technologies in areas where internet connectivity may be sparse. The weather and topography issues encountered in our study, which may limit travel to allow for regular data uploads, present an additional complication and consideration for our study. Berry et al [23] described an approach for population health surveillance in Bangladesh using random digit dialing instead of the door-to-door approach used in our work. Other studies in south Asia and Sub-Saharan Africa have also demonstrated the feasibility of using SMS text messaging for basic data capture and sending reminders to improve adherence to maternal and child health regimens [24-26]. However, as noted by one of our participants, cell phone towers have only recently been constructed in many villages (as of 2019); therefore, cell phones are not yet a ubiquitous, equitable means for communicating with the population of the Chin State. Our results may be particularly helpful for regions with less developed cellular infrastructure as well as those facing climate- and geography-related challenges.



**Table 3.** Considerations derived from interviews implemented in subsequent efforts.

Work system domain and design considerations	Implemented approach
Environmental considerations derived implemented in pilot study	<ul style="list-style-type: none"> <li>Considerations derived implemented in pilot study</li> </ul>
Ongoing organizational training and support	<ul style="list-style-type: none"> <li>Developed videos for data collection forms to providing ongoing training</li> </ul>
<b>Task to support</b>	
Considerations derived implemented in pilot study	<ul style="list-style-type: none"> <li>Considerations derived implemented in pilot study</li> </ul>
Secure data collection and storage	<ul style="list-style-type: none"> <li>Obtained software that may be locked (with a passcode) and wiped remotely</li> </ul>
Ability to transmit information to database	<ul style="list-style-type: none"> <li>Purchased cords for data transfer so data could be stored and uploaded later in villages with poor cell service</li> </ul>
<b>People</b>	
Easy to learn and use	<ul style="list-style-type: none"> <li>Iterative participatory design</li> </ul>
Multiple languages	<ul style="list-style-type: none"> <li>Added text or audio from additional local languages (eg, Mara)</li> </ul>
Limited mobile device experience	<ul style="list-style-type: none"> <li>Incorporated mobile device basics into data collection training</li> </ul>
Motivating data collection	<ul style="list-style-type: none"> <li>Connected data collection to larger purpose of the work in training</li> <li>Providing needed supplies to community health workers</li> <li>Created a data plausibility review process by NGO<sup>a</sup> staff</li> </ul>

<sup>a</sup>NGO: nongovernmental organization.

## Limitations

Our study has limitations related to (1) sample size for the iterative design process, (2) field study, (3) transferability of findings, (4) ongoing COVID-19 pandemic, and (5) military coup in the region. First, our interviews involved 5 participants and 7 total interviews, which are smaller than sample sizes recommended in some qualitative guidelines but appropriate for our approach—an iterative, formative design process [15,16,27]. Our goal was not to meet qualitative saturation but to pragmatically sample to gain enough information for our iterative design process. Previous studies on technology design have indicated that  $\geq 5$  participants may be sufficient for early-stage formative design [16,28]. Our initial studies focused on the clinical staff of the NGO, and future studies will involve community health workers who will also eventually use mHealth data collection forms to provide more timely information from each village to the NGO's central headquarters. Second, the field study did not formally evaluate feasibility (eg, via survey) or data quality; rather, it was a proof of concept. We plan to include ratings of data plausibility by trained staff members and retraining as needed to record implausible values in future efforts, and to assess the timeliness of data upload to detect the effect of connectivity issues in certain areas [29]. Third, data were collected from a single NGO in a specific Myanmar region. Although we sought to determine general design considerations for extremely low-resource settings, it is unclear whether our results will be transferable to other settings. Fourth, we

originally planned to conduct formative interviews and field assessments with members of our US-based research team traveling to Myanmar for in-person data collection. Owing to safety precautions and travel restrictions related to the COVID-19 pandemic, this was not possible at the time of the study. Finally, on February 1, 2021, the democratically elected government of Myanmar was overthrown by a military coup and the country was in a state of emergency; as such, the NGO operations were suspended.

## Conclusions

We conducted formative, qualitative interviews with key stakeholders to develop an mHealth-based system for population health surveillance in a low-resource setting in Myanmar. Design considerations involved various work system domains including physical environment, organization-, task-, technology-, people-, and information-related issues. Members of our NGO partners were able to use the mHealth-based population health surveillance tool to establish an initial proof of concept in a feasibility study of 33 villages. Future work will involve evaluating the plausibility of the data collected and expanding data collection beyond simple census information to disease and maternal or child health outcome information. The February 2021 military coup in Myanmar presented a difficult political environment for the employment of this approach in our continued work that will also need to be addressed, moving forward regarding safety and security reasons.

## Acknowledgments

This work was supported by the Anna-Maria and Stephen Kellen Foundation and the Eunice Kennedy Shriver National Institute of Child Health and Human Development for their support. (R21HD103053; principal investigator: RMC)

## Conflicts of Interest

None declared.

## References

1. Grembowski D. The Practice of Health Program Evaluation. Thousand Oaks, CA, USA: Sage Publications; 2001.
2. Nsubuga P, White ME, Thacker SB, Anderson MA, Blount SB, Broome CV, et al. Public health surveillance: a tool for targeting and monitoring interventions. In: Jamison DT, Breman JG, Measham AR, Alleyne G, Claeson M, Evans DB, et al, editors. Disease Control Priorities in Developing Countries. 2nd edition. Washington, DC, USA: The World Bank; 2006.
3. Hashmi M, Beane A, Murthy S, Dondorp AM, Haniffa R. Leveraging a cloud-based critical care registry for COVID-19 pandemic surveillance and research in low- and middle-income countries. *JMIR Public Health Surveill* 2020 Nov 23;6(4):e21939 [FREE Full text] [doi: [10.2196/21939](https://doi.org/10.2196/21939)] [Medline: [33147162](https://pubmed.ncbi.nlm.nih.gov/33147162/)]
4. McIntosh S, Pérez-Ramos J, Demment MM, Vélez Vega C, Avendaño E, Ossip DJ, et al. Development and implementation of culturally tailored offline mobile health surveys. *JMIR Public Health Surveill* 2016 Jun 02;2(1):e28 [FREE Full text] [doi: [10.2196/publichealth.5408](https://doi.org/10.2196/publichealth.5408)] [Medline: [27256208](https://pubmed.ncbi.nlm.nih.gov/27256208/)]
5. McCool J, Dobson R, Whittaker R, Paton C. Mobile health (mHealth) in low- and middle-income countries. *Annu Rev Public Health* 2022 Apr 05;43:525-539. [doi: [10.1146/annurev-publhealth-052620-093850](https://doi.org/10.1146/annurev-publhealth-052620-093850)] [Medline: [34648368](https://pubmed.ncbi.nlm.nih.gov/34648368/)]
6. Bartolomeos KK. The case for investing in public health surveillance in low- and middle-income countries. *Afr J Emerg Med* 2018 Dec;8(4):127-128 [FREE Full text] [doi: [10.1016/j.afjem.2018.07.005](https://doi.org/10.1016/j.afjem.2018.07.005)] [Medline: [30534514](https://pubmed.ncbi.nlm.nih.gov/30534514/)]
7. Jayatilleke K. Challenges in implementing surveillance tools of high-income countries (HICs) in low middle income countries (LMICs). *Curr Treat Options Infect Dis* 2020;12(3):191-201 [FREE Full text] [doi: [10.1007/s40506-020-00229-2](https://doi.org/10.1007/s40506-020-00229-2)] [Medline: [32874140](https://pubmed.ncbi.nlm.nih.gov/32874140/)]
8. Lupu N, Michelitch K. Advances in survey methods for the developing world. *Annu Rev Polit Sci* 2018 May 11;21(1):195-214. [doi: [10.1146/annurev-polisci-052115-021432](https://doi.org/10.1146/annurev-polisci-052115-021432)]
9. Nam KY, Cham WM, Halili PR. Developing Myanmar's information and communication technology sector toward inclusive growth. ADB Economics Working Paper Series. 2015 Dec 30. URL: <https://www.adb.org/sites/default/files/publication/176518/ewp-462.pdf> [accessed 2021-08-14]
10. WHO Guideline: Recommendations on digital interventions for health system strengthening. World Health Organization. Geneva, Switzerland: World Health Organization; 2019 Jun 6. URL: <https://www.who.int/publications/i/item/9789241550505> [accessed 2021-08-14]
11. Nijland N, van Gemert-Pijnen JE, Kelders SM, Brandenburg BJ, Seydel ER. Factors influencing the use of a Web-based application for supporting the self-care of patients with type 2 diabetes: a longitudinal study. *J Med Internet Res* 2011 Sep 30;13(3):e71 [FREE Full text] [doi: [10.2196/jmir.1603](https://doi.org/10.2196/jmir.1603)] [Medline: [21959968](https://pubmed.ncbi.nlm.nih.gov/21959968/)]
12. Habibović M, Cuijpers P, Alings M, van der Voort P, Theuns D, Bouwels L, et al. Attrition and adherence in a WEB-based distress management program for implantable cardioverter defibrillator patients (WEBCARE): randomized controlled trial. *J Med Internet Res* 2014 Feb 28;16(2):e52 [FREE Full text] [doi: [10.2196/jmir.2809](https://doi.org/10.2196/jmir.2809)] [Medline: [24583632](https://pubmed.ncbi.nlm.nih.gov/24583632/)]
13. Latest Human Development Index Ranking. United Nations Development Programme. 2020. URL: <http://hdr.undp.org/en/content/latest-human-development-index-ranking> [accessed 2021-06-07]
14. The 2014 Myanmar Population and Housing Census: Thematic Report on Mortality. Department of Population, Ministry of Labour, Immigration and Population. 2016 Sep. URL: [https://myanmar.unfpa.org/sites/default/files/pub-pdf/4B\\_Mortality.pdf](https://myanmar.unfpa.org/sites/default/files/pub-pdf/4B_Mortality.pdf) [accessed 2021-08-14]
15. Ancker JS, Benda NC, Reddy M, Unertl KM, Veinot T. Guidance for publishing qualitative research in informatics. *J Am Med Inform Assoc* 2021 Nov 25;28(12):2743-2748 [FREE Full text] [doi: [10.1093/jamia/ocab195](https://doi.org/10.1093/jamia/ocab195)] [Medline: [34537840](https://pubmed.ncbi.nlm.nih.gov/34537840/)]
16. Kirk D, McClelland I, Suri JF. Involving people in design research. In: Wilson JR, Sharples S, editors. Evaluation of Human Work. Boca Raton, FL, USA: CRC Press; 2015:282-331.
17. Corbin J, Strauss A. Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory. 4th edition. Thousand Oaks, CA, USA: Sage Publications; 2014.
18. Benda NC, Masterson Creber RM. Underserved populations: integrating social determinants of health into the study of patient work. In: Valdez RS, Holden RJ, editors. The Patient Factor: Applications of Patient Ergonomics. Boca Raton, FL, USA: CRC Press; 2021:163-186.
19. Valdez RS, Holden RJ, Novak LL, Veinot TC. Transforming consumer health informatics through a patient work framework: connecting patients to context. *J Am Med Inform Assoc* 2015 Jan;22(1):2-10 [FREE Full text] [doi: [10.1136/amiajnl-2014-002826](https://doi.org/10.1136/amiajnl-2014-002826)] [Medline: [25125685](https://pubmed.ncbi.nlm.nih.gov/25125685/)]

20. Carayon P, Wooldrige A, Hoonakker P, Hundt AS, Kelly MM. SEIPS 3.0: human-centered design of the patient journey for patient safety. *Appl Ergon* 2020 Apr;84:103033 [FREE Full text] [doi: [10.1016/j.apergo.2019.103033](https://doi.org/10.1016/j.apergo.2019.103033)] [Medline: [31987516](https://pubmed.ncbi.nlm.nih.gov/31987516/)]
21. Saldana J. *The Coding Manual For Qualitative Researchers*. 2nd edition. Thousand Oaks, CA, USA: Sage Publications; 2012.
22. Holden RJ, Carayon P, Gurses AP, Hoonakker P, Hundt AS, Ozok AA, et al. SEIPS 2.0: a human factors framework for studying and improving the work of healthcare professionals and patients. *Ergonomics* 2013;56(11):1669-1686 [FREE Full text] [doi: [10.1080/00140139.2013.838643](https://doi.org/10.1080/00140139.2013.838643)] [Medline: [24088063](https://pubmed.ncbi.nlm.nih.gov/24088063/)]
23. Berry I, Mangtani P, Rahman M, Khan IA, Sarkar S, Naureen T, et al. Population health surveillance using mobile phone surveys in low- and middle-income countries: methodology and sample representativeness of a cross-sectional survey of live poultry exposure in Bangladesh. *JMIR Public Health Surveill* 2021 Nov 12;7(11):e29020 [FREE Full text] [doi: [10.2196/29020](https://doi.org/10.2196/29020)] [Medline: [34766914](https://pubmed.ncbi.nlm.nih.gov/34766914/)]
24. Kazi AM, Ali M, Zubair K, Kalimuddin H, Kazi AN, Iqbal SP, et al. Effect of mobile phone text message reminders on routine immunization uptake in Pakistan: randomized controlled trial. *JMIR Public Health Surveill* 2018 Mar 07;4(1):e20 [FREE Full text] [doi: [10.2196/publichealth.7026](https://doi.org/10.2196/publichealth.7026)] [Medline: [29514773](https://pubmed.ncbi.nlm.nih.gov/29514773/)]
25. Kazi AM, Carmichael JL, Hapanna GW, Wangoo PG, Karanja S, Wanyama D, et al. Assessing mobile phone access and perceptions for texting-based mHealth interventions among expectant mothers and child caregivers in remote regions of Northern Kenya: a survey-based descriptive study. *JMIR Public Health Surveill* 2017 Jan 30;3(1):e5 [FREE Full text] [doi: [10.2196/publichealth.5386](https://doi.org/10.2196/publichealth.5386)] [Medline: [28137702](https://pubmed.ncbi.nlm.nih.gov/28137702/)]
26. Choudhury A, Asan O, Choudhury MM. Mobile health technology to improve maternal health awareness in tribal populations: mobile for mothers. *J Am Med Inform Assoc* 2021 Oct 12;28(11):2467-2474 [FREE Full text] [doi: [10.1093/jamia/ocab172](https://doi.org/10.1093/jamia/ocab172)] [Medline: [34459478](https://pubmed.ncbi.nlm.nih.gov/34459478/)]
27. Nielsen J. The usability engineering life cycle. *Computer* 1992 Mar;25(3):12-22. [doi: [10.1109/2.121503](https://doi.org/10.1109/2.121503)]
28. Morse JM. The significance of saturation. *Qual Health Res* 1995 May;5(2):147-149. [doi: [10.1177/104973239500500201](https://doi.org/10.1177/104973239500500201)]
29. Kahn MG, Callahan TJ, Barnard J, Bauck AE, Brown J, Davidson BN, et al. A harmonized data quality assessment terminology and framework for the secondary use of electronic health record data. *EGEMS (Wash DC)* 2016 Sep 11;4(1):1244 [FREE Full text] [doi: [10.13063/2327-9214.1244](https://doi.org/10.13063/2327-9214.1244)] [Medline: [27713905](https://pubmed.ncbi.nlm.nih.gov/27713905/)]

## Abbreviations

**mHealth:** mobile health

**NGO:** nongovernmental organization

**SEIPS:** Systems Engineering Initiative for Patient Safety

*Edited by A Mavragani; submitted 28.01.22; peer-reviewed by A Johnson, C Weerth; comments to author 29.06.22; revised version received 20.07.22; accepted 23.08.22; published 14.10.22.*

*Please cite as:*

Benda NC, Zawtha S, Anderson K, Sharma MM, Lin PB, Zawtha B, Masterson Creber R

*Developing Population Health Surveillance Using mHealth in Low-Resource Settings: Qualitative Assessment and Pilot Evaluation*  
*JMIR Form Res* 2022;6(10):e36260

URL: <https://formative.jmir.org/2022/10/e36260>

doi: [10.2196/36260](https://doi.org/10.2196/36260)

PMID: [36239997](https://pubmed.ncbi.nlm.nih.gov/36239997/)

©Natalie C Benda, Sakie Zawtha, Kyrie Anderson, Mohit Manoj Sharma, Phoe Be Lin, Beichoza Zawtha, Ruth Masterson Creber. Originally published in *JMIR Formative Research* (<https://formative.jmir.org>), 14.10.2022. 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 for Telemedicine's Ongoing Transformation of Health Care Delivery Since the Onset of COVID-19: Retrospective Observational Study

Soumik Mandal<sup>1,2</sup>, PhD; Batia M Wiesenfeld<sup>3</sup>, PhD; Devin Mann<sup>1,4</sup>, MD, MS; Katharine Lawrence<sup>1,4</sup>, MD; Rumi Chunara<sup>5,6</sup>, PhD; Paul Testa<sup>4</sup>, MD; Oded Nov<sup>1,2</sup>, PhD

<sup>1</sup>Department of Population Health, New York University Grossman School of Medicine, New York University, New York, NY, United States

<sup>2</sup>Department of Technology Management & Innovation, New York University Tandon School of Engineering, New York University, New York, NY, United States

<sup>3</sup>New York University Leonard N Stern School of Business, New York University, New York, NY, United States

<sup>4</sup>Medical Center Information Technology, New York University Langone Health, New York University, New York, NY, United States

<sup>5</sup>Computer Science & Engineering, New York University Tandon School of Engineering, New York University, New York, NY, United States

<sup>6</sup>Biostatistics, New York University School of Global Public Health, New York University, New York, NY, United States

**Corresponding Author:**

Soumik Mandal, PhD

Department of Population Health

New York University Grossman School of Medicine

New York University

180 Madison Ave

New York, NY, 10016

United States

Phone: 1 6465013442

Email: [mandal.soumik@gmail.com](mailto:mandal.soumik@gmail.com)

## Abstract

**Background:** The surge of telemedicine use during the early stages of the COVID-19 pandemic has been well documented. However, scarce evidence considers the use of telemedicine in the subsequent period.

**Objective:** This study aims to evaluate use patterns of video-based telemedicine visits for ambulatory care and urgent care provision over the course of recurring pandemic waves in 1 large health system in New York City (NYC) and what this means for health care delivery.

**Methods:** Retrospective electronic health record (EHR) data of patients from January 1, 2020, to February 28, 2022, were used to longitudinally track and analyze telemedicine and in-person visit volumes across ambulatory care specialties and urgent care, as well as compare them to a prepandemic baseline (June–November 2019). Diagnosis codes to differentiate suspected COVID-19 visits from non–COVID-19 visits, as well as evaluating COVID-19–based telemedicine use over time, were compared to the total number of COVID-19–positive cases in the same geographic region (city level). The time series data were segmented based on change-point analysis, and variances in visit trends were compared between the segments.

**Results:** The emergence of COVID-19 prompted an early increase in the number of telemedicine visits across the urgent care and ambulatory care settings. This use continued throughout the pandemic at a much higher level than the prepandemic baseline for both COVID-19 and non–COVID-19 suspected visits, despite the fluctuation in COVID-19 cases throughout the pandemic and the resumption of in-person clinical services. The use of telemedicine-based urgent care services for COVID-19 suspected visits showed more variance in response to each pandemic wave, but telemedicine visits for ambulatory care have remained relatively steady after the initial crisis period. During the Omicron wave, the use of all visit types, including in-person activities, decreased. Patients between 25 and 34 years of age were the largest users of telemedicine-based urgent care. Patient satisfaction with telemedicine-based urgent care remained high despite the rapid scaling of services to meet increased demand.

**Conclusions:** The trend of the increased use of telemedicine as a means of health care delivery relative to the pre–COVID-19 baseline has been maintained throughout the later pandemic periods despite fluctuating COVID-19 cases and the resumption of in-person care delivery. Overall satisfaction with telemedicine-based care is also high. The trends in telemedicine use suggest that telemedicine-based health care delivery has become a mainstream and sustained supplement to in-person-based ambulatory

care, particularly for younger patients, for both urgent and nonurgent care needs. These findings have implications for the health care delivery system, including practice leaders, insurers, and policymakers. Further investigation is needed to evaluate telemedicine adoption by key demographics, identify ongoing barriers to adoption, and explore the impacts of sustained use of telemedicine on health care outcomes and experience.

(*JMIR Form Res* 2022;6(10):e38661) doi:[10.2196/38661](https://doi.org/10.2196/38661)

## KEYWORDS

digital health; telemedicine; urgent care; COVID-19; health care delivery

## Introduction

Before the COVID-19 pandemic, the use of telemedicine as a care delivery modality was limited; only 8% of Americans reported using telemedicine for medical care in 2019 [1]. Barriers to scaled adoption and use included limited reimbursement, patients' and providers' lack of comfort with telemedicine technologies, and a strong cultural norm of in-person care [2]. The dynamic interactions between these individual factors [3] often lead to nonadoption and abandonment of telemedicine technologies by their intended users [4-6]. However, beginning in March 2020, the telemedicine landscape in the United States changed rapidly, as the World Health Organization (WHO) declared COVID-19 a global pandemic, and a nationwide health care emergency was declared in the United States [7]. Prior to the large-scale availability of vaccines and effective therapies, social distancing and quarantine were the only widely accepted approaches to minimizing viral spread, creating a compelling (and often compulsory) pressure to find alternatives to in-person care [8]. To help maintain existing health care operations while meeting the new demands imposed by rising COVID-19 cases, health care systems quickly turned to telemedicine solutions for care provision, with many experiencing early exponential growth in telemedicine adoption [8]. To ensure the pace of scaling telemedicine capacity matched the growing demand, implementation with rapid iterative improvements was preferred over perfect execution [9]. Where possible, existing technology and vendors were used instead of investing time into procuring brand-new technology. As a result, telemedicine infrastructure often spanned multiple technologies and platforms, supported different modalities (voice based over the telephone, or video based) rather than any standardized implementation, and evolved rapidly in a short period, all of which could negatively affect patients' satisfaction [10] and patients' continued use of telemedicine.

Although research has documented the enthusiastic adoption of telemedicine technology in the early stage of the pandemic, little subsequent research has explored whether this migration has been sustained in the postacute pandemic periods. Additionally, prior reports relating the growing prevalence of telemedicine to a steady decline in in-person clinic visit volumes have suggested that telemedicine is at least partially replacing clinic visits [8]. Since the early phase of the pandemic in the spring of 2020, improved public health measures, breakthrough developments in vaccine research, and widespread vaccine and treatment protocols have made the resumption of in-person activities possible, including the provision of in-person medical

care; currently, it is unclear whether the rise in telemedicine will be sustained as the US health care system transitions to more "regular" operations. There is a growing general literature on the long-term sustainability of technology-supported change in health care services [11], but studies on sustainability of telehealth services remain sparse [12]. This study, drawn from a large academic health care system in New York City (NYC), aims to explore patterns in patients' use of telemedicine during the recurring waves of the pandemic.

The research question being answered in this study is, What were the trends in the use of video-based telemedicine visits for ambulatory care and urgent care provision over the course of recurring pandemic waves?

## Methods

### Study Setting

In this study, we used data from the New York University Langone Health (NYULH) system, a large urban and suburban academic health care system in NYC whose operations were significantly impacted by the COVID-19 pandemic and that responded by developing a robust telemedicine infrastructure to provide patient care during the period of clinic closures and disruptions.

The NYULH network consists of over 8000 health care providers across 4 hospitals and more than 350 ambulatory care locations in urban and suburban settings, all connected to a single electronic health record (EHR) system (Epic, Verona, WI). To enable its telemedicine services (known in the health system as "virtual health"), the NYULH uses a single instance of the Epic health record with more than 8.17 million active patients leveraging an integrated video visit platform. Prior to the COVID-19 pandemic, the NYULH implemented telemedicine capabilities in approximately 25 locations via its "virtual urgent care" (VUC) service, a video visit experience tightly integrated into its enterprise EHR and patient portal, offering same-day virtual appointments with emergency medical physicians for acute nonemergent health concerns (eg, new cough, fever). Virtual nonurgent care or ambulatory care, such as virtual primary care, was subsequently developed, offering a more comprehensive set of services, including chronic disease management, interdisciplinary care with specialists and ancillary care (eg, dietitians, therapists), and preventive care, with care handled by internal medicine or specialty clinicians.

Patients access the virtual services through the NYULH app built upon the Epic MyChart suite of patient tools and using standard Application Programming Interfaces (APIs) made available by Epic. During a telemedicine encounter, patients

can begin their video visits directly through their patient portal app; the provider has to simply click a link in their EHR system to launch the visit. The provider's click action opens a browser for the video that can be seen in tandem with the EHR in the same manner as an in-person visit. In addition, the NYULH has deployed native open scheduling technologies as well as custom features enabling simplified telemedicine access and matriculation. The NYULH uses Q-Reviews (New York), a real-time hospital review digital engagement platform to collect feedback from patients on their VUC visits.

### Study Design

In this study, we used patients' visit information from the EHR data to characterize visit types from January 1, 2020, to February 28, 2022, representing the period of recurring waves of pandemic intensity. We used heterogeneous sources of data, including encounters, visits, diagnoses, patient satisfaction, and patients' age, to identify the age groups that accessed care through telemedicine or in-person visits during this period. To categorize whether a telemedicine visit happened in ambulatory care or urgent care, the visit type, location, and specialty information were used.

To evaluate whether telemedicine use was skewed toward COVID-19 suspected visits, we evaluated *International Classification of Diseases 10th Revision (ICD-10)* diagnosis codes containing relevant primary respiratory and primary nonrespiratory symptoms via partial matching with 34 keywords (Table 1) [8,13,14]. This included COVID-19-related diagnosis codes, which were frequently used in the health system used prior to updated COVID-19 coding recommendations in 2020 and 2021 [15]. COVID-19 suspected visits were compared to total COVID-19 cases per day in NYC during similar periods to evaluate whether the prevalence of COVID-19-related illness among NYULH patients compared to the larger NYC population [16]. Descriptive statistics were computed to estimate rates of telemedicine visits in urgent care and nonurgent care settings. Telemedicine use for COVID-19 suspected and non-COVID-19 suspected visits were evaluated independently to assess for relationships between visit type and telemedicine use preference. A change-point detection analysis with binary segmentation [17] was used to identify changes in the visit trends over time and locate mean shifts in combined telemedicine and in-person visits. The change-point indexes were used to segment the 26 months' visit data; each segment represented a change in the

distribution of the time-ordered visit counts with respect to preceding and subsequent segments [18]. Finally, statistical properties (mean, variance) of visit counts were computed and compared between the segments for each visit type independently. The Levene test was used to evaluate the equality of variances in visit counts among time segments.

Prior studies that analyzed the demographics of patients during the initial surge in the use of telemedicine reported that telemedicine use was mostly confined to young patients [8,10]. To assess potential changes in telemedicine patient demographics throughout the study period and whether the expansion of telemedicine facilities has galvanized telemedicine adoption across a range of age groups, we also evaluated the age group of patients participating in telemedicine visits in our data. For each telemedicine visit record, we determined the patient's age at the time of the visit and combined the records for patients who were from similar age groups. We compared the telemedicine use with the baseline population estimate from the 2020 US Census Bureau data for NYC [19] for each age group.

In addition to data collected from EHRs, patient satisfaction and engagement were captured and evaluated via a brief text message survey disseminated via Q-Reviews at the close of VUC telemedicine encounters. The survey assessed various domains, including satisfaction with the visit, likelihood to use telemedicine again, and how well the visit addressed/managed the patient's medical needs, on 5-point scales (5=most satisfied); see Table 2. Satisfaction was assessed based on the responses to these 3 questions ( $\alpha=.87$ ), and trends in patients' satisfaction were analyzed. The survey also asked respondents to estimate time costs/savings relative to in-person visits and how likely they would be to recommend VUC to a friend or colleague. Finally, average visits per patient were measured based on the count of unique patient identifiers in the data. Patients' average telemedicine-based visits and in-person visits were compared with the prepandemic baseline.

To assess whether virtual health care delivery supplements or replaces in-person care, we calculated the average number of in-person and virtual visits per patient in 3 periods: a prepandemic baseline of June–November 2019, June–November 2020, and a postacute pandemic comparison of June–November 2021.

**Table 1.** Keywords used to identify COVID-19 suspected cases from ICD-10<sup>a</sup> diagnostic codes.

Symptom type	Keywords
Primary respiratory	(1) COVID, (2) respiratory distress, (3) flu, (4) sore throat, (5) congestion, (6) URI, (7) pneumonia, (8) shortness of breath, (9) cough, (10) dyspnea, (11) pharyngitis, (12) bronchitis, (13) sinusitis, (14) ARDS, (15) lung infiltrates, (16) hypoxia, (17) tachypnea, (18) opacities, (19) wheezing
Primary nonrespiratory	(20) chest pain, (21) muscle pain, (22) joint pain, (23) stress, (24) headache, (25) fever, (26) bleeding, (27) swelling, (28) rash, (29) skin lesion, (30) insomnia, (31) malaise, (32) constipation, (33) anxiety, (34) depression

<sup>a</sup>ICD-10: *International Classification of Diseases 10th Revision*.

**Table 2.** Survey of patients' satisfaction with VUC<sup>a</sup>.

Survey question	Scale
How satisfied were you with your VUC visit?	• 1-5
How well did the VUC visit address/manage your medical needs?	• 1-5
How likely are you to use VUC again?	• 1-5
How much time did you save by using VUC, including travel time?	<ul style="list-style-type: none"> <li>• &lt;1 hour</li> <li>• 1-2 hours</li> <li>• 2-3 hours</li> <li>• 3-4 hours</li> <li>• &gt;4 hours</li> <li>• N/A<sup>b</sup></li> </ul>
How likely are you to recommend VUC to a friend or colleague?	• 1-10

<sup>a</sup>VUC: virtual urgent care.

<sup>b</sup>N/A: not applicable.

### Data Exclusion

Other than the EHR-integrated platform, the NYULH also used Webex by Cisco and telephone calls for a brief period for providing telemedicine services (<1% of all telemedicine visits), which are not included in this report.

### Ethical Considerations

We submitted the study proposal to the NYULH Institutional Review Board (IRB), for which exemption was awarded (#s21-01207). Further clarification with regard to the policies and terms of reference can be obtained from the IRB.

## Results

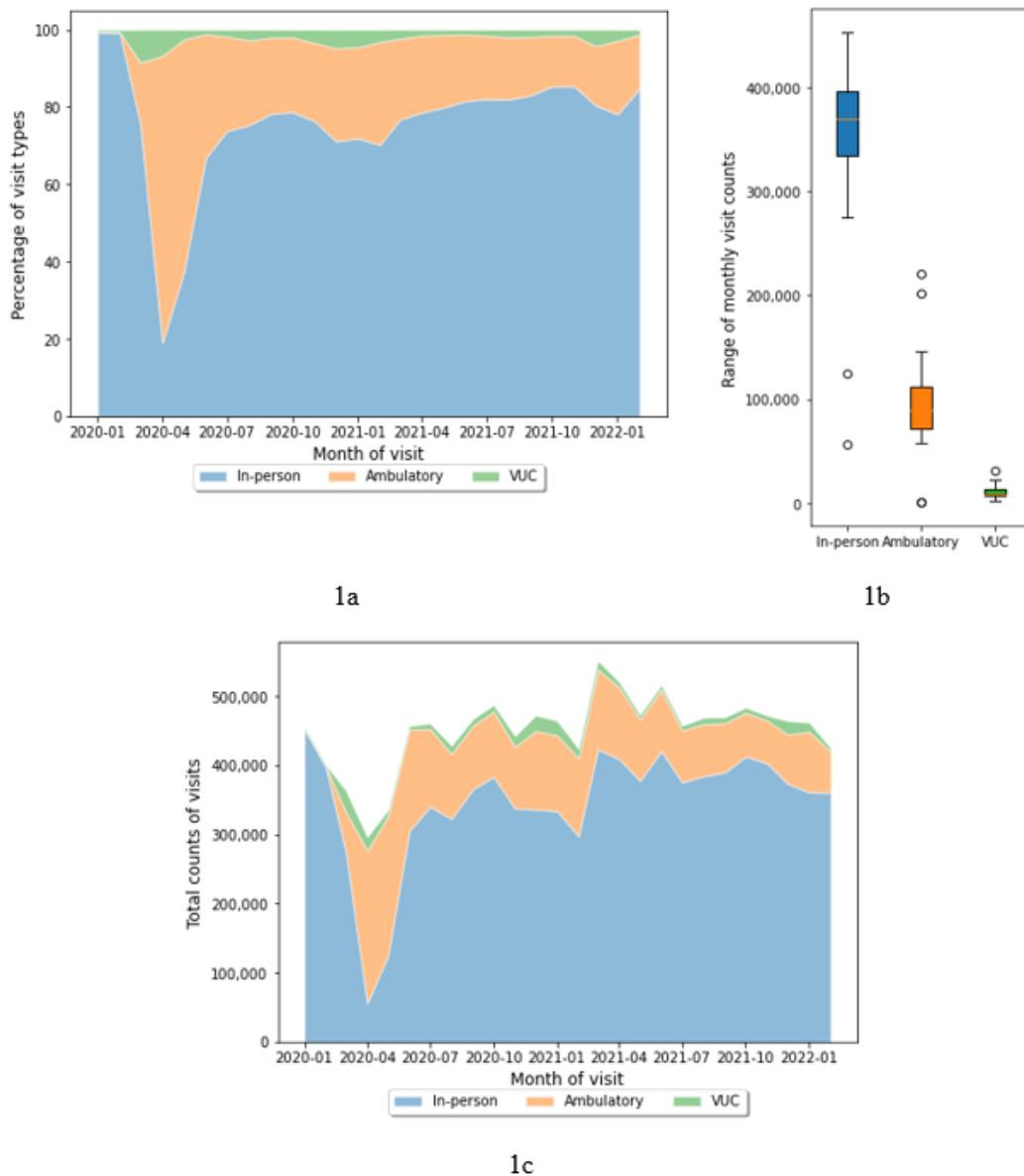
### Overall Trends in Telemedicine Use

During the 26-month pandemic period, a total of 2,748,635 telemedicine visits were recorded, measuring nearly one-third (30.45%) the volume of in-person visits (N=9,025,553) in the same time frame. The use of ambulatory nonurgent care (eg, virtual primary care) was much higher than VUC (see [Figure 1b](#)). Nearly 89.26% of all video visits (n=2,409,003) were for ambulatory nonurgent care (eg, virtual primary care), with the remaining 10.74% (n=289,836) visits for VUC. Overall, the visit trends showed that volumes of telemedicine visits peaked in the acute pandemic phase and continued at a higher rate than before the pandemic; telemedicine volume between January and February 2020 was <100 visits per day and subsequently peaked during the month of April 2020 (n=240,356, 80.98%) with simultaneous declines in in-person visits ([Figures 1a](#) and [1c](#)). This shift from in-person visits to telemedicine was particularly evident during the acute pandemic period (March and April 2020) and during periods distinguished by the spread of newer strains of the virus, Delta (October 2020-January 2021) and Omicron (November 2021-January 2022). The Pearson coefficient (r) showed that volumes of in-person and VUC visits

per month were negatively correlated ( $r=-0.421$ ,  $P=.03$ ). Additionally, the distribution of telemedicine visits demonstrated higher use of telemedicine by patients for nonurgent ambulatory care needs than for urgent care (see [Figure 1b](#)). Overall, the visit trends showed that volumes of telemedicine visits peaked in the acute pandemic phase, declined as in-person visits resumed, but then remained at a rate much higher than before the pandemic and with less fluctuation from July 2020 to February 2022 (monthly telemedicine visits ranged from 64,570 to 136,181 across the period). Further details of the VUC, ambulatory care telemedicine, and in-person visit data during the period are provided in [Multimedia Appendix 1](#).

The change-point analysis detected 4 change points or mean shifts in the combined telemedicine and in-person visit trends. Based on the change-point indexes, the 26-month time series was divided into the following 5 segments: first segment (until April 2020), second segment (May-September 2020), third segment (October 2020-February 2021), fourth segment (March-September 2021), and fifth segment (October 2021-February 2022). Descriptive statistics were computed for each combination of time segment and visit type independently, and the results are provided in [Table 3](#). Overall, the result showed the highest variations (IQR) in data on the visits per month observed in the first segment for all 3 visit types. The average number of monthly ambulatory care visits peaked in the second segment (mean 129,406, SD 46,281), which coincided with the lowest in-person visits (mean 291,829, SD 96,115). The third segment witnessed the most use (mean 16,269, SD 5351) of VUC services but with large variations (IQR 7416). The fifth and final segment was characterized by declines in both telemedicine and in-person visit types. The Levene test result found variance in visits among time segments to be significant for in-person visits ( $F_{4,21}=3.56$ ,  $P=.02$ ) and VUC ( $F_{4,21}=6.30$ ,  $P=.001$ ) but not for ambulatory care ( $F_{4,21}=2.57$ ,  $P=.07$ ).

**Figure 1.** Trends in visits in telemedicine-based urgent care (VUC), nonurgent care (ambulatory), and in-person care. (a) Percentages of visit types, (b) total counts of visits per visit type, and (c) total counts for visit types per month. VUC: virtual urgent care.



**Table 3.** Telemedicine use per month by segment.

Visit type per month	First segment (until April 2020)			Second segment (May-September 2020)			Third segment (October 2020-February 2021)			Fourth segment (March-September 2021)			Fifth segment (October 2021-February 2022)		
	Mean (SD)	Median	IQR	Mean (SD)	Median	IQR	Mean (SD)	Median	IQR	Mean (SD)	Median	IQR	Mean (SD)	Median	IQR
In person	295,607 (175,893)	336,531	191,153	291,829 (96,115)	322,933	35,107	337,981 (30,520)	336,339	3698	401,529 (23,189)	409,320	43,008	383,774 (20,086)	384,503	29,301
Ambulatory care	70,140 (103,669)	29,599	97,418	129,406 (46,281)	112,360	52,660	103,956 (11,241)	109,711	18,493	94,292 (15,399)	88,877	14,626	70,023 (9,786)	71,155	10,749
VUC <sup>a</sup>	13,729 (14,129)	11,094	20,543	8611 (2,409)	8476	1479	16,269 (5,351)	15,068	7416	8028 (2,693)	6802	1322	10,053 (4,906)	8587	3,906

<sup>a</sup>VUC: virtual urgent care.



## Trends in Telemedicine Service Use for COVID-19 Suspected Cases

Of all visits recorded (in person and telemedicine) in this period ( $N=11,774,188$ ), 1,264,487 (~10.74%) reported at least 1 COVID-19-related symptom in the diagnosis code, representing a COVID-19 suspected case. Among these cases, 766,548 (60.62%) were recorded in in-person facilities versus 497,939 (39.38%) telemedicine visits. For COVID-19 suspected cases, the percentage of telemedicine to total visits (39.38%) was higher than the overall percentage of telemedicine to all visits ( $n=2,748,635$ , 30.45%) in the data.

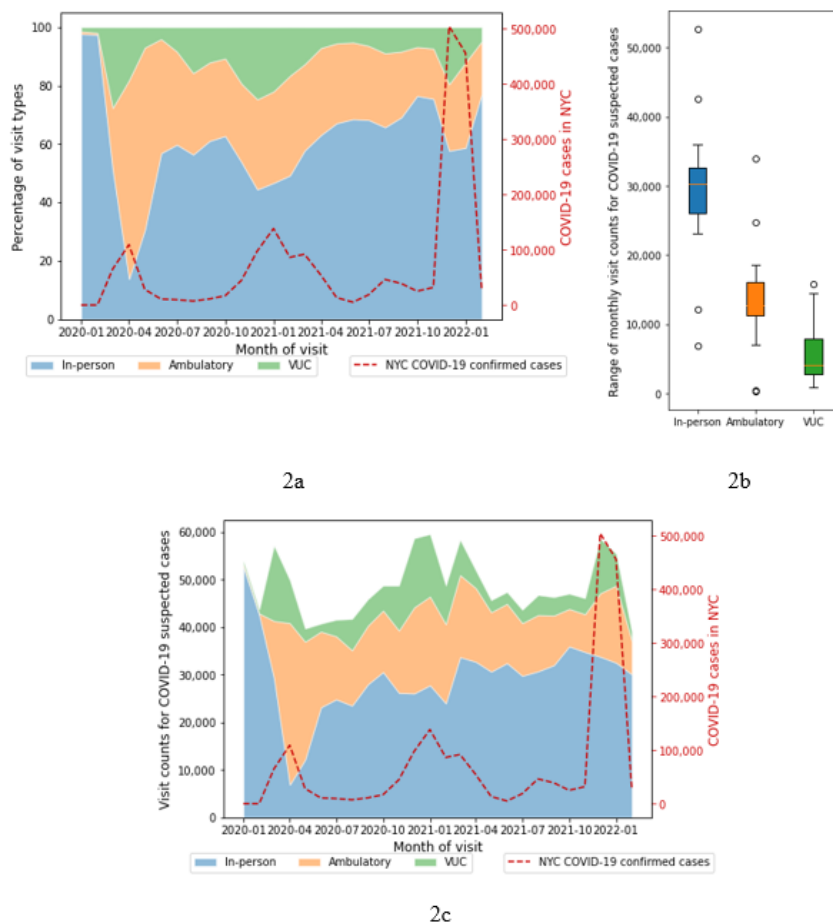
The distribution of video visit types for urgent and ambulatory care showed greater use of urgent care services for COVID-19 symptoms (see [Figures 1b](#) and [2b](#)) than what was witnessed overall: the virtual visit volumes for COVID-19 suspected cases were far more evenly distributed between VUC and nonurgent facilities than what was witnessed for all recorded visits. From all COVID-19 suspected telemedicine visits, 150,735 (30.27%) were reported in urgent care facilities and the rest 347,204 (69.73%) visits in ambulatory care.

We further compared the distributions of COVID-19 suspected visit types with confirmed COVID-19 cases in NYC [20] in the same period (see [Figures 2a](#) and [2c](#)) to evaluate the relationship between telemedicine use and surges of COVID-19 cases during recurring waves of the pandemic. The distributions demonstrated that increases in COVID-19 cases coincided with increased telemedicine visits, especially to urgent care facilities, and at the same time decreased in-person visits. This was evident in

the first (March and April 2020), second (November 2020-February 2021), and third (November 2021-January 2022) waves of the pandemic, when COVID-19 cases surged in NYC. Overall, the Pearson coefficient ( $r$ ) showed that counts of confirmed COVID-19 cases in NYC were negatively correlated with in-person visit volumes ( $r=-0.230$ ) and almost entirely unrelated to nonurgent care visit counts ( $r=0.086$ ). In contrast, urgent care visit volumes and confirmed COVID-19 numbers in NYC, which were strongly correlated (Pearson  $r=0.727$ ) until November 2021, were less correlated when the Omicron outbreak was considered ( $r=0.393$ ). Overall, the fraction of telemedicine visits changed more dynamically for urgent care (mean 0.15, SD 0.28 after normalization) than for ambulatory care (mean 0.85, SD 0.20) among COVID-19 suspected cases.

[Table 4](#) shows distributions of both in-person and telemedicine service use for COVID-19 suspected cases among the 5 sequential time segments based on the change-point analysis reported before. Similar to overall visit trends, the highest variation (IQR) in visits was observed in the first segment (until April 2020) for all 3 visit types. Among the remaining 4 time segments, the highest variation in telemedicine visits was observed in the third time segment (October 2020-February 2021) for both ambulatory care (IQR 5016) and VUC (IQR 4988). Using the Levene test, the variance in COVID-19 suspected visits was found to be significant among the time segments for in-person visits ( $F_{4,21}=6.55$ ,  $P=.001$ ) and ambulatory care ( $F_{4,21}=2.85$ ,  $P=.05$ ) but not for VUC ( $F_{4,21}=2.86$ ,  $P=.06$ ).

**Figure 2.** Trends in visit types for suspected COVID-19 cases and confirmed COVID-19 cases in NYC. (a) Percentages of visit types, (b) total counts of visits by visit type, and (c) total counts for visit types per month. NYC: New York City; VUC: virtual urgent care.



**Table 4.** Telemedicine use trends for COVID-19 suspected cases by period.

Visit type per month	First segment (until April 2020)			Second segment (May-September 2020)			Third segment (October 2020-February 2021)			Fourth segment (March-September 2021)			Fifth segment (October 2021-February 2022)		
	Mean (SD)	Median	IQR	Mean (SD)	Median	IQR	Mean (SD)	Median	IQR	Mean (SD)	Median	IQR	Mean (SD)	Median	IQR
In person	32,846 (19,777)	35,936	21,397	22,310 (5999)	23,497	1683	26,912 (2449)	26,186	1723	31,850 (1606)	32,442	2068	32,828 (2135)	32,518	2922
Ambulatory care	11,683 (15,828)	6241	17042	15,572 (5399)	13,234	3602	15,874 (2724)	16,627	5016	13,735 (2531)	12,475	3001	10,651 (3365)	10,422	4691
VUC <sup>a</sup>	6,628 (7252)	4953	9,910	4014 (2015)	3485	2776	10,088 (3767)	9454	4988	3795 (2086)	2814	1172	4960 (3217)	3875	2161

<sup>a</sup>VUC: virtual urgent care.

**Trends in Telemedicine Service Use for Non-COVID-19 Suspected Cases**

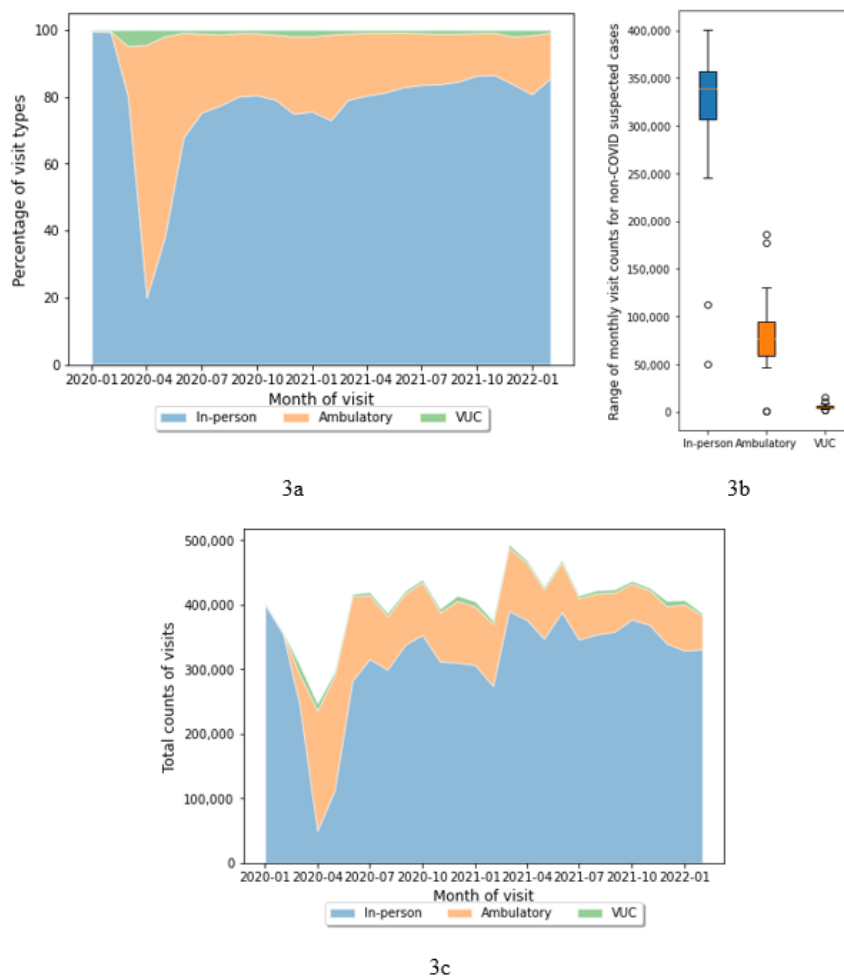
To examine the use of telemedicine beyond COVID-19 needs, we analyzed visit types for non-COVID-19 suspected cases (n=10,459,905 visits) separately (see Figure 3). Since non-COVID-19 suspected visits accounted for almost 90% of all visits recorded, their distributions across visit types (see Figures 3a and 3c) were near identical to those of all visits. For non-COVID-19 suspected cases, the proportion of telemedicine use was more skewed toward ambulatory care compared to COVID-19 suspected cases (see Figures 2b and 3b). Among all

non-COVID-19 suspected telemedicine visits, nearly 2,061,799 (93.68%) cases were from nonurgent care. In the same period, only 139,101 (6.32%) non-COVID-19 suspected cases were recorded for urgent care. In addition, 8,259,005 non-COVID-19 suspected visits were in person, representing nearly 78.96% of all non-COVID-19 suspected visits from the same period. Overall, these distributions suggest that although COVID-19 prompted rapid scaling and use of telemedicine, its use grew and then remained steady at a higher level than the pre-COVID-19 baseline (Figure 3c) for non-COVID-19 suspected cases as well.

Table 5 shows variations in telemedicine use for non-COVID-19 suspected cases among the 5 segments. The distributions of monthly visits showed that among telemedicine services, the use of ambulatory care services reached its peak in the second segment (mean 113,834, SD 41,001), declined slowly between the third (mean 88,082, SD 8790) and fourth (mean 80,556, SD 12,900) segments, and witnessed the lowest use in the fifth segment (mean 59,372, SD 6667). Compared to ambulatory care, VUC use peaked earlier in the first segment (mean 7100, SD 6920), had overall been in decline, but saw increases during

the third (mean 6180, SD 1594) and fifth (mean 5092, SD 1712) segments compared to immediate previous segments. The IQR values showed that overall, variations in visits per month data shrunk in later time segments, which was particularly noticeable in the fourth time segment for urgent care (IQR 87) and the fifth segment for ambulatory care (IQR 7502). Levene test results on visits per month among time segments found the variance to be significant for in-person visits ( $F_{4,21}=3.28, P=.03$ ) and urgent care visits ( $F_{4,21}=17.02, P<.001$ ) but not for ambulatory care visits ( $F_{4,21}=2.55, P=.07$ ).

Figure 3. Trends in visit types for non-COVID-19 suspected cases. (a) Percentages of visit types, (b) total counts of visit types, and (c) total counts for visit types per month. VUC: virtual urgent care.



**Table 5.** Telemedicine use per month for non-COVID-19 suspected cases by time segment.

Visit type per month	First segment (until April 2020)			Second segment (May-September 2020)			Third segment (October 2020-February 2021)			Fourth segment (March-September 2021)			Fifth segment (October 2021-February 2022)		
	Mean (SD)	Median	IQR	Mean (SD)	Median	IQR	Mean (SD)	Median	IQR	Mean (SD)	Median	IQR	Mean (SD)	Median	IQR
In person	262,761 (15,630)	300,595	169,756	269,518 (90,160)	299,436	33,424	311,069 (28,209)	310,285	5,289	369,679 (21,669)	376,610	41,208	350,946 (18,623)	353,797	28,118
Ambulatory care	58,457 (87,918)	23,360	80,374	113,834 (41,001)	99,126	48,358	88,082 (8,790)	91,066	14,275	80,556 (12,900)	76,402	11,625	59,372 (6,667)	57,845	7,502
VUC <sup>a</sup>	7100 (6,920)	6140	10,633	4596 (888)	4598	1206	6180 (1,594)	5614	2428	4233 (613)	3988	87	5092 (1,712)	4712	1744

<sup>a</sup>VUC: virtual urgent care.

### Trends in Telemedicine Service Use by Age Group

**Table 6** decomposes telemedicine use by age group in our data. Across virtual visit types, the 25-34-year age group accounted for the largest proportion of telemedicine visits, peaking at 40,251 (16.74%) in the month of April 2020. This pattern of higher telemedicine use for those aged 25-34 years was even stronger for VUC visits, where this age group was responsible for a total of 112,247 (38.03%) urgent care visits in the entire period. The use of telemedicine for urgent care needs was the lowest for children and young adolescents aged less than 15 years (n=7420 visits, 2.56%) despite this age group being among the largest in the NYC population [21]. The distribution further showed that although telemedicine adoption for nonurgent care needs was relatively evenly distributed across age groups, patients between 25 and 44 years old were responsible for a disproportionate share of telemedicine-based urgent care visits.

**Figures 4a** and **4b** illustrate the trends in virtual health visits among different age groups throughout the entire period considered in this study. **Figure 4b** shows that telemedicine

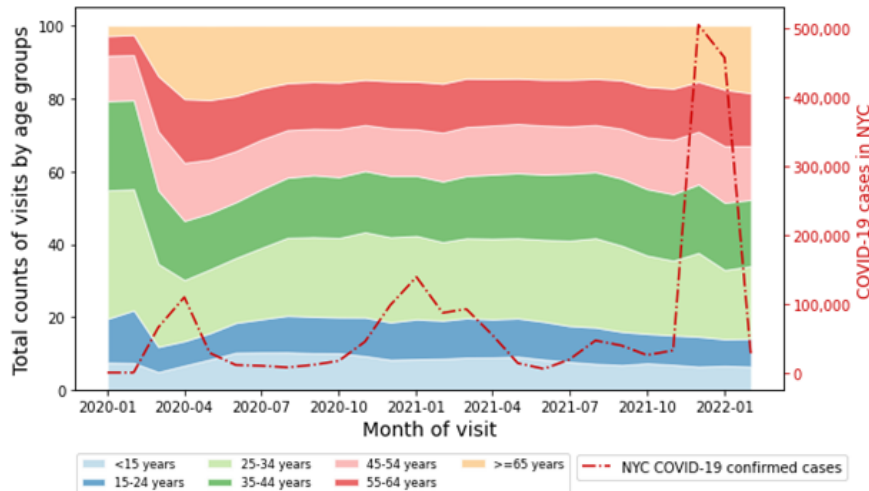
adoption was the highest at the beginning of the pandemic (between March and May 2020) for most age groups. Although the number of telemedicine visits by the largest contributing age group (25-34 years) decreased from its peak of 40,251 per month in April 2020 to 12,948 in February 2022, the average visits per month (n=21,949, 21.34%) remained consistently higher than the prepandemic level of 1062 in February 2020. **Figure 4a** shows that although telemedicine use of the largest contributing age group (25-34 years) was high throughout the period after April 2020, the other age groups' use of telemedicine grew each time the number of COVID-19 cases surged. Our analysis also found that patients from the 65 years and older age group remained consistent users of telemedicine (maximum 20,799, 15.21%, visits per month; minimum 11,927, 16.75%, visits per month) from June 2020 to February 2021. Overall, the distribution of telemedicine visits among age groups (**Table 6**) showed that although the use of telemedicine for nonurgent care among older patients increased (n=413,517, 15.37%) relative to prior reports [16], the use of telemedicine for urgent care remained quite low (n=11,630, 4.01%).

**Table 6.** Percentage distribution of telemedicine visits by age group and baseline population figures in NYC<sup>a</sup> from the US Census Bureau data of 2020.

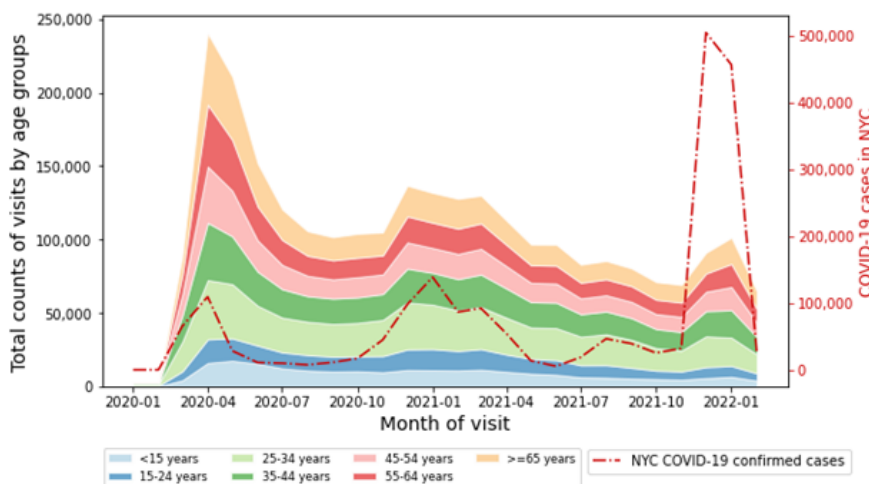
Age (years)	Population in NYC (%)	All telemedicine care (%)	Urgent care (%)	Nonurgent care (%)
<15	17.53	8.14	2.56	8.71
15-24	11.66	10.42	15.88	9.71
25-34	17.81	21.34	38.03	19.40
35-44	13.64	16.61	20.93	16.15
45-54	12.54	14.14	11.18	14.52
55-64	11.87	13.98	7.40	14.76
65 and above	14.95	15.37	4.01	16.75

<sup>a</sup>NYC: New York City.

**Figure 4.** Trends in telemedicine use by age group. (a) Stacked area graph of percentage of visits and (b) total counts of visits. NYC: New York City.



4a



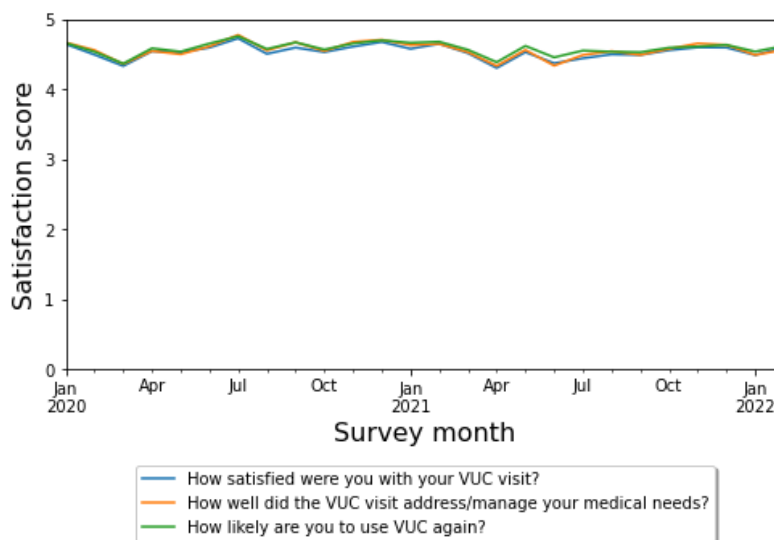
4b

**Patient Satisfaction**

In total, 13,669 patients who used VUC across the 26-month study period responded to the satisfaction survey. Despite the inexperience of providers who adopted telemedicine rapidly, patients’ satisfaction with VUC remained unchanged during the acute pandemic phase (pre-COVID-19: n=847, 6.20%, mean satisfaction 4.38/5; acute COVID-19: n=1693, 12.39%, mean satisfaction 4.38/5). Q-Reviews data on patients’ satisfaction indicated that patients were consistently highly satisfied (n=13,669, 100%, mean satisfaction 4.53/5, minimum=4.31, maximum=4.78) with VUC visits (see Figure 5), despite 2173 (15.9%) patients reporting technical issues. In addition, 10,719

(78.41%) patients were highly satisfied with their VUC visits, and only 856 (6.26%) patients were least satisfied. More than 74% of patients (105 of 141) felt they saved at least an hour of time (including travel time) by using virtual care services and would likely recommend the services to a friend or colleague.

Finally, average video visits per patient increased from 0.013 in the prepandemic baseline to 0.827 between June and February 2020, before experiencing a slight decline, and then stabilized at 0.588 between June 2021 and February 2022. During the same periods, the average number of in-person visits slightly declined from 2.928 to 2.670 at first, followed by a steady increase to the prepandemic level of 2.894 per patient.

**Figure 5.** Trends in patients' satisfaction with VUC visits (Q-Reviews) during the COVID-19 pandemic. VUC: virtual urgent care.

## Discussion

### Principal Findings

The COVID-19 pandemic prompted many health care systems to rapidly expand telemedicine services in response to significant disruptions in in-person care provisions [8,9]. However, the extent to which higher rates of telemedicine use have been maintained in the subsequent period is not yet clear. This study, which evaluated the extent to which telemedicine use has been sustained throughout the pandemic in 1 large health system in NYC, shows that although the early pandemic catalyzed rapid growth in telemedicine use, the inverse relationship between the volume of telemedicine visits and in-person emergency department visits [8] continued in subsequent periods of the pandemic, for both COVID-19-related care needs as well as routine care, such as preventive medicine, chronic condition management, and ambulatory specialty care. These findings suggest that the transition to telemedicine use in the manner of care delivery is at least partially lasting and not bounded by the end of the COVID-19 pandemic.

The detected change points in the time series data mostly coincided with the emergence of new variants and subsequent surges in COVID-19 cases in the United States. The first and second time segments match the timeline of the acute pandemic phase and the postacute pandemic phase from prior reports [8], while the subsequent 3 segments have overlaps between the timeline of the Delta and Omicron surges reported in the United States [22]. The analysis of time segments in general supports the characterization that telemedicine is part of the new norm in health care delivery. Specifically, telemedicine-based urgent care services increased alongside decreases in in-person emergency room visits with each recurring pandemic wave; these data substantiate the critical role of telemedicine in expanding emergency care capacity and services during a period of significant emergency service strain. In the ambulatory care setting, our evidence suggests a delayed, but even more pronounced, shift to telemedicine. Overall, although urgent care visits opened the door for wider adoption of telemedicine during the pandemic, it is nonurgent video visits that are currently

driving the continued prevalence of telemedicine use. The analyses of visit trends from the time segments further suggest that although the use of telemedicine for both urgent care and ambulatory care services has gradually declined since its peak in the acute pandemic phase, the simultaneous decrease in variations in the monthly visit distribution further hints that telemedicine use is heading toward an equilibrium phase. Our result also suggests that during the latest Omicron wave, the use of telemedicine-based urgent care and ambulatory care demonstrated contradictory trends; although visits to urgent care for both COVID-19 and non-COVID-19 suspected cases increased, the use of ambulatory care decreased. Despite the massive increase in COVID-19 cases, why the use of ambulatory services decreased needs to be further investigated, and any potential barriers to wider adoption of virtual-ambulatory services need to be identified.

The trends in telemedicine visits suggest a proportionately larger role for urgent care facilities, particularly for COVID-19 care. Although the correlation between urgent care visits and COVID-19 confirmed cases was lower in the last wave of the pandemic (Omicron variant), we posit this result is due to the lower risk of severe outcomes from Omicron infection than in previous waves, especially the Delta variant [22,23]. With the emergence of new COVID variants [24], and the strong correlation between VUC visits and COVID-19 confirmed cases, the demand for VUC is not expected to decrease in the near future. More importantly, our observation of a new pattern of sustained demand for nonurgent, non-COVID-19-related telemedicine has enormous implications for health care delivery and equity. For patients, the high and steady level of satisfaction with virtual visits indicates their acceptance and willingness to persist with telemedicine services in the future. Whether broad reimbursement of virtual visits will continue [25] will be 1 of the factors determining the future of telemedicine as a mainstream mode of health care delivery in the United States. Analysis of patients' demographics shows telemedicine use, especially for urgent care, was more frequent among younger patients. Combined with high satisfaction among telemedicine-based urgent care users, this indicates their acceptance and willingness to persist with telemedicine services

in the future. Lower rates of telemedicine adoption among older adults may be due to their preference for emergency department visits, lower rates of technology adoption [26], and other reasons. Although recent reports suggest that smartphone adoption and internet use have more than doubled in the past 7 years among older adults [27], there remains a notable digital divide between younger and older Americans in telemedicine use. This divide is further skewed in the use of telemedicine for urgent care needs. Whether increased technology adoption will translate into telemedicine use uptake among seniors remains to be seen.

Furthermore, combined telemedicine and in-person visits increased by 18% from the pre-COVID-19 era (2019) to recent times (2021), and telemedicine was responsible (106%) for this increase, which suggests that virtual care delivery supplements rather than replaces in-person care. This may be a consequence of the enhanced access to care that telemedicine provides, allowing people with geographic, logistic, or other barriers to in-person care to more regularly access care. Telemedicine may be unlocking unmet needs of underserved patient populations and may potentially improve health equity and reduce health disparities if made accessible to inclusive populations. Although prior studies have found evidence that telemedicine access disparities mirror those in in-person health care access [16], whether telemedicine access disparities have reduced over time remains to be investigated. Nonetheless, evidence suggests health care organizations need to allocate additional resources to telemedicine, which should not come at the expense of in-person care. For providers, the transition means quickly developing and adjusting skills in virtual rapport building, empathy, diagnosis, and counseling.

### Comparison With Prior Work

To the best of our knowledge, this is 1 of the first studies to explore the longitudinal trends in telemedicine use throughout the pandemic. Other studies have explored various aspects of telemedicine during the COVID-19 pandemic, particularly its impressive expansion during the earliest phases. In 1 of the first case studies on telemedicine's early growth, Mann et al [8] and Sherwin et al [28] described the exponential growth of telemedicine visits within our health system during the first wave of the pandemic, outlining the health system's operational response as well. This work is complemented by a large volume of telemedicine-specific publications in 2020 and 2021, with the majority reporting on data and experiences from the early 2020 period (a review of PubMed literature on "telemedicine adoption" and "COVID-19" returns over 8000 papers, including case studies, opinion pieces, and reviews, from both US health systems and worldwide). Importantly, a number of articles on telemedicine during COVID-19 have called attention to new or growing disparities in the access and use of this technology, and its impact on health inequity [16,29-31]. A related systematic review on the use of digital health tools during the pandemic by Golinelli et al [31] revealed growth in the use of numerous digital health tools, including wearable devices, artificial intelligence (AI)-supported computing and clinical

decision support, blockchain technology, and the internet of things (IOT), largely for the purposes of diagnosis, managing, and monitoring COVID-19-related disease. Our findings contribute to this growing body of literature by expanding our understanding of the longitudinal patterns of telemedicine use and its potential sustainable impact on care delivery.

### Limitations

Although there are many strengths of this study, we note the following limitations that can be addressed in future research. First, we used keyword matching to identify COVID-19 suspected cases from the diagnosis data, and the list of keywords were limited to the most common COVID-19-related symptoms to minimize the number of false-positive identifications. Additionally, most keywords were related to respiratory issues, which were the most common symptoms during the early waves of COVID-19 [32]. More recent studies have reported nonrespiratory symptoms of COVID-19 [33-36] that we incorporated, but we were unable to use a more accurate method, such as COVID-19 test results to evaluate how the recurring pandemic waves relate to telemedicine use. Additionally, satisfaction data were only available for VUC visits. Although we currently do not have similar systemwide patient satisfaction data for ambulatory care, recent reports from our maternal-fetal medicine practices suggest high satisfaction among patients who used telemedicine for nonurgent care, corroborating our findings [21]. Our data may not generalize to all contexts. For example, remote and rural patient populations were not well represented. Finally, with respect to demographics, we reported telemedicine usage by age group without correcting for the baseline proportion of the population in each age group, which may not be evenly distributed. In addition, this study did not consider any demographics other than age when evaluating patient populations that are telemedicine users. Prior studies have reported evidence of disparities for Black, male patients when accessing telemedicine [16]. Future research should consider race, gender, socioeconomic status, and geographic location when evaluating the demographics of telemedicine adopters.

### Conclusion

In conclusion, data show that the transition to telemedicine care in major health care systems prompted by the early phases of the pandemic has been sustained throughout the later phases of the pandemic [37]. This has been driven by a variety of telemedicine care seeking, including urgent care, primary care, and ambulatory specialty care, as well as both COVID-19-related and non-COVID-19-related complaints. Those most likely to use telemedicine are younger patients, with patients reporting high levels of satisfaction with telemedicine-based services. Overall, this suggests that telemedicine-based care has high acceptability for patients and potential sustainability as an important modality of care delivery. More research is needed to understand patterns of telemedicine use across different types of health systems, patients, and health concerns, as well as addressing ongoing challenges in telemedicine access and equity.

## Acknowledgments

The authors would like to thank Dr. Simon A. Jones for his thoughtful review and assistance in manuscript preparation. This work was supported by National Science Foundation (NSF) grants (nos. 1928614 and 2129076).

## Data Availability

The primary data on visit records used for analysis came from the New York University (NYU) Langone Health (NYULH) system's electronic health record (EHR) patient data containing protected health information (PHI) and cannot be made public. We included tabularized visit data per month in [Multimedia Appendix 2](#). In addition, a partial de-identified version of the visit-level data might be available from Cosmos, Epic's de-identified Health Insurance Portability and Accountability Act (HIPAA)-limited patient database. The data on COVID-19 numbers in New York City (NYC) are available as open source in Lau et al [9].

## Authors' Contributions

All authors made substantial contributions to conception and design, acquisition of data, or analysis and interpretation of data; were involved in drafting the manuscript or revising it critically for important intellectual content; gave final approval of the version published; and agreed to be accountable for all aspects of the work.

## Conflicts of Interest

None declared.

### Multimedia Appendix 1

Overall trends in telemedicine and in-person visits in the 26-month period.

[[PNG File , 25 KB - formative\\_v6i10e38661\\_app1.png](#) ]

### Multimedia Appendix 2

Data on telehealth and in-person visits for COVID-19 suspected and non-COVID-19 cases between January 2020 and February 2022.

[[DOCX File , 21 KB - formative\\_v6i10e38661\\_app2.docx](#) ]

## References

1. AmericanWell. Amwell Telehealth Index Consumer Survey 2019. 2019. URL: <https://static.americanwell.com/app/uploads/2019/07/American-Well-Telehealth-Index-2019-Consumer-Survey-eBook2.pdf> [accessed 2020-03-30]
2. Dorsey ER, Topol EJ. State of telehealth. *N Engl J Med* 2016 Jul 14;375(2):154-161. [doi: [10.1056/nejmra1601705](https://doi.org/10.1056/nejmra1601705)]
3. Wade VA, Elliott JA, Hiller JE. Clinician acceptance is the key factor for sustainable telehealth services. *Qual Health Res* 2014 May 31;24(5):682-694. [doi: [10.1177/1049732314528809](https://doi.org/10.1177/1049732314528809)] [Medline: [24685708](https://pubmed.ncbi.nlm.nih.gov/24685708/)]
4. Greenhalgh T, Robert G, Macfarlane F, Bate P, Kyriakidou O. Diffusion of innovations in service organizations: systematic review and recommendations. *Milbank Q* 2004 Dec;82(4):581-629. [doi: [10.1111/j.0887-378x.2004.00325.x](https://doi.org/10.1111/j.0887-378x.2004.00325.x)]
5. Bentley CL, Powell LA, Orrell A, Mountain GA. Addressing design and suitability barriers to telecare use: has anything changed? *Technol Disabil* 2015 Mar 13;26(4):221-235. [doi: [10.3233/tad-150421](https://doi.org/10.3233/tad-150421)]
6. Clark J, McGee - Lennon M. A stakeholder - centred exploration of the current barriers to the uptake of home care technology in the UK. *J Assist Technol* 2011 Mar 18;5(1):12-25. [doi: [10.5042/jat.2011.0097](https://doi.org/10.5042/jat.2011.0097)]
7. Centers for Disease Control and Prevention (CDC). CDC Museum COVID-19 Timeline. URL: <https://www.cdc.gov/museum/timeline/covid19.html> [accessed 2022-06-10]
8. Mann D, Chen J, Chunara R, Testa P, Nov O. COVID-19 transforms health care through telemedicine: evidence from the field. *J Am Med Assoc* 2020 Jul 01;323(7):1132-1135 [FREE Full text] [doi: [10.1093/jamia/ocaa072](https://doi.org/10.1093/jamia/ocaa072)] [Medline: [32324855](https://pubmed.ncbi.nlm.nih.gov/32324855/)]
9. Lau J, Knudsen J, Jackson H, Wallach AB, Bouton M, Natsui S, et al. Staying connected in the covid-19 pandemic: telehealth at the largest safety-net system in the United States. *Health Aff (Millwood)* 2020 Aug 01;39(8):1437-1442. [doi: [10.1377/hlthaff.2020.00903](https://doi.org/10.1377/hlthaff.2020.00903)] [Medline: [32525705](https://pubmed.ncbi.nlm.nih.gov/32525705/)]
10. J.D. Power. U.S. Telehealth Satisfaction Study. URL: <https://www.jdpower.com/business/healthcare/us-telehealth-satisfaction-study> [accessed 2022-09-16]
11. Grin J, Rotmans J, Schot J. *Transitions to Sustainable Development: New Directions in the Study of Long Term Transformative Change*. New York, NY: Routledge; 2010.
12. Greenhalgh T, Wherton J, Papoutsis 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/)]



13. Parikh S, Davoudi A, Yu S, Giraldo C, Schriver E, Mowery D. Lexicon development for covid-19-related concepts using open-source word embedding sources: an intrinsic and extrinsic evaluation. *JMIR Med Inform* 2021 Feb 22;9(2):e21679 [FREE Full text] [doi: [10.2196/21679](https://doi.org/10.2196/21679)] [Medline: [33544689](https://pubmed.ncbi.nlm.nih.gov/33544689/)]
14. Wang L, Foer D, MacPhaul E, Lo Y, Bates DW, Zhou L. PASClex: a comprehensive post-acute sequelae of COVID-19 (PASC) symptom lexicon derived from electronic health record clinical notes. *J Biomed Inform* 2022 Jan;125:103951 [FREE Full text] [doi: [10.1016/j.jbi.2021.103951](https://doi.org/10.1016/j.jbi.2021.103951)] [Medline: [34785382](https://pubmed.ncbi.nlm.nih.gov/34785382/)]
15. 2020 ICD-10-CM. URL: <https://www.cms.gov/Medicare/Coding/ICD10/2020-ICD-10-CM> [accessed 2022-09-16]
16. Chunara R, Zhao Y, Chen J, Lawrence K, Testa P, Nov O, et al. Telemedicine and healthcare disparities: a cohort study in a large healthcare system in New York City during COVID-19. *J Am Med Inform Assoc* 2021 Jan 15;28(1):33-41 [FREE Full text] [doi: [10.1093/jamia/ocaa217](https://doi.org/10.1093/jamia/ocaa217)] [Medline: [32866264](https://pubmed.ncbi.nlm.nih.gov/32866264/)]
17. Truong C, Oudre L, Vayatis N. Selective review of offline change point detection methods. *Signal Process* 2020 Feb;167:107299. [doi: [10.1016/j.sigpro.2019.107299](https://doi.org/10.1016/j.sigpro.2019.107299)]
18. Cho H, Fryzlewicz P. Multiple-change-point detection for high dimensional time series via sparsified binary segmentation. *J R Stat Soc B* 2014 Jul 24;77(2):475-507. [doi: [10.1111/rssb.12079](https://doi.org/10.1111/rssb.12079)]
19. United States Census Bureau. Age and Sex. URL: <https://www.census.gov/topics/population/age-and-sex.html> [accessed 2022-09-16]
20. NYC. COVID-19: Data. URL: <https://www1.nyc.gov/site/doh/covid/covid-19-data.page> [accessed 2022-09-16]
21. Tozour JN, Bandremer S, Patberg E, Zavala J, Akerman M, Chavez M, et al. Application of telemedicine video visits in a maternal-fetal medicine practice at the epicenter of the COVID-19 pandemic. *Am J Obstet Gynecol MFM* 2021 Nov;3(6):100469 [FREE Full text] [doi: [10.1016/j.ajogmf.2021.100469](https://doi.org/10.1016/j.ajogmf.2021.100469)] [Medline: [34450341](https://pubmed.ncbi.nlm.nih.gov/34450341/)]
22. Duong BV, Larpruenrudee P, Fang T, Hossain SI, Saha SC, Gu Y, et al. Is the SARS CoV-2 omicron variant deadlier and more transmissible than delta variant? *Int J Environ Res Public Health* 2022 Apr 11;19(8):4586 [FREE Full text] [doi: [10.3390/ijerph19084586](https://doi.org/10.3390/ijerph19084586)] [Medline: [35457468](https://pubmed.ncbi.nlm.nih.gov/35457468/)]
23. Nyberg T, Ferguson NM, Nash SG, Webster HH, Flaxman S, Andrews N, et al. Comparative analysis of the risks of hospitalisation and death associated with SARS-CoV-2 omicron (B.1.1.529) and delta (B.1.617.2) variants in England: a cohort study. *Lancet* 2022 Apr;399(10332):1303-1312. [doi: [10.1016/s0140-6736\(22\)00462-7](https://doi.org/10.1016/s0140-6736(22)00462-7)]
24. Karim SSA, Karim QA. Omicron SARS-CoV-2 variant: a new chapter in the COVID-19 pandemic. *Lancet* 2021 Dec;398(10317):2126-2128. [doi: [10.1016/s0140-6736\(21\)02758-6](https://doi.org/10.1016/s0140-6736(21)02758-6)]
25. Medicare Telemedicine Health Care Provider Fact Sheet. URL: <https://www.cms.gov/newsroom/fact-sheets/medicare-tele-medicine-health-care-provider-fact-sheet> [accessed 2022-09-16]
26. Pew Research Center. Tech Adoption Climbs Among Older Adults. URL: <https://www.pewresearch.org/internet/2017/05/17/tech-adoption-climbs-among-older-adults/> [accessed 2021-12-08]
27. Pew Research Center. Americans 60 and Older Are Spending More Time in Front of Their Screens than a Decade Ago. URL: <https://www.pewresearch.org/fact-tank/2019/06/18/americans-60-and-older-are-spending-more-time-in-front-of-their-screens-than-a-decade-ago/> [accessed 2021-12-08]
28. Sherwin J, Lawrence K, Gragnano V, Testa PA. Scaling virtual health at the epicentre of coronavirus disease 2019: a case study from NYU Langone Health. *J Telemed Telecare* 2020 Jul 19;28(3):224-229. [doi: [10.1177/1357633x20941395](https://doi.org/10.1177/1357633x20941395)]
29. Katzow M, Steinway C, Jan S. Telemedicine and health disparities during covid-19. *Pediatrics* 2020 Aug;146(2):e20201586. [doi: [10.1542/peds.2020-1586](https://doi.org/10.1542/peds.2020-1586)] [Medline: [32747592](https://pubmed.ncbi.nlm.nih.gov/32747592/)]
30. Ortega G, Rodriguez JA, Maurer LR, Witt EE, Perez N, Reich A, et al. Telemedicine, COVID-19, and disparities: policy implications. *Health Policy Technol* 2020 Sep;9(3):368-371 [FREE Full text] [doi: [10.1016/j.hlpt.2020.08.001](https://doi.org/10.1016/j.hlpt.2020.08.001)] [Medline: [32837888](https://pubmed.ncbi.nlm.nih.gov/32837888/)]
31. Golinelli D, Boetto E, Carullo G, Nuzzolese AG, Landini MP, Fantini MP. Adoption of digital technologies in health care during the covid-19 pandemic: systematic review of early scientific literature. *J Med Internet Res* 2020 Nov 06;22(11):e22280 [FREE Full text] [doi: [10.2196/22280](https://doi.org/10.2196/22280)] [Medline: [33079693](https://pubmed.ncbi.nlm.nih.gov/33079693/)]
32. Yuki K, Fujiogi M, Koutsogiannaki S. COVID-19 pathophysiology: a review. *Clin Immunol* 2020 Jun;215:108427 [FREE Full text] [doi: [10.1016/j.clim.2020.108427](https://doi.org/10.1016/j.clim.2020.108427)] [Medline: [32325252](https://pubmed.ncbi.nlm.nih.gov/32325252/)]
33. Sobh E, Abuarrh E, Abdelsalam KG, Awad SS, Badawy MA, Fathelbab MA, et al. Novel coronavirus disease 2019 (COVID-19) non-respiratory involvement. *Egypt J Bronchol* 2020 Oct 01;14(1):1-6. [doi: [10.1186/s43168-020-00030-1](https://doi.org/10.1186/s43168-020-00030-1)]
34. AlSamman M, Caggiula A, Ganguli S, Misak M, Pourmand A. Non-respiratory presentations of COVID-19, a clinical review. *Am J Emerg Med* 2020 Nov;38(11):2444-2454 [FREE Full text] [doi: [10.1016/j.ajem.2020.09.054](https://doi.org/10.1016/j.ajem.2020.09.054)] [Medline: [33039218](https://pubmed.ncbi.nlm.nih.gov/33039218/)]
35. Zhou A, Song Q, Peng Y, Liao X, Huang P, Liu W, et al. Symptoms at disease onset predict prognosis in COVID-19 disease. *Libyan J Med* 2022 Dec 21;17(1):2010338 [FREE Full text] [doi: [10.1080/19932820.2021.2010338](https://doi.org/10.1080/19932820.2021.2010338)] [Medline: [34931582](https://pubmed.ncbi.nlm.nih.gov/34931582/)]
36. Bello-Chavolla O, Antonio-Villa N, Vargas-Vázquez A, Fermín-Martínez CA, Márquez-Salinas A, Bahena-López JP. Profiling cases with nonrespiratory symptoms and asymptomatic severe acute respiratory syndrome coronavirus 2 infections in Mexico City. *Clin Infect Dis* 2021 May 18;72(10):e655-e658 [FREE Full text] [doi: [10.1093/cid/ciaa1288](https://doi.org/10.1093/cid/ciaa1288)] [Medline: [32857829](https://pubmed.ncbi.nlm.nih.gov/32857829/)]

37. Lawrence K, Nov O, Mann D, Mandal S, Iturrate E, Wiesenfeld B. The impact of telemedicine on physicians' after-hours electronic health record "work outside work" during the covid-19 pandemic: retrospective cohort study. *JMIR Med Inform* 2022 Jul 28;10(7):e34826 [FREE Full text] [doi: [10.2196/34826](https://doi.org/10.2196/34826)] [Medline: [35749661](https://pubmed.ncbi.nlm.nih.gov/35749661/)]

## Abbreviations

**EHR:** electronic health record

**ICD-10:** International Classification of Diseases 10th Revision

**NYC:** New York City

**NYU:** New York University

**NYULH:** New York University Langone Health

**VUC:** virtual urgent care

*Edited by A Mavragani; submitted 11.04.22; peer-reviewed by E Borycki, K Adapa, B Wolfe-Piggott; comments to author 06.05.22; revised version received 14.06.22; accepted 02.09.22; published 14.10.22.*

*Please cite as:*

*Mandal S, Wiesenfeld BM, Mann D, Lawrence K, Chunara R, Testa P, Nov O*

*Evidence for Telemedicine's Ongoing Transformation of Health Care Delivery Since the Onset of COVID-19: Retrospective Observational Study*

*JMIR Form Res* 2022;6(10):e38661

URL: <https://formative.jmir.org/2022/10/e38661>

doi: [10.2196/38661](https://doi.org/10.2196/38661)

PMID: [36103553](https://pubmed.ncbi.nlm.nih.gov/36103553/)

©Soumik Mandal, Batia M Wiesenfeld, Devin Mann, Katharine Lawrence, Rumi Chunara, Paul Testa, Oded Nov. Originally published in *JMIR Formative Research* (<https://formative.jmir.org>), 14.10.2022. 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 the COVID-19 Pandemic on the Implementation of Mobile Health to Improve the Uptake of Hydroxyurea in Patients With Sickle Cell Disease: Mixed Methods Study

Sherif M Badawy<sup>1,2\*</sup>, MD, MBBCh, MS; Lisa DiMartino<sup>3\*</sup>, MPH, PhD; Donald Brambilla<sup>3</sup>, PhD; Lisa Klesges<sup>4</sup>, PhD; Ana Baumann<sup>4</sup>, PhD; Ebony Burns<sup>5</sup>, MA; Terri DeMartino<sup>5</sup>, BA; Sara Jacobs<sup>3</sup>, PhD; Hamda Khan<sup>6</sup>, MA; Chinonyelum Nwosu<sup>6</sup>, MPH; Nirmish Shah<sup>5</sup>, MD, MS; Jane S Hankins<sup>6,7</sup>, MD, MS; Sickle Cell Disease Implementation Consortium<sup>8\*</sup>

<sup>1</sup>Division of Hematology, Oncology and Stem Cell Transplant, Ann & Robert H. Lurie Children's Hospital of Chicago, Chicago, IL, United States

<sup>2</sup>Department of Pediatrics, Northwestern University Feinberg School Medicine, Chicago, IL, United States

<sup>3</sup>RTI International, Research Triangle Park, NC, United States

<sup>4</sup>Department of Surgery, Washington University in St. Louis, St. Louis, MO, United States

<sup>5</sup>Division of Hematology, Duke University School of Medicine, Durham, NC, United States

<sup>6</sup>Department of Hematology, St. Jude Children's Research Hospital, Memphis, TN, United States

<sup>7</sup>University of Tennessee Health Science Center, Memphis, TN, United States

<sup>8</sup>See Acknowledgments

\* these authors contributed equally

**Corresponding Author:**

Jane S Hankins, MD, MS

Department of Hematology

St. Jude Children's Research Hospital

262 Danny Thomas Place, Mail Stop 800

Memphis, TN, 38105

United States

Phone: 1 901 5954153

Fax: 1 901 5952952

Email: [jane.hankins@stjude.org](mailto:jane.hankins@stjude.org)

## Abstract

**Background:** Hydroxyurea therapy is effective for reducing complications related to sickle cell disease (SCD) and is recommended by National Health Lung and Blood Institute care guidelines. However, hydroxyurea is underutilized, and adherence is suboptimal. We wanted to test a multilevel mobile health (mHealth) intervention to increase hydroxyurea adherence among patients and improve prescribing among providers in a multicenter clinical trial. In the first 2 study sites, participants were exposed to the early phases of the COVID-19 pandemic, which included disruption to their regular SCD care.

**Objective:** We aimed to describe the impact of the COVID-19 pandemic on the implementation of an mHealth behavioral intervention for improving hydroxyurea adherence among patients with SCD.

**Methods:** The first 2 sites initiated enrollment 3 months prior to the start of the pandemic (November 2019 to March 2020). During implementation, site A clinics shut down for 2 months and site B clinics shut down for 9 months. We used the reach, effectiveness, adoption, implementation, and maintenance (RE-AIM) framework to evaluate the implementation and effectiveness of the intervention. mHealth implementation was assessed based on patients' daily app use. Adherence to hydroxyurea was calculated as the proportion of days covered (PDC) from prescription records over the first 12 and 24 weeks after implementation. A linear model examined the relationship between app usage and PDC change, adjusting for baseline PDC, lockdown duration, and site. We conducted semistructured interviews with patients, health care providers, administrators, and research staff to identify factors associated with mHealth implementation and effectiveness. We used a mixed methods approach to investigate the convergence of qualitative and quantitative findings.

**Results:** The percentage of patients accessing the app decreased after March 15, 2020 from 86% (n=55) to 70% (n=45). The overall mean PDC increase from baseline to week 12 was 4.5% (P=.32) and to week 24 was 1.5% (P=.70). The mean PDC change was greater at site A (12 weeks: 20.9%; P=.003; 24 weeks: 16.7%; P=.01) than site B (12 weeks: -8.2%; P=.14; 24 weeks:

–10.3%;  $P=.02$ ). After adjustment, PDC change was 13.8% greater in those with increased app use after March 15, 2020. Interview findings indicated that site B's closure during COVID-19 had a greater impact, but almost all patients reported that the *InCharge Health* app helped support more consistent medication use.

**Conclusions:** We found significant impacts of the early clinic lockdowns, which reduced implementation of the mHealth intervention and led to reduced patient adherence to hydroxyurea. However, disruptions were lower among participants who experienced shorter clinic lockdowns and were associated with higher hydroxyurea adherence. Investigation of added strategies to mitigate the effects of care interruptions during major emergencies (eg, patient coaching and health navigation) may “insulate” the implementation of interventions to increase medication adherence.

**Trial Registration:** ClinicalTrials.gov NCT04080167; <https://clinicaltrials.gov/ct2/show/NCT04080167>

**International Registered Report Identifier (IRRID):** RR2-10.2196/16319

(*JMIR Form Res* 2022;6(10):e41415) doi:[10.2196/41415](https://doi.org/10.2196/41415)

## KEYWORDS

sickle cell anemia; implementation science; adherence; hydroxycarbamide; self-efficacy

## Introduction

Sickle cell disease (SCD) is a chronic blood disorder in which acute painful acute events occur on the background of progressive organ dysfunction, leading to premature mortality [1]. SCD disproportionately affects low-income Americans who face access barriers to evidence-based treatments [2]. Hydroxyurea therapy is effective in reducing SCD-related complications, including acute pain episodes, and is recommended by National Health Lung and Blood Institute (NHLBI) care guidelines [1,3]. However, hydroxyurea is underutilized and adherence is suboptimal because providers often have a limited understanding of the optimal use of hydroxyurea [4,5], and many patients lack motivation and general knowledge about hydroxyurea and fear complications or side effects [6,7]. These barriers further increase health disparities in the care of the SCD population.

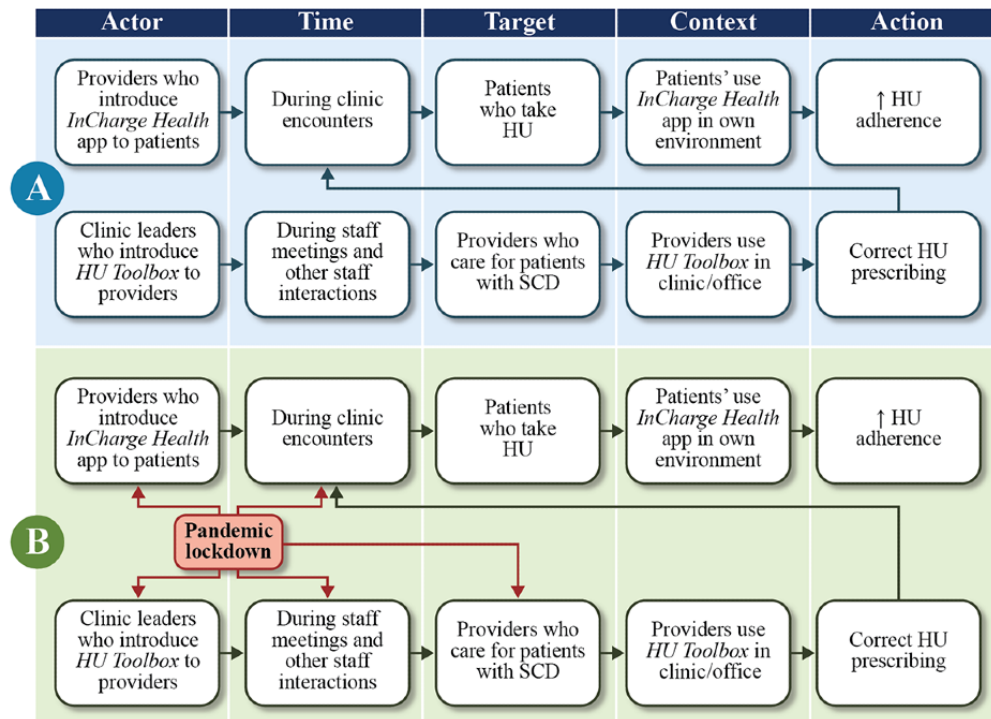
Mobile health (mHealth) apps, for both patients and providers, can be used as a strategy to incorporate behavioral change interventions that can potentially improve medication adoption and effectiveness [8-10]. Two months prior to when the COVID-19 pandemic reached US soil, we initiated a multicenter study as part of the NHLBI-funded Sickle Cell Disease Implementation Consortium (SCDIC) [11,12] to investigate the effectiveness and measure the implementation outcomes of a 2-level intervention using mHealth to support hydroxyurea adherence among patients (*InChargeHealth* app) [13] and hydroxyurea prescribing among providers (*HU Toolbox* app). This multilevel strategy focused on increasing hydroxyurea use, by targeting the determinants involved in medication adherence (ie, motivation, knowledge, self-efficacy, and social support) among patients (*InCharge Health* app) and those involved in appropriate prescribing (eg, knowledge, attitude, and self-efficacy) among providers (*HU Toolbox* app). To facilitate the description and identification of the characteristics of the mHealth intervention, we specified them according to the action, actor, context, target, and time (AACTT) framework [14].

Providers (actor) introduce the *InCharge Health* app during clinic encounters (time) to patients (target), who then use the app in their own environment (context) to improve hydroxyurea adherence behavior (action) (Figure 1A). Clinic leaders (actor) introduce the *HU Toolbox* app to providers who care for patients with SCD (target) during clinic staff interactions (time), who then use the toolbox in their offices or clinics (context) to improve correct hydroxyurea prescribing behavior (action) (Figure 1A). In clinical practice, as providers prescribe and counsel patients on the benefits of hydroxyurea during regular visits, providers are both the actors and targets of our multilevel intervention.

The COVID-19 pandemic has disrupted health care access in unprecedented ways, including reducing patient-provider contact during health maintenance care and reducing medication adherence for chronic diseases [15-17]. In particular, many SCD patients are often seen in outpatient clinics every 2 to 4 months if they are on a disease-modifying therapy, such as hydroxyurea, or every 6 months if they are not. This close monitoring of their clinical condition and treatment plan was deeply affected by the COVID-19 pandemic. Additionally, patients with SCD had worse outcomes from COVID-19 infection compared with race-matched individuals without SCD [18]. Leveraging mHealth interventions to facilitate health care delivery, including the use of evidence-based hydroxyurea, is underscored by the COVID-19 pandemic [15]. However, how the pandemic may affect mHealth use for medication adherence in SCD has not been investigated and remains unclear. Hydroxyurea use can be potentially amplified by mHealth, but disruptions during the COVID-19 pandemic may threaten its implementation.

Because the COVID-19 pandemic lockdown restrictions disrupted the care of the study participants, we sought to evaluate how the implementation and preliminary effectiveness of the patient *InCharge Health* app were affected. We hope our lessons learned will inform future studies disrupted by unplanned emergencies, such as pandemics, by anticipating possible required study adaptations.

**Figure 1.** Implementation intervention specification. Specification is done according to the action, actor, context, target, and time (AACTT) framework [14]. (A) Flow before the COVID-19 pandemic. (B) After the start of the COVID-19 pandemic, clinic lockdown measures were put in place, which led to reductions in leadership-staff interactions and patient-provider interactions. Introduction of the respective apps and hydroxyurea (HU) prescribing were, thereby, reduced, leading to lower HU adherence among patients. SCD: sickle cell disease.



## Methods

### Study Setting and Participants

In this report, we describe the results of the implementation of the *InCharge Health* app in the first 2 (out of 7) SCDIC study sites, to represent the early impact of the pandemic response. The study has been registered at ClinicalTrials.gov (NCT04080167). Both sites were in the south region of the United States, and they initiated study enrollment 3 months prior to the start of the pandemic (November 2019 to March 2020) as part of a planned staggered intervention design [11]. As a response to the COVID-19 pandemic and to reduce virus spread, both sites temporarily suspended nonemergent in-person clinical activities. During implementation, site A clinics closed for 2 months (March 15, 2020, to May 15, 2020) and quickly initiated telemedicine visits, which were already standard practice at this institution, but were scaled up during the pandemic. Site B clinics were closed for 9 months (March 15, 2020, to December 15, 2020) and did not initiate telehealth visits until year 2 of the pandemic due to delays in training staff and distributing equipment. The 5 remaining sites started enrollment later in the pandemic when the clinic lockdowns were no longer in effect, and therefore, their results are not being reported here.

Patient participants were individuals with a diagnosis of SCD between the ages of 15 and 45 years treated with hydroxyurea and receiving care at the 2 initial participating sites [12]. Patients with SCD receiving chronic blood transfusions or using another mHealth modality for medication adherence were not eligible. To best represent the impact of both site's pandemic response

lockdown, only patient participants who enrolled prior to March 15, 2020, and were followed up until after March 15, 2020, were included in this analysis, allowing the contrast of their study behavior before and after the start of the pandemic. The date of enrollment varied among subjects, but follow-up time was the same for all subjects in this analysis (24 weeks). Therefore, the time over which subjects were exposed to the shutdown varied, a point that is considered in more detail below. Provider participants were physicians and advance care practitioners (nurse practitioners and physician assistants) who cared for at least 1 patient with SCD within the participating sites.

### Ethics Approval

This study was approved by the Institutional Review Board at St Jude Children's Research Hospital (19-0159) and Duke University (Pro00073506), and all participants (or their legal guardians) signed consent prior to study participation.

### Study Design and Key Measures

The methods of the study have been published [11]. Briefly, all patients enrolled were asked to download the *InCharge Health* app on their cell phones and use it for at least 24 weeks. Providers were asked to download the *HU Toolbox* app and use it for at least 9 months. Study visits occurred 12 and 24 weeks after enrollment, and data regarding app use and hydroxyurea refills were collected.

As part of the planned approach for the multisite study, we used the reach, effectiveness, adoption, implementation, and maintenance (RE-AIM) framework to inform the evaluation of the implementation and effectiveness of the *InCharge Health*

app [19]. mHealth implementation was assessed based on user mHealth engagement, which was classified according to patients' daily app use (number of days accessed during study participation). App use was categorized as follows: low (<25% days/month), medium-low (25% - 49% days/month), medium-high (50%-74% days/month), and high (75%-100% days/month) [11].

Patient-level data on hydroxyurea adherence provided a measure of effectiveness. Hydroxyurea adherence was measured by calculating the proportion of days covered (PDC), which is the ratio of the number of days the patient is covered by the medication to the number of days in the treatment period. In other words, the PDC is the number of days covered by prescriptions that were filled divided by the length of the study interval. If a prescription was filled just before the start of the study interval, the days between the prescription fill date and start of the interval were excluded. If a prescription was filled near the end of the study interval, the part of the interval covered by the prescription that was after the end of the study interval was also excluded. The PDC was calculated over a 24-week baseline interval and over the first 12 and first 24 weeks after implementation. Provider app use was classified as low ( $\leq 1$  time on average monthly during the study) and high ( $> 1$  time on average monthly during the study).

### Quantitative Analysis

We compared PDC change over 12 and 24 weeks after mHealth implementation using 1-sample  $t$  tests. The relation between the amount of app usage and PDC change was calculated using linear models of PDC change as a function of app use, treating app use as continuous in some models and as a 4-level categorical variable in others. The choice of linear models was based on prior experience showing that linear models were appropriate for analyzing changes in the PDC. To ensure the appropriateness of this method, we verified the fit of the models by examining the distributions of residuals and the relationships between residuals and predicted values. Site was included in some models to allow for variation among sites in changes in the PDC that might have been induced by site-to-site variations in the responses to the pandemic. By definition, the PDC is  $\geq 0\%$  and  $\leq 100\%$ . It is very common for change in a bounded measure, such as PDC, to be negatively correlated with the baseline value. Baseline PDC was therefore included in most models to avoid confounding between baseline PDC and other predictors of interest. The results are presented as the differences between changes in PDC at the 2 sites, adjusted for baseline PDC. While power calculations are not provided for this study, the large ongoing trial power calculation has been published [11]. The reach of the *InCharge Health* app was considered as the proportion of eligible patients who were enrolled in the study, and of those enrolled, the proportion who downloaded and used the app at least once. Adoption of *HU Toolbox* was considered as the proportion of eligible providers who were enrolled, and of those enrolled, the proportion who downloaded and used the app at least once. App use was measured as the proportion of follow-up days on which the app was accessed at least once per day. Use was calculated over the entire follow-up for each patient participant in some analyses and separately for the periods from enrollment up through March 15, 2020, and after

March 15, 2020, in others. Given that the length of follow-up was the same for all subjects, the number of days from enrollment through March 15, 2020, was a measure of the proportion of the follow-up interval the subjects experienced before March 15, 2020 (ie, the start of the lockdown period). Logistic regression was employed to identify the predictors of increased app use after March 15, 2020, in those with follow-up time both before and after that date.

The association between the use of the *InCharge Health* app and PDC change from baseline through follow-up was examined in linear models of PDC change as a function of baseline PDC, time from March 15, 2020, to the end of each subject's follow-up, an indicator for site, and app use during the follow-up interval. The following 2 measures of app use were considered in separate sets of models: (1) the number of follow-up days on which the app was accessed at least once and (2) a binary indicator for whether app use increased or decreased after March 15, 2020. Interaction terms were also considered, particularly interactions between site and the other predictors, to determine whether the effects of any predictors differed at the 2 sites.

### Qualitative Analysis

We used the RE-AIM framework to qualitatively identify factors that may have influenced mHealth implementation and effectiveness during the initial phases of the COVID-19 pandemic [20]. Semistructured interviews were conducted with patients to better understand the contextual factors associated with *InCharge Health* app implementation and effectiveness. Health care providers, administrators, and research staff were also interviewed to provide qualitative data regarding the impact of the COVID-19 pandemic on app use. Interviews were conducted by research coordinators between June 2020 and March 2021. Semistructured interview guides were developed using the RE-AIM framework to understand participants' engagement and experiences with the apps, and included several questions specific to COVID-19 (Multimedia Appendix 1) [20,21]. For example, participants were asked to describe how using the app changed the way they took hydroxyurea during COVID-19 (effectiveness) and how COVID-19 impacted use of the app (implementation). For interviews with research staff, questions were asked about the challenges related to COVID-19 encountered at the site. A purposive sample of participants was interviewed based on app use frequency. Data were transcribed and entered into NVivo 12.0 (QSR International) for qualitative data analysis. Data were coded and analyzed with the goal of achieving theme saturation. Results were grouped into themes and mapped to the effectiveness and implementation RE-AIM domains by 2 study team members. We used mixed methods to investigate the possible convergence of qualitative and quantitative findings, through data triangulation [22]. Through this analysis, we sought to corroborate and expand quantitative findings using qualitative data.

## Results

### Participant Characteristics and Study Participation

A total of 75 patients (out of 508 eligible) and 42 providers (out of 55 eligible) were enrolled between November 2019 and September 2020 in the first 2 participating SCDIC sites. To

characterize the influence of the COVID-19 pandemic shutdowns on app use, we only included 64 patients and all 42 providers enrolled prior to March 15, 2022. Among the patients, 28 were enrolled at site A and 36 at site B. Half (32/64, 50%)

were young adults (18-25 years), with almost even gender distribution (Table 1). Among the providers, most (29/42, 69%) were between 26 and 45 years of age, and majority were female (30/42, 71%) and physicians (24/42, 59%) (Table 1).

**Table 1.** Participant characteristics.

Characteristic	Patients, n (%)			Providers, n (%)		
	All patients (N=64)	Site A (N=28)	Site B (N=36)	All providers (N=42)	Site A (N=15)	Site B (N=27)
<b>Age range (years)</b>						
15-17	7 (11)	0 (0)	7 (19)	0 (0)	0 (0)	0 (0)
18-25	32 (50)	10 (36)	22 (61)	0 (0)	0 (0)	0 (0)
26-45	25 (39)	18 (64)	7 (19)	29 (69)	9 (60)	20 (74)
46-64	0 (0)	0 (0)	0 (0)	12 (29)	5 (33)	7 (26)
>65	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Missing	0 (0)	0 (0)	0 (0)	1 (2)	1 (7)	0 (0)
<b>Gender</b>						
Male	31 (48)	15 (54)	16 (44)	12 (29)	2 (13)	10 (37)
Female	33 (52)	13 (46)	20 (56)	30 (71)	13 (87)	17 (63)
<b>Race</b>						
Black	64 (100)	28 (100)	36 (100)	8 (19)	5 (33)	3 (11)
White	0 (0)	0 (0)	0 (0)	24 (59)	7 (47)	17 (63)
Asian	0 (0)	0 (0)	0 (0)	9 (22)	2 (13)	7 (26)
Missing	0 (0)	0 (0)	0 (0)	1 (2)	1 (7)	0 (0)
<b>Ethnicity</b>						
Not Hispanic	64 (100)	28 (100)	36 (100)	41 (98)	14 (93)	27 (100)
Hispanic	0 (0)	0 (0)	0 (0)	1 (2)	1 (7)	0 (0)
<b>Sickle cell disease genotype</b>						
HbSS/HbS $\beta^0$ -thalassemia	56 (88)	27 (96)	29 (81)	N/A <sup>a</sup>	N/A	N/A
HbSC/HbS $\beta^+$ -thalassemia/other	8 (12)	1 (4)	7 (19)	N/A	N/A	N/A
<b>Provider type</b>						
Physician	N/A	N/A	N/A	24 (59)	9 (60)	15 (56)
Nurse practitioner or physician assistant	N/A	N/A	N/A	17 (41)	6 (40)	11 (41)
Missing	N/A	N/A	N/A	1 (2)	0 (0)	1 (4)
<b>InCharge Health app use level<sup>b</sup></b>						
High	6 (9)	6 (21)	0 (0)	N/A	N/A	N/A
Medium-high	8 (13)	5 (18)	3 (8)	N/A	N/A	N/A
Medium-low	9 (14)	2 (7)	7 (19)	N/A	N/A	N/A
Low	41 (64)	15 (54)	26 (72)	N/A	N/A	N/A
<b>HU Toolbox app use level<sup>c,d</sup></b>						
High	N/A	N/A	N/A	21 (52)	8 (53)	13 (48)
Low	N/A	N/A	N/A	19 (48)	5 (33)	14 (52)

<sup>a</sup>N/A: not applicable.

<sup>b</sup>The *InCharge Health* app use level for patients was categorized based on the percentage of days used per month as follows: low, <25%; medium-low, 25%-50%; medium-high, 51%-74%; and high, 75%-100%.

<sup>c</sup>The *HU-Toolbox* app use level for providers was categorized as low (<1 app usage per month) and high ( $\geq 1$  app usage per month) over a 9-month period.

<sup>d</sup>Two providers were removed from the study (moved to a new institution or requested to be withdrawn).



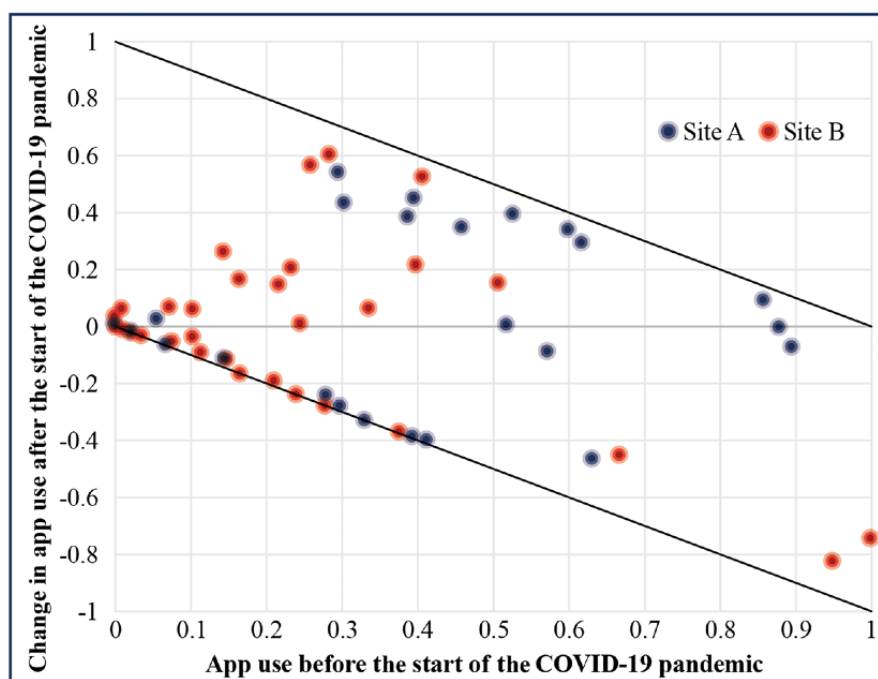
## App Use

All 64 patient participants downloaded the app, and 58 participants used it at least once during the 6-month study period, representing a 91% reach. On average, patients accessed the app on 42.7 (25.5%) days throughout the 6 months of the study period, and 24 (38%) of the 64 patients accessed it on  $\geq 25\%$  of the total days over 6 months. The percentage of participants accessing the app decreased after March 15, 2020, from 86% ( $n=55$ ) before that date to 70% ( $n=45$ ) after that date. However, the average change in app use was very close to 0 (mean change:  $-0.0016$ ;  $P=.96$ ), which means that reductions in use by some participants were balanced by increases in others. It is important to note, however, that there appeared to be 2 distinct subgroups at each site, one with increased app use after March 15, 2020, and the other with decreased app use after March 15, 2020 (Figure 2). A logistic regression of the probability of increasing app use after March 15, 2020, indicated that the probability decreased with increasing time between

March 15, 2020, and the end of follow-up, but did not differ between sites (OR  $-0.0276$ ;  $P=.004$ ). The odds ratio for increased app use was 0.67 when the time from March 15, 2020, to the end of follow-up increased by 14.7 days, and it decreased to 0.5 when the time increased to 25 days. Thus, longer exposure to the shutdown was associated with a reduction in app use. Other predictors were considered, including demographic variables, such as gender, age, income, and education, and measures of pain frequency, pain intensity, and recent use of hydroxyurea, but none made statistically significant contributions to the model for increased app use after March 15, 2020.

Of the 42 providers enrolled, 41 downloaded and used the *HU Toolbox* app at least once (adoption 98%). Overall app use among providers averaged around a day per month (1.1 days per month) prior to the pandemic, but declined to less than a day per month (0.2 days per month) during the lockdown.

**Figure 2.** Change in InCharge Health app use relative to the COVID-19 pandemic lockdown. March 15, 2020, corresponds to the date when both sites went on lockdown in response to the COVID-19 pandemic. The black diagonal lines represent the boundaries for the maximum that app use can change after March 15, 2020, given app use before March 15, 2020. Since app use is expressed as a proportion of days on which the app is accessed, app use must be  $\geq 0$  and  $\leq 1.0$ . As app use prior to March 15, 2020, increases, the maximum amount by which it can drop after March 15, 2020, increases, while the amount by which it can increase after March 15, 2020, decreases. For example, if app use is 0.25 (25% of days) before March 15, 2020, it can drop by a maximum of 0.25 or increase by a maximum of 0.75, whereas if app use is 0.75 (75% of days) before March 15, 2020, it can drop by a maximum of 0.75 or increase by a maximum of 0.25. There were 2 subgroups. The diagonal line of points along the lower black boundary line indicates the first subgroup consisting of participants whose app use dropped from some use to little or no use after March 15, 2020. On the other hand, the cloud of points from both sites above the line of zero change indicates the second subgroup consisting of patients whose app use increased after March 15, 2020.



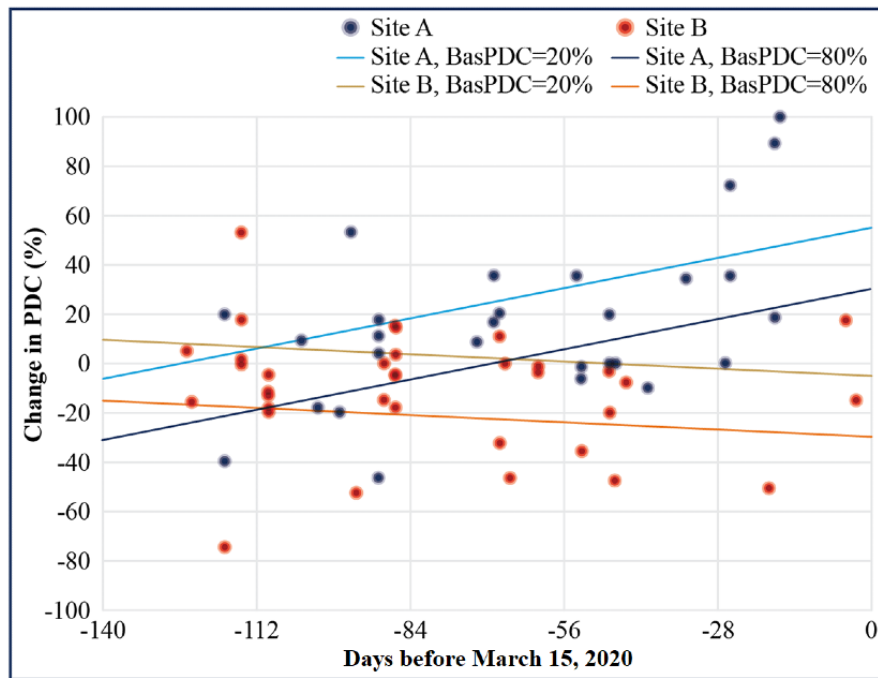
## Hydroxyurea Adherence

The mean increase in the PDC was 4.5% ( $P=.32$ ) on comparing the first 12 weeks of follow-up to the baseline interval and was 1.5% ( $P=.70$ ) on comparing 24 weeks of follow-up to the baseline interval. However, PDC changes differed between sites. Site A had a significant mean increase in the PDC (20.9% at 12 weeks;  $P=.003$  and 16.7% at 24 weeks;  $P=.01$ ). At site B, the mean PDC did not change significantly at 12 weeks ( $-8.2\%$ ;  $P=.14$ ) but declined over 24 weeks ( $-10.3\%$ ;  $P=.02$ ).

Additionally, changes in the PDC were negatively correlated with baseline PDC (24 weeks:  $r=-0.55$ ;  $P<.001$ ; 12 weeks:  $r=-0.52$ ;  $P<.001$ ), reflecting mainly positive PDC changes at lower baseline PDC and mainly negative changes at the highest baseline PDC. Importantly, PDC change, adjusted for baseline PDC, varied with the proportion of the follow-up interval that occurred after March 15, 2020, but did so differently at the 2 sites. The PDC change from baseline through follow-up increased as the proportion of follow-up days after March 15,

2020, at site A increased, but decreased with an increasing proportion after March 15, 2020, at site B (Figure 3).

**Figure 3.** Proportion of days covered (PDC) change at 24 weeks of follow-up. PDC increases were observed at site A and PDC decreases were observed at site B, but a lower baseline PDC was associated with a higher PDC change at 24 weeks at both sites. The duration of time from March 15, 2020, to the end of each participant’s follow-up was associated with greater PDC increases at site A (where the lockdown duration after March 15, 2020, was shorter) and greater decreases at site B (where the lockdown duration after March 15, 2020, was longer). BasPDC: baseline proportion of days covered.



### Medication Adherence Relative to App Use

App use measured by the number of follow-up days on which the app was used at least once was not statistically significant ( $P=.46$ ) when added to a model of PDC change that also included baseline PDC, time from March 15, 2020, to the end of each participant’s follow-up, site, and the interaction between site and time from March 15, 2020, to the end of follow-up. However, the indicator for increased app use after March 15, 2020, was statistically significant when it replaced app use over

the follow-up interval in this model (Table 2). After adjusting for the other predictors, PDC change was 13.8% greater in those with increased app use after March 15, 2020. In other words, those with increased app use after March 15, 2020, showed either a smaller drop in the PDC or a greater gain in the PDC (depending on baseline PDC, site, and days after March 15, 2020), while those with decreased app use showed a greater reduction in the PDC where the PDC declined in both groups, or a smaller increase where it increased in both groups.

**Table 2.** Linear model of the change in the proportion of days covered from baseline to 24 weeks of follow-up.

Parameter <sup>a</sup>	Estimate	SE	P value
Intercept	4.4493	11.0926	.69
App use increased after March 15, 2020	13.7584	6.1096	.03
App use decreased after March 15, 2020	0	N/A <sup>b</sup>	N/A
Baseline PDC <sup>c</sup>	-0.3928	0.0862	<.001
Days from enrollment through March 15, 2020	-0.0116	0.1247	.93
Site A	53.3618	14.3925	<.001
Site B	0	N/A	N/A
Days from enrollment through March 15, 2020, at site A	-0.4695	0.1803	.01
Days from enrollment through March 15, 2020, at site B	0	N/A	N/A

<sup>a</sup>Model variables included baseline proportion of days covered, site, time from March 15, 2020, to the end of each participant’s follow-up, the interaction between site and time from March 15, 2020, to the end of follow-up, and an indicator for increased app use after March 15, 2020.

<sup>b</sup>N/A: not applicable.

<sup>c</sup>PDC: proportion of days covered.

## Qualitative Data

Eleven patients (mean age, 26.4 years; 64% [7/11] males; 100% [11/11] Black; 73% [8/11] HbSS; 45% [5/11] low app users) completed interviews across the 2 sites. Site B's closure during COVID-19 had a greater impact on patients, who had difficulty obtaining hydroxyurea and reaching their providers and the clinic for nonurgent or emergent reasons. One low user from site B stated:

*Before COVID-19, I could just call my clinic or doctor and ask if I could come in and it would be a 'yes', but now, its [COVID-19] cut down on the days the clinic is open and the time the clinic is open. It's harder to get in.*

However, almost all reported that the *InCharge Health* app helped support more consistent daily medication use ([Multimedia Appendix 2](#)). One high user from site A stated:

*I can appreciate it [the app]. It helped me. I think it's a good thing. I think it makes me better with my hydroxyurea.*

Consistent with patients, providers and administrators reported that clinic shutdowns during COVID-19 negatively impacted the ability to care for patients. For example, because fewer patients were coming to the clinic, there was a reduction in the need to use the *HU Toolbox* app as an aid for hydroxyurea prescribing ([Multimedia Appendix 3](#)). One provider from site B reported:

*It [COVID-19] definitely impacted [app use]. As follows, we were not coming to the clinic as often for at least two to three months. So, I didn't happen to think about the app or just didn't have an opportunity to use it.*

Research staff at both sites also reported that reduced in-person clinic visits was a barrier for implementing the study in general. One staff member stated:

*It has been quite difficult during the pandemic. It was easier for us when we were in person. We had that carved out time when [patients] weren't doing anything else, they were specifically focused on what we were doing.*

## Discussion

Hydroxyurea is an evidence-based therapy in SCD, with proven clinical benefits, but its uptake is low. In a multicenter NHLBI-funded study, we tested the use of mHealth to improve hydroxyurea use among adolescents and adults with SCD. At the first 2 study sites, participants were exposed to the early phases of the COVID-19 pandemic, which included disruption in their regular SCD care. While the ubiquitous access to mobile technology among patients with SCD represents a unique opportunity to leverage mHealth interventions to support clinical care, the contextual changes, such as those during global emergencies, can affect its implementation. Our study is the first to assess, among individuals with SCD, the impact of the COVID-19 pandemic on the implementation of an mHealth behavioral intervention aimed at improving medication

adherence. In the 2 clinical trial sites where study activities happened during the early phases of the pandemic, we found evidence of significant reductions in the implementation of the app relative to the duration of the clinic lockdown in response to the COVID-19 pandemic. While low baseline adherence levels predicted higher improvements in adherence, the pandemic disruptions also affected the adherence to hydroxyurea, which was proportionally reduced to the duration of the clinic lockdown. However, we also found evidence of the benefit of mHealth to improve adherence. Among patients whose mHealth use increased after the start of the lockdown, improvements in hydroxyurea adherence were also observed. Our findings highlight the influence of unplanned contextual changes on the implementation of mHealth behavioral interventions and the potential benefits of investing in strategies to sustain use. These data are key for the future implementation of mHealth behavioral interventions, for both patients and providers, in clinical settings during pandemics or other similar situations.

Earlier studies have demonstrated the potential efficacy of mHealth interventions for enhancing hydroxyurea adherence among patients with SCD [23,24]. It is worth noting that not receiving hydroxyurea, along with other factors, is predictive of mortality in SCD patients with COVID-19 infection [25], supporting the additional clinical benefits of hydroxyurea use, particularly during the pandemic. However, it is possible that the effect of mHealth may be mediated or moderated by ongoing contacts with health care providers, as demonstrated by the lower app use (and consequent lack of an effect for improving adherence) at site A, where disruptions in patient-provider contact were prolonged, and the higher app use at site B, where lockdown was shorter and telehealth was implemented quicker.

In our study, the usage of the *InCharge Health* app varied among patients. We also found different barriers to app implementation, especially those related to access to the health system, with fewer in-person clinic visits, and low contact between patients and providers, potentially reducing hydroxyurea adherence. Further, providers and administrators reported that the *HU Toolbox* app was not used often due to clinic lockdown, which may have contributed to reduced gains in the PDC. To conceptualize how the lockdown disruptions influenced mHealth use and consequently medication adherence, we described the influence of the lockdown on the different targets of the mHealth intervention ([Figure 1B](#)). This model is supported by our qualitative data that validate the disruptions in care leading to decreased patient-provider contact, decreased hydroxyurea prescribing, decreased patient and provider mHealth use, and decreased hydroxyurea adherence.

During the COVID-19 pandemic, reliance on telehealth exponentially increased, and for some chronic conditions, it not only facilitated care delivery but also improved health outcomes [26,27]. In our study, although mHealth may have supported adherence, this effect was moderated by the duration of the lockdown, which negatively impacted app use over time and consequently affected hydroxyurea adherence. The early use of telemedicine at site A might have helped support the use of the patient app, as it maintained patient-provider contact, and might have mitigated the clinical care disruptions. A full evaluation

of the impact of telemedicine is planned when the final results of the trial are available.

The finding that reduced in-person visits was a barrier to study implementation was not surprising, as the impact of COVID-19 on clinical trials is well recognized [28]. Various strategies were suggested to mitigate some of these effects, such as (1) remote enrollment and follow-up and completion of study procedures when possible; (2) prioritization of primary outcomes; (3) alternative approaches for outcome assessment; (4) obtaining three or more phone numbers and email addresses for patients and relatives or friends; and (5) using different ways to contact patients and families, including text messaging, phone calls, emails, or social media [28]. Although these strategies have been reported, we were unable to document their impact on study implementation, as use of telehealth, for instance, was limited in the first year of the pandemic at one of the sites.

Our study has limitations. Data included in this analysis were from 2 study sites with relatively small sample sizes, which limits the generalizability of our findings. However, they do reflect the impact of the early institutional responses to the COVID-19 pandemic, which also occurred in other health institutions worldwide. Additionally, the results presented are not representative of the full study results, as this study is currently ongoing. Additional data regarding the effectiveness of mHealth for hydroxyurea adherence is, therefore, forthcoming. We also were not able to conduct interviews with all SCD participants to better understand specific barriers to hydroxyurea adherence during the pandemic, but our interview sample was purposefully selected based on the participants' app use levels and achieved theme saturation. Because of the nature of the pandemic, we were not able to measure the mental health impact of the pandemic in the initial months and how it would have affected app use. Further, there are other possible variables beyond patient-level barriers or characteristics, such as system-level ones, that may have affected the implementation

of our app during COVID, including clinics being shut down and limited use of telehealth at one site versus the other. Finally, although the PDC is an indirect measure of adherence, it is considered reliable and reflective of real-world settings (as opposed to adherence measured during clinical trials), and it has been used in many published research studies on SCD and other chronic medical conditions [29].

In conclusion, mHealth apps are promising tools for improving hydroxyurea adherence among adolescents and adults with SCD. In this preliminary analysis, we found significant impacts of the early clinic lockdowns, which reduced the implementation of the mHealth intervention for increasing hydroxyurea uptake. This disruption led to reduced patient adherence to hydroxyurea. However, disruptions to mHealth implementation were lower among participants who experienced shorter clinic lockdowns and among those who increased mHealth use during the pandemic, and evidence of the benefit was provided by higher hydroxyurea adherence. In qualitative analysis, we found concordance between low app use and perceived barriers to obtaining care early on during the pandemic. Triangulation of our findings suggests the benefit of mHealth for improving medication adherence and indicates that its use may be influenced by frequent contact with health care providers. Patients' barriers to care access might have hindered app implementation, potentially reducing medication adherence. Investigation of added strategies to mitigate the effects of imposed care interruptions during major emergencies, particularly greater patient touchpoints (eg, patient coaching and health navigation), may "insulate" the implementation of interventions for increasing medication adherence. Future studies are essentially needed to better understand both patient- and system-level barriers in the context of pandemics or other similar situations. A focus on removing barriers to mHealth use during care disruptions will likely improve app implementation and medication adherence, ultimately reducing health inequities for vulnerable populations.

---

## Acknowledgments

The Sickle Cell Disease Implementation Consortium ([Multimedia Appendix 4](#)) has been supported by the US Federal Government cooperative agreements HL133948, HL133964, HL133990, HL133996, HL133994, HL133997, HL134004, HL134007, and HL134042 from the National Heart, Lung, and Blood Institute (NHLBI) and the National Institute on Minority Health and Health Disparities of the National Institutes of Health (NIH) (Bethesda, Maryland), and the Bridge grant from the American Society of Hematology. This project was also partially supported by a grant from the NHLBI of the NIH (K23HL150232, principal investigator: SMB) for the protected time to be able to contribute to this research and manuscript writing. The content of this manuscript is solely the responsibility of the authors and does not necessarily represent the views of the NIH.

---

## Authors' Contributions

JSH, NS, and LK designed the research study; LD and DB analyzed the data; SMB, LD, DB, AB, LK, NS, and JSH interpreted the data; SMB and LD drafted the paper; and DB, AB, EB, TDM, SJ, HK, LK, CN, NS, and JSH critically revised the paper. All authors approved the submitted final version of the paper.

---

## Conflicts of Interest

None declared.

---

## Multimedia Appendix 1

Interview questions to describe app effectiveness and implementation during COVID-19.

[[DOC File , 33 KB - formative\\_v6i10e41415\\_app1.doc](#) ]

#### Multimedia Appendix 2

Thematic analysis of patient experiences with hydroxyurea adherence and health care access during COVID-19.

[[DOCX File , 13 KB - formative\\_v6i10e41415\\_app2.docx](#) ]

#### Multimedia Appendix 3

Thematic analysis of provider, administrator, and research staff experiences with app implementation during COVID-19.

[[DOCX File , 14 KB - formative\\_v6i10e41415\\_app3.docx](#) ]

#### Multimedia Appendix 4

Sickle Cell Disease Implementation Consortium Members.

[[DOCX File , 27 KB - formative\\_v6i10e41415\\_app4.docx](#) ]

## References

1. Kato GJ, Piel FB, Reid CD, Gaston MH, Ohene-Frempong K, Krishnamurti L, et al. Sickle cell disease. *Nat Rev Dis Primers* 2018 Mar 15;4:18010. [doi: [10.1038/nrdp.2018.10](#)] [Medline: [29542687](#)]
2. Adams-Graves P, Bronte-Jordan L. Recent treatment guidelines for managing adult patients with sickle cell disease: challenges in access to care, social issues, and adherence. *Expert Rev Hematol* 2016 Jun;9(6):541-552. [doi: [10.1080/17474086.2016.1180242](#)] [Medline: [27098013](#)]
3. Yawn BP, Buchanan GR, Afenyi-Annan AN, Ballas SK, Hassell KL, James AH, et al. Management of sickle cell disease: summary of the 2014 evidence-based report by expert panel members. *JAMA* 2014 Sep 10;312(10):1033-1048. [doi: [10.1001/jama.2014.10517](#)] [Medline: [25203083](#)]
4. Lanzkron S, Haywood C, Hassell KL, Rand C. Provider barriers to hydroxyurea use in adults with sickle cell disease: a survey of the Sickle Cell Disease Adult Provider Network. *J Natl Med Assoc* 2008 Aug;100(8):968-973. [Medline: [18717150](#)]
5. Smeltzer MP, Howell KE, Treadwell M, Preiss L, King AA, Glassberg JA, Sickle Cell Disease Implementation Consortium. Identifying barriers to evidence-based care for sickle cell disease: results from the Sickle Cell Disease Implementation Consortium cross-sectional survey of healthcare providers in the USA. *BMJ Open* 2021 Nov 17;11(11):e050880 [FREE Full text] [doi: [10.1136/bmjopen-2021-050880](#)] [Medline: [34789492](#)]
6. Hodges JR, Phillips SM, Norell S, Nwosu C, Khan H, Luo L, et al. Intentional and unintentional nonadherence to hydroxyurea among people with sickle cell disease: a qualitative study. *Blood Adv* 2020 Sep 22;4(18):4463-4473 [FREE Full text] [doi: [10.1182/bloodadvances.2020001701](#)] [Medline: [32941646](#)]
7. Haywood C, Beach MC, Bediako S, Carroll CP, Lattimer L, Jarrett D, et al. Examining the characteristics and beliefs of hydroxyurea users and nonusers among adults with sickle cell disease. *Am J Hematol* 2011 Jan;86(1):85-87 [FREE Full text] [doi: [10.1002/ajh.21883](#)] [Medline: [21117058](#)]
8. Anglada-Martinez H, Riu-Viladoms G, Martin-Conde M, Rovira-Illamola M, Sotoca-Momblona JM, Codina-Jane C. Does mHealth increase adherence to medication? Results of a systematic review. *Int J Clin Pract* 2015 Jan;69(1):9-32. [doi: [10.1111/ijcp.12582](#)] [Medline: [25472682](#)]
9. Gandapur Y, Kianoush S, Kelli HM, Misra S, Urrea B, Blaha MJ, et al. The role of mHealth for improving medication adherence in patients with cardiovascular disease: a systematic review. *Eur Heart J Qual Care Clin Outcomes* 2016 Oct 01;2(4):237-244 [FREE Full text] [doi: [10.1093/ehjqcco/qcw018](#)] [Medline: [29474713](#)]
10. Kong T, Scott MM, Li Y, Wichelman C. Physician attitudes towards-and adoption of-mobile health. *Digit Health* 2020;6:2055207620907187 [FREE Full text] [doi: [10.1177/2055207620907187](#)] [Medline: [32128235](#)]
11. Hankins JS, Shah N, DiMartino L, Brambilla D, Fernandez ME, Gibson RW, Sickle Cell Disease Implementation Consortium. Integration of mobile health into sickle cell disease care to increase hydroxyurea utilization: Protocol for an efficacy and implementation study. *JMIR Res Protoc* 2020 Jul 14;9(7):e16319 [FREE Full text] [doi: [10.2196/16319](#)] [Medline: [32442144](#)]
12. DiMartino LD, Baumann AA, Hsu LL, Kanter J, Gordeuk VR, Glassberg J, Sickle Cell Disease Implementation Consortium. The sickle cell disease implementation consortium: Translating evidence-based guidelines into practice for sickle cell disease. *Am J Hematol* 2018 Dec;93(12):E391-E395 [FREE Full text] [doi: [10.1002/ajh.25282](#)] [Medline: [30203558](#)]
13. Alberts NM, Badawy SM, Hodges J, Estep JH, Nwosu C, Khan H, et al. Development of the InCharge Health mobile app to improve adherence to hydroxyurea in patients with sickle cell disease: User-centered design approach. *JMIR Mhealth Uhealth* 2020 May 08;8(5):e14884 [FREE Full text] [doi: [10.2196/14884](#)] [Medline: [32383683](#)]
14. Presseau J, McCleary N, Lorencatto F, Patey AM, Grimshaw JM, Francis JJ. Action, actor, context, target, time (AACTT): a framework for specifying behaviour. *Implement Sci* 2019 Dec 05;14(1):102 [FREE Full text] [doi: [10.1186/s13012-019-0951-x](#)] [Medline: [31806037](#)]
15. Moynihan R, Sanders S, Michaleff ZA, Scott AM, Clark J, To EJ, et al. Impact of COVID-19 pandemic on utilisation of healthcare services: a systematic review. *BMJ Open* 2021 Mar 16;11(3):e045343 [FREE Full text] [doi: [10.1136/bmjopen-2020-045343](#)] [Medline: [33727273](#)]

16. Subathra GN, Rajendrababu SR, Senthilkumar VA, Mani I, Udayakumar B. Impact of COVID-19 on follow-up and medication adherence in patients with glaucoma in a tertiary eye care centre in south India. *Indian J Ophthalmol* 2021 May;69(5):1264-1270 [FREE Full text] [doi: [10.4103/ijo.IJO\\_164\\_21](https://doi.org/10.4103/ijo.IJO_164_21)] [Medline: [33913874](https://pubmed.ncbi.nlm.nih.gov/33913874/)]
17. Alshareef R, Al Zahrani A, Alzahrani A, Ghandoura L. Impact of the COVID-19 lockdown on diabetes patients in Jeddah, Saudi Arabia. *Diabetes Metab Syndr* 2020;14(5):1583-1587 [FREE Full text] [doi: [10.1016/j.dsx.2020.07.051](https://doi.org/10.1016/j.dsx.2020.07.051)] [Medline: [32947759](https://pubmed.ncbi.nlm.nih.gov/32947759/)]
18. Singh A, Brandow AM, Panepinto JA. COVID-19 in individuals with sickle cell disease/trait compared with other Black individuals. *Blood Adv* 2021 Apr 13;5(7):1915-1921 [FREE Full text] [doi: [10.1182/bloodadvances.2020003741](https://doi.org/10.1182/bloodadvances.2020003741)] [Medline: [33792626](https://pubmed.ncbi.nlm.nih.gov/33792626/)]
19. Glasgow RE, Vogt TM, Boles SM. Evaluating the public health impact of health promotion interventions: the RE-AIM framework. *Am J Public Health* 1999 Sep;89(9):1322-1327. [doi: [10.2105/ajph.89.9.1322](https://doi.org/10.2105/ajph.89.9.1322)] [Medline: [10474547](https://pubmed.ncbi.nlm.nih.gov/10474547/)]
20. Forman J, Heisler M, Damschroder LJ, Kaselitz E, Kerr EA. Development and application of the RE-AIM QuEST mixed methods framework for program evaluation. *Prev Med Rep* 2017 Jun;6:322-328 [FREE Full text] [doi: [10.1016/j.pmedr.2017.04.002](https://doi.org/10.1016/j.pmedr.2017.04.002)] [Medline: [28451518](https://pubmed.ncbi.nlm.nih.gov/28451518/)]
21. Holtrop JS, Rabin BA, Glasgow RE. Qualitative approaches to use of the RE-AIM framework: rationale and methods. *BMC Health Serv Res* 2018 Mar 13;18(1):177 [FREE Full text] [doi: [10.1186/s12913-018-2938-8](https://doi.org/10.1186/s12913-018-2938-8)] [Medline: [29534729](https://pubmed.ncbi.nlm.nih.gov/29534729/)]
22. de Haan M, van Eijk-Hustings Y, Vrijhoef HJ. Using mixed methods in health services research: A review of the literature and case study. *J Health Serv Res Policy* 2021 Apr;26(2):141-147. [doi: [10.1177/1355819620955223](https://doi.org/10.1177/1355819620955223)] [Medline: [32957813](https://pubmed.ncbi.nlm.nih.gov/32957813/)]
23. Estep JH, Winter B, Johnson M, Smeltzer MP, Howard SC, Hankins JS. Improved hydroxyurea effect with the use of text messaging in children with sickle cell anemia. *Pediatr Blood Cancer* 2014 Nov;61(11):2031-2036. [doi: [10.1002/psc.25177](https://doi.org/10.1002/psc.25177)] [Medline: [25132074](https://pubmed.ncbi.nlm.nih.gov/25132074/)]
24. Badawy SM, Cronin RM, Hankins J, Crosby L, DeBaun M, Thompson AA, et al. Patient-centered eHealth interventions for children, adolescents, and adults with sickle cell disease: Systematic review. *J Med Internet Res* 2018 Jul 19;20(7):e10940 [FREE Full text] [doi: [10.2196/10940](https://doi.org/10.2196/10940)] [Medline: [30026178](https://pubmed.ncbi.nlm.nih.gov/30026178/)]
25. Minniti CP, Zaidi AU, Nourai M, Manwani D, Crouch GD, Crouch AS, et al. Clinical predictors of poor outcomes in patients with sickle cell disease and COVID-19 infection. *Blood Adv* 2021 Jan 12;5(1):207-215 [FREE Full text] [doi: [10.1182/bloodadvances.2020003456](https://doi.org/10.1182/bloodadvances.2020003456)] [Medline: [33570644](https://pubmed.ncbi.nlm.nih.gov/33570644/)]
26. Bezerra GMF, de Lucena Feitosa ES, Vale Catunda JG, Nogueira Sales Graça C, Lucena de Aquino P, Bezerra Neto AG, et al. Telemedicine application and assessment during the COVID-19 pandemic. *Stud Health Technol Inform* 2022 Jun 06;290:854-857. [doi: [10.3233/SHTI220200](https://doi.org/10.3233/SHTI220200)] [Medline: [35673139](https://pubmed.ncbi.nlm.nih.gov/35673139/)]
27. Nittari G, Savva D, Tomassoni D, Tayebati SK, Amenta F. Telemedicine in the COVID-19 era: A narrative review based on current evidence. *Int J Environ Res Public Health* 2022 Apr 22;19(9):5101 [FREE Full text] [doi: [10.3390/ijerph19095101](https://doi.org/10.3390/ijerph19095101)] [Medline: [35564494](https://pubmed.ncbi.nlm.nih.gov/35564494/)]
28. McDermott MM, Newman AB. Preserving clinical trial integrity during the coronavirus pandemic. *JAMA* 2020 Jun 02;323(21):2135-2136. [doi: [10.1001/jama.2020.4689](https://doi.org/10.1001/jama.2020.4689)] [Medline: [32211830](https://pubmed.ncbi.nlm.nih.gov/32211830/)]
29. PQA Adherence Measures. Pharmacy Quality Alliance. URL: <https://www.pqaalliance.org/adherence-measures> [accessed 2022-09-24]

## Abbreviations

**mHealth:** mobile health

**NHLBI:** National Health Lung and Blood Institute

**PDC:** proportion of days covered

**RE-AIM:** reach, effectiveness, adoption, implementation, and maintenance

**SCD:** sickle cell disease

**SCDIC:** Sickle Cell Disease Implementation Consortium

*Edited by A Mavragani; submitted 26.07.22; peer-reviewed by J Seethal, B Padhi, F Amenta, J Hodges; comments to author 20.08.22; revised version received 19.09.22; accepted 20.09.22; published 14.10.22.*

*Please cite as:*

Badawy SM, DiMartino L, Brambilla D, Klesges L, Baumann A, Burns E, DeMartino T, Jacobs S, Khan H, Nwosu C, Shah N, Hankins JS, Sickle Cell Disease Implementation Consortium

*Impact of the COVID-19 Pandemic on the Implementation of Mobile Health to Improve the Uptake of Hydroxyurea in Patients With Sickle Cell Disease: Mixed Methods Study*

*JMIR Form Res* 2022;6(10):e41415

URL: <https://formative.jmir.org/2022/10/e41415>

doi: [10.2196/41415](https://doi.org/10.2196/41415)

PMID: [36240004](https://pubmed.ncbi.nlm.nih.gov/36240004/)

©Sherif M Badawy, Lisa DiMartino, Donald Brambilla, Lisa Klesges, Ana Baumann, Ebony Burns, Terri DeMartino, Sara Jacobs, Hamda Khan, Chinonyelum Nwosu, Nirmish Shah, Jane S Hankins, Sickle Cell Disease Implementation Consortium. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 14.10.2022. 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

# An Accessible Communication System for Population-Based Genetic Testing: Development and Usability Study

Tara Coffin<sup>1</sup>, BA, MEd, PhD; Deborah Bowen<sup>1</sup>, PhD; Elizabeth Swisher<sup>1</sup>, MD; Karen Lu<sup>2</sup>, MD; Nadine Rayes<sup>2</sup>, MS; Barbara Norquist<sup>1</sup>, MD; Stephanie Blank<sup>3</sup>, MD; Douglas Levine<sup>4</sup>, MD; Jamie Bakkum-Gamez<sup>5</sup>, MD; Gini Fleming<sup>6</sup>, MD; Olufunmilayo Olopade<sup>6</sup>, MD; Alan D'Andrea<sup>7</sup>, MD; Denise Nebgen<sup>2</sup>, MD, PhD; Christine Peterson<sup>2</sup>, PhD; Mark Munsell<sup>2</sup>, MS; Kathleen Gavin<sup>8</sup>, MPH; Rebecca Lechner<sup>8</sup>; Jamie Crase<sup>1</sup>; Deborah Polinsky<sup>9</sup>, MS; Iris Romero<sup>6</sup>, MD

<sup>1</sup>University of Washington, Seattle, WA, United States

<sup>2</sup>MD Anderson Cancer Center, Houston, TX, United States

<sup>3</sup>Mount Sinai Medical Center, New York, NY, United States

<sup>4</sup>NYU Langone Health, New York, NY, United States

<sup>5</sup>Mayo, Rochester, MN, United States

<sup>6</sup>University of Chicago, Chicago, IL, United States

<sup>7</sup>Harvard, Cambridge, MA, United States

<sup>8</sup>Minnesota Ovarian Cancer Alliance, Minneapolis, MN, United States

<sup>9</sup>SHARE, New York City, NY, United States

**Corresponding Author:**

Tara Coffin, BA, MEd, PhD

University of Washington

1400 NE Campus Parkway

Seattle, WA, 98195-4550

United States

Phone: 1 206 543 2100

Email: [tbcoffin@uw.edu](mailto:tbcoffin@uw.edu)

## Abstract

**Background:** Genetic testing uptake is low, despite the well-established connection between pathogenic variants in certain cancer-linked susceptibility genes and ovarian cancer risk. Given that most major insurers cover genetic testing for those with a family history suggestive of hereditary cancer, the issue may lie in access to genetic testing. Remotely accessible web-based communication systems may improve awareness, and uptake, of genetic testing services.

**Objective:** This study aims to present the development and formative evaluation of the multistep web-based communication system required to support the implementation of, and access to, genetic testing.

**Methods:** While designing the multistep web-based communication system, we considered various barriers and facilitators to genetic testing, guided by dimensions of accessibility. In addition to conducting usability testing, we performed ongoing assessments focusing on the function of the web-based system and participant response rates, with the goal of continuing to make modifications to the web-based communication system as it is in use.

**Results:** The combined approach of usability testing and expert user experience consultation resulted in several modifications to the multistep web-based communication system, including changes that related to imagery and content, web accessibility, and general organization of the web-based system. All recommendations were made with the goal of improving the overall accessibility of the web-based communication system.

**Conclusions:** A multistep web-based communication system appears to be an effective way to address many potential barriers to access, which may otherwise make genetic testing difficult for at-risk individuals to participate in. Importantly, some dimensions of access were easy to assess before study recruitment, but other aspects of the communication system required ongoing assessment during the implementation process of the Making Genetic Testing Accessible study.

(JMIR Form Res 2022;6(10):e34055) doi:[10.2196/34055](https://doi.org/10.2196/34055)



**KEYWORDS**

genetic testing; internet; social media; accessibility

## Introduction

### Background

The association between pathogenic variants in certain cancer-linked susceptibility genes, such as *BRCA1* and *BRCA2*, and an increased risk of ovarian cancer is well established [1-3]. Women with a personal or family history of breast or epithelial ovarian cancer are encouraged to undergo genetic counseling [4]. Current recommendations state that genetic testing leaves at-risk women better prepared to make decisions about cancer prevention, early detection, and treatment [5].

Despite these recommendations, only 15% to 30% of eligible patients, women with a personal history of ovarian cancer, are offered these services in clinical settings [4,6-10], and <20% of women with a first-degree relative diagnosed with breast or ovarian cancer are ever offered genetic testing [11,12]. As a result, it is estimated that <5% and 10% of women at elevated or high risk, respectively, ever receive appropriate genetic testing services [12]. Importantly, although nonclinical direct-to-consumer genetic testing, such as 23andMe and Ancestry DNA, is widely accessible, this service does not take the place of clinical grade genetic testing, which is still a required step in the verification process of commercial DNA test findings. As most major insurers offer coverage for testing for individuals who meet the US Preventative Services Task Force guidelines for genetic testing, including those with a personal or family history suggestive of hereditary cancer, the obstacle may lie with another facet of testing accessibility.

Health care access is a broad and multidimensional concept: these dimensions include approachability, acceptability, availability, affordability, and appropriateness [13]. Each dimension is critical in creating a truly accessible health resource. According to this model of accessibility, for genetic testing to be approachable, the service must exist, be reachable, and impact health outcomes [13]. Furthermore, the acceptability of current methods of genetic counseling and testing will impact the overall likelihood of testing [13]. Availability is dependent on the individual physically reaching genetic testing services in a timely manner [13]. Genetic testing must be universally affordable to everyone who may benefit from it, in that those who can potentially benefit from testing must be able to pay for it without a catastrophic expenditure of resources, including time and money. Finally, for genetic testing services to be appropriate, the services should meet the needs of the intended community, be perceived as appropriate by that population, and be provided at a time they can access it.

### Objective

As accessibility to genetic testing for at-risk individuals continues to lag, public health professionals face an urgent need to explore alternative approaches to genomic service implementation [14,15]. With this urgency in mind, we designed the Making Genetic Testing Accessible (MAGENTA) study to evaluate a new model of providing genetic testing that will

potentially address the low uptake of genetic testing for women at risk for ovarian cancer. The MAGENTA study was a 4-arm noninferiority trial, using a multistep web-based communication system to deliver a combination of pre- and posttest electronic and phone-based genetic education and counseling and genetic testing results. Through this system, the MAGENTA study sought to better understand how to make genetic testing more accessible, increasing women's ability to obtain genetic testing for ovarian cancer risk assessment. By providing genetic testing information and services on the web, individuals who can benefit from these services are arguably more likely to speak with their physicians or reach out to service providers and improve subsequent testing uptake.

The study team opted to conduct usability testing of the web-based communication system before opening study recruitment. The usability testing of MAGENTA's web-based communication system was required to ensure that the system functioned as it was intended to, to identify bottlenecks in the enrollment process and to address general accessibility of the system itself, ensuring that participants could physically access information about the study. The MAGENTA study laid out a schema for increasing awareness of genetic education and counseling, potentially improving access by optimizing the usability of the web-based communication system before implementation. Here, we present the development and formative evaluation of the multistep web-based communication system required to support the implementation of genetic testing through web-based platforms that can be accessed anywhere there is an internet connection.

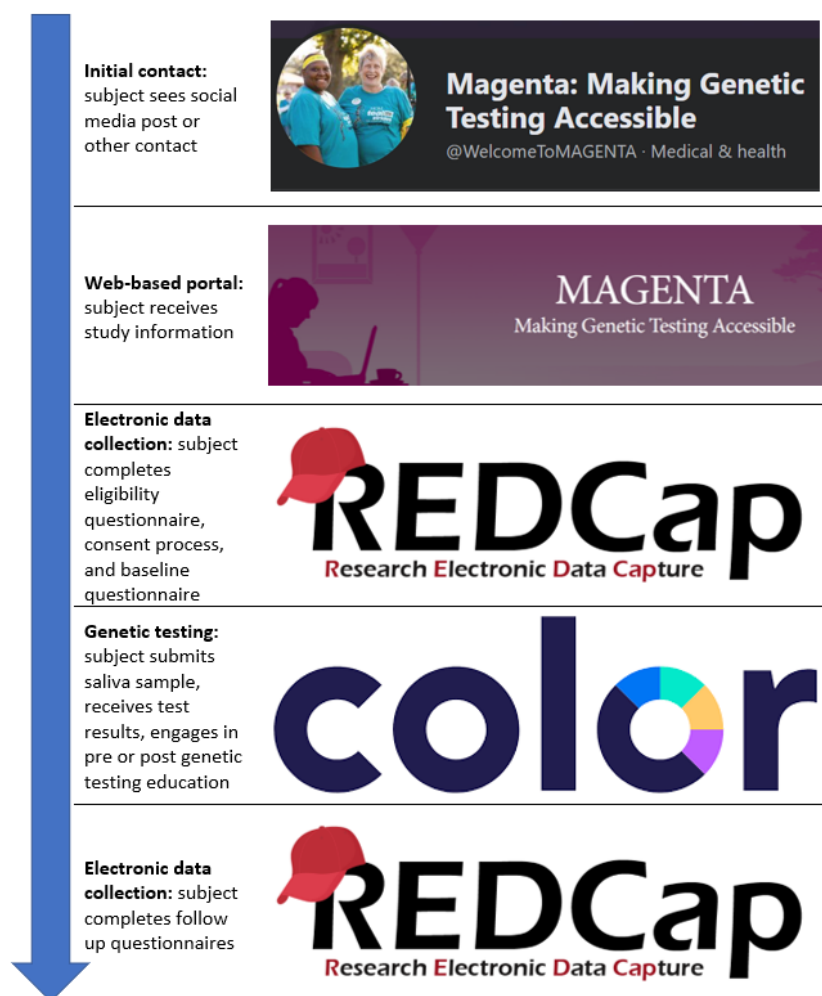
## Methods

### Overview

The MAGENTA multistep web-based communication system sought to provide background information about the study, determine participant eligibility, guide participants through a web-based consent process, collect baseline data, walk participants through the web-based genetic testing protocol, and collect follow-up data. Figure 1 presents a flowchart of these components, emphasizing accessibility at each stage. Participants learned about the study on the web, in a clinical setting, or via a media outlet (eg, news article or story, and radio station). From there, they visited the MD Anderson MAGENTA website to review study information, learn more about eligibility criteria, and proceed to the eligibility questionnaire via a link to an electronic data collection tool. Eligible participants completed the informed consent process over an electronic data collection tool and received a follow-up email sending them back to the electronic data collection tool to complete baseline questionnaires. Once complete, participants were referred to Color Genomics, via a link over the email, where they begin the process of genetic testing. Participants receive their testing kit in the mail, engage in web-based genetic education with or without pre- and posttest genetic counseling over the phone, and receive their results on the web. At 3, 12, and 24 months

following genetic testing, participants are prompted to return questionnaires [16].  
to the electronic data collection tool to complete follow-up

**Figure 1.** Overview of the web-based communication system for Making Genetic Testing Accessible (MAGENTA). This provides an overview of the flow for the web-based communication system developed and evaluated by the MAGENTA study team, with accessibility in mind.



## Usability Testing

When using a web-based communication system that required many moving parts, it was critical to evaluate each component individually and as a system with the dimensions of accessibility in mind. In addition to an expert user experience (UX) consultation and ongoing collaboration with ovarian cancer advocates and clinicians, we conducted extensive usability testing. Usability testing, the process of evaluating the UX of interacting with a computer system, was implemented to assess the multistep web-based communication system such as this [17].

We recruited 10 cisgender women aged between 30 and 55 years, who had access to a computer, spoke English, and were located within driving distance of the University of Washington, to participate in usability testing. The eligibility criteria for usability testing mirrored the inclusion criteria for the MAGENTA study with the exception of geographic location; however, a personal or family history of cancer was not required. Most usability testing participants self-identified as non-Hispanic White. Participants responded to a flyer about the

research and contacted the research team if they were interested in participating. If eligible, they were scheduled for a one-time in-person visit. At the start of this visit, participants signed a consent form and were informed about the usability study. Web-based components, including the initial study webpage and varying REDCap (Research Electronic Data Capture; Vanderbilt University) pages involved with the web-based communication system, were reviewed individually and when interacting with one another, capturing transitions between steps in the web-based communication system (Figure 1). Color Genomics was excluded from usability testing and expert UX consultation based out of the University of Washington. The Color Genomics system was not available for modification, so this portion of the web-based communication system was not assessed during usability testing. Color Genomics has extensive experience providing direct-to-consumer genetic testing services on the web and has previously undergone rigorous usability testing to assess its web-based platforms.

During testing, participants were prompted to “try to enroll in the study” or to “show me how you would go about learning more.” Components of the web-based communication system

were reviewed in different formats, including how they would be viewed from a handheld mobile device or a computer, to assess the usability experience from different platforms. As participants responded to each component, they were encouraged to use the think-aloud approach, talking out loud about their expectations for how the web-based artifact functioned, prompted with follow-up questions encouraging them to speak more to their actions or thoughts concerning the system [17-20]. All participant-computer interactions were recorded using a screen-capture tool. At the end of the session, participants completed an adapted Post-Study System Usability Questionnaire developed by Lewis [21], which uses a 7-point Likert scale, where 1 indicates low satisfaction and 7 indicates high satisfaction, to measure user satisfaction with materials undergoing usability testing. The Post-Study System Usability Questionnaire developed by Lewis [21] examines system ease of use and efficiency, information quality, and quality of the system interface, in addition to the overall perception of the system. Participants who completed the questionnaire once were not asked to complete the questionnaire in subsequent usability testing. A total of 7 participants completed the poststudy usability questionnaire.

The resulting video data were coded in Atlas.ti (ATLAS.ti Scientific Software Development GmbH), a program well suited to code visual data. Data were coded using a content analysis approach. We used a deductive approach to coding, drawing on the dimensions of accessibility conceptualized by Levesque et al [13] to inform codebook development and analysis. The MAGENTA study team used the cumulative results from all evaluations to address usability issues, anticipate potential bottlenecks, and improve system accessibility. In the following

section, we will review the process of designing and evaluating each of the components of the MAGENTA web-based communication system, a process guided by the dimensions of accessibility by Levesque et al [13].

## Dimensions of Accessibility

### Overview

Access to health care is often reduced to a focus on financial and geographic barriers, hurdles that can be resolved with insurance coverage and reimbursement for transportation. These are not the only dimensions of accessibility that impact the uptake and use of health care resources, as illustrated by Levesque et al [13]. In addition to the availability of the services, which encompasses physical proximity of services, and affordability (cost), Levesque et al [13] described the approachability of a health care service, the acceptability of this service, and the appropriateness of the health care service in question, dimensions that are entangled with cultural competency. The concept of approachability is closely tied to awareness and health communication. An approachable health care service is one that a patient knows exists and understands how this service may impact their health and well-being [13]. Acceptability focuses on social factors that may influence uptake, including an individual's comfort with a specific health care provider [13]. Finally, appropriateness speaks to the patients' actual need, versus the need the service addresses [13]. We have discussed how each of these dimensions was evaluated during the development of the web-based communication system and through subsequent usability testing. In addition, [Table 1](#) provides an overview of how the dimensions of accessibility were evaluated through each step of the assessment process.

**Table 1.** Summary of dimensions of accessibility and what methods were used to evaluate them.

Method and dimension	Assessment	
<b>Usability testing and patient advocate review</b>		
Approachability	Usability test participants and patient advocates provided feedback on the approachability of unpaid social media posts and paid Facebook advertisement, the MAGENTA <sup>a</sup> study homepage, the REDCap <sup>b</sup> study questionnaires and consent process, and all scheduled MAGENTA email communications	
Acceptability	Usability test participants and patient advocates provided feedback on the acceptability of images and text implemented across all written study components; familiarity of the components of the study, including the name of organizations associated with the research; and the REDCap study system, including consent process	
Availability	Usability test participants and patient advocates provided feedback on the availability of the general functionality of each component of the web-based communication system, including buttons and links; and the ability to locate study communications, including email	
Affordability	Usability test participants provided feedback on content related to cost of services	
Appropriateness	Usability test participants and patient advocates provided feedback on the appropriateness of images and text implemented across all written study components	
Post-Study System Usability Questionnaire	Availability	Usability test participants completed the Post-Study System Usability Questionnaire developed by Lewis [21], after completing usability testing, providing general feedback concerning the usability and functionality of the web-based communication system
<b>Expert UX<sup>c</sup></b>		
Approachability	UX experts assessed the web-based communication system, focusing on web accessibility	
Availability	The UX consultation provided feedback on high-level workflows, page layout, and the usability of each web-based component	
Ongoing system assessment	Availability	The study team tracked web analytics on an ongoing basis, across each web-based component in hopes of identifying and addressing any issues or bottlenecks before they impacted availability

<sup>a</sup>MAGENTA: Making Genetic Testing Accessible.

<sup>b</sup>REDCap: Research Electronic Data Capture.

<sup>c</sup>UX: user experience.

### Approachability

We provided multiple points of entry, over multiple platforms, to make information about the study widely approachable. We generated a media kit that included internet-based outreach materials, including unpaid social media posts, paid Facebook advertisements, and other traditional outreach tools (eg, press release, flyers, and news articles). We opted to focus our web-based promotion efforts on Facebook because of its popularity as the most widely used web-based social media platform at the time of study recruitment [22]. Paid Facebook advertisements and unpaid social media posts included information about the study, a rotation of relevant imagery, and a link to the MAGENTA website. Unpaid social media posts included material shaped for use over Facebook, Twitter, YouTube, Instagram, and web-based blog sites.

In designing paid social media advertisements and unpaid social media posts, we drew on content and tone that would be more approachable to potential participants. Current research speaks to different approaches motivating people in varying ways [23]. For example, some may be more motivated to engage in a health behavior if they are doing it for their family, whereas others may have more self-serving interests [23]. Similarly, content

that paints an image of teamwork and collaboration may facilitate engagement for some, but others may be more interested in acting independently [23]. We built posts and advertisements with this research in mind, in hopes of having a selection of materials that would resonate with different groups.

Unpaid posts were published to existing social media pages and groups by page administrators, on behalf of MAGENTA. These groups included Facebook pages run by breast and ovarian cancer advocacy organizations, gynecological cancer professional groups, and genetic testing companies. These groups were identified through keyword searches and through referrals from patient advocates. By focusing on organizations with a vested interest in ovarian cancer genetics and an established community of followers, MAGENTA investigators aimed to leverage existing social networks, using familiar and trusted names to make information about the study more approachable [24].

Paid advertisements were posted to Facebook. All Facebook advertising campaigns targeted English-speaking women aged >30 years who live in the United States. After an initial soft launch period, we published a series of unpaid Facebook posts

before moving to paid Facebook advertisements. We also used other traditional outreach efforts, including emails to clinicians and flyers in clinical settings.

It was critical that additional information was easy to locate and review. Knowing that participants may have found the study via several routes, we built a study website to introduce the multistep enrollment and participation process. All outreach efforts directed people to this website for study information and details about eligibility criteria. The MAGENTA study website also referred participants to the next step in the communication process via a link to the electronic data collection system.

We ran social media materials, the MAGENTA website, and each component of the electronic data collection system through usability testing to evaluate approachability. Usability testing focuses on whether a system is aesthetically and functionally accessible to the end user, evaluating if users could approach study materials as intended. We introduced participants to the study through Facebook posts and advertisements, prompting them to explain what they would do to learn more about the study. Participants also reviewed the MAGENTA website and the electronic data collection system in this fashion. This included a series of emails participants receive when they interacted with the communication system. Participants were prompted to explain the perceived goal of each component, what they noticed about the artifact, and what they would need to do next if they wanted to learn more or enroll in the study. During usability testing, we noted instances when participants did not know what to do next or when they reported an inaccurate take-away, informing modifications to system approachability.

To assess approachability for individuals with disabilities, we requested a review by an expert UX consultation service at the University of Washington. During the UX consultation, a team of UX experts assessed each component of the web-based communication system, with web accessibility in mind.

### **Acceptability**

MAGENTA addressed the acceptability of genetic testing by enabling web-based participation over an internet-based communication system. By using a web-based communication system, the importance of provider identity diminished compared with in-person services, improving the acceptability of ovarian cancer genetic testing along the way. This has been the case for health services targeting a variety of stigmatized issues [25]. Web-based communication also might have made it easier to adapt and change messaging on an ongoing basis, modifying content to fit the needs of a diverse community. With this in mind, we opted to rotate through a diverse array of imagery and content when constructing web-based materials.

Web-based services invite an opportunity for branding. The MD Anderson MAGENTA website in particular served as a platform for building consumer trust and illustrating ties between the MAGENTA study and collaborating organizations, including familiar names such as MD Anderson and Stand Up To Cancer, sponsors of the study. Although potential participants may not have recognized MD Anderson, they may have heard of Stand Up To Cancer and may be more likely to consume and engage

with information from a familiar source [26,27]. In favor of branding, we also chose to create a study specific Facebook page, lending credence to our social media presence. This page was linked to our paid Facebook advertising campaigns. Importantly, the study team also designed a logo for the MAGENTA study. This logo was used across all study materials to build familiarity and brand awareness.

Finally, we considered user privacy when designing the web-based communication system. Presumably, users prefer web-based tools with appropriate privacy protection. This was especially relevant when it came time to consider options for where to house data from study questionnaires and consent forms. With this goal in mind, we selected REDCap as the electronic data collection system [28]. REDCap is a Health Insurance Portability and Accountability Act-compliant platform, which is also compliant with Part 11 electronic signature regulations for the purposes of e-consent, features that made REDCap a natural fit for the MAGENTA study, addressing concerns about the management of confidential information.

All imagery and content were assessed through usability testing, with acceptability in mind. Participants reviewed images of the social media posts and advertisements. These materials were displayed on the computer screen and participants were asked to reflect on the images and content used and comment on what they noticed. Conversation was recorded alongside the screen recording, capturing what was being viewed. We also asked participants to identify collaborating institutions, or any familiar names associated with the study (eg, Stand Up To Cancer). We noted instances when participants made a positive or negative comment about the imagery or text used, as well as noting if they could identify the organizations behind the research effort. This information was used to inform modifications to images and content, changes tested in a second round of usability testing. Finally, we asked ovarian cancer advocates on the study team to review each web-based component. Insights from usability testing and advocates helped shape web-based materials.

### **Availability**

We chose to use a web-based communication system to improve service availability [6]. Instead of requiring someone to have the means necessary to travel to a clinic for testing, the web-based communication system made it possible to learn about the study and participate in genetic testing via an internet connection. Participants could proceed through each web-based step at their own pace, without requirement to travel to any specific office. REDCap used automatic email reminders, making it easier to remind participants to complete various steps, simplifying longitudinal follow-up and potentially improving study attrition rates.

Similarly, the MAGENTA study team chose to offer genetic testing services through Color Genomics, a web-based genetic testing company, favoring the availability of web-based services over the more traditional clinic-based alternatives. Color Genomics is a company facilitating at-home genetic testing, along with web-based or phone-based counseling and education. Each step of this process, from ordering a saliva-based test kit

to reviewing results, occurred on the web at the participant's convenience. By giving participants the opportunity to complete genetic testing, including pre- and posttest genetic counseling via telehealth, MAGENTA effectively expanded the service range of genetic testing. Under this mode of provision, genetic testing was not just for those who live near a clinic offering this specialty service, it was available for anyone with an internet connection.

With availability in mind, we took precautions to ensure that each web-based component had the capacity to support high web traffic. Without this assurance, the MAGENTA study risked page crashes, reduced system availability, and, similarly, reduced availability of genetic testing services.

When dealing with a web-based system, availability extends beyond the physical availability of the service and into the usability of the system. In other words, a web-based communication system is only available to a user if it functions the way it is supposed to. To assess usability, we conducted usability testing of the web-based communication system. As part of the usability testing, participants were required to complete a web-based version of the Post-Study System Usability Questionnaire developed by Lewis [21]. This questionnaire provided quantitative data concerning system usability. We also referred the MD Anderson MAGENTA website and the electronic data collection system to expert UX consultation. The UX consultation focused on providing feedback on high-level workflows, page layout, and the usability of each web-based component. To encourage continued assessment of system usability, web analytics were tracked, on an ongoing basis across each web-based component in hopes of identifying and addressing any issues or bottlenecks before they impacted availability.

### **Affordability**

For MAGENTA's ovarian cancer genetic testing services to be widely affordable to women who may benefit, they must be available at no or low cost. Affordability is not just restricted to financial limitations but includes opportunity cost or the cost an individual is burdened with in exchange for accessing a service [29]. Additional barriers, such as cost of transportation, childcare, and work time lost, also play a role in the cost associated with the service, therefore reducing access. By offering genetic testing services on the web, we aim to reduce time lost and address secondary costs associated with in-person clinic visits, making genetic testing more affordable in terms of time and money. To ensure access, we also had to make sure that potential participants knew it was affordable. We tried to ensure that this information was included across study outreach efforts and other components of the web-based communication

system, focusing on messaging such as "at no cost to you" and "genetic testing from your living room."

Usability testing participants reviewed the web-based study components, focusing on outreach and background information (eg, social media posts or advertisements and study website), and were asked to reflect on the cost and location of research-related services, checking for comprehension and message clarity.

### **Appropriateness**

MAGENTA staff created multiple images with diverse situations, appropriate for all social media outlets. This allowed us to rotate through a diverse array of imagery across each component, with the goal of creating an evolving narrative that resonated across different populations. We assessed imagery and content used throughout all social media advertisements, posts, and the study website for appropriateness through usability testing. During usability testing, we asked participants to talk through their initial reactions to imagery and posts, prompting them to discuss whether these components resonated with them, or to hypothesize who they might resonate for. That said, appropriateness invited ongoing assessment and modifications. Moving forward, the study team planned to use analytics to track the efficacy of different web-based components, informing modifications, and allowing the opportunity to tailor each component to ensure the overall system meets the needs of the targeted population.

### **Ethics Approval**

This research was reviewed by University of Washington Institutional Review Board, and determined to meet the criteria for exemption from the institutional review board review.

## **Results**

### **Overview**

The combined approach of usability testing and expert UX consultation resulted in a total of (1) 12 recommended changes to the social media advertisements and posts, (2) 34 recommended changes to the REDCap components of the system, and (3) 9 recommended changes across the MD Anderson MAGENTA website. Of these 55 recommendations, 36 (65%) addressed content (eg, imagery and text), 11 (20%) were related to web accessibility (eg, font and contrast), and 8 (15%) were related to page organization. All recommendations and design considerations were made with the goal of improving the overall accessibility of the web-based communication system, addressing potential barriers to the acceptability, availability, approachability, affordability, and appropriateness of the system, as defined by Levesque et al [13]. Table 2 summarizes these changes by accessibility component.

**Table 2.** Summary of accessibility concerns with current genetic testing practices and how Making Genetic Testing Accessible study addresses each concern.

Dimension	Barriers	Assessment	Decisions and adjustments
Approachability	Genetic testing is not approachable	Assess all advertisements and posts	<ul style="list-style-type: none"> <li>Outreach routes should include social media, in-person (clinical settings and flyers), emails, and collaborations with trusted organizations (including ovarian cancer advocacy groups)</li> <li>Outreach materials should include physical flyers, paid social media advertisements, and unpaid posts across different social media channels</li> </ul>
Acceptability	Genetic testing is not universally acceptable	Assess imagery and text used across all written study components	<ul style="list-style-type: none"> <li>Use web-based system that (1) facilitates anonymity, (2) identifies study collaborators, (3) facilitates branding, (4) builds associations between study and related trusted organizations, (5) encourages responsive messaging, (6) facilitates rotating imagery and content, and (7) enables ongoing assessment</li> <li>Use an electronic data collection tool</li> </ul>
Availability	Genetic testing is not universally available	Assess usability of each component and the whole system (buttons and links)	<ul style="list-style-type: none"> <li>Use a web-based system that (1) meets accessibility standards (including font and device accessibility); (2) makes all web components physically available (eg, the Get Started button), (3) simplifies the email verification process, (4) enables reminders (including when to check email), and (5) includes a video explaining the study</li> <li>Use a web-based genetic testing service to make genetic testing itself physically available</li> </ul>
Affordability	Genetic testing is not affordable to everyone	Assess content related to cost of services	<ul style="list-style-type: none"> <li>Genetic testing is free to participants</li> </ul>
Appropriateness	Genetic testing is not appropriate for everyone	Assess imagery and text used across all components	<ul style="list-style-type: none"> <li>Use a diverse array of imagery (including families, uplifting, women aged &gt;30 years, and candid images)</li> <li>Avoid fear-based language</li> <li>Make modifications to outreach materials based on performance</li> </ul>

## Approachability

Usability testing participants were unclear about what service was being offered when just viewing the advertisements. A participant noted, “I don’t know if this is about research or they’re trying to sell me something,” and another said, “what is Facebook doing thinking about my ovaries?” With these comments in mind, we updated the advertisement content to reflect the research goals more appropriately. We monitored responses to the advertisement and adjusted as needed.

Participants also noted where the font was difficult to read or decode, issues that were also identified through our expert UX consultation, which provided substantial insight regarding web accessibility changes. Many of the web accessibility recommendations were not able to be supported in REDCap. Although we could update the font color over REDCap, other changes, such as removing the font resize function built into REDCap surveys, were not possible. Although REDCap supported the confidential collection of data, enhancing the acceptability of the system, our assessments indicated that REDCap has room for improvement in terms of system approachability, particularly regarding web accessibility.

## Acceptability

In terms of acceptability, several usability testing participants mentioned that they did not know who *MD Anderson Cancer Center* or *Stand Up To Cancer* was. Others were unsure what MAGENTA was when seeing this name connected to Facebook and social media outreach. Many usability testing participants commented that they felt unsure about who was behind the study and felt uneasy about relinquishing their genetic data to a web-based entity. This observation led to some content revisions, with the goal of highlighting the organizations behind the research more clearly on the MD Anderson MAGENTA website. We tracked traffic across each component of the multistep web-based communication system, via web analytics. This enabled us to identify which components are working well and where there might be problems.

## Availability

In general, participants reported that the progression from social media posts, or advertisements, to the MD Anderson website was easy to understand. Although the progression was clear, participants still made some recommendations addressing system availability. For example, several participants noted that it was difficult to figure out what they needed to do next on the MD Anderson MAGENTA website and suggested that the *Do I qualify?* button be relabeled and moved to a more prominent

position. This suggestion was also brought up during our expert UX consultation. We addressed this problem by relabeling the button *Get Started* and by moving the button to a more central location that was easier to find.

Expert user consultation combined with usability testing also led to several recommendations that addressed the usability and flow of the electronic data collection system. System usability was also captured via the Post-Study System Usability Questionnaire developed by Lewis [21]. Participants who completed the questionnaire (N=7) indicated that they were moderately satisfied with the system ease of use and efficiency (5.06, where 1 indicates low satisfaction and 7 indicates high satisfaction). When asked about information quality, meaning they had the information they needed to complete the task or interact with the system, participants reported moderate satisfaction again (4.88). Participants reported lower levels of satisfaction when asked about the quality of the interface (4.56). Overall, participants were moderately satisfied with the system, with a mean of 4.91 (SD 0.35) across all items.

### Affordability

When assessing the affordability of the system, usability testing participants could accurately identify that the services rendered through study participation were free of charge and available to complete on their own time and at their own pace. To demonstrate this, participants commented on genetic testing being available at “no cost.”

### Appropriateness

When assessing the appropriateness of the system, usability test participants noted that images with a more natural appearance, as opposed to those that were clearly posed, were easier to relate to. When speaking about a more candid image, a participant said, “I like that image, because it looks more realistic.” Another participant commented on the age of a mother-child dyad featured in an image, stating, “I’m unclear why there is a mom and a baby, when this probably should be for women with teenagers.” When looking at an image featuring a mother with a young baby, another participant pointed out that a mother-infant dyad such as this might exclude women without children, who may still be at risk of ovarian cancer. Most participants who noted the diversity across images remarked positively about the representation, with one noting, “I like that I see a lot of diversity in these images.”

Usability test participants and ovarian cancer advocates spoke favorably about content that framed ovarian cancer as something that is relevant to anyone with a family history. They felt that using a picture of a family in this context suggested that the actions of one individual could benefit others in their family. Insight gained from this discussion led to the decision to focus on images that featured >1 person or families, as well as images that appeared to be more candid. The ovarian cancer advocates also noted that the *fear*-based language present in earlier iterations of the paid advertisements and unpaid posts, such as “Are you at risk for ovarian cancer?”, would be less likely to inspire them into action, when compared with something more positive or hopeful, such as a message imploring people to work as a team toward a solution. This note was echoed among

ovarian cancer advocates that we consulted with and informed changes to content.

## Discussion

### Principal Findings

The accessibility of genetic testing continues to be a challenge, leaving many eligible candidates faced with barriers to participation. Current research suggests that the internet may address many of the components of accessibility outlined by Levesque et al [13]. With a paucity of research dedicated to assessing the efficacy of multistep web-based communication systems for facilitating genetic testing, careful planning and consideration were required. The recommendations generated through this assessment focused primarily on the organization and usability of the system, and there was continued assessment over the course of recruitment. The success of the multistep web-based communication system could help facilitate genetic testing implementation in the future, increasing the use of the internet for physician-mediated genetic counseling and genetic testing, ultimately leading to increased access to genetic testing, particularly among populations currently underserved by genetic testing.

Applications of telehealth, such as those evaluated through this research, address critical barriers and improve accessibility [30]. Telehealth has become even more important during the COVID-19 pandemic, with the halt of in-person activities in research and health care settings, in the interest of disease mitigation [30]. The web-based communication system evaluated for use in the MAGENTA study was fully remote, leveraging some of the positive attributes of telehealth services. Although the usability testing described in this paper was conducted in person, the described methods can be implemented remotely, improving access among adults with an internet connection and providing an opportunity for a more diverse subject population. This accommodation does not address access among those who do not have access to the internet, and additional research is needed to identify innovative approaches for reaching these individuals.

### Limitations

We assessed the strengths and weaknesses of the evaluation process along the way. Although we attempted to address issues with the evaluation process as they emerged, usability testing had its limitations. For parts of usability testing, we had to rely on screenshots of systems or use systems that were not functional. Instead of proceeding through the multistep system in these situations, we had participants talk about what they expected or what they would interact with if the system was active. Thus, we were unable to determine how long it took users to complete certain tasks, both in time elapsed and number of clicks. This was one reason the study team, in partnership with patient advocates, remain dedicated to the ongoing assessment of the multistep web-based communication system. By continuing the assessment and inviting modification as needed, focusing on web analytics to identify bottlenecks and attrition, we aim to address any issues that usability testing failed to identify, which may arise through further system use



and increased traffic to different components of the web-based communication system.

Participants spoke favorably of the use of diverse imagery, featuring Black and indigenous people of color. Diverse imagery is known to be an important feature to ensure a message resonates with a wider research participant population. Although MAGENTA outreach materials ultimately featured people from different racial-ethnic backgrounds, the research recruited a predominantly White study participant population [31]. With this observation in mind, it is clear that diverse imagery alone cannot address the widespread and complex issue of underrepresentation in research. Racial and ethnic minorities are often underrepresented in genetic testing studies and services [32,33].

The eligibility criteria implemented in usability testing contributed to another limitation related to representation. The MAGENTA study specifically recruited women with at least one ovary, effectively excluding transgender men who may also be at an increased risk of ovarian cancer. The decision to exclude transgender men from participating in the MAGENTA study was made in response to the different health needs exhibited between cisgender women and transgender men. In addition, the standardized instruments used in the larger MAGENTA study were not validated across transgender and nonbinary populations. As the eligibility criteria for usability testing loosely mirrored the larger study, we made the decision to specifically include cisgender women. The exclusion of transgender and nonbinary individuals from reproductive cancer research, and subsequent health care services is a significant limitation that contributes to adverse health outcomes for these individuals, perpetuates health disparities, and requires attention.

Similarly, usability testing did not recruit participants based on health insurance status, reflecting the eligibility criteria in the MAGENTA study. MAGENTA did not formally require participants to have insurance; however, all research participants were required to provide the name and contact information of a health care provider. This step may have functionally excluded uninsured individuals. Although the MAGENTA study addressed the issue of affordability to a certain extent by providing genetic testing to participants at no cost, usability testing arguably did not fully evaluate this dimension of accessibility, as financial barriers do not stop with the cost of testing but also encompass the cost associated with increased cancer surveillance that a positive genetic test result may catalyze. Further assessment is needed to evaluate web-based communication platforms, with these underrepresented populations in mind.

Usability testing is poised to help address this trend, given the focus on identifying communication and system barriers. To effectively address barriers for underrepresented racial and ethnic groups, usability testing needs to be conducted with a representative population that includes racial and ethnic minorities. Given the predominantly White population enrolled in this usability study and the parallel research participant population enrolled through the MAGENTA study, it appears that the approachability and appropriateness dimensions in this usability study were limited. Future usability testing for web-based genetic testing research should focus on assessing barriers that are specific to underrepresented groups, including underinsured and uninsured individuals, as well as Black and indigenous people of color, in the interest of better assessing the accessibility barriers.

---

## Acknowledgments

This work was supported by the Stand Up To Cancer-Ovarian Cancer Research Alliance-National Ovarian Cancer Coalition Ovarian Cancer Dream Team Translational Research Grant (grant SU2C-AACR-DT16-15). Stand Up To Cancer is a division of the Entertainment Industry Foundation. Research grants are administered by the American Association for Cancer Research, the Scientific Partner of Stand Up To Cancer.

---

## Conflicts of Interest

OO has received research funding from Color Genomics and is the cofounder of CancerIQ. OO also serves on the scientific advisory board of Tempus and is on the board of directors of 54gene.

---

## References

1. Schrader K, Hurlburt J, Kalloger S, Hansford S, Young S, Huntsman D, et al. Germline BRCA1 and BRCA2 mutations in ovarian cancer. *Obstetrics Gynecol* 2012;120(2, Part 1):235-240 [FREE Full text] [doi: [10.1097/aog.0b013e31825f3576](https://doi.org/10.1097/aog.0b013e31825f3576)]
2. Pal T, Permuth-Wey J, Betts JA, Krischer JP, Fiorica J, Arango H, et al. BRCA1 and BRCA2 mutations account for a large proportion of ovarian carcinoma cases. *Cancer* 2005 Dec 15;104(12):2807-2816 [FREE Full text] [doi: [10.1002/cncr.21536](https://doi.org/10.1002/cncr.21536)] [Medline: [16284991](https://pubmed.ncbi.nlm.nih.gov/16284991/)]
3. Easton D, Ford D, Bishop DT. Breast and ovarian cancer incidence in BRCA1-mutation carriers. Breast cancer linkage consortium. *Am J Hum Genet* 1995 Jan;56(1):265-271 [FREE Full text] [Medline: [7825587](https://pubmed.ncbi.nlm.nih.gov/7825587/)]
4. Moyer V. Risk assessment, genetic counseling, and genetic testing for BRCA-related cancer in women: U.S. Preventive services task force recommendation statement. *Ann Intern Med* 2014 Feb 18;160(4):271-281 [FREE Full text] [doi: [10.7326/m13-2747](https://doi.org/10.7326/m13-2747)]
5. National Cancer Institute. 2022. URL: <https://www.cancer.gov/about-cancer/causes-prevention/genetics/genetic-testing-fact-sheet#what-are-some-of-the-benefits-of-genetic-testing-for-inherited-cancer-susceptibility-syndromes> [accessed 2022-03-16]

6. George A, Riddell D, Seal S, Talukdar S, Mahamdallie S, Ruark E, et al. Implementing rapid, robust, cost-effective, patient-centred, routine genetic testing in ovarian cancer patients. *Sci Rep* 2016 Jul 13;6(1):29506 [FREE Full text] [doi: [10.1038/srep29506](https://doi.org/10.1038/srep29506)] [Medline: [27406733](https://pubmed.ncbi.nlm.nih.gov/27406733/)]
7. Febbraro T, Robison K, Wilbur JS, Laprise J, Bregar A, Lopes V, et al. Adherence patterns to National Comprehensive Cancer Network (NCCN) guidelines for referral to cancer genetic professionals. *Gynecol Oncol* 2015 Jul;138(1):109-114 [FREE Full text] [doi: [10.1016/j.ygyno.2015.04.029](https://doi.org/10.1016/j.ygyno.2015.04.029)] [Medline: [25933682](https://pubmed.ncbi.nlm.nih.gov/25933682/)]
8. Ricci M, Sciallero S, Mammoliti S, Gismondi V, Franiuk M, Bruzzi P, et al. Referral of ovarian cancer patients for genetic counselling by oncologists: need for improvement. *Public Health Genomics* 2015;18(4):225-232 [FREE Full text] [doi: [10.1159/000431352](https://doi.org/10.1159/000431352)] [Medline: [26111740](https://pubmed.ncbi.nlm.nih.gov/26111740/)]
9. Slade I, Riddell D, Turnbull C, Hanson H, Rahman N, programme M. Development of cancer genetic services in the UK: a national consultation. *Genome Med* 2015;7(1):18 [FREE Full text] [doi: [10.1186/s13073-015-0128-4](https://doi.org/10.1186/s13073-015-0128-4)] [Medline: [25722743](https://pubmed.ncbi.nlm.nih.gov/25722743/)]
10. Norquist B, Pennington KP, Agnew KJ, Harrell MI, Pennil CC, Lee MK, et al. Characteristics of women with ovarian carcinoma who have BRCA1 and BRCA2 mutations not identified by clinical testing. *Gynecol Oncol* 2013 Mar;128(3):483-487 [FREE Full text] [doi: [10.1016/j.ygyno.2012.12.015](https://doi.org/10.1016/j.ygyno.2012.12.015)] [Medline: [23262210](https://pubmed.ncbi.nlm.nih.gov/23262210/)]
11. Baer HJ, Brawarsky P, Murray MF, Haas JS. Familial risk of cancer and knowledge and use of genetic testing. *J Gen Intern Med* 2010 Jul 2;25(7):717-724 [FREE Full text] [doi: [10.1007/s11606-010-1334-9](https://doi.org/10.1007/s11606-010-1334-9)] [Medline: [20361271](https://pubmed.ncbi.nlm.nih.gov/20361271/)]
12. Bellcross CA, Peipins LA, McCarty FA, Rodriguez JL, Hawkins NA, Hensley Alford S, et al. Characteristics associated with genetic counseling referral and BRCA1/2 testing among women in a large integrated health system. *Genet Med* 2015 Jan;17(1):43-50 [FREE Full text] [doi: [10.1038/gim.2014.68](https://doi.org/10.1038/gim.2014.68)] [Medline: [24946155](https://pubmed.ncbi.nlm.nih.gov/24946155/)]
13. Levesque J, Harris MF, Russell G. Patient-centred access to health care: conceptualising access at the interface of health systems and populations. *Int J Equity Health* 2013 Mar 11;12(1):18 [FREE Full text] [doi: [10.1186/1475-9276-12-18](https://doi.org/10.1186/1475-9276-12-18)] [Medline: [23496984](https://pubmed.ncbi.nlm.nih.gov/23496984/)]
14. Schwartz M, Valdimarsdottir H, Peshkin B, Mandelblatt J, Nusbaum R, Huang A, et al. Randomized noninferiority trial of telephone versus in-person genetic counseling for hereditary breast and ovarian cancer. *J Clin Oncol* 2014 Mar 01;32(7):618-626 [FREE Full text] [doi: [10.1200/JCO.2013.51.3226](https://doi.org/10.1200/JCO.2013.51.3226)] [Medline: [24449235](https://pubmed.ncbi.nlm.nih.gov/24449235/)]
15. Bradbury A, Patrick-Miller L, Egleston B, Sands C, Li T, Schmidheiser H, et al. Parent opinions regarding the genetic testing of minors for BRCA1/2. *J Clin Oncol* 2010 Jul 20;28(21):3498-3505 [FREE Full text] [doi: [10.1200/jco.2009.27.2971](https://doi.org/10.1200/jco.2009.27.2971)]
16. Rayes N, Bowen DJ, Coffin T, Nebgen D, Peterson C, Munsell MF, et al. MAGENTA (Making Genetic testing accessible): a prospective randomized controlled trial comparing online genetic education and telephone genetic counseling for hereditary cancer genetic testing. *BMC Cancer* 2019 Jul 02;19(1):648 [FREE Full text] [doi: [10.1186/s12885-019-5868-x](https://doi.org/10.1186/s12885-019-5868-x)] [Medline: [31266460](https://pubmed.ncbi.nlm.nih.gov/31266460/)]
17. Kushniruk AW, Patel VL. Cognitive and usability engineering methods for the evaluation of clinical information systems. *J Biomed Inform* 2004 Feb;37(1):56-76 [FREE Full text] [doi: [10.1016/j.jbi.2004.01.003](https://doi.org/10.1016/j.jbi.2004.01.003)] [Medline: [15016386](https://pubmed.ncbi.nlm.nih.gov/15016386/)]
18. Contextual Design: Defining Customer-Centered Systems (Interactive Technologies). Burlington, Massachusetts: Morgan Kaufmann; 1997.
19. Makri S, Blandford A, Gow J, Rimmer J, Warwick C, Buchanan G. A library or just another information resource? A case study of users' mental models of traditional and digital libraries. *J Am Soc Inf Sci* 2007 Feb 01;58(3):433-445. [doi: [10.1002/asi.20510](https://doi.org/10.1002/asi.20510)]
20. Charters E. The use of think-aloud methods in qualitative research an introduction to think-aloud methods. *Brock Educ J* 2003 Jul 01;12(2) [FREE Full text] [doi: [10.26522/broeked.v12i2.38](https://doi.org/10.26522/broeked.v12i2.38)]
21. Lewis J. Psychometric evaluation of the PSSUQ using data from five years of usability studies. *Intl J Human Comput Interact* 2002 Sep;14(3-4):463-488. [doi: [10.1080/10447318.2002.9669130](https://doi.org/10.1080/10447318.2002.9669130)]
22. Social media fact sheet. Pew Research Center. URL: <http://www.pewinternet.org/fact-sheet/social-media/> [accessed 2022-03-06]
23. Batterham P. Recruitment of mental health survey participants using internet advertising: content, characteristics and cost effectiveness. *Int J Methods Psychiatr Res* 2014 Jun;23(2):184-191 [FREE Full text] [doi: [10.1002/mpr.1421](https://doi.org/10.1002/mpr.1421)] [Medline: [24615785](https://pubmed.ncbi.nlm.nih.gov/24615785/)]
24. Brickman Bhutta C. Not by the book. *Sociol Methods Res* 2012 Mar 21;41(1):57-88 [FREE Full text] [doi: [10.1177/0049124112440795](https://doi.org/10.1177/0049124112440795)]
25. Thomas N, McLeod B, Jones N, Abbott J. Developing internet interventions to target the individual impact of stigma in health conditions. *Internet Interventions* 2015 Sep;2(3):351-358. [doi: [10.1016/j.invent.2015.01.003](https://doi.org/10.1016/j.invent.2015.01.003)]
26. Steelman TA, McCaffrey SM, Velez AK, Briefel JA. What information do people use, trust, and find useful during a disaster? Evidence from five large wildfires. *Nat Hazards* 2014 Nov 16;76(1):615-634. [doi: [10.1007/s11069-014-1512-x](https://doi.org/10.1007/s11069-014-1512-x)]
27. Lindell M, Perry RW. Warning mechanisms in emergency response systems. *Int J Mass Emerg Disasters* 1987;5(2):137-153.
28. 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](https://doi.org/10.1016/j.jbi.2008.08.010)] [Medline: [18929686](https://pubmed.ncbi.nlm.nih.gov/18929686/)]
29. Cairns J. Using cost-effectiveness evidence to inform decisions as to which health services to provide. *Health Syst Reform* 2016 Jan 02;2(1):32-38. [doi: [10.1080/23288604.2015.1124172](https://doi.org/10.1080/23288604.2015.1124172)] [Medline: [31514650](https://pubmed.ncbi.nlm.nih.gov/31514650/)]

30. Hirko K, Kerver JM, Ford S, Szafranski C, Beckett J, Kitchen C, et al. Telehealth in response to the COVID-19 pandemic: implications for rural health disparities. *J Am Med Inform Assoc* 2020 Nov 01;27(11):1816-1818 [FREE Full text] [doi: [10.1093/jamia/ocaa156](https://doi.org/10.1093/jamia/ocaa156)] [Medline: [32589735](https://pubmed.ncbi.nlm.nih.gov/32589735/)]
31. Swisher EM, Rayes N, Bowen D, Peterson CB, Coffin T, Norquist B, et al. Results from MAGENTA: a national randomized four-arm noninferiority trial evaluating pre- and post-test genetic counseling during online testing for breast and ovarian cancer genetic risk. *J Clin Oncol* 2020 May 20;38(15\_suppl):1506. [doi: [10.1200/jco.2020.38.15\\_suppl.1506](https://doi.org/10.1200/jco.2020.38.15_suppl.1506)]
32. Manriquez E, Chapman JS, Mak J, Blanco AM, Chen L. Disparities in genetics assessment for women with ovarian cancer: can we do better? *Gynecol Oncol* 2018 Apr;149(1):84-88. [doi: [10.1016/j.ygyno.2017.10.034](https://doi.org/10.1016/j.ygyno.2017.10.034)] [Medline: [29605055](https://pubmed.ncbi.nlm.nih.gov/29605055/)]
33. Langford AT, Hawley ST, Stableford S, Studts JL, Byrne MM. Development of a plain language decision support tool for cancer clinical trials: blending health literacy, academic research, and minority patient perspectives. *J Cancer Educ* 2020 Jun 09;35(3):454-461. [doi: [10.1007/s13187-019-1482-5](https://doi.org/10.1007/s13187-019-1482-5)] [Medline: [30739270](https://pubmed.ncbi.nlm.nih.gov/30739270/)]

## Abbreviations

**MAGENTA:** Making Genetic Testing Accessible

**REDCap:** Research Electronic Data Capture

**UX:** user experience

*Edited by A Mavragani; submitted 05.10.21; peer-reviewed by S Nelson, H Kim; comments to author 21.12.21; revised version received 27.01.22; accepted 28.01.22; published 17.10.22.*

*Please cite as:*

*Coffin T, Bowen D, Swisher E, Lu K, Rayes N, Norquist B, Blank S, Levine D, Bakkum-Gamez J, Fleming G, Olopade O, D'Andrea A, Nebgen D, Peterson C, Munsell M, Gavin K, Lechner R, Crase J, Polinsky D, Romero I*

*An Accessible Communication System for Population-Based Genetic Testing: Development and Usability Study*

*JMIR Form Res* 2022;6(10):e34055

URL: <https://formative.jmir.org/2022/10/e34055>

doi: [10.2196/34055](https://doi.org/10.2196/34055)

PMID: [36251350](https://pubmed.ncbi.nlm.nih.gov/36251350/)

©Tara Coffin, Deborah Bowen, Elizabeth Swisher, Karen Lu, Nadine Rayes, Barbara Norquist, Stephanie Blank, Douglas Levine, Jamie Bakkum-Gamez, Gini Fleming, Olufunmilayo Olopade, Alan D'Andrea, Denise Nebgen, Christine Peterson, Mark Munsell, Kathleen Gavin, Rebecca Lechner, Jamie Crase, Deborah Polinsky, Iris Romero. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 17.10.2022. 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

# Readiness for Telemedical Services in Patients With Cardiovascular Diseases: Cross-sectional Study

Barbara Kalańska-Łukasik<sup>1</sup>; Aleksandra Gładyś<sup>2</sup>; Tomasz Jadczyk<sup>1,3</sup>, MD, PhD; Monika Gruz-Kwapisz<sup>1</sup>; Wojciech Wojakowski<sup>1</sup>, MD, PhD; Malgorzata Kowalska<sup>2</sup>, MD, PhD

<sup>1</sup>Department of Cardiology and Structural Heart Diseases, Medical University of Silesia, Katowice, Poland

<sup>2</sup>Department of Epidemiology, Faculty of Medical Sciences in Katowice, Medical University of Silesia, Katowice, Poland

<sup>3</sup>International Clinical Research Center, Interventional Cardiac Electrophysiology Group, St Anne's University Hospital, Brno, Czech Republic

**Corresponding Author:**

Barbara Kalańska-Łukasik

Department of Cardiology and Structural Heart Diseases

Medical University of Silesia

Ziołowa 47

Katowice, 40-635

Poland

Phone: 1 32 359 86 90

Fax: 1 32 359 88 87

Email: [b.kalanska.lukasik@gmail.com](mailto:b.kalanska.lukasik@gmail.com)

## Abstract

**Background:** Telemedicine solutions, especially in the face of epidemiological emergencies such as the COVID-19 pandemic, played an important role in the remote communication between patients and medical providers. However, the implementation of modern technologies should rely on patients' readiness toward new services to enable effective cooperation with the physician. Thus, successful application of patient-centric telehealth services requires an in-depth analysis of users' expectations.

**Objective:** This study aimed to evaluate factors determining readiness for using telehealth solutions among patients with cardiovascular diseases.

**Methods:** We conducted a cross-sectional study based on an investigator-designed, validated questionnaire that included 19 items (demographics, health status, medical history, previous health care experience, expected telehealth functionalities, and preferred remote communication methods). Multivariate logistic regression was applied to assess the relationship between readiness and their determinants.

**Results:** Of the 249 respondents, 83.9% (n=209) consented to the use of telemedicine to contact a cardiologist. The nonacceptance of using telemedicine was 2 times more frequent in rural dwellers (odds ratio [OR] 2.411, 95% CI 1.003-5.796) and patients without access to the internet (OR 2.432, 95% CI 1.022-5.786). In comparison to participants living in rural areas, city dwellers demonstrated a higher willingness to use telemedicine, including following solutions: issuing e-prescriptions (19/31, 61.3% vs 141/177, 79.7%;  $P=.02$ ); alarming at the deterioration of health (18/31, 58.1% vs 135/177, 76.3%;  $P=.03$ ); and arranging or canceling medical visits (16/31, 51.6% vs 126/176, 71.6%;  $P=.03$ ). Contact by mobile phone was preferred by younger patients (OR 2.256, 95% CI 1.058-4.814), whereas older patients and individuals who had no previous difficulties in accessing physicians preferred landline phone communication.

**Conclusions:** During a nonpandemic state, 83.9% of patients with cardiovascular diseases declared readiness to use telemedicine solutions.

(*JMIR Form Res* 2022;6(10):e33769) doi:[10.2196/33769](https://doi.org/10.2196/33769)

**KEYWORDS**

telemedicine; readiness; patient-cardiologist contact; telehealth; acceptance

## Introduction

Telemedicine, as an integral component of modern health care, has been used for nearly 30 years [1]. It offers a wide range of services based on remote communication between patients and clinicians for virtual care including education, monitoring, and therapeutic interventions [2,3]. Importantly, telehealth is recognized by professional medical societies such as the European Society of Cardiology and the American College of Cardiology [4], which emphasizes its role in everyday clinical practice. The emergence of new guidelines supports the implementation of virtual care for patients with cardiovascular diseases (CVD) [5], with additional focus on web-based medical education [6] and training for health care professionals (ie, undergraduate educational modulus in the field of telemedicine and telehealth licenses) [7,8]. The aforementioned examples illustrate the growing importance of telehealth, which has an impact on the health care ecosystem including patients, providers, and medical institutions.

Telemedicine covers a wide range of synchronous (live) and asynchronous (store and forward) services, including home telemonitoring, remote measurements of vital signs, medical consultations over a telephone (traditional landline and mobile phones), communication via email, and dedicated patient's portals, as well as the implementation of voice technology and smart speakers [9,10]. However, telemedical care should be tailored to patients' needs, reflecting the complex aspects of readiness and willingness to use medical technologies. Accordingly, user experience translates into the preferences of individuals whose health status is affected by remote care. This reasoning is why telemedicine should answer the needs of specific medical domains, in which patients with CVD constitute a heterogeneous group in terms of diseases as well as electronic literacy. Of note, there are smartphone-proficient older adults as well as patients with many comorbidities for whom being digitally connected might be challenging [11]. Furthermore, potential barriers associated with telehealth (geographic: limited internet access in rural areas; financial: expensive internet connection; and sociological: no face-to-face contact and a lack of equipment compatibility) may create real-world problems [12]. Accordingly, understanding patients' capabilities and, consequently, readiness for the novel technology used in the telemedical practice is a crucial step toward the development of successful remote care programs. Notably, the COVID-19 pandemic created a kind of "telemedicine boom" that was caused by the forced need to implement temporary telemedicine solutions, which were used by almost 100% of US patients [13]. Therefore, to exclude bias and focus on long-term sustainable applications, it seems necessary to assess patients' opinions toward virtual medical services during a nonpandemic state. In the aim to provide a basis for a successful patient-specific telehealth design, we evaluated factors affecting preferences and readiness for virtual care in patients with CVD.

## Methods

### Study Design

Between March 2019 and January 2020, 249 patients were enrolled in this epidemiological cross-sectional study, representing approximately 10% of the yearly hospitalized population at the Department of Cardiology and Structural Heart Diseases, Medical University of Silesia in Katowice, Poland. Respondents completed questionnaires while being assisted by a physician. The inclusion criterion was informed consent to participate in the study. Other factors, such as multiple morbidities, reasons for hospitalization, age, and gender, did not affect recruitment. The exclusion criteria were the lack of consent due to personal preferences or severe clinical condition. The patients were informed about the possibility to take part in the study on the first day of hospitalization.

### Instrumentation

To evaluate the willingness and preferences of patients with CVD toward telemedicine, a 19-item validated questionnaire was designed at the Department of Epidemiology, Medical University of Silesia in Katowice, Poland. A detailed description of the research tool was presented in a publication by Kowalska et al [14]. Briefly, the questionnaire included questions about sociodemographic data, medical history, and potential hindrances while contacting cardiologists remotely. The obtained results of the validation procedure confirmed the usefulness of the questionnaire as the key questions had high repeatability, ranging from 80% to 100% (Cohen  $\kappa$  statistics ranged from 0.419 to 0.920).

### Ethics Approval

Ethics approval for this study was received from the Bioethical Committee of the Medical University of Silesia in Katowice, Poland (KNW/0022/KB1/160/1617) on February 3, 2017.

### Statistical Analysis

Statistical analysis was performed using the Statistica software (version 13.0; Dell Software Inc). The missing values were removed from the final database. The qualitative variables were presented by frequency and percentage. Simple tests (chi-square or Fisher test) were used to assess the differences between independent groups of patients. In the interpretation of the results,  $P$  values  $<.05$  were considered statistically significant. Finally, the relevant relationships between particular variables were verified in the multivariable analysis (logistic regression models with Hooke-Jeeves and quasi-Newton estimation). Only the statistically significant variables obtained in bivariate analyses were included in the models. The result section presents adequately the goodness of fit of the used model (confirmed by chi-square test and its  $P$  value).

## Results

The total studied group included 249 patients aged 65.3 (SD 13.8) years; more than half ( $n=158$ , 63.5%) were male, and the majority ( $n=211$ , 84.7%) were city dwellers. The vast majority ( $n=209$ , 83.9%) of patients reported readiness for telemedicine solutions, whereas 34 (13.6%) patients were opposed, and 6

(2.4%) did not respond to the question. Further multivariate analysis was carried out on a group of 202 (81.1%) patients for which a complete set of answers was obtained.

Men and people with previous difficulties in accessing medical doctors more frequently declared readiness for telemedicine solutions ( $P=.006$  and  $P<.001$ , respectively). The other independent variables had no statistically significant impact on readiness for telemedicine (Table 1). The analysis deliberately omitted race and ethnicity because the study group was homogeneously of the White race.

In comparison to participants living in rural areas, city dwellers demonstrated a higher willingness to use telemedicine solutions in a particular form such as issuing e-prescriptions (19/31, 61.3%

vs 141/177, 79.7%;  $P=.02$ ); alarming at the deterioration of health (18/31, 58.1% vs 135/177, 76.3%;  $P=.03$ ); and arranging or canceling medical visits (16/31, 51.6% vs 126/176, 71.6%;  $P=.03$ ). Furthermore, a significant correlation was found between the level of education and willingness to use specific telehealth services; patients reporting a secondary level of education showed almost complete (58/64, 90.6%) compliance with solutions to control blood pressure, temperature, and bodyweight ( $P=.03$ ), as well as e-prescriptions service ( $P=.01$ ). In turn, respondents without internet access showed the lowest interest in arranging or canceling medical visits (39/67, 58.2% vs 102/138, 73.9%;  $P=.02$ ) or issuing e-prescriptions (44/67, 65.7% vs 115/139, 82.7%;  $P=.006$ ; Table 2).

**Table 1.** Frequency of positive patients' declaration toward telemedicine according to particular determinants ( $P$  value in chi-square test).

Determinant	Positive patient declaration, n/N (%)	$P$ value
<b>Sex</b>		.006
Female	66/85 (77.6)	
Male	142/157 (90.4)	
<b>Age</b>		.053
Older (aged $\geq 68$ years)	98/120 (81.7)	
Younger (aged $< 68$ years)	111/123 (90.2)	
<b>Place of residence</b>		.68
City	177/205 (86.3)	
Rural areas	31/37 (83.8)	
<b>Education</b>		.20
Primary	88/108 (81.5)	
Secondary	62/70 (88.6)	
Higher	57/63 (90.5)	
<b>Internet access</b>		.16
Yes	138/156 (88.5)	
No	68/83 (81.9)	
<b>Previous difficulties in accessing medical doctors</b>		<.001
Yes	123/131 (93.4)	
No	85/109 (77.9)	

**Table 2.** Factors influencing the readiness for telemedicine services of patients with cardiovascular diseases (percentage of declaration and *P* value of chi-square test).

Independent variable	Telemedicine services accepted by patients											
	Remote contact with a cardiologist, n/N (%)	<i>P</i> value	Telemonitoring of vital signs (blood pressure, temperature, and body-weight), n/N (%)	<i>P</i> value	Issuing e-prescriptions, n/N (%)	<i>P</i> value	Alarming health status deterioration, n/N (%)	<i>P</i> value	Scheduling and managing of medical visits, n/N (%)	<i>P</i> value	Medication reminder, n/N (%)	<i>P</i> value
<b>Sex</b>		.17		.30		.87		.08		.29		.18
Female (n=90)	53/67 (79.1)		51/67 (76.1)		52/67 (77.6)		44/67 (65.7)		42/66 (63.6)		29/66 (43.9)	
Male (n=158)	122/141 (86.5)		116/141 (82.3)		108/141 (76.6)		109/141 (77.3)		100/141 (70.9)		76/141 (53.9)	
<b>Previous difficulties accessing cardiologists</b>		.42		<.001		.22		.12		.99		.39
Yes (n=133)	107/124 (86.3)		109/124 (87.9)		99/124		96/124 (77.4)		85/123 (69.1)		66/123 (53.7)	
No (n=112)	69/84 (82.1)		58/84 (69)		61/84 (72.6)		57/84 (67.9)		58/84 (69)		40/84 (47.6)	
<b>Living with family</b>		.61		.81		.10		.38		.67		.30
Yes (n=193)	141/168 (83.9)		135/168 (80.4)		133/168 (79.2)		122/168 (72.6)		114/167 (68.3)		83/167 (49.7)	
No (n=53)	34/39 (87.2)		32/39 (82.0)		26/39 (66.7)		31/39 (79.5)		28/39 (71.8)		23/39 (58.9)	
<b>Age</b>		.64		.89		.18		.40		.22		.21
Younger (aged <68 years; n=126)	93/109 (85.3)		88/109 (80.7)		88/109 (80.7)		83/109 (76.1)		79/109 (72.5)		51/109 (46.8)	
Older (aged ≥68 years; n=123)	83/100 (83)		80/100 (80)		73/100 (73)		71/100 (71)		64/99 (64.6)		55/99 (55.6)	
<b>Place of residence</b>		.10		.16		.02		.03		.03		.06
City (n=211)	155/177 (85.9)		145/177 (81.9)		141/177 (79.7)		135/177 (76.3)		126/176 (71.6)		95/176 (53.9)	
Rural areas (n=37)	23/31 (74.2)		22/31 (70.9)		19/31 (61.3)		18/31 (58.1)		16/31 (51.6)		11/31 (35.5)	
<b>Educational level</b>		.79		.03		.01		.16		.26		.08
Primary (n=111)	71/86 (82.6)		63/86 (73.3)		59/86 (68.6)		60/86 (69.8)		58/86 (67.4)		48/86 (55.8)	
Secondary (n=73)	55/64 (90.6)		58/64 (90.6)		57/64 (89.1)		53/64 (82.8)		49/64 (76.6)		36/64 (43.7)	
Higher (n=63)	50/58 (81)		47/58 (81)		44/58 (75.9)		41/58 (70.7)		36/57 (63.2)		35/57 (38.6)	

Independent variable	Telemedicine services accepted by patients											
	Remote contact with a cardiologist, n/N (%)	<i>P</i> value	Telemonitoring of vital signs (blood pressure, temperature, and body-weight), n/N (%)	<i>P</i> value	Issuing e-prescriptions, n/N (%)	<i>P</i> value	Alarming health status deterioration, n/N (%)	<i>P</i> value	Scheduling and managing of medical visits, n/N (%)	<i>P</i> value	Medication reminder, n/N (%)	<i>P</i> value
<b>Internet access</b>		.61		.71		.006		.04		.02		.62
Yes (n=158)	118/139 (84.9)		113/139 (81.3)		115/139 (82.7)		108/139 (77.7)		102/138 (73.9)		69/138 (50)	
No (n=87)	55/67 (82.1)		53/67 (79.1)		44/67 (65.7)		43/67 (64.2)		39/67 (58.2)		36/67 (53.7)	

Table S1 in [Multimedia Appendix 1](#) presents the results of the multivariate analysis of the relationship between readiness for virtual care applications and demographic or socioeconomic determinants. Patients who had no previous difficulties in contact with a cardiologist were more than 3 times less likely to use telemedicine for vital signs measurement (odds ratio [OR] 3.596, 95% CI 1.681-7.690). The lack of acceptance for issuing e-prescriptions was 2 times more frequent in rural dwellers (OR 2.411, 95% CI 1.003-5.796) and in patients with no access to the internet (OR 2.432, 95% CI 1.022-5.786). Similarly, participants living in rural areas and individuals without internet connection were 2 times less likely to implement telemedicine for alarming the deterioration of health and for managing medical visits. On the contrary, city dwellers and patients with a lower level of education reported willingness to use telemedicine for medication reminder.

[Table 3](#) presents patients' declarations of readiness for communication modalities in particular groups of subjects defined by sociodemographic determinants. Patients' preferences for face-to-face contact with physicians significantly differed based on the age of respondents; older individuals (aged  $\geq 68$  years) accepted this solution more frequently than younger participants (aged  $< 68$  years)—34.3% (34/99) versus 12.1% (13/107), respectively. Similarly, older people and patients with a lower level of education preferred contact by landline phone. Mobile phone contact was preferred by younger patients and individuals with access to the internet. Younger patients and individuals living in the city were more likely to use email and web services to contact cardiologists.

The results of the multivariate analysis revealed that particular opinions about telehealth communication solutions varied on different determinants (Table S2 in [Multimedia Appendix 1](#)).



**Table 3.** Factors influencing the readiness for telemedicine communication modalities of patients with cardiovascular diseases (percentage of declaration and *P* value of chi-square or Fisher test).

Independent variable	Preferred type of contact with physician									
	Face-to-face, n/N (%)	<i>P</i> value	Landline phone, n/N (%)	<i>P</i> value	Mobile phone, n/N (%)	<i>P</i> value	Email contact, n/N (%)	<i>P</i> value	Web page, n/N (%)	<i>P</i> value
<b>Sex</b>		.77		.98		.29		.01		.08
Male (n=90)	14/66 (21.2)		35/66 (53)		51/66 (77.3)		7/66 (10.6)		4/66 (6.1)	
Female (n=158)	32/139 (23)		74/139 (53.2)		116/139 (83.4)		36/139 (25.9)		19/139 (13.7)	
<b>Previous difficulties accessing cardiologists</b>		.20		.06		.049		.44		.42
Yes (n=133)	32/123 (26)		72/123 (58.5)		105/123 (85.4)		28/123 (22.8)		12/123 (9.8)	
No (n=112)	15/82 (18.3)		37/82 (45.1)		61/82 (74.4)		15/82 (18.3)		11/82 (13.4)	
<b>Living with family</b>		.62		.28		.78		.34		.30
Yes (n=193)	39/165 (23.6)		90/165 (54.5)		133/165 (80.6)		36/165 (21.8)		20/165 (12.1)	
No (n= 53)	8/40 (20)		18/40 (45)		33/40 (82.5)		6/40 (15)		3/40 (7.5)	
<b>Age</b>		<.001		.03		.003		<.001		.001 <sup>a</sup>
Younger (aged <68 years; n=126)	13/107 (12.1)		49/107 (45.8)		95/107 (88.8)		35/107 (32.7)		19/107 (17.8)	
Older (aged ≥68 years; n=123)	34/99 (34.3)		60/99 (60.6)		72/99 (72.7)		8/99 (8.1)		4/99 (4)	
<b>Place of residence</b>		.68		.94		.52		.03 <sup>a</sup>		.02 <sup>a</sup>
City (n=211)	41/175 (23.4)		92/175 (52.6)		143/175 (81.7)		40/175 (22.9)		23/175 (13.1)	
Rural areas (n=37)	6/30 (20)		16/30 (53.3)		23/30 (76.7)		2/30 (6.7)		0/30 (0)	
<b>Educational level</b>		.39		.01		.29		<.001		<.001
Primary (n=111)	23/88 (26.1)		49/88 (55.7)		68/88 (77.3)		7/88 (7.9)		1/88 (1.1)	
Secondary (n=73)	15/63 (23.8)		40/63 (87.3)		55/63 (87.3)		18/63 (28.6)		10/63 (15.9)	
Higher (n=63)	9/55 (16.4)		20/55 (36.4)		44/55 (80)		18/55 (32.7)		12/55 (21.8)	
<b>Internet access</b>		.047		.52		.005		<.001 <sup>a</sup>		<.001 <sup>a</sup>
Yes (n=158)	25/135 (18.5)		69/135 (51.1)		117/135 (86.7)		41/135 (30.4)		23/135 (17)	
No (n=87)	21/68 (30.9)		38/68 (55.9)		48/68 (70.6)		1/68 (1.5)		0/68 (0)	

<sup>a</sup>Result of Fisher test.

## Discussion

### Principal Findings

The main group of patients enrolled the study included individuals aged >65 years, and those participants largely accepted telemedicine solutions (83.9%). Importantly, our findings are in line with previous findings reported in the literature [15,16]. The observed upward trend in the acceptance

and, consequently, the consent of older adults for the daily use of a wide range of telemedicine devices was strongly emphasized [12,16]. Importantly, factors such as technical adaptation (previous training and patients' own experiences) is closely associated with the acceptance of telemedicine [17]. Therefore, gerontechnology (technology for aging populations) must be well designed and suited for health self-controlling [18], especially as improvement in health status is not always felt by patients using telemedicine services. As an example,

patients with heart failure who received telehealth during the COVID-19 pandemic did not improve their health condition [19]. Thus, it is crucial to understand patients' expectations and realistic opportunities to implement virtual care in clinical practice.

We report that patients who had no previous difficulties in accessing cardiologists were against using the home telemonitoring of vital signs (blood pressure and weight measurement). Interestingly, the results of studies conducted during the COVID-19 pandemic showed that remote monitoring is warranted, and the patient's results at home (eg, a 6-minute walk test on a smart watch and the self-measurement of blood pressure) were comparable to those obtained during medical appointments [20,21]. Notably, the patients responded to our survey at a time when there was no epidemiological threat in Poland (before January 2020). Perhaps, the COVID-19 sanitary restrictions and severely limited access to health services leveraged patients' support for telemedicine.

We observed that rural dwellers and patients without access to the internet were 2 times more frequently opposed to using telemedicine for issuing e-prescriptions. Given territorial sociodemographics, the juxtaposition of these 2 patient groups is not coincidental, especially in Poland where the division between rural and urban areas continues to matter (eg, limits concerning economy, education, and access to the internet) [22]. In Poland, only 50% of rural households have access to the internet, with an approximate rate of 70% in the European Union [23,24]. On the contrary, the study by Shemesh and Barnoy [20] conducted on Israel's population proved that there are no significant sociodemographic differences in the use of mobile health apps. However, the authors strongly emphasized the specificity of this region (a technologically developed country with a high acceptance rate of innovation across the sociodemographic gradient) that could have impacted the results of the study [25].

In our study, there was a lack of acceptance for telemedicine solutions for the alarming the provider in the case of health status deterioration and in the case of arranging or canceling medical visits. The answers given in our survey, despite the effort put into explaining each option of telemedicine services (except self-testing), may be associated with a lack of understanding and awareness of the functions that a telemedicine device can perform. Moreover, the patients might have subconsciously chosen a lower number of possibilities in the survey, as they were scared of technological difficulties in handling a large number of applications.

We reported that medicine reminders are significantly more frequent in urban areas and in patients with lower levels of education. Many studies have proved that irregular adherence to drug use is an important problem, and the most common reason for failing to achieve a therapeutic effect is the drug's spontaneous reduction, which occurs in about 30% to 35% of patients [26,27]. Nowadays, not only rural but also urban areas are significantly associated with a higher percentage of people with lower levels of education, which may constitute a higher need for self-control in their treatments [28]. As the previous

study shows [7], a lower level of education does not lead to a lack of acceptance, but it is associated with even better satisfaction after the introduction of telemedicine services.

Patients' preferences for direct contact with physicians are statistically significant and the most frequent in older people (aged  $\geq 68$  years). Patients who gained trust in medical personnel (physicians and nurses) were more willing to use telemedical devices as observed in the previous study [29]. Dario et al [18] reported that older patients tend to trust clinicians with whom they are already familiar. This finding is one of the reasons why during the COVID-19 pandemic, there was an urgent need to train doctors in the field of telemedicine, so that they would help their patients in servicing eHealth [30]. Trust in the doctor was crucial in the process of learning about new telecommunications.

Similarly, older people and patients with no previous difficulties accessing cardiologists prefer contact by landline phone. Traditional landlines, although currently not popular, are still used by older adults due to several advantages; for example, they are not sensitive to network coverage, are easy to use, and are cheaper in the case of international phone calls. Analogously, the previous studies show the importance of easy-to-use technology [16], whereas Scheibe et al [31] conclude that design features such as a simple intuitive menu, large icons, and high color contrast are especially important for older users. On the contrary, contact by mobile phone was preferred by younger patients and people who already have access to a cardiologist. Our findings are consistent with previous studies, in which young individuals were more eager to use telemedicine as they are more familiar with new technologies. Therefore, it could be concluded that experience in the use of mobile technologies, not age, is the main indicator of the willingness to use virtual care solutions [11].

### Limitations

The number of respondents participating in the study accounted for approximately 10% of the total number of patients hospitalized yearly at the Department of Cardiology, which is the main limitation of the presented results. However, the group of respondents is representatives of patients with CVD who statistically share common characteristics (age and gender). Furthermore, in the future, it is worth extending the survey to include the opinions of respondents who experienced telehealth during the COVID-19 pandemic and can verify which type of telemedicine solution is the most convenient.

### Conclusion

Patients with CVD are ready to accept remote solutions to contact with a cardiologist in clinical practice. They seem to be at least mostly aware of the needs and ready-made solutions that can make their everyday life easier. This finding confirms the fact that patients with CVD—mainly older adults—are often familiar with modern communication systems, which have become a natural component of daily life. Therefore, after identifying patients' preferences associated with telehealth, the possibility of implementing user-friendly and well-designed telecommunication methods should be further explored.

## Conflicts of Interest

None declared.

Multimedia Appendix 1

Multivariate analyses.

[[DOCX File, 25 KB - formative\\_v6i10e33769\\_app1.docx](#)]

## References

1. Office for the Advancement of Telehealth. Health Resources & Services Administration. 2022 Jul. URL: <https://www.hrsa.gov/rural-health/topics/telehealth> [accessed 2022-09-27]
2. Shaw DK. Overview of telehealth and its application to cardiopulmonary physical therapy. *Cardiopulm Phys Ther J* 2009 Jun;20(2):13-18 [FREE Full text] [doi: [10.1097/01823246-200920020-00003](https://doi.org/10.1097/01823246-200920020-00003)] [Medline: [20467533](#)]
3. Masson M. Benefits of TED Talks. *Can Fam Physician* 2014 Dec;60(12):1080 [FREE Full text] [Medline: [25500595](#)]
4. ASCVD telehealth management. American College of Cardiology. URL: <https://www.acc.org/Tools-and-Practice-Support/Clinical-Toolkits/Telehealth-Workbook> [accessed 2022-09-27]
5. Vidal-Perez R. eHealth and COVID-19: time for telecardiology. *e-Journal of Cardiology Practice* 2021 Oct 20;21:6 [FREE Full text]
6. Task Force for the management of COVID-19 of the European Society of Cardiology. ESC guidance for the diagnosis and management of cardiovascular disease during the COVID-19 pandemic: part 2-care pathways, treatment, and follow-up. *Eur Heart J* 2022 Mar 14;43(11):1059-1103 [FREE Full text] [doi: [10.1093/eurheartj/ehab697](https://doi.org/10.1093/eurheartj/ehab697)] [Medline: [34791154](#)]
7. Keswani A, Brooks JP, Khoury P. The future of telehealth in allergy and immunology training. *J Allergy Clin Immunol Pract* 2020 Jul;8(7):2135-2141 [FREE Full text] [doi: [10.1016/j.jaip.2020.05.009](https://doi.org/10.1016/j.jaip.2020.05.009)] [Medline: [32426217](#)]
8. Jumreornvong O, Yang E, Race J, Appel J. Telemedicine and medical education in the age of COVID-19. *Acad Med* 2020 Dec;95(12):1838-1843 [FREE Full text] [doi: [10.1097/ACM.0000000000003711](https://doi.org/10.1097/ACM.0000000000003711)] [Medline: [32889946](#)]
9. Barrios V, Cosín-Sales J, Bravo M, Escobar C, Gámez JM, Huelmos A, Cequier. Telemedicine consultation for the clinical cardiologists in the era of COVID-19: present and future. consensus document of the Spanish Society of Cardiology. Article in Spanish. *Rev Esp Cardiol* 2020 Nov;73(11):910-918 [FREE Full text] [doi: [10.1016/j.recesp.2020.06.027](https://doi.org/10.1016/j.recesp.2020.06.027)] [Medline: [32836664](#)]
10. Jadczyk T, Kiwic O, Khandwalla RM, Grabowski K, Rudawski S, Magaczewski P, et al. Feasibility of a voice-enabled automated platform for medical data collection: CardioCube. *Int J Med Inform* 2019 Sep;129:388-393 [FREE Full text] [doi: [10.1016/j.ijmedinf.2019.07.001](https://doi.org/10.1016/j.ijmedinf.2019.07.001)] [Medline: [31445282](#)]
11. Hong Z, Li N, Li D, Li J, Li B, Xiong W, et al. Telemedicine during the COVID-19 pandemic: experiences from Western China. *J Med Internet Res* 2020 May 08;22(5):e19577 [FREE Full text] [doi: [10.2196/19577](https://doi.org/10.2196/19577)] [Medline: [32349962](#)]
12. Phadke NA, Del Carmen MG, Goldstein SA, Vagle J, Hidrue MK, Botti ES, et al. Trends in ambulatory electronic consultations during the COVID-19 pandemic. *J Gen Intern Med* 2020 Oct;35(10):3117-3119 [FREE Full text] [doi: [10.1007/s11606-020-05878-z](https://doi.org/10.1007/s11606-020-05878-z)] [Medline: [32424786](#)]
13. Demeke HB, Pao LZ, Clark H, Romero L, Neri A, Shah R, et al. Telehealth practice among health centers during the COVID-19 pandemic - United States, July 11-17, 2020. *MMWR Morb Mortal Wkly Rep* 2020 Dec 18;69(50):1902-1905 [FREE Full text] [doi: [10.15585/mmwr.mm6950a4](https://doi.org/10.15585/mmwr.mm6950a4)] [Medline: [33332297](#)]
14. Kowalska M, Gładys A, Kalańska-Łukasik B, Gruz-Kwapisz M, Wojakowski W, Jadczyk T. Readiness for voice technology in patients with cardiovascular diseases: cross-sectional study. *J Med Internet Res* 2020 Dec 17;22(12):e20456 [FREE Full text] [doi: [10.2196/20456](https://doi.org/10.2196/20456)] [Medline: [33331824](#)]
15. Petrovcic A, Boot W, Burnik T, Dolnicar V. Improving the measurement of older adults' mobile device proficiency: results and implications from a study of older adult smartphone users. *IEEE Access* 2019 Oct 16;7:150412-150422 [FREE Full text] [doi: [10.1109/access.2019.2947765](https://doi.org/10.1109/access.2019.2947765)]
16. Harst L, Lantzsch H, Scheibe M. Theories predicting end-user acceptance of telemedicine use: systematic review. *J Med Internet Res* 2019 May 21;21(5):e13117 [FREE Full text] [doi: [10.2196/13117](https://doi.org/10.2196/13117)] [Medline: [31115340](#)]
17. Koivumäki T, Pekkarinen S, Lappi M, Väisänen J, Juntunen J, Pikkarainen M. Consumer adoption of future MyData-based preventive eHealth services: an acceptance model and survey study. *J Med Internet Res* 2017 Dec 22;19(12):e429 [FREE Full text] [doi: [10.2196/jmir.7821](https://doi.org/10.2196/jmir.7821)] [Medline: [29273574](#)]
18. Dario C, Luisotto E, Dal Pozzo E, Mancin S, Aletras V, Newman S, et al. Assessment of patients' perception of telemedicine services using the Service User Technology Acceptability Questionnaire. *Int J Integr Care* 2016 Jun 01;16(2):13 [FREE Full text] [doi: [10.5334/ijic.2219](https://doi.org/10.5334/ijic.2219)] [Medline: [27616968](#)]
19. Ploux S, Strik M, Abu-Alrub S, Ramirez FD, Buliard S, Marchand H, et al. Remote monitoring of patients with heart failure during the first national lockdown for COVID-19 in France. *Eur Heart J Dig Health* 2021 Sep;2(3):487-493 [FREE Full text] [doi: [10.1093/ehjdh/ztab044](https://doi.org/10.1093/ehjdh/ztab044)]

20. Shemesh T, Barnoy S. Assessment of the intention to use mobile health applications using a Technology Acceptance Model in an Israeli adult population. *Telemed J E Health* 2020 Sep;26(9):1141-1149. [doi: [10.1089/tmj.2019.0144](https://doi.org/10.1089/tmj.2019.0144)] [Medline: [31930955](https://pubmed.ncbi.nlm.nih.gov/31930955/)]
21. Gijón-Conde T, Rubio E, Gorostidi M, Vinyoles E, Armario P, Rodilla E, et al. 2021 Spanish Society of Hypertension position statement about telemedicine. *Hipertens Riesgo Vasc* 2021 Oct;38(4):186-196 [FREE Full text] [doi: [10.1016/j.hipert.2021.03.003](https://doi.org/10.1016/j.hipert.2021.03.003)] [Medline: [33888438](https://pubmed.ncbi.nlm.nih.gov/33888438/)]
22. Komunikat z badań: Korzystanie z internetu. Research communication: internet use. Centrum Badania Opinii Społecznej. 2020. URL: [https://www.cbos.pl/SPISKOM.POL/2020/K\\_085\\_20.PDF](https://www.cbos.pl/SPISKOM.POL/2020/K_085_20.PDF) [accessed 2022-09-27]
23. The Digital Economy and Society Index (DESI). European Commission. URL: <https://digital-strategy.ec.europa.eu/en/policies/desi> [accessed 2022-09-27]
24. The Digital Economy and Society Index (DESI) of the European Union. eTrade for all. 2019 Jun 11. URL: <https://etradeforall.org/news/the-digital-economy-and-society-index-desi-of-the-european-union/> [accessed 2022-09-27]
25. Greiwe J. Telemedicine lessons learned during the COVID-19 pandemic. *Curr Allergy Asthma Rep* 2022 Jan;22(1):1-5 [FREE Full text] [doi: [10.1007/s11882-022-01026-1](https://doi.org/10.1007/s11882-022-01026-1)] [Medline: [35061150](https://pubmed.ncbi.nlm.nih.gov/35061150/)]
26. Derkacz M, Flisek J, Wicha J, Tarach J. Jak pacjenci stosują się do zaleceń lekarskich dotyczących przyjmowania zaordynowanych leków? How do patients follow the doctor's instructions as far as taking the prescribed medicines is concerned? Article in Polish. *Family Medicine & Primary Care Review* 2014 Apr;16(2):91-93 [FREE Full text]
27. Wang J, Li Y, Chia Y, Cheng H, Minh HV, Siddique S, Hypertension Cardiovascular Outcome Prevention, Evidence (HOPE) Asia Network. Telemedicine in the management of hypertension: evolving technological platforms for blood pressure telemonitoring. *J Clin Hypertens (Greenwich)* 2021 Mar;23(3):435-439 [FREE Full text] [doi: [10.1111/jch.14194](https://doi.org/10.1111/jch.14194)] [Medline: [33484617](https://pubmed.ncbi.nlm.nih.gov/33484617/)]
28. Singh A, Mountjoy N, McElroy D, Mittal S, Al Hemyari B, Coffey N, et al. Patient perspectives with telehealth visits in cardiology during COVID-19: online patient survey study. *JMIR Cardio* 2021 Jan 22;5(1):e25074 [FREE Full text] [doi: [10.2196/25074](https://doi.org/10.2196/25074)] [Medline: [33385234](https://pubmed.ncbi.nlm.nih.gov/33385234/)]
29. Chang CP. The Technology Acceptance Model and its application in a telehealth program for the elderly with chronic illnesses. Article in Chinese. *Hu Li Za Zhi* 2015 Jun;62(3):11-16. [doi: [10.6224/JN.62.3.11](https://doi.org/10.6224/JN.62.3.11)] [Medline: [26073951](https://pubmed.ncbi.nlm.nih.gov/26073951/)]
30. Brørs G, Norman C, Norekvål TM. Accelerated importance of eHealth literacy in the COVID-19 outbreak and beyond. *Eur J Cardiovasc Nurs* 2020 Aug;19(6):458-461 [FREE Full text] [doi: [10.1177/1474515120941307](https://doi.org/10.1177/1474515120941307)] [Medline: [32667217](https://pubmed.ncbi.nlm.nih.gov/32667217/)]
31. Scheibe M, Reichelt J, Bellmann M, Kirch W. Acceptance factors of mobile apps for diabetes by patients aged 50 or older: a qualitative study. *Med 2.0* 2015 Mar 02;4(1):e1 [FREE Full text] [doi: [10.2196/med20.3912](https://doi.org/10.2196/med20.3912)] [Medline: [25733033](https://pubmed.ncbi.nlm.nih.gov/25733033/)]

## Abbreviations

**CVD:** cardiovascular diseases

**OR:** odds ratio

*Edited by A Mavragani; submitted 22.09.21; peer-reviewed by MDG Pimentel, D Madhusudhan; comments to author 21.12.21; revised version received 03.05.22; accepted 27.06.22; published 18.10.22.*

*Please cite as:*

*Kalańska-Łukasik B, Gładyś A, Jadczyk T, Gruz-Kwapisz M, Wojakowski W, Kowalska M  
Readiness for Telemedical Services in Patients With Cardiovascular Diseases: Cross-sectional Study  
JMIR Form Res 2022;6(10):e33769  
URL: <https://formative.jmir.org/2022/10/e33769>  
doi: [10.2196/33769](https://doi.org/10.2196/33769)  
PMID: [36256834](https://pubmed.ncbi.nlm.nih.gov/36256834/)*

©Barbara Kalańska-Łukasik, Aleksandra Gładyś, Tomasz Jadczyk, Monika Gruz-Kwapisz, Wojciech Wojakowski, Małgorzata Kowalska. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 18.10.2022. 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 Smartphone Physical Activity App for Patients in Alcohol Treatment: Single-Arm Feasibility Trial

Ana M Abrantes<sup>1,2</sup>, PhD; Lidia Z Meshesha<sup>3</sup>, PhD; Claire E Blevins<sup>1,2†</sup>, PhD; Cynthia L Battle<sup>1,2</sup>, PhD; Clifford Lindsay<sup>4</sup>, PhD; Eliza Marsh<sup>1</sup>, BA; Sage Feltus<sup>1</sup>, BA; Matthew Buman<sup>5,6</sup>, PhD; Emmanuel Agu<sup>6</sup>, PhD; Michael Stein<sup>7</sup>, MD

<sup>1</sup>Butler Hospital, Providence, RI, United States

<sup>2</sup>Alpert Medical School of Brown University, Providence, RI, United States

<sup>3</sup>Department of Psychology, University of Central Florida, Orlando, FL, United States

<sup>4</sup>UMass Chan Medical School, Worcester, MA, United States

<sup>5</sup>Arizona State University, Tempe, AZ, United States

<sup>6</sup>Worcester Polytechnic Institute, Worcester, MA, United States

<sup>7</sup>Boston University, Boston, MA, United States

†deceased

**Corresponding Author:**

Ana M Abrantes, PhD

Butler Hospital

345 Blackstone Blvd

Providence, RI, 02906

United States

Phone: 1 4014556440

Email: [Ana\\_Abrantes@Brown.edu](mailto:Ana_Abrantes@Brown.edu)

## Abstract

**Background:** Alcohol use disorder (AUD) is a significant public health concern worldwide. Alcohol consumption is a leading cause of death in the United States and has a significant negative impact on individuals and society. Relapse following treatment is common, and adjunct intervention approaches to improve alcohol outcomes during early recovery continue to be critical. Interventions focused on increasing physical activity (PA) may improve AUD treatment outcomes. Given the ubiquity of smartphones and activity trackers, integrating this technology into a mobile app may be a feasible, acceptable, and scalable approach for increasing PA in individuals with AUD.

**Objective:** This study aims to test the Fit&Sober app developed for patients with AUD. The goals of the app were to facilitate self-monitoring of PA engagement and daily mood and alcohol cravings, increase awareness of immediate benefits of PA on mood and cravings, encourage setting and adjusting PA goals, provide resources and increase knowledge for increasing PA, and serve as a resource for alcohol relapse prevention strategies.

**Methods:** To preliminarily test the Fit&Sober app, we conducted an open pilot trial of patients with AUD in early recovery (N=22; 13/22, 59% women; mean age 43.6, SD 11.6 years). At the time of hospital admission, participants drank 72% of the days in the last 3 months, averaging 9 drinks per drinking day. The extent to which the Fit&Sober app was feasible and acceptable among patients with AUD during early recovery was examined. Changes in alcohol consumption, PA, anxiety, depression, alcohol craving, and quality of life were also examined after 12 weeks of app use.

**Results:** Participants reported high levels of satisfaction with the Fit&Sober app. App metadata suggested that participants were still using the app approximately 2.5 days per week by the end of the intervention. Pre-post analyses revealed small-to-moderate effects on increase in PA, from a mean of 5784 (SD 2511) steps per day at baseline to 7236 (SD 3130) steps per day at 12 weeks (Cohen  $d=0.35$ ). Moderate-to-large effects were observed for increases in percentage of abstinent days (Cohen  $d=2.17$ ) and quality of life (Cohen  $d=0.58$ ) as well as decreases in anxiety (Cohen  $d=-0.71$ ) and depression symptoms (Cohen  $d=-0.58$ ).

**Conclusions:** The Fit&Sober app is an acceptable and feasible approach for increasing PA in patients with AUD during early recovery. A future randomized controlled trial is necessary to determine the efficacy of the Fit&Sober app for long-term maintenance of PA, ancillary mental health, and alcohol outcomes. If the efficacy of the Fit&Sober app could be established, patients with

AUD would have a valuable adjunct to traditional alcohol treatment that can be delivered in any setting and at any time, thereby improving the overall health and well-being of this population.

**Trial Registration:** ClinicalTrials.gov NCT02958280; <https://www.clinicaltrials.gov/ct2/show/NCT02958280>

(*JMIR Form Res* 2022;6(10):e35926) doi:[10.2196/35926](https://doi.org/10.2196/35926)

## KEYWORDS

alcohol use disorder; AUD; physical activity; smartphone app; Fitbit; feasibility study; mobile phone

## Introduction

### Background

Alcohol use disorder (AUD) is a public health problem that has a significant negative impact on both individuals and society [1,2]. Data from the US National Epidemiological Survey on Alcohol and Related Conditions III indicate a lifetime prevalence rate of 29.1% for the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition, diagnoses of AUD [3]. Alcohol consumption is a leading cause of death in the United States and is associated with a significant economic burden [4,5]. Despite the existing treatments for AUD, relapse rates in the first year following treatment are high [6-8]. Therefore, adjunct intervention approaches to improve alcohol outcomes during early recovery are critical.

Over the last 2 decades, physical exercise has emerged as an adjunct intervention for alcohol treatment. There are a number of biological mechanisms that can explain the potential effect of exercise on alcohol treatment outcomes. For example, exercise may normalize disrupted dopaminergic signaling in patients with AUD [9]. Through its influence on these neural reward pathways, exercise may serve as a competing reinforcer that diminishes alcohol use [10]. Furthermore, decreases in attentional bias to alcohol cues after bouts of exercise may acutely decrease alcohol cravings [11]. Another potential mechanism involves the effects of exercise on cognitive functioning, particularly executive functioning, both in the acute and long term [12,13]. Improved neurocognitive functioning can aid in treatment retention and decision-making, both of which affect alcohol relapse [14].

Owing to these potential mechanisms, there has been increased attention toward the development and testing of physical activity (PA) interventions for individuals with AUD. Indeed, this study has demonstrated that increasing PA during early recovery from AUD can be beneficial for a number of important reasons. First, patients undergoing AUD treatment express a strong interest in increasing PA as a means of helping to support their recovery [15,16]. Second, patients with AUD have significant alcohol-related health concerns (eg, hypertension, diabetes, and liver disease [17,18]) that are exacerbated by low levels of PA and ameliorated with increase in PA [19,20]. Third, negative affect associated with comorbid internalizing disorders and stressful life events is common among patients with AUD and increases the risk of self-medication and drinking to cope [21]. Both acute bouts of PA and long-term engagement in a PA program have been associated with decreases in negative affect, depression, and anxiety [22,23], including in patients with AUD [24,25]. Finally, although empirical support for the effect of

PA interventions on drinking outcomes in patients with AUD has been limited [26], evidence for the acute effect of bouts of PA, even those of relatively brief durations, on reductions in alcohol cravings has been demonstrated across a number of studies [11,27]. Therefore, strategies to increase PA in early recovery are likely to be well received and potentially helpful for patients' physical and mental health, as well as potentially promoting abstinence.

Although prior studies examining PA as an intervention approach for individuals in AUD treatment demonstrated promising outcomes in terms of cardiorespiratory and psychological functioning, and in a few cases, alcohol outcomes [26], there are also limitations associated with this existing work. Many of these studies were conducted more than 2 decades ago, and the interventions were conducted primarily in residential treatment facilities, although the majority of AUD treatments were delivered in outpatient settings [28]. Most of the interventions tested involved supervised, structured exercise programs without consideration of the participants' PA preferences. Although these types of approaches have their advantages (eg, social support), scheduling conflicts, transportation difficulties, and a restricted range of PA type may impact PA adherence. High dropout rates are common [26]. Finally, few studies have incorporated theoretically informed cognitive or behavioral features to increase motivation for PA [29]. Therefore, the development of PA interventions that can address these key limitations may lead to better outcomes in patients with AUD.

Using technology such as activity monitors and smartphones to support PA interventions has numerous advantages. For example, self-monitoring of PA is one of the most effective strategies for increasing PA [30] but is burdensome for individuals. Activity monitors and smartphone apps may ease this burden (ie, people usually have their phones with them) [31] in an efficient, interactive, and tailored manner [32]. Activity monitors and smartphone apps provide device-based PA feedback that can be used to produce individualized goal setting. Another advantage of using mobile apps to deliver PA interventions is the decreased need for formal training of providers on how to conduct PA counseling in individuals with AUD [33]. Furthermore, as smartphones are owned by over 85% of the population across most demographics [34], smartphone apps are a potentially cost-effective approach that can reach a wide number of individuals with AUD, with the ability to convey standardized therapeutic information within the app.

Although many smartphone-based PA apps are available, very few are theoretically informed or have been empirically

evaluated [35,36] and none, beyond this study, have targeted a population with AUD. Most people are motivated to exercise for health enhancement, weight loss, and appearance [37]. However, long-term engagement in PA continues to be a significant public health challenge, including for those with AUD [26]. Social cognitive theory [38] posits that behavior is increased or sustained when it is immediately reinforced, and self-determination theory [39] identifies intrinsic goals and motivation as key for long-term adherence to PA. Indeed, personally meaningful goals (rather than culturally or societally driven expectations of exercise) are critical for sustaining behavior change [40]. A PA smartphone app may help patients with AUD in early recovery to identify and support unique intrinsic sobriety-related goals and motivations that, if integrated with their PA goals, could lead to the adoption and maintenance of an exercise program. For example, by using a mobile app that facilitates self-monitoring and provides feedback, patients with AUD could increase their awareness of the immediate benefits of PA bouts on negative affect and alcohol cravings [24] and, in turn, be more likely to persist with PA.

### Objectives

The purpose of this study was to determine the feasibility and acceptability of a recently developed PA smartphone app called Fit&Sober for patients with AUD. To do so, we conducted an open trial in which the Fit&Sober app was pilot-tested for 12 weeks in a small sample of individuals with AUD in early recovery (N=22). The details on the formative work conducted in developing the Fit&Sober app are available in the study by Abrantes et al [41]. We hypothesized that participants would report that using the Fit&Sober app was feasible and acceptable, while also demonstrating increase in PA and quality of life and decrease in alcohol consumption, anxiety, depression, and alcohol craving after 12 weeks of app use.

## Methods

### Procedure and Study Design

The Alcohol and Drug Partial (ADP) hospitalization program runs every day of the week from 9 AM to 3:30 PM. ADP is an abstinence-based, relapse prevention program focused on

increasing cognitive behavioral skills for sobriety. The patients participated in daily group and individual therapy, received medication management, and were discharged with aftercare plans in place. The length of stay ranged from 5 to 10 days. Upon treatment admission, the patients' medical records were screened, and those who met the study criteria were provided with brief information about the study. Interested participants underwent a brief screen (5-10 minutes) to determine their PA levels and smartphone ownership. If eligible, the participants were scheduled for a more comprehensive baseline assessment (approximately 90 minutes) on the following day. Informed consent was obtained, and assessment measures to confirm eligibility were conducted at the baseline appointment. The study physician reviewed the patient's medical history and results of a routine physical examination conducted during partial hospitalization. Upon receiving medical clearance to exercise, participants were enrolled into the open pilot trial and scheduled for an app orientation session. All procedures were performed within 5 to 10 days of partial hospitalization. Participants were contacted at 2, 6, and 12 weeks after discharge from ADP hospitalization to provide feedback on their use of the Fit&Sober app and to complete clinical outcome measures (at 12 weeks only).

### Participants

The inclusion and exclusion criteria for the study are presented in [Textbox 1](#).

A total of 48 patients in an ADP hospitalization program in the northeast region of the United States were screened for study eligibility from December 2017 to June 2018. Of these, 42% (20/48) of participants did not meet the eligibility criteria for the following reasons: they were too physically active (11/20, 55%), did not own a smartphone (6/20, 30%), were diagnosed with a manic episode in the past 6 months (1/20, 5%), were diagnosed with a moderate or severe substance use disorder (1/20, 5%), or were not medically cleared (1/20, 5%). In addition, 8% (4/48) of participants declined to participate, and 4% (2/48) of participants did not complete all baseline procedures, leaving 22 participants who were fully eligible and enrolled in the study.

**Textbox 1.** Inclusion and exclusion criteria.**Inclusion criteria**

- Aged between 18 and 65 years
- Met the Diagnostic and Statistical Manual of Mental Disorders (DSM), Fifth Edition, criteria for alcohol use disorder as assessed by the Structured Clinical Interview for the DSM-Patient Version
- Reported low physical activity (ie, <90 minutes of moderate-intensity exercise per week for the past 6 months)
- Engaged in alcohol treatment
- Owned a smartphone

**Exclusion criteria**

- Clinical diagnosis from a medical record review of current moderate-severe substance use disorder (except nicotine), anorexia, bulimia nervosa, or mania
- History of psychotic disorder or had current psychotic symptoms
- Endorsed current suicidality or homicidality
- Physical or medical problem that would not allow safe participation in a program of moderate-intensity physical activity
- Currently pregnant or expressed intent to become pregnant during the next 12 weeks

**Ethical Considerations**

All procedures performed in the studies involving human participants were in accordance with the ethical standards of the institutional review board of Butler Hospital in Providence, Rhode Island, United States (protocol 1604-003). Informed consent was obtained from all the participants included in the study. This study was registered at ClinicalTrials.gov (NCT02958280).

**Intervention****App Orientation Session**

In preparation for the app orientation session, the research staff downloaded the Fit&Sober and Fitbit apps onto the participants' smartphones. The orientation session was conducted by a doctoral-level study clinician and lasted for approximately 30 to 40 minutes. It included brief advice on increasing PA during early recovery, providing an overview and walkthrough of Fit&Sober app components, and connecting a wrist-worn Fitbit device (Charge HR) to the Fit&Sober app. Brief advice provided participants with information on the public health guidelines for PA, the benefits of PA for physical and mental health, as well as sobriety, strategies for getting started, and instruction on gradually increasing PA. Then, participants were guided through the Fit&Sober app's setup that included entry of their sobriety date, selection of values-driven statements on reasons for increasing PA during early recovery (eg, "Being physically active will help me stay sober"), names of social support for PA and recovery, inputting of initial PA goals (ie, steps per day and minutes per week of PA), and present-moment ratings of mood and alcohol cravings. Finally, with the participants' permission, their Fitbit accounts were connected to the Fit&Sober app so that their PA data migrated to the Fit&Sober app and populated PA graphs on the Fit&Sober app's dashboard in real time. Participants were instructed to use the Fit&Sober app daily for the next 12 weeks.

**Fit&Sober App**

The development of the Fit&Sober app was guided by social cognitive theory [38] and self-determination theory [39], and a detailed description of its features and components can be found in the study by Abrantes et al [41], along with specific links between theory concepts and app features. In brief, the goals of the Fit&Sober app were to facilitate self-monitoring of PA engagement and daily mood and alcohol cravings, increase awareness of the immediate benefits of PA on mood and cravings, encourage setting and adjusting PA goals, provide resources and increase knowledge for increasing PA, and serve as a resource for alcohol relapse prevention strategies.

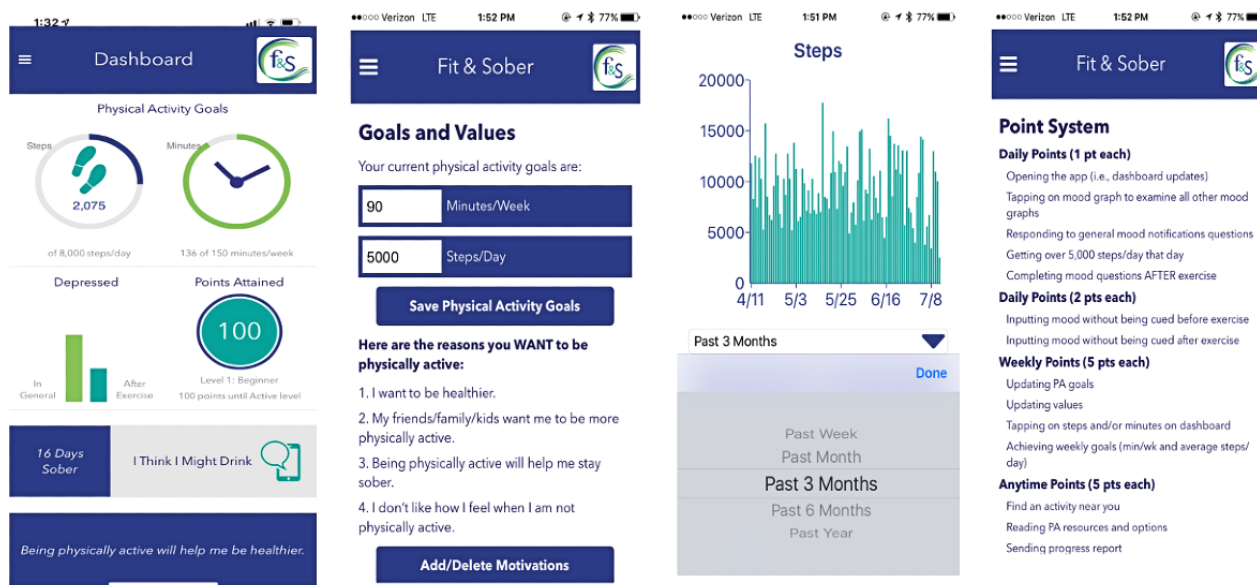
Therefore, the Fit&Sober app was designed as a readily accessible tool to aid patients in early alcohol recovery and improve treatment outcomes through increased engagement in PA. To that end, the following functionalities were available in the app: (1) real-time graphical displays of daily progress toward step per day goal and weekly progress toward minutes per week of PA; (2) personalized, values-based messages on reasons for engaging in PA during early recovery that are refreshed each time the user opens the app; (3) daily notifications to input mood and cravings ratings in the morning and after bouts of PA are detected that are then graphically displayed on the dashboard to reflect changes in mood and cravings after engaging in PA; (4) the ability to directly contact social supports for PA and recovery directly from the app (eg, send an SMS text message), when needed; (5) the ability to find local community PA resources by inputting zip codes; (6) alcohol relapse prevention resources that include strategies for managing cravings, a link to the meeting finder on the Alcoholics Anonymous website, and ready access to sobriety support through SMS text messaging or phone call to their identified network of supports; and (7) dashboard view of total days sober, with the ability to change sobriety date. Finally, efforts to support engagement with the Fit&Sober app included gamification and inactivity notices. Specifically, participants were able to collect points as they engaged in the app (eg, completed mood ratings and



updated PA goals) and were able to move up in *rank* based on the number of points they collect. Figure 1 presents screenshots

of the Fit&Sober app.

**Figure 1.** Screenshots of the Fit&Sober app.



## Measures

### Alcohol Use

The Timeline Follow Back (TLFB) [42] was administered at baseline and at the end of treatment (EOT) to assess the frequency of alcohol use over the previous 90 days. The TLFB uses anchor dates such as holidays to prompt participant recall. Using data from the TLFB, we calculated 2 indices of alcohol use for our analyses: days of alcohol use and percentage of abstinent days.

### Alcohol Craving

The Penn Alcohol Craving Scale [43] is a 5-item measure designed to assess recent craving. Participants were asked to rate their experiences over the past week on a scale of 0 (low) to 6 (high). In this study, Cronbach  $\alpha$  reliability coefficients were excellent, with an  $\alpha$  of .89 at baseline and .98 at EOT.

### Self-report PA

Consistent with adding exercise as a “vital sign” [44], self-reported levels of PA were assessed by asking participants (1) Over the last 3 months, on average how many days per week did you exercise? and (2) On those days, on average how many minutes per day did you exercise? The weekly time spent exercising was calculated by multiplying these 2 responses.

### Objectively Measured PA

At baseline, participants wore a GT3x accelerometer (ActiGraph). The duration of accelerometer use was dependent on the timing of study enrollment relative to hospital discharge. Considering the varying length of stay, it was not possible for all participants to wear the accelerometer for the desired 7 days, the gold standard for in vivo accelerometry assessments. For the purposes of this study, baseline step counts were calculated for participants with at least one valid day, defined here as 8 hours per day of accelerometer wear time. Wear time validation

was conducted using Actilife Software (ActiGraph, LLC), using the Choi algorithm [45].

### Depressive Symptoms

Depressive symptoms were assessed using the Center for Epidemiologic Studies Depression Scale-Revised [46]. This scale is a 20-item measure that assesses depressive symptoms on a 4-point scale (0-3). After reverse coding appropriate items, the depressive symptom score was obtained by summing the 20 items, with higher scores indicating higher levels of depressive symptoms (Cronbach  $\alpha$ =.82 at baseline and Cronbach  $\alpha$ =.86 at EOT).

### Anxiety Symptoms

The Generalized Anxiety Disorder-7 [47] was used as a measure of generalized anxiety disorder symptomatology. The Generalized Anxiety Disorder-7 asks participants to rate how often over the last 2 weeks they have experienced 7 symptoms of anxiety on a scale of 0 (not at all) to 3 (nearly every day;  $\alpha$ =.86 at baseline and  $\alpha$ =.94 at EOT).

### Quality of Life

This was measured using the Quality of Life Enjoyment and Satisfaction Questionnaire-Short Form [48]. This questionnaire is a 16-item self-report measure of enjoyment and satisfaction in various domains of life, including physical health, mood, work, leisure, and social activities. Participants rated their satisfaction in the past week on a 5-point scale (1=very poor to 5=very good). In this study,  $\alpha$  reliability coefficients were excellent ( $\alpha$ =.95 at baseline and  $\alpha$ =.96 at EOT).

### Usability and Acceptability

Participants completed the 8-item Client Satisfaction Questionnaire [49] at EOT that assessed the level of satisfaction with the app and overall program on a scale of 1 to 4, with higher numbers indicating greater satisfaction.

The participants completed the 10-item System Usability Scale (SUS) [50,51] at the 12-week EOT assessment. The range of scores on the SUS was 0 to 100, and scores of  $\geq 70$  indicated good usability and satisfaction [52]. We also evaluated the attractiveness and detail of the 8 design features on the app on a scale of 1 (very unattractive or not detailed at all) to 5 (very attractive or very detailed). Participants were then asked to rate the usefulness of the proposed app components on a scale of 1 (not useful at all) to 5 (very useful).

Participants were asked to rate their experience with the Fitbit tracker using the 19-item Participant Experience Questionnaire of Wearable Activity Trackers [53], with items rated on a 5-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). Sample items included, "I found the Fitbit activity tracker to be clear and understandable to use" and "I have the technology necessary to use the Fitbit activity tracker."

To determine the helpfulness of the specific Fit&Sober app components, participants were asked to rate a series of statements on a scale of 0 (not helpful) to 10 (extremely helpful). This measure contained 13 items and included items such as, "I am able to identify my social support for exercise (eg, someone to exercise with me)" and "I am able to see how my cravings or urges for alcohol change with exercise."

### App Use

Participants were asked to estimate how many days per week they used the app and which components of the app they used at 2, 6, and 12 weeks after discharge. In addition, the frequency of days when the Fit&Sober app was opened was collected from the app metadata.

### Data Analysis Plan

We presented descriptive statistics for participant characteristics at baseline, including demographic information, self-reported alcohol use, PA, depression, anxiety, and quality of life. Baseline demographic and clinical characteristics were compared between participants who completed the EOT assessments and those who did not, using chi-square and 2-tailed *t* tests. Descriptive statistics were presented to evaluate the participants' perceptions of acceptability and feasibility. Given the developmental nature of this project, we conducted an open pilot study with a small sample ( $N=22$ ), although sufficient to provide relatively stable group means for the dependent measures of interest. Therefore, we refrained from conducting significance testing and instead calculated effect sizes [54] in the form of Cohen *d* [55].

## Results

### Participant Characteristics

Participants ( $N=22$ ) ranged in age from 20 to 61 years, with a mean age of 43.64 (SD 11.57) years, including more than half of the sample being women (13/22, 59%) and most being White (21/22, 95%) individuals. At baseline, participants drank an average of 65.09 (SD 19.79) drinks in the past 90 days, and when they drank, they averaged 9.2 (SD 3.8) drinks per drinking day. Half of the participants (11/22, 50%) used an iPhone and the

other half (11/22, 50%) used an Android phone. At the 3-month follow-up, 86% (19/22) of patients provided data.

### Feasibility and Acceptability

App use was determined by self-report at the 2-, 6-, and 12-week assessments and by examination of the Fit&Sober metadata. Table 1 presents the mean number of days per week that participants reported opening and interacting with the app, as well as the percentage of participants who interacted with specific app features at each time point. Table 2 presents the app metadata on the percentage of the sample who used the app (ie, opening the app on a given day) and the average number of days per week of use across the 12-week intervention period. On average, at the end of the intervention period, participants self-reported using the app 4.33 (SD 2.72) days per week. In comparison, the Fit&Sober app metadata suggested that participants used the app, defined as opening the app, 2.55 (SD 1.68) days per week. At the EOT, on average, the SUS total score was 74.64 (SD 14.90) for the Fit&Sober app. Further, the app's average attractiveness was 3.20 (SD 1.16), and its usefulness was 3.11 (SD 1.07) on a scale of 1 to 5. Satisfaction ratings on the Client Satisfaction Questionnaire were generally high, with an average of 3.40 (SD 0.43) for the full scale (on a 1-4 scale). Individual item-level mean ratings were as follows: quality of program 3.44 (SD 0.51), the kind of program participant wanted 3.31 (SD 0.60), program meeting participant needs 3.13 (SD 0.81), recommend to a friend 3.63 (SD 0.62), amount of help received 3.44 (SD 0.89), helped with increasing PA 3.44 (SD 0.63), overall satisfaction 3.38 (SD 0.81), would come back to program in the future 3.44 (SD 0.51), and a supplemental question of whether PA helped with alcohol recovery 3.31 (SD 0.87).

Overall, participants found the app helpful, 6.26 (SD 2.11) on a 1 to 10 scale. Table 3 details the item-level ratings of how helpful the participants found the various components of the app. The components that received a high rating of  $\geq 7$  were keeping track of their PA, capability to see the number of sober days, observe mood change with PA, and the app's ability to communicate with the Fitbit tracker. The components that received a lower rating of  $\leq 5$  were finding social support for PA, sobriety, and seeing motivational quotes every day. Finally, participants wore the Fitbit for an average of 9.5 (SD 3.7) weeks (ie, 79%) of the intervention period) and reported being generally satisfied with the tracker, with an average rating of 4.04 (SD 1.16) at the EOT.

The mean of number of days of valid accelerometer wear time was 2.4 (SD 1.4). Accelerometry data collected at baseline were available for 95% (21/22) of participants. The baseline wear time for this sample ( $N=22$ ) was as follows: 1 day (4/22, 18%), 2 days (9/22, 41%), 3 days (4/22, 18%), 4 days (2/22, 9%), 5 days (1/22, 5%), and 6 days (1/22, 5%). Daily step counts were collected throughout the 12-week intervention objectively via the Fitbit, with average daily step counts calculated for participants who had at least 8 weeks of Fitbit data, with Fitbit devices worn at least 8 hours per day (18/22, 82%).

**Table 1.** Self-reported use of the Fit&Sober app.

App use features	Weeks 1 and 2 (n=15)	Weeks 3-6 (n=14)	Weeks 7-12 (n=19)
Opened the app (days per week), mean (SD)	5.33 (1.95)	4.33 (2.6)	4.21 (2.7)
Interacted with the app <sup>a</sup> (days per week), mean (SD)	5.04 (2.42)	3.95 (2.81)	3.63 (2.89)
<b>Sample who interacted with each of the following Fit&amp;Sober app features, n (%)</b>			
Updated PA <sup>b</sup> goals	10 (67)	8 (57)	5 (26)
Responded to notifications the app sent	12 (79)	9 (62)	13 (68)
Entered mood ratings after a bout of PA	13 (87)	10 (71)	9 (47)
Reviewed mood ratings graphs	9 (60)	11 (77)	11 (58)
Updated social supports for PA and sobriety	1 (7)	0 (0)	0 (0)
Updated sobriety date	0 (0)	2 (14)	0 (0)
Updated reasons for PA	1 (7)	1 (7)	1 (5)
Reviewed the alcohol relapse prevention strategies	5 (27)	5 (36)	5 (26)
Looked for ideas for increasing PA	6 (40)	7 (50)	4 (21)

<sup>a</sup>Do something on the app.

<sup>b</sup>PA: physical activity.

**Table 2.** Objective app use from the Fit&Sober app metadata.

Intervention week	1	2	3	4	5	6	7	8	9	10	11	12
Sample who used the app, n (%)	18 (82)	17 (77)	18 (82)	16 (73)	14 (64)	14 (64)	14 (64)	16 (73)	12 (55)	11 (50)	12 (55)	10 (45)
App use (days per week), mean (SD)	4.4 (1.5)	3.7 (2.5)	3.4 (2.5)	2.4 (2.1)	1.9 (1.9)	2.7 (2.3)	2.8 (2.4)	2.8 (2.5)	1.8 (2.1)	1.8 (2.2)	1.7 (1.8)	1.6 (2.2)

**Table 3.** Item-level ratings of the self-reported helpfulness of various components of the Fit&Sober app<sup>a</sup>.

App feature	Value, mean (SD)
1. I am able to set my own exercise goals	6.29 (2.89)
2. I am able to see my progress toward achieving my exercise goal	6.59 (2.85)
3. I am able to track my physical activity	7.29 (2.57)
4. I am able to keep track of how many days I have been sober	8.59 (2.00)
5. I am able to identify my social supports for exercise	4.71 (3.35)
6. I am able to identify my social supports for sobriety	4.59 (3.39)
7. I am able to identify my values and the reasons I am increasing my physical activity	6.06 (3.09)
8. I am able to see how my mood changes with exercise	7.18 (2.86)
9. I am able to see how my cravings or urges for alcohol change with exercise	6.71 (2.87)
10. I am able to view motivational quotes every day	4.88 (3.44)
11. I am able to receive specific advice if I feel an urge to drink	5.24 (3.47)
12. I am able to view resources and ideas for ways I can be more physically active	5.47 (3.12)
13. The app is able to communicate with the Fitbit tracker	7.82 (2.30)

<sup>a</sup>The range of items is 1 to 10 with higher ratings indicating greater helpfulness.

## PA Outcomes

PA assessments included self-reported average minutes per week of exercise (assessed with the Exercise as a Vital Sign questions) over the last 3 months, 7-day accelerometry-derived

steps per day, and daily steps per day from the Fitbit during the 12 weeks of the intervention. Baseline steps per day collected from wearing the GT3x accelerometer averaged 5783 (SD=2511) steps per day. Among the participants who wore Fitbit for at least 8 weeks (18/22, 82%), the average steps per

day over the course of the 12-week intervention was 7236 (SD 3130). Self-report of average minutes per week of exercise at baseline was 76.13 (SD 124.11) and at the EOT was 160.82 (SD 149.76); a moderate effect of increase in self-reported PA (Cohen  $d=0.65$ , 95% CI 0.11-1.16). Self-reported PA over the 12-week intervention and the average, objective measurement

of steps per day on the Fitbit during the intervention were highly correlated, although the 95% CI was large ( $r=0.51$ , 95% CI  $-0.03$  to  $0.82$ ). Table 4 presents the means and effect sizes, and Figure 2 presents the changes in steps per day over the course of each week of the intervention.

**Table 4.** Intervention outcome and effect sizes.

Intervention outcomes	Baseline (n=22), mean (SD)	End of treatment (n=19), mean (SD)	Cohen $d$ (95% CI)
<b>Alcohol outcomes</b>			
Number of days drinking alcohol in the preceding 90 days	65.09 (19.79)	7.47 (19.92)	-2.27 (-3.12 to -1.40)
Percentage of abstinent days	27.68 (21.99)	91.10 (23.72)	2.17 (1.33 to 2.99)
Alcohol craving	13.41 (7.01)	9.56 (8.90)	-0.36 (-0.86 to 0.16)
<b>Mental health outcomes</b>			
Anxiety (GAD-7 <sup>a</sup> )	13.86 (4.73)	9.31 (7.14)	-0.71 (-1.25 to -0.15)
Depression (CES-D <sup>b</sup> )	29.09 (9.28)	25.19 (11.81)	-0.58 (-1.10 to -0.40)
Quality of life	38.47 (12.95)	44.54 (15.33)	0.58 (-0.21 to 1.16)
<b>PA<sup>c</sup> outcomes</b>			
Self-reported minutes of PA during the 12-week intervention	76.13 (124.11)	160.82 (149.76)	0.65 (0.11 to 1.16)
Objectively measured steps per day <sup>d</sup>	5784 (2511)	7236 (3130)	0.35 (-0.15 to 0.83)

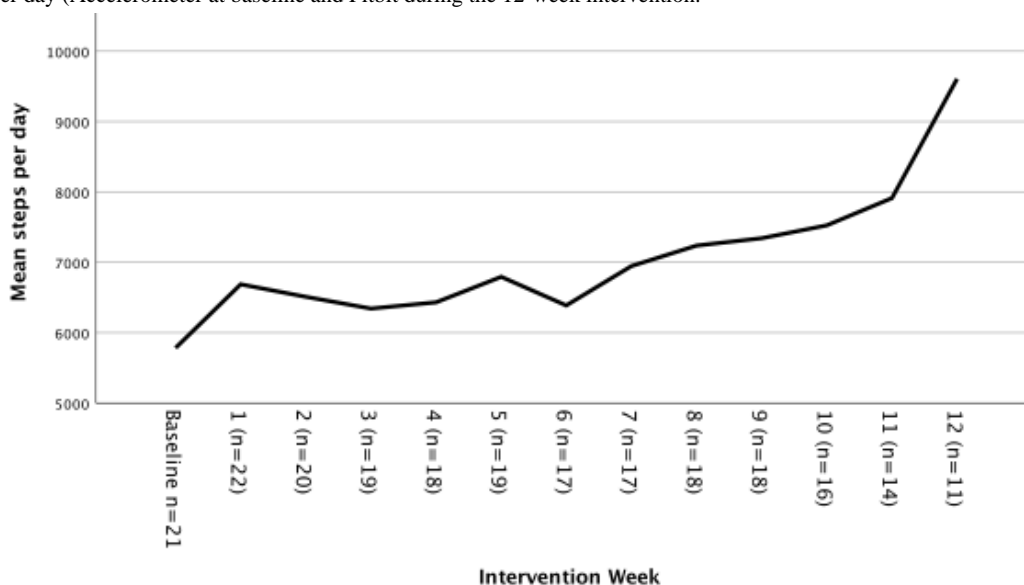
<sup>a</sup>GAD-7: Generalized Anxiety Disorder-7.

<sup>b</sup>CES-D: Center for Epidemiologic Studies Depression Scale-Revised.

<sup>c</sup>PA: physical activity.

<sup>d</sup>Accelerometer at baseline; average steps per day over 12-weeks with Fitbit (among participants with 8 weeks of Fitbit data; n=18).

**Figure 2.** Steps per day (Accelerometer at baseline and Fitbit during the 12-week intervention).



**Alcohol Outcomes**

Among participants who completed the EOT assessment (n=19), 10 (53%) reported complete abstinence over the course of the 12-week intervention. Among the 9 individuals who consumed alcohol, 7 (78%) did so for 6 days or fewer, 1 (11%) consumed alcohol for 47 days, and 1 (11%) consumed alcohol on 77 days

in the past 90 days. Overall, the mean reduction in percentage of abstinent days was a Cohen  $d$  of 2.77. In addition, participants demonstrated lower levels of alcohol craving at EOT (mean 56, SD 8.90) than at baseline (mean 13.41, SD 7.01, Cohen  $d=0.48$ ). Table 4 presents the means and effect sizes.

## Mental Health Outcomes

Participants who completed the EOT assessment reported decrease in anxiety (Cohen  $d=-0.71$ , 95% CI  $-1.25$  to  $-0.15$ ) and depressive symptoms (Cohen  $d=-0.58$ , 95% CI  $-1.10$  to  $-0.40$ ) from baseline to the EOT. There were also improvements in the quality-of-life scores (Cohen  $d=0.58$ , 95% CI  $-0.21$  to  $1.16$ ).

## Discussion

### Principal Findings

This open pilot trial examined the use of Fit&Sober, a theoretically informed PA smartphone app, targeted toward patients with AUD, in early recovery. The results of this study show high levels of feasibility and acceptability of the Fit&Sober app over a 12-week period. In addition, large effect sizes were observed for changes in alcohol outcomes, small-to-moderate effects for increases in PA and quality of life, and reductions in anxiety and depression from baseline to the EOT. These findings are promising and suggest that future research to determine the efficacy of the Fit&Sober app compared with a control condition is warranted.

### Comparison With Prior Work

Despite the high ratings on measures of app satisfaction, usability, and helpfulness, app use declined over time, such that fewer than half of the participants were using it at the end of the 12-week intervention period. However, these levels of app use are quite comparable and, in some cases, more favorable than what has been reported in previous studies on both PA and alcohol apps. For example, in testing a well-designed, theoretically informed smartphone app for increasing PA and decreasing sedentary behavior, Direito et al [56] found that participants opened the app an average of 11 days during the entire 8-week intervention. In addition, over the course of 5 months, use of 3 different PA apps showed declines ranging from 45% to 75% in use [57]. Similarly, among at-risk drinkers who were interested in reducing their alcohol consumption, 50% had stopped using an app by the end of the first month [58]. In another study of a Veteran population engaged in alcohol misuse, an alcohol self-management app was used by 96% of the sample in week 1, which decreased to 55% by week 4 [59]. Indeed, long-term app use is quite low for health apps—the vast majority of health app users stop using the app after 10 visits and 26% use it only once after downloading it [60]. Our results also suggest that participants may self-report higher use than actual use. Given the social desirability of reporting greater app use observed in our study, future research should consider including objective indicators of app use, such as app metadata.

Although app use consistently wanes over time, it appears that the extent of app use may not be related to either increase in PA or decrease in drinking [61,62]. In other words, even when marked decreases in app use were observed, significant increase in PA and decrease in drinking behaviors also occurred [56,59,63]. Thus, some individuals may be able to increase their PA and sustain this change over time with only a few weeks of app use. In future studies, typologies of user engagement with

mobile health (mHealth) apps can be identified and then examined in relation to changes in the desired behavioral outcome (eg, PA or drinking), and the utility of these apps in changing and maintaining specific behaviors can be better determined.

The participants in this study demonstrated a 25% increase in objectively determined steps per day (approximately 1400 steps per day from baseline). Increase of 26.9% in steps per day have been shown to result in significant improvements in health, including decreased blood pressure [64]. In a recent systematic review of PA smartphone apps, there was a mean increase of 476 steps per day in those assigned to the app conditions [65]. Therefore, the observed increase in our study are consistent with those of prior research. However, our baseline accelerometer-determined steps per day was limited by less-than-optimal wear time. Indeed, it was challenging to collect a full 7 days of accelerometer wear time, the gold standard in PA research, given the varied and short duration of partial hospitalization admissions and the goal of orienting patients to the app before their discharge from treatment. However, adherence to wearing the Fitbit tracker was good and similar to other patient populations (eg, 39.6%-85.7% in a systematic review of adherence to PA monitoring devices in adults with cardiovascular disease [66]). Given that Fitbit devices are valid for the measurement of steps [67], are feasibly worn for long durations, and considered satisfactory by users, future studies may consider their use for objectively determining long-term engagement in PA, which has, to date, been limited to self-report.

Certain features of the app were used by a greater number of participants than the other features. Specifically, updating PA goals, monitoring changes in mood, and responding to notifications were more popular than updating social support, sobriety dates, and updating reasons for being physically active. Daily app notifications were targeted toward reminding users to self-monitor mood and cravings and update PA goals, which correspond to the features most used by participants. As such, sending notifications to users is an effective strategy for increasing engagement with specific features of the app. Future studies may benefit from the inclusion of innovative designs that allow for microrandomization messaging and notifications (ie, randomizing messages or notifications at the daily level) to identify those most likely to lead to increases in app engagement and behavior change [68].

The last decade has witnessed an increase in the number of smartphone apps targeting individuals with AUD [69]. The acceptability of smartphone apps among at-risk drinkers is high, but their effect on changes in drinking outcomes has been mixed [70]. Common features of these apps include tracking of drinking, alcohol-related consequences, and locating 12-step meetings and treatment programs [69]. Several of the Fit&Sober app features were directly relevant to alcohol recovery including tracking sobriety date and alcohol cravings, locating 12-step meetings, identifying sober supports, and strategies for preventing relapse. Integrating these alcohol-specific features into an app designed to increase PA is a novel approach that may provide an opportunity to synergistically impact multiple health behavior changes. However, PA apps that also include

dietary components are not as effective at increasing PA as those focusing only on PA [65]. Therefore, future research that can test the effectiveness of specific combinations of app components may be necessary to optimize app use.

### Strengths and Limitations

This study has several strengths. Our ability to collect objective metadata from app use allowed us to identify the most used app features. As mHealth apps are typically multicomponent, data regarding the usability of these features are critical for the continued optimization of the apps. In addition, we objectively measured PA outcomes. Indeed, overreliance on self-reported PA has been a limitation of prior PA studies with individuals with AUD. Our study demonstrated the feasibility of these passive data collection modalities and, in turn, increased the methodological rigor of this study.

There are also several limitations that merit further discussion. First, this was a small, open pilot study designed to determine the feasibility of the Fit&Sober app during early alcohol recovery. Although increases in PA and quality of life and decreases in anxiety and depression were observed, these effects cannot be attributed to the Fit&Sober app. It is possible that simply being in early recovery and engaging in these related activities (eg, attendance at 12-step meetings, therapy, and avoiding triggers) would lead to the same outcomes. Therefore, a future randomized controlled trial that accounts for these factors is necessary. Second, the sample lacked racial and ethnic heterogeneity. It is important to understand the feasibility and acceptability of the Fit&Sober app across a more diverse sample of patients. Third, the baseline accelerometry measurement of steps per day must be interpreted with caution, given the limited device wear time. The cost-benefit relationship of whether to delay a PA intervention to collect a rigorous assessment of

objective PA measurement for individuals who could most benefit from immediately increasing PA (owing to higher levels of craving in the early periods of abstinence) needs to be examined in future research. Finally, although the 12-week duration of the intervention maps onto the first 90 days of recovery being the highest risk for relapse, any future study on the maintenance of PA would require longer study durations.

### Conclusions

Individuals receiving treatment for AUD were willing to participate in an open pilot trial on the feasibility and acceptability of using a PA smartphone app tailored for early alcohol recovery. Participants used the app at rates consistent with other PA and AUD apps. They reported that utilizing the Fit&Sober app was helpful in increasing PA during the early stages of alcohol recovery. The observed increase in PA in this sample approximate those associated with significant improvements in physical health indicators, a critical finding, given the significant concomitant physical health problems experienced by individuals with AUD. Participants most often used the app to self-monitor PA goals, mood, and alcohol cravings, suggesting that these should be key features of any future efforts to optimize the Fit&Sober app or the development of other mHealth interventions for this population. Within-person changes in mental health and drinking outcomes were promising, although an important next step will be to conduct a randomized controlled trial to determine the efficacy of the Fit&Sober app in improving these alcohol treatment outcomes. In conclusion, if the efficacy of the Fit&Sober app can be established, patients with AUD can be provided with a valuable adjunct to traditional alcohol treatment that can be delivered in any setting and at any time, thereby improving the overall health and well-being of this population.

### Acknowledgments

This work was supported by a grant from the National Institute on Alcohol Abuse and Alcoholism (R21 AA024295; principal investigator: AMA). LZM was supported by the National Institute on Alcohol Abuse and Alcoholism grant, K23 AA028269.

### Conflicts of Interest

None declared.

### References

1. Rehm J, Dawson D, Frick U, Gmel G, Roerecke M, Shield KD, et al. Burden of disease associated with alcohol use disorders in the United States. *Alcohol Clin Exp Res* 2014 Apr;38(4):1068-1077 [FREE Full text] [doi: [10.1111/acer.12331](https://doi.org/10.1111/acer.12331)] [Medline: [24428196](https://pubmed.ncbi.nlm.nih.gov/24428196/)]
2. Rehm J, Mathers C, Popova S, Thavorncharoensap M, Teerawattananon Y, Patra J. Global burden of disease and injury and economic cost attributable to alcohol use and alcohol-use disorders. *Lancet* 2009 Jun 27;373(9682):2223-2233. [doi: [10.1016/S0140-6736\(09\)60746-7](https://doi.org/10.1016/S0140-6736(09)60746-7)] [Medline: [19560604](https://pubmed.ncbi.nlm.nih.gov/19560604/)]
3. Grant BF, Goldstein RB, Saha TD, Chou SP, Jung J, Zhang H, et al. Epidemiology of DSM-5 alcohol use disorder: results from the national epidemiologic survey on alcohol and related conditions III. *JAMA Psychiatry* 2015 Aug;72(8):757-766 [FREE Full text] [doi: [10.1001/jamapsychiatry.2015.0584](https://doi.org/10.1001/jamapsychiatry.2015.0584)] [Medline: [26039070](https://pubmed.ncbi.nlm.nih.gov/26039070/)]
4. Alcohol and public health: alcohol related disease impact (ARDI). Centers for Disease Control and Prevention. 2012. URL: [https://nccd.cdc.gov/DPH\\_ARDI/default/default.aspx#:~:text=ARDI%20is%20an%20online%20application,and%20sex%20for%202015%E2%80%932019](https://nccd.cdc.gov/DPH_ARDI/default/default.aspx#:~:text=ARDI%20is%20an%20online%20application,and%20sex%20for%202015%E2%80%932019) [accessed 2021-10-15]
5. Sacks JJ, Gonzales KR, Bouchery EE, Tomedi LE, Brewer RD. 2010 National and state costs of excessive alcohol consumption. *Am J Prev Med* 2015 Nov;49(5):e73-e79. [doi: [10.1016/j.amepre.2015.05.031](https://doi.org/10.1016/j.amepre.2015.05.031)] [Medline: [26477807](https://pubmed.ncbi.nlm.nih.gov/26477807/)]

6. Hunt WA, Barnett LW, Branch LG. Relapse rates in addiction programs. *J Clin Psychol* 1971 Oct;27(4):455-456. [doi: [10.1002/1097-4679\(197110\)27:4<455::aid-jclp2270270412>3.0.co;2-r](https://doi.org/10.1002/1097-4679(197110)27:4<455::aid-jclp2270270412>3.0.co;2-r)] [Medline: [5115648](#)]
7. Xie H, McHugo GJ, Fox MB, Drake RE. Substance abuse relapse in a ten-year prospective follow-up of clients with mental and substance use disorders. *Psychiatr Serv* 2005 Oct;56(10):1282-1287. [doi: [10.1176/appi.ps.56.10.1282](https://doi.org/10.1176/appi.ps.56.10.1282)] [Medline: [16215196](#)]
8. Anton RF, O'Malley SS, Ciraulo DA, Cisler RA, Couper D, Donovan DM, COMBINE Study Research Group. Combined pharmacotherapies and behavioral interventions for alcohol dependence: the COMBINE study: a randomized controlled trial. *JAMA* 2006 May 03;295(17):2003-2017. [doi: [10.1001/jama.295.17.2003](https://doi.org/10.1001/jama.295.17.2003)] [Medline: [16670409](#)]
9. Lynch WJ, Peterson AB, Sanchez V, Abel J, Smith MA. Exercise as a novel treatment for drug addiction: a neurobiological and stage-dependent hypothesis. *Neurosci Biobehav Rev* 2013 Sep;37(8):1622-1644 [FREE Full text] [doi: [10.1016/j.neubiorev.2013.06.011](https://doi.org/10.1016/j.neubiorev.2013.06.011)] [Medline: [23806439](#)]
10. Abrantes AM, Blevins CE. Exercise in the context of substance use treatment: key issues and future directions. *Curr Opin Psychol* 2019 Dec;30:103-108. [doi: [10.1016/j.copsyc.2019.04.001](https://doi.org/10.1016/j.copsyc.2019.04.001)] [Medline: [31079012](#)]
11. Taylor AH, Oh H, Cullen S. Acute effect of exercise on alcohol urges and attentional bias towards alcohol related images in high alcohol consumers. *Ment Health Phys Act* 2013 Oct;6(3):220-226. [doi: [10.1016/j.mhpa.2013.09.004](https://doi.org/10.1016/j.mhpa.2013.09.004)]
12. Ludyga S, Gerber M, Brand S, Holsboer-Trachsler E, Pühse U. Acute effects of moderate aerobic exercise on specific aspects of executive function in different age and fitness groups: a meta-analysis. *Psychophysiology* 2016 Nov;53(11):1611-1626. [doi: [10.1111/psyp.12736](https://doi.org/10.1111/psyp.12736)] [Medline: [27556572](#)]
13. Mandolesi L, Polverino A, Montuori S, Foti F, Ferraioli G, Sorrentino P, et al. Effects of physical exercise on cognitive functioning and wellbeing: biological and psychological benefits. *Front Psychol* 2018 Apr 27;9:509 [FREE Full text] [doi: [10.3389/fpsyg.2018.00509](https://doi.org/10.3389/fpsyg.2018.00509)] [Medline: [29755380](#)]
14. Aharonovich E, Nunes E, Hasin D. Cognitive impairment, retention and abstinence among cocaine abusers in cognitive-behavioral treatment. *Drug Alcohol Depend* 2003 Aug 20;71(2):207-211 [FREE Full text] [doi: [10.1016/s0376-8716\(03\)00092-9](https://doi.org/10.1016/s0376-8716(03)00092-9)] [Medline: [12927659](#)]
15. Abrantes AM, Battle CL, Strong DR, Ing E, Dubreuil ME, Gordon A, et al. Exercise preferences of patients in substance abuse treatment. *Ment Health Phys Act* 2011 Dec;4(2):79-87 [FREE Full text] [doi: [10.1016/j.mhpa.2011.08.002](https://doi.org/10.1016/j.mhpa.2011.08.002)] [Medline: [22125581](#)]
16. Stoutenberg M, Warne J, Vidot D, Jimenez E, Read JP. Attitudes and preferences towards exercise training in individuals with alcohol use disorders in a residential treatment setting. *J Subst Abuse Treat* 2015 Feb;49:43-49. [doi: [10.1016/j.jsat.2014.08.008](https://doi.org/10.1016/j.jsat.2014.08.008)] [Medline: [25242191](#)]
17. Rehm J, Gmel G, Sempos CT, Trevisan M. Alcohol-related morbidity and mortality. *Alcohol Res Health* 2003;27(1):39-51 [FREE Full text] [Medline: [15301399](#)]
18. Baan R, Straif K, Grosse Y, Secretan B, El Ghissassi F, Bouvard V, WHO International Agency for Research on Cancer Monograph Working Group. Carcinogenicity of alcoholic beverages. *Lancet Oncol* 2007 Apr;8(4):292-293. [doi: [10.1016/s1470-2045\(07\)70099-2](https://doi.org/10.1016/s1470-2045(07)70099-2)] [Medline: [17431955](#)]
19. Tsunoda K, Kai Y, Uchida K, Kuchiki T, Nagamatsu T. Physical activity and risk of fatty liver in people with different levels of alcohol consumption: a prospective cohort study. *BMJ Open* 2014 Aug 05;4(8):e005824 [FREE Full text] [doi: [10.1136/bmjopen-2014-005824](https://doi.org/10.1136/bmjopen-2014-005824)] [Medline: [25095878](#)]
20. Soedamah-Muthu SS, De Neve M, Shelton NJ, Tielemans SM, Stamatakis E. Joint associations of alcohol consumption and physical activity with all-cause and cardiovascular mortality. *Am J Cardiol* 2013 Aug 01;112(3):380-386. [doi: [10.1016/j.amjcard.2013.03.040](https://doi.org/10.1016/j.amjcard.2013.03.040)] [Medline: [23647792](#)]
21. Crum RM, Mojtabai R, Lazareck S, Bolton JM, Robinson J, Sareen J, et al. A prospective assessment of reports of drinking to self-medicate mood symptoms with the incidence and persistence of alcohol dependence. *JAMA Psychiatry* 2013 Jul;70(7):718-726 [FREE Full text] [doi: [10.1001/jamapsychiatry.2013.1098](https://doi.org/10.1001/jamapsychiatry.2013.1098)] [Medline: [23636710](#)]
22. Cooney GM, Dwan K, Greig CA, Lawlor DA, Rimer J, Waugh FR, et al. Exercise for depression. *Cochrane Database Syst Rev* 2013 Sep 12(9):CD004366. [doi: [10.1002/14651858.CD004366.pub6](https://doi.org/10.1002/14651858.CD004366.pub6)] [Medline: [24026850](#)]
23. Hunt-Shanks T, Blanchard C, Reid RD. Gender differences in cardiac patients: a longitudinal investigation of exercise, autonomic anxiety, negative affect and depression. *Psychol Health Med* 2009 May;14(3):375-385. [doi: [10.1080/13548500902866939](https://doi.org/10.1080/13548500902866939)] [Medline: [19444715](#)]
24. Brown RA, Prince MA, Minami H, Abrantes AM. An exploratory analysis of changes in mood, anxiety and craving from pre- to post-single sessions of exercise, over 12 weeks, among patients with alcohol dependence. *Ment Health Phys Act* 2016 Oct;11:1-6 [FREE Full text] [doi: [10.1016/j.mhpa.2016.04.002](https://doi.org/10.1016/j.mhpa.2016.04.002)] [Medline: [29606975](#)]
25. Brown RA, Abrantes AM, Minami H, Read JP, Marcus BH, Jakicic JM, et al. A preliminary, randomized trial of aerobic exercise for alcohol dependence. *J Subst Abuse Treat* 2014 Jul;47(1):1-9 [FREE Full text] [doi: [10.1016/j.jsat.2014.02.004](https://doi.org/10.1016/j.jsat.2014.02.004)] [Medline: [24666811](#)]
26. Hallgren M, Vancampfort D, Giesen ES, Lundin A, Stubbs B. Exercise as treatment for alcohol use disorders: systematic review and meta-analysis. *Br J Sports Med* 2017 Jul;51(14):1058-1064. [doi: [10.1136/bjsports-2016-096814](https://doi.org/10.1136/bjsports-2016-096814)] [Medline: [28087569](#)]

27. Ussher M, Sampuran AK, Doshi R, West R, Drummond DC. Acute effect of a brief bout of exercise on alcohol urges. *Addiction* 2004 Dec;99(12):1542-1547. [doi: [10.1111/j.1360-0443.2004.00919.x](https://doi.org/10.1111/j.1360-0443.2004.00919.x)] [Medline: [15585045](https://pubmed.ncbi.nlm.nih.gov/15585045/)]
28. Treatment Episode Data Set (TEDS): 2018. Admissions to and Discharges From Publicly Funded Substance Use Treatment. Substance Abuse and Mental Health Services Administration, Center for Behavioral Health Statistics and Quality. Rockville, MD, USA: Substance Abuse and Mental Health Services Administration; 2020. URL: [https://www.samhsa.gov/data/sites/default/files/reports/rpt31097/2018\\_TEDS/2018\\_TEDS.html](https://www.samhsa.gov/data/sites/default/files/reports/rpt31097/2018_TEDS/2018_TEDS.html) [accessed 2022-08-26]
29. Roessler KK, Bramsen RH, Dervisevic A, Bilberg R. Exercise based interventions for alcohol use disorder: a comment on motivational aspects of participation. *Scand J Psychol* 2017 Feb;58(1):23-28. [doi: [10.1111/sjop.12334](https://doi.org/10.1111/sjop.12334)] [Medline: [27727463](https://pubmed.ncbi.nlm.nih.gov/27727463/)]
30. Michie S, Abraham C, Whittington C, McAteer J, Gupta S. Effective techniques in healthy eating and physical activity interventions: a meta-regression. *Health Psychol* 2009 Nov;28(6):690-701. [doi: [10.1037/a0016136](https://doi.org/10.1037/a0016136)] [Medline: [19916637](https://pubmed.ncbi.nlm.nih.gov/19916637/)]
31. Kirwan M, Duncan MJ, Vandelanotte C, Mummery WK. Design, development, and formative evaluation of a smartphone application for recording and monitoring physical activity levels: the 10,000 Steps "iStepLog". *Health Educ Behav* 2013 Apr;40(2):140-151. [doi: [10.1177/1090198112449460](https://doi.org/10.1177/1090198112449460)] [Medline: [22984196](https://pubmed.ncbi.nlm.nih.gov/22984196/)]
32. Ritterband LM, Thorndike FP, Cox DJ, Kovatchev BP, Gonder-Frederick LA. A behavior change model for Internet interventions. *Ann Behav Med* 2009 Aug;38(1):18-27 [FREE Full text] [doi: [10.1007/s12160-009-9133-4](https://doi.org/10.1007/s12160-009-9133-4)] [Medline: [19802647](https://pubmed.ncbi.nlm.nih.gov/19802647/)]
33. Phongsavan P, Merom D, Bauman A, Wagner R. Mental illness and physical activity: therapists' beliefs and practices. *Aust N Z J Psychiatry* 2007 May;41(5):458-459. [doi: [10.1080/00048670701266862](https://doi.org/10.1080/00048670701266862)] [Medline: [17464739](https://pubmed.ncbi.nlm.nih.gov/17464739/)]
34. Pew R. Mobile Fact Sheet. Pew Research Center. 2021 Apr 7. URL: <https://www.pewresearch.org/internet/fact-sheet/mobile/> [accessed 2021-12-15]
35. Bort-Roig J, Gilson ND, Puig-Ribera A, Contreras RS, Trost SG. Measuring and influencing physical activity with smartphone technology: a systematic review. *Sports Med* 2014 May;44(5):671-686. [doi: [10.1007/s40279-014-0142-5](https://doi.org/10.1007/s40279-014-0142-5)] [Medline: [24497157](https://pubmed.ncbi.nlm.nih.gov/24497157/)]
36. Cowan LT, Van Wagenen SA, Brown BA, Hedin RJ, Seino-Stephan Y, Hall PC, et al. Apps of steel: are exercise apps providing consumers with realistic expectations?: a content analysis of exercise apps for presence of behavior change theory. *Health Educ Behav* 2013 Apr;40(2):133-139. [doi: [10.1177/1090198112452126](https://doi.org/10.1177/1090198112452126)] [Medline: [22991048](https://pubmed.ncbi.nlm.nih.gov/22991048/)]
37. Segar ML, Eccles JS, Peck SC, Richardson CR. Midlife women's physical activity goals: sociocultural influences and effects on behavioral regulation. *Sex Roles* 2007 Oct 4;57(11-12):837-849. [doi: [10.1007/s11199-007-9322-1](https://doi.org/10.1007/s11199-007-9322-1)]
38. Bandura A. Health promotion from the perspective of social cognitive theory. *Psychol Health* 1998 Jul;13(4):623-649. [doi: [10.1080/08870449808407422](https://doi.org/10.1080/08870449808407422)]
39. Ryan RM, Deci EL. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *Am Psychol* 2000 Jan;55(1):68-78. [doi: [10.1037//0003-066x.55.1.68](https://doi.org/10.1037//0003-066x.55.1.68)] [Medline: [11392867](https://pubmed.ncbi.nlm.nih.gov/11392867/)]
40. Segar ML, Eccles JS, Richardson CR. Rebranding exercise: closing the gap between values and behavior. *Int J Behav Nutr Phys Act* 2011 Aug 31;8:94 [FREE Full text] [doi: [10.1186/1479-5868-8-94](https://doi.org/10.1186/1479-5868-8-94)] [Medline: [21884579](https://pubmed.ncbi.nlm.nih.gov/21884579/)]
41. Abrantes AM, Blevins C, Lindsay C, Battle CL, Buman MP, Agu E, et al. Formative work in the development of a physical activity smartphone app targeted for patients with alcohol use disorders. *Psychol Sport Exerc* 2019 Mar;41:162-171. [doi: [10.1016/j.psychsport.2018.02.007](https://doi.org/10.1016/j.psychsport.2018.02.007)]
42. Sobell LC, Sobell MB. Timeline follow-back. In: Litten RZ, Allen JP, editors. *Measuring Alcohol Consumption: Psychosocial and Biochemical Methods*. New York, NY, USA: Humana Totowa; 1992:41-72.
43. Flannery BA, Volpicelli JR, Pettinati HM. Psychometric properties of the Penn Alcohol Craving Scale. *Alcohol Clin Exp Res* 1999 Aug;23(8):1289-1295. [Medline: [10470970](https://pubmed.ncbi.nlm.nih.gov/10470970/)]
44. Coleman KJ, Ngor E, Reynolds K, Quinn VP, Koebnick C, Young DR, et al. Initial validation of an exercise "vital sign" in electronic medical records. *Med Sci Sports Exerc* 2012 Nov;44(11):2071-2076. [doi: [10.1249/MSS.0b013e3182630ec1](https://doi.org/10.1249/MSS.0b013e3182630ec1)] [Medline: [22688832](https://pubmed.ncbi.nlm.nih.gov/22688832/)]
45. Choi L, Liu Z, Matthews CE, Buchowski MS. Validation of accelerometer wear and nonwear time classification algorithm. *Med Sci Sports Exerc* 2011 Feb;43(2):357-364 [FREE Full text] [doi: [10.1249/MSS.0b013e3181ed61a3](https://doi.org/10.1249/MSS.0b013e3181ed61a3)] [Medline: [20581716](https://pubmed.ncbi.nlm.nih.gov/20581716/)]
46. Radloff LS. The CES-D scale: a self-report depression scale for research in the general population. *Appl Psychol Meas* 1977 Jun 1;1(3):385-401. [doi: [10.1177/014662167700100306](https://doi.org/10.1177/014662167700100306)]
47. Spitzer RL, Kroenke K, Williams JB, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch Intern Med* 2006 May 22;166(10):1092-1097. [doi: [10.1001/archinte.166.10.1092](https://doi.org/10.1001/archinte.166.10.1092)] [Medline: [16717171](https://pubmed.ncbi.nlm.nih.gov/16717171/)]
48. Stevanovic D. Quality of Life Enjoyment and Satisfaction Questionnaire-short form for quality of life assessments in clinical practice: a psychometric study. *J Psychiatr Ment Health Nurs* 2011 Oct;18(8):744-750. [doi: [10.1111/j.1365-2850.2011.01735.x](https://doi.org/10.1111/j.1365-2850.2011.01735.x)] [Medline: [21896118](https://pubmed.ncbi.nlm.nih.gov/21896118/)]
49. Attkisson CC, Zwick R. The client satisfaction questionnaire. Psychometric properties and correlations with service utilization and psychotherapy outcome. *Eval Program Plann* 1982;5(3):233-237. [doi: [10.1016/0149-7189\(82\)90074-x](https://doi.org/10.1016/0149-7189(82)90074-x)] [Medline: [10259963](https://pubmed.ncbi.nlm.nih.gov/10259963/)]



50. Brooke J. SUS: a 'quick and dirty' usability scale. In: Jordan PW, Thomas B, McClelland IL, Weerdmeester B, editors. *Usability Evaluation in Industry*. London, UK: Taylor and Francis; 1996:189-194.
51. Bangor A, Kortum PT, Miller JT. An empirical evaluation of the system usability scale. *Int J Human Comput Interact* 2008 Jul 30;24(6):574-594. [doi: [10.1080/10447310802205776](https://doi.org/10.1080/10447310802205776)]
52. Bangor A, Kortum P, Miller J. Determining what individual SUS scores mean: adding an adjective rating scale. *J Usability Stud* 2009 May;4(3):114-123.
53. Mercer K, Giangregorio L, Schneider E, Chilana P, Li M, Grindrod K. Acceptance of commercially available wearable activity trackers among adults aged over 50 and with chronic illness: a mixed-methods evaluation. *JMIR Mhealth Uhealth* 2016 Jan 27;4(1):e7 [FREE Full text] [doi: [10.2196/mhealth.4225](https://doi.org/10.2196/mhealth.4225)] [Medline: [26818775](https://pubmed.ncbi.nlm.nih.gov/26818775/)]
54. Leon AC, Davis LL, Kraemer HC. The role and interpretation of pilot studies in clinical research. *J Psychiatr Res* 2011 May;45(5):626-629 [FREE Full text] [doi: [10.1016/j.jpsychires.2010.10.008](https://doi.org/10.1016/j.jpsychires.2010.10.008)] [Medline: [21035130](https://pubmed.ncbi.nlm.nih.gov/21035130/)]
55. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. 2nd edition. Mahwah, NJ, USA: Lawrence-Erlbaum Associates; 1988.
56. Direito A, Tooley M, Hinbarji M, Albatal R, Jiang Y, Whittaker R, et al. Tailored daily activity: an adaptive physical activity smartphone intervention. *Telemed J E Health* 2020 Apr;26(4):426-437. [doi: [10.1089/tmj.2019.0034](https://doi.org/10.1089/tmj.2019.0034)] [Medline: [31063038](https://pubmed.ncbi.nlm.nih.gov/31063038/)]
57. Herrmann LK, Kim J. The fitness of apps: a theory-based examination of mobile fitness app usage over 5 months. *Mhealth* 2017 Jan 30;3:2 [FREE Full text] [doi: [10.21037/mhealth.2017.01.03](https://doi.org/10.21037/mhealth.2017.01.03)] [Medline: [28293619](https://pubmed.ncbi.nlm.nih.gov/28293619/)]
58. Bell L, Garnett C, Qian T, Perski O, Williamson E, Potts HW. Engagement with a behavior change app for alcohol reduction: data visualization for longitudinal observational study. *J Med Internet Res* 2020 Dec 11;22(12):e23369 [FREE Full text] [doi: [10.2196/23369](https://doi.org/10.2196/23369)] [Medline: [33306026](https://pubmed.ncbi.nlm.nih.gov/33306026/)]
59. Malte CA, Dulin PL, Baer JS, Fortney JC, Danner AN, Lott AM, et al. Usability and acceptability of a mobile app for the self-management of alcohol misuse among veterans (step away): pilot cohort study. *JMIR Mhealth Uhealth* 2021 Apr 08;9(4):e25927 [FREE Full text] [doi: [10.2196/25927](https://doi.org/10.2196/25927)] [Medline: [33830064](https://pubmed.ncbi.nlm.nih.gov/33830064/)]
60. Yang X, Ma L, Zhao X, Kankanhalli A. Factors influencing user's adherence to physical activity applications: a scoping literature review and future directions. *Int J Med Inform* 2020 Feb;134:104039. [doi: [10.1016/j.ijmedinf.2019.104039](https://doi.org/10.1016/j.ijmedinf.2019.104039)] [Medline: [31865054](https://pubmed.ncbi.nlm.nih.gov/31865054/)]
61. Newton Jr RL, Carter L, St Romain J, Jerrod T, Griffith DM, Myers V. Development of a mobile phone app to maintain physical activity in African American men: MobileMen. *Mhealth* 2019 Jun 13;5:16 [FREE Full text] [doi: [10.21037/mhealth.2019.05.03](https://doi.org/10.21037/mhealth.2019.05.03)] [Medline: [31380408](https://pubmed.ncbi.nlm.nih.gov/31380408/)]
62. Garnett C, Perski O, Tombor I, West R, Michie S, Brown J. Predictors of engagement, response to follow up, and extent of alcohol reduction in users of a smartphone app (Drink less): secondary analysis of a factorial randomized controlled trial. *JMIR Mhealth Uhealth* 2018 Dec 14;6(12):e11175 [FREE Full text] [doi: [10.2196/11175](https://doi.org/10.2196/11175)] [Medline: [30552081](https://pubmed.ncbi.nlm.nih.gov/30552081/)]
63. Loeckx M, Rabinovich RA, Demeyer H, Louvaris Z, Tanner R, Rubio N, et al. Smartphone-based physical activity telecoaching in chronic obstructive pulmonary disease: mixed-methods study on patient experiences and lessons for implementation. *JMIR Mhealth Uhealth* 2018 Dec 21;6(12):e200 [FREE Full text] [doi: [10.2196/mhealth.9774](https://doi.org/10.2196/mhealth.9774)] [Medline: [30578215](https://pubmed.ncbi.nlm.nih.gov/30578215/)]
64. Bravata DM, Smith-Spangler C, Sundaram V, Gienger AL, Lin N, Lewis R, et al. Using pedometers to increase physical activity and improve health: a systematic review. *JAMA* 2007 Nov 21;298(19):2296-2304. [doi: [10.1001/jama.298.19.2296](https://doi.org/10.1001/jama.298.19.2296)] [Medline: [18029834](https://pubmed.ncbi.nlm.nih.gov/18029834/)]
65. Romeo A, Edney S, Plotnikoff R, Curtis R, Ryan J, Sanders I, et al. Can smartphone apps increase physical activity? Systematic review and meta-analysis. *J Med Internet Res* 2019 Mar 19;21(3):e12053 [FREE Full text] [doi: [10.2196/12053](https://doi.org/10.2196/12053)] [Medline: [30888321](https://pubmed.ncbi.nlm.nih.gov/30888321/)]
66. Marin TS, Kourbelis C, Foote J, Newman P, Brown A, Daniel M, et al. Examining adherence to activity monitoring devices to improve physical activity in adults with cardiovascular disease: a systematic review. *Eur J Prev Cardiol* 2019 Mar;26(4):382-397. [doi: [10.1177/2047487318805580](https://doi.org/10.1177/2047487318805580)] [Medline: [30322268](https://pubmed.ncbi.nlm.nih.gov/30322268/)]
67. Feehan LM, Geldman J, Sayre EC, Park C, Ezzat AM, Yoo JY, et al. Accuracy of Fitbit devices: systematic review and narrative syntheses of quantitative data. *JMIR Mhealth Uhealth* 2018 Aug 09;6(8):e10527 [FREE Full text] [doi: [10.2196/10527](https://doi.org/10.2196/10527)] [Medline: [30093371](https://pubmed.ncbi.nlm.nih.gov/30093371/)]
68. Klasnja P, Smith S, Seewald NJ, Lee A, Hall K, Luers B, et al. Efficacy of contextually tailored suggestions for physical activity: a micro-randomized optimization trial of HeartSteps. *Ann Behav Med* 2019 May 03;53(6):573-582 [FREE Full text] [doi: [10.1093/abm/kay067](https://doi.org/10.1093/abm/kay067)] [Medline: [30192907](https://pubmed.ncbi.nlm.nih.gov/30192907/)]
69. Tofighi B, Chemi C, Ruiz-Valcarcel J, Hein P, Hu L. Smartphone apps targeting alcohol and illicit substance use: systematic search in in commercial app stores and critical content analysis. *JMIR Mhealth Uhealth* 2019 Apr 22;7(4):e11831 [FREE Full text] [doi: [10.2196/11831](https://doi.org/10.2196/11831)] [Medline: [31008713](https://pubmed.ncbi.nlm.nih.gov/31008713/)]
70. Colbert S, Thornton L, Richmond R. Smartphone apps for managing alcohol consumption: a literature review. *Addict Sci Clin Pract* 2020 May 07;15(1):17 [FREE Full text] [doi: [10.1186/s13722-020-00190-x](https://doi.org/10.1186/s13722-020-00190-x)] [Medline: [32381062](https://pubmed.ncbi.nlm.nih.gov/32381062/)]

## Abbreviations

**ADP:** Alcohol and Drug Partial  
**AUD:** alcohol use disorder  
**EOT:** end of treatment  
**mHealth:** mobile health  
**PA:** physical activity  
**SUS:** System Usability Scale  
**TLFB:** Timeline Follow Back

*Edited by A Mavragani; submitted 23.12.21; peer-reviewed by S Badawy, T Leung; comments to author 04.05.22; revised version received 20.06.22; accepted 21.06.22; published 19.10.22.*

*Please cite as:*

*Abrantes AM, Meshesha LZ, E Blevins C, Battle CL, Lindsay C, Marsh E, Feltus S, Buman M, Agu E, Stein M  
A Smartphone Physical Activity App for Patients in Alcohol Treatment: Single-Arm Feasibility Trial  
JMIR Form Res 2022;6(10):e35926  
URL: <https://formative.jmir.org/2022/10/e35926>  
doi: [10.2196/35926](https://doi.org/10.2196/35926)  
PMID: [36260381](https://pubmed.ncbi.nlm.nih.gov/36260381/)*

©Ana M Abrantes, Lidia Z Meshesha, Claire E Blevins, Cynthia L Battle, Clifford Lindsay, Eliza Marsh, Sage Feltus, Matthew Buman, Emmanuel Agu, Michael Stein. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 19.10.2022. 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

# Shared Decision-Making for Drug-Drug Interactions: Formative Evaluation of an Anticoagulant Drug Interaction

Ainhoa Gomez Lumbreras<sup>1\*</sup>, MD, PhD; Thomas J Reese<sup>2\*</sup>, PhD, PharmD; Guilherme Del Fiol<sup>3\*</sup>, MD, PhD; Malinda S Tan<sup>1</sup>, PharmD; Jorie M Butler<sup>3</sup>, PhD; Jason T Hurwitz<sup>4</sup>, PhD; Mary Brown<sup>5</sup>, PhD; Kensaku Kawamoto<sup>3</sup>, MHS, MD, PhD; Henrik Thies<sup>6</sup>, MS; Maria Wright<sup>1</sup>; Daniel C Malone<sup>1</sup>, RPh, PhD

<sup>1</sup>Department of Pharmacotherapy, Skaggs College of Pharmacy, University of Utah, Salt Lake City, UT, United States

<sup>2</sup>Department of Biomedical Informatics, Vanderbilt University, Nashville, TN, United States

<sup>3</sup>Department of Biomedical Informatics, University of Utah, Salt Lake City, UT, United States

<sup>4</sup>Center for Health Outcomes and Pharmacoeconomic Research, University of Arizona, Tucson, AZ, United States

<sup>5</sup>University of Arizona, Tucson, AZ, United States

<sup>6</sup>University of Heidelberg, Heidelberg, Germany

\* these authors contributed equally

**Corresponding Author:**

Daniel C Malone, RPh, PhD

Department of Pharmacotherapy

Skaggs College of Pharmacy

University of Utah

30 S 2000 East, Room 4922

Salt Lake City, UT, 84112

United States

Phone: 1 801 581 8054

Email: [dan.malone@utah.edu](mailto:dan.malone@utah.edu)

## Abstract

**Background:** Warnings about drug-drug interactions (DDIs) between warfarin and nonsteroidal anti-inflammatory drugs (NSAIDs) within electronic health records indicate potential harm but fail to account for contextual factors and preferences. We developed a tool called DDInteract to enhance and support shared decision-making (SDM) between patients and physicians when both warfarin and NSAIDs are used concurrently. DDInteract was designed to be integrated into electronic health records using interoperability standards.

**Objective:** The purpose of this study was to conduct a formative evaluation of a DDInteract that incorporates patient and product contextual factors to estimate the risk of bleeding.

**Methods:** A randomized formative evaluation was conducted to compare DDInteract to usual care (UC) using physician-patient dyads. Using case vignettes, physicians and patients on warfarin participated in simulated virtual clinical encounters where they discussed the use of taking ibuprofen and warfarin concurrently and determined an appropriate therapeutic plan based on the patient's individualized risk. Dyads were randomized to either DDInteract or UC. Participants completed a postsession interview and survey of the SDM process. This included the 9-item Shared Decision-Making Questionnaire (SDM-Q-9), tool usability and workload National Aeronautics and Space Administration (NASA) Task Load Index, Unified Theory of Acceptance and Use of Technology (UTAUT), Perceived Behavioral Control (PBC) scale, System Usability Scale (SUS), and Decision Conflict Scale (DCS). They also were interviewed after the session to obtain perceptions on DDInteract and UC resources for DDIs.

**Results:** Twelve dyad encounters were performed using virtual software. Most (n=11, 91.7%) patients were over 50 years of age, and 9 (75%) had been taking warfarin for more than 2 years (75%). Regarding scores on the SDM-Q-9, participants rated DDInteract higher than UC for questions pertaining to helping patients clarify the decision ( $P=.03$ ), involving patients in the decision ( $P=.01$ ), displaying treatment options ( $P<.001$ ), identifying advantages and disadvantages ( $P=.01$ ), and facilitating patient understanding ( $P=.01$ ) and discussion of preferences ( $P=.01$ ). Five of the 8 UTAUT constructs showed differences between the 2 groups, favoring DDInteract ( $P<.05$ ). Usability ratings from the SUS were significantly higher ( $P<.05$ ) for physicians using DDInteract compared to those in the UC group but showed no differences from the patient's perspective. No differences in patient responses were observed between groups using the DCS. During the session debrief, physicians indicated little concern for the

additional time or workload entailed by DDInteract use. Both clinicians and patients indicated that the tool was beneficial in simulated encounters to understand and mitigate the risk of harm from this DDI.

**Conclusions:** Overall, DDInteract may improve encounters where there is a risk of bleeding due to a potential drug-drug interaction involving anticoagulants. Participants rated DDInteract as logical and useful for enhancing SDM. They reported that they would be willing to use the tool for an interaction involving warfarin and NSAIDs.

(*JMIR Form Res* 2022;6(10):e40018) doi:[10.2196/40018](https://doi.org/10.2196/40018)

## KEYWORDS

decision making, shared; decision support systems, clinical; decision making; decision support; user-centered design; patient-centered care; risk management; drug interaction; pharmacotherapy; pharmacy; pharmaceutical; warfarin; unified theory of acceptance and use of technology; UTAUT; NSAID; anti-inflammatory; non-steroidal

## Introduction

Approximately 8 million people in the United States received anticoagulants, including warfarin, in 2019 [1]. Warfarin is well known for having a multitude of drug-drug interactions (DDIs), with many avoided through health care provider awareness. However, studies found that up to 24% of people taking warfarin also received a prescription for nonsteroidal anti-inflammatory drugs (NSAIDs), and almost 50% of patients on warfarin have some form of ongoing pain [2,3]. Concomitant use of a NSAID and warfarin increases the risk of bleeding [4,5]. The risk of gastrointestinal (GI) bleeding is nearly 2-fold greater with this combination compared to using warfarin alone [5]. Furthermore, because ibuprofen and naproxen sodium are NSAIDs that are available over the counter (OTC), it is critical that patients understand the risks of taking the 2 medications concurrently. While the use of warfarin has declined since the approval of direct oral anticoagulants in 2013, it is the only anticoagulant indicated for mechanical valve replacement and is used by at least half of the patients taking oral anticoagulants [6-8]. Furthermore, patients' knowledge of the adverse effects of NSAIDs is incomplete [9,10].

To improve health outcomes, stakeholders have advocated for the incorporation of shared decision-making (SDM) that encourages physicians to involve patients in selecting therapeutic treatments, among others [11]. SDM improves patient behavior and results in higher adherence to physician guidance, leading to improved health outcomes [12,13]. Currently, there are SDM tools available to facilitate decisions regarding nonvalvular atrial fibrillation anticoagulant treatment [14-18]. However, none of these SDM tools focus on or incorporate anticoagulant DDIs and their associated risks [19].

Current DDI alerting systems within electronic health records (EHRs) are clinician centric, suffer from high override rates, and do not account for patient and drug attributes that affect risk of harm from the interaction [20]. Furthermore, the primary source of DDI information (online DDI checkers) do not provide any personalized quantification of the risk of harm [21-24]. Accordingly, we developed DDInteract, a tool to facilitate SDM for patients receiving warfarin that calculates evidence-based risk of GI bleeding by considering an individual patient's risk factors such as age, history of previous GI bleeding, and other

medications [25]. We previously reported on the user-centered design and usability assessment of DDInteract. The purpose of this study was to conduct a formative evaluation comparing DDInteract to usual care (UC).

## Methods

### Study Design

This manuscript follows SUNDAE (Standards for Universal Reporting of Decision Aid Evaluations) guidelines [26]. We conducted a randomized formative evaluation of DDInteract compared to UC from March 2021 to August 2021. Formative evaluation is a research methodology used to rigorously examine factors that might influence the progress during the implementation of health care innovations [27]. We used a mixed methods approach, as we collected both quantitative and qualitative data.

### Ethics Approval

This project was approved by University of Utah Institutional Review Board (00127062). All participants signed the informed consent before participating in this study.

### Intervention

We designed DDInteract to be a free SDM tool that is integrated with the EHR to mitigate risk of harm due to DDIs involving warfarin and NSAIDs [25]. When launched from within an EHR, the tool retrieves risk factors from the patient's record (eg, history of GI bleeding, age, medications) and calculates the patient's risk of bleeding. A dynamic risk array displays and updates the risk of GI bleeding in real time based on clinician input, patient history, and medication choices (Figure 1). When an NSAID is selected, the risk array changes to reflect the increase in risk due to the combined effects of both agents on risk of bleeding. The tool also (1) provides information on nonmedication treatments for pain, such as acupuncture and physical therapy; (2) includes strategies to reduce risk of GI bleeding through use of proton pump inhibitors; (3) facilitates the medication prescription process; (4) automates clinical documentation of the SDM discussion; and (5) provides patient education in an "after visit summary." Further details about DDInteract functionality are described elsewhere, and the tool is available for demonstration purposes on the DDInteract website [28].

Figure 1. DDInteract tool (partial screen).

**Shared decision aid to reduce gastrointestinal (stomach) bleed risk**

Increased risk of **gastrointestinal (stomach) bleed** with **warfarin-NSAID interaction**

**Risk Profile for Patient**

Major Risk Factors

- Previous gastrointestinal bleed (Gastrointestinal hemorrhage (02/03/2018))
- Older than 65
- On warfarin (warfarin (Coumadin) #30 (08/03/2020))
- On aspirin or clopidogrel
- On Selective Serotonin Reuptake Inhibitor (SSRI) (fluoxetine (Prozac) #30 (08/03/2020))
- On systemic corticosteroid
- On spironolactone or eplerenone

Estimated Gastrointestinal Bleeds (100 patients)

48 potential bleeds with this Risk Profile and celecoxib

16 fewer bleeds with celecoxib compared to ibuprofen

Points to discuss in making this treatment decision:

- Goal is to reduce pain AND risk of stomach bleeding
- Options have different benefits and risks
- Best option may depend on the patient's preferences
- Summarize the decision and plan for follow up

Answer these questions to help the patient decide:

How do you prefer to treat your pain?

Medication  Non-medication

What type of pain medication do you prefer?

Oral NSAID  Other medication

Less risk

- celecoxib (Celebrex) 100mg Take 1 capsule every 12 hours... \$\$\$
- diclofenac (Voltaren) 75mg Take 1 tablet every 12 hours... \$\$
- ibuprofen (Motrin) 600mg Take 1 tablet every 8 hours... \$
- naproxen (Aleve) 250mg Take 1 tablet every 8 hours... \$

## Study Settings

We randomized physician-patient dyads to participate in simulated clinical encounters using either DDInteract or conventional drug information resources (UC) to treat a patient on warfarin who is also seeking relief of pain by using an NSAID. Physicians in both groups (DDInteract or UC) received access to various online DDI resources (eg, Micromedex), while those in the intervention group also had access to DDInteract.

## Recruitment: Inclusion and Exclusion Criteria

We recruited physicians through electronic communications, presentations to medical groups, and snowball sampling. The inclusion criteria for physicians required experience with treating patients on oral anticoagulants. We recruited patients from the University of Utah Health System, including the Division of Cardiology and the Anticoagulation Service. Patient participants needed to be older than 21 years of age, fluent in English, and currently taking an oral anticoagulant.

## Encounters

We integrated and deployed DDInteract with the Logica sandbox EHR environment using the Substitutable Medical Applications and Reusable Technologies (SMART) guidelines on Fast Health Interoperability Resources (FHIR) standards [29]. Details on the technical architecture including interoperability approach are available elsewhere [30]. Physician-patient dyads participated in a simulated virtual encounter using an online meeting software and provided consent prior to starting the simulated encounter. We instructed both physicians and patients to conduct themselves as they would in an encounter where an anticoagulated patient wants to use ibuprofen to treat their pain

but has concerns about bleeding. Participants permitted us to record the encounters. Physicians and patients received instructions separately via breakout rooms and asked study investigators any questions prior to the simulated visit.

Physicians randomized to the DDInteract group received a link to a video demonstrating the functionality of DDInteract, while those randomized to the UC group received information on the main topic and of online drug information sources that can screen for drug interactions (see Table S1 in the [Multimedia Appendix 1](#)). Physicians in both groups received access to online resources for DDIs to use during the encounter at their discretion. After the simulated encounters, both physicians and patients participated in a semistructured interview with a member of the research team to discuss the encounter and use of DDInteract and online drug information resources. Upon concluding the simulated encounter, participants completed an online survey containing the instruments described in the sequent section.

## Outcomes

The primary outcome was quality of SDM according to the 9-item Shared Decision-Making Questionnaire (SDM-Q-9) questionnaire assessed by physicians. The physician postsimulation survey included items in 6 groupings: (1) SDM-Q-9 scale [31]; (2) adapted System Usability Scale (SUS) [32]; (3) National Aeronautics and Space Administration (NASA) Task Load Index (NASA-TLX) [33]; (4) Unified Theory of Acceptance and Use of Technology (UTAUT) [34], including the performance expectancy, effort expectancy, attitude toward using technology, social influence, facilitating conditions, self-efficacy, anxiety, and behavior constructs; (5)

Perceived Behavioral Control (PBC) scale [35]; and (6) questions pertaining to participants' demographics and professional experience. The patients' survey included 6 items adapted from the SUS scale, 11 items adapted from the Decisional Conflict Scale (DCS) [36], and questions about their demographics and anticoagulation treatment.

### Statistical Analysis

We calculated the mean and standard deviation for each item/construct from the various scales and instruments. For SDM quality, which was the primary outcome, we conducted a 2-group Student *t*-test assessing differences for DDInteract versus UC as rated by physicians. A similar approach was used to test for differences in secondary measures including the SUS, NASA-TLX, and UTAUT. The PBC scale rated by physicians was also analyzed using the same approach, as well as items on the DCS. The total time of the simulated encounter for both groups was compared using Student *t*-test. All statistical comparisons used a 2-tailed test with an alpha <.05. Given the relatively small sample size, no adjustments were made for multiple comparisons. Since we could not assume that the populations were normally distributed, we estimated the Mann-Whitney U test for all the outcomes results reported. We also reviewed transcriptions of the post encounter semistructured interviews and highlighted the most insightful comments. All transcriptions were reviewed, and the principal investigators extracted the most relevant and repeated comments. Data were analyzed using Microsoft Excel for Mac 2016 (IBM Corp) and Stata (version 17, basic edition).

## Results

### DDInteract Tool and UC Survey Results

A total of 12 physician-patient dyads completed the formative evaluation. We randomized 6 dyads to DDInteract and 6 to UC. Participant demographics appear in Table 1. The UC group had 1 more female physician and 2 more female patients than the DDInteract group. All patients and 9 (75%) of the physicians reported being Caucasian. There was no difference in physicians' self-reported experience managing patients with warfarin between the 2 groups (DDInteract 68.8, SD 31.5 vs UC 77.7, SD 20.6, from a 0-100 scale,  $P=.58$ ).

The primary outcome of interest (SDM), as measured by the SDM-Q-9, appears in Table 2. Perceptions of SDM differed significantly between DDInteract and UC groups for 6 of the 9 attributes ( $P<.05$ ) and include: facilitated the discussion on preferences, assisted patient comprehension, presented different options, presented advantages and disadvantages of the different treatment options, involved the patient, and provided clarity on the decision to be made. Not surprisingly, the largest difference between the groups was with respect to presentation of different treatment options (mean 5.7, SD 0.5 vs mean 1.8, SD 1.2,  $P<.001$ ) and presentation of advantages and disadvantages

(mean 5.0, SD 1.1 vs mean 2.8, SD 1.2,  $P=.008$ ) for DDI and UC, respectively.

The task load assessed by the NASA-TLX index, scaled from zero (low/easy/successful) to 10 (high/demanding/unsuccessful), was lower for the DDInteract group than the UC group; it was not statistically significant but trended toward significance (mean 2.0, SD 1.2 vs mean 4.2, SD 2.1,  $P=.08$ , respectively) (Table 3). Participants indicated that DDInteract took significantly less effort to use than the tools in UC (mean 0.8, SD 0.7 vs mean 7.3, SD 3.7,  $P=.008$ , respectively), but none of the other constructs differed significantly between the study groups.

Patient-reported scores on the adapted SUS scale did not differ between groups, but physician ratings differed significantly, with DDInteract perceived as more logical, efficient, and helpful/effective than UC and SDM perceived as more valuable than using traditional DDI tools in UC ( $P<.05$ ) (Table 4).

Physicians' responses on a Likert scale ranging from 1 (strongly disagree) to 7 (strongly agree) in the DDInteract group differed from those in the UC group on the following UTAUT constructs: performance expectancy (mean 5.6, SD 1.2 vs mean 3.4, SD 1.7,  $P<.001$ ), effort expectancy (mean 6.8, SD 0.5 vs mean 5.2, SD 1.6,  $P<.001$ ), attitude toward using technology (mean 5.7, SD 1.1 vs mean 3.9, SD 1.8,  $P<.001$ ), social influence (mean 5.5, SD 0.8 vs mean 4.1, SD 1.8,  $P<.001$ ), and anxiety (mean 1.4, SD 0.5 vs mean 2.5, SD 1.5,  $P<.001$ ), respectively (Table 5). In general, it appears that physicians perceived DDInteract to perform better, require less effort, improve attitudes toward technology, be supported by administration and colleagues, and reduce perceived anxiety more compared to UC.

Table 6 shows the physicians' assessment of PBC. Physicians exposed to DDInteract were more likely to indicate using SDM in the clinic compared to physicians randomized to UC (mean 6.8, SD 0.4 vs mean 6.0, SD 0.6,  $P=.02$ , respectively). DDInteract physicians were also more likely to perceive conducting SDM without extending the duration of the visit than UC physicians (mean 5.5, SD 0.5 vs mean 3.7, SD 1.5,  $P=.03$ ). However, the mean duration of the encounters increased 5 minutes on average for the DDInteract, but this increase did not significantly differ between the DDInteract and UC groups (mean 17.6, SD 5.4 minutes vs mean 12.7, SD 4.8 minutes,  $P=.13$ ). No other PBC items differed significantly between the 2 groups. There were no differences between the groups in patients' ratings on the DCS (Table S2 in Multimedia Appendix 1).

The results of the nonparametric tests for the different outcomes (SDM-Q-9, UTAUT, SUS, NASA-TLX, PBC, and DCS) did not differ in terms of statistical significance from the ones presented above (Tables S3-S9 in Multimedia Appendix 1). Qualitative narrative data were collected to gain insight into participants thoughts on SDM and DDI.

**Table 1.** Participants' sociodemographic, professional, and treatment characteristics.

Characteristics	Physicians			Patients		
	DDInteract (N=6), n (%)	UC <sup>a</sup> (N=6), n (%)	P value	DDInteract (N=6), n (%)	UC (N=6), n (%)	P value
<b>Age</b>			.34			.39
21-29	1 (16.7)	—		—	—	
30-39	2 (33.3)	1 (16.7)		—	—	
40-49	3 (50)	2 (33.3)		1 (16.7)	—	
50-59	—	1(16.7)		3 (50)	2 (33.3)	
≥60	—	2 (33.3)		2 (33.3)	4 (66.7)	
Female sex	2 (33.3)	3 (50)	.56	3 (50)	5 (83.3)	.22
<b>Race</b>			.37			>.99
American Indian/ Alaska Native	—	—		—	—	
Asian/Pacific Islander	—	1 (16.7)		—	—	
African American	—	—		—	—	
Hispanic	1 (16.7)	—		—	—	
White/ Caucasian	4 (66.7)	5 (83.3)		6 (100)	6 (100)	
Multiple ethnicity/other	1 (16.7)	—		—	—	
<b>Clinical experience</b>			.64			—
<5 years	1 (16.7)	1 (16.7)		—	—	
6-10 years	2 (33.3)	1 (16.7)		—	—	
11-15 years	2 (33.3)	1 (16.7)		—	—	
≥16 years	1 (16.7)	3 (50)		—	—	
<b>Level of education</b>			—			.99
College graduate	—	—		2 (33.3)	3 (50)	
Attended college	—	—		1 (16.7)	1 (16.7)	
High school graduate	—	—		1 (16.7)	1 (16.7)	
Completed graduate school	—	—		2 (33.3)	1 (16.7)	
<b>Time on anticoagulant</b>			—			.50
≤6 months	—	—		2 (33.3)	1 (16.7)	
>6 months ≤2 years	—	—		0	0	
>2 years	—	—		4 (66.7)	5	
<b>Time in clinical care</b>			.08			—
≤20% week	—	2 (33.3)		—	—	
21%-40% week	1 (16.7)	1 (16.7)		—	—	
41%-60% week	—	2 (33.3)		—	—	
61%-80% week	—	—		—	—	
≥81% week	5 (83.3)	1 (16.7)		—	—	
<b>Specialty</b>			.56			—
Internal medicine	3 (50)	3 (50)		—	—	
Family medicine	1 (16.7)	2 (33.3)		—	—	
Geriatric medicine	1 (16.7)	—		—	—	
Other	1 (16.7)	1 (16.7)		—	—	

<sup>a</sup>UC: usual care.

**Table 2.** The 9-item Shared Decision-Making Questionnaire (SDM-Q-9).

	Physicians, mean (SD)		<i>P</i> value
	DDInteract	UC <sup>a</sup>	
Clarity of the SDM <sup>b</sup>	4.5 (0.8)	3.3 (0.8)	.03
Involve the patient	5.5 (0.5)	3.2 (1.3)	.01
Different option presented	5.7 (0.5)	1.8 (1.2)	<.001
Presented advantages/disadvantages	5.0 (1.1)	2.8 (1.2)	.01
Assisted patient comprehension	4.8 (0.4)	2.7 (1.4)	.01
Facilitated discussion on preferences	5.2 (0.8)	3.0 (1.3)	.01
Evaluated options	5.0 (0.0)	4.7 (0.8)	.36
Co-selected a treatment option	5.2 (0.4)	5.0 (0.9)	.69
Reach an agreement	5.7 (0.5)	5.3 (0.8)	.42

<sup>a</sup>UC: usual care.

<sup>b</sup>SDM: shared decision-making.

**Table 3.** Adapted National Aeronautics and Space Administration Task Load Index (NASA-TLX)<sup>a</sup>.

NASA <sup>b</sup> Task Load Index construct	Questions	Physicians, mean (SD)		<i>P</i> value
		DDInteract	UC <sup>c</sup>	
Mental demand	How much mental effort was required to decide on the patient's treatment?	3.8 (1.0)	4.7 (3.2)	.57
Physical effort	Was using the DDInteract decision tool easy or demanding?	0.8 (0.7)	7.3 (3.7)	.008
Temporal demand	How much time did it take to investigate the drug interaction during the simulated visit?	1.8 (1.6)	3.6 (3.5)	.28
Effort	How hard did you have to work to make a decision with the patient?	2.3 (1.0)	3.6 (2.6)	.30
Performance	How successful do you think you were in making a shared decision with the DDInteract decision tool?	1 (0.9)	1.7 (1.8)	.45
Total average		2.0 (1.2)	4.2 (2.1)	.08

<sup>a</sup>Likert scale from 0 (low/easy/successful) to 10 (high, demanding, unsuccessful).

<sup>b</sup>NASA: National Aeronautics and Space Administration.

<sup>c</sup>UC: usual care.

**Table 4.** Modified System Usability Scale (SUS).

Items	Patients, mean (SD)		<i>P</i> value	Physicians, mean (SD)		<i>P</i> value
	DDInteract	UC <sup>a</sup>		DDInteract	UC	
It was logical	4.7 (0.5)	4.8 (0.4)	.65	4.8 (0.4)	3.7 (1)	.03
It was efficient	4.5 (0.5)	4.4 (0.5)	.77	4.7 (0.5)	3.5 (1)	.04
Helpful/effective in the decision-making process	4.7 (0.5)	4.4 (0.9)	.57	4.5 (0.5)	2.8 (1.2)	.02
The SDM <sup>b</sup> using the tool was valuable	4.5 (0.5)	4.8 (0.4)	.26	4.7 (0.5)	2.8 (0.7)	<.001
The tool was valuable	N/A <sup>c</sup>	N/A	N/A	4.3 (0.5)	2.8 (1.2)	.02
Easy to use	4.5 (0.5)	4.4 (0.9)	.83	4.8 (0.4)	3.2 (1.5)	.04
Enjoyed the experience	4.7 (0.5)	4.8 (0.4)	.55	N/A	N/A	N/A
Learned something from this experience	4.8 (0.4)	4.7 (0.8)	.67	4.5 (0.5)	2.8 (1.7)	.06

<sup>a</sup>UC: usual care.

<sup>b</sup>SDM: shared decision-making.

<sup>c</sup>N/A: not applicable.



**Table 5.** Unified Theory of Acceptance and Use of Technology (UTAUT) responses.

	Physicians, mean (SD)		P value
	DDInteract	UC <sup>a</sup>	
Performance expectancy	5.6 (1.2)	3.4 (1.8)	<.001
Effort expectancy	6.8 (0.5)	5.2 (1.6)	<.001
Attitude toward using technology	5.7 (1.1)	3.9 (1.8)	<.001
Social influence	5.5 (0.9)	4.1 (1.8)	<.001
Self-efficacy	5.3 (2.1)	4.6 (2.2)	.26
Facilitating conditions	5.1 (1.9)	5.4 (1.7)	.58
Anxiety	1.4 (0.5)	2.5 (1.5)	<.001
Behavioral	5.5 (1.0)	4.7 (2.9)	.53

<sup>a</sup>UC: usual care.

**Table 6.** Perceived Behavioral Control (PBC) scale.

Item	Physicians, mean (SD)		P value
	DDInteract	UC <sup>a</sup>	
I am convinced that I can share decision - making in the clinic	6.8 (0.4)	6 (0.6)	.02
I have control over the level of SDM <sup>b</sup> that is accomplished in the clinic	6.2 (0.8)	5.7 (0.5)	.21
I can perform SDM without extending the duration of the consultation	5.5 (0.5)	3.7 (1.5)	.03
Knowledge about SDM is important in order to apply SDM	6.5 (0.5)	5.7 (1.4)	.20
Communication skills are important for SDM	7 (0)	6.7 (0.5)	.17
Patients are motivated to participate in SDM	6 (0.9)	5.5 (1.2)	.44
In general patients have enough knowledge, intelligence and understanding needed for SDM	6 (0.9)	5 (1.7)	.23

<sup>a</sup>UC: usual care.

<sup>b</sup>SDM: shared decision-making.

### Physician and Patient Perceptions Regarding the DDInteract Tool and Online DDI Resources

After the virtual encounters, both clinicians and patients discussed the simulation using a semistructured interview format to elicit their impressions of DDInteract and online DDI resources. Below are select quotes from the participants. In general, DDInteract was well received by the physicians, as indicated by the following statements:

*I like it; it helps patients visualize the risk instead of me just talking statistics, they see the risk in the more obvious way. I like the intervention that shows the risk of bleeding, you can increase more risk factors for bleeding, the idea is excellent.* [Physician #4, male]

*I really like it. Whenever I can, I like to show something visual while talking to the patient. It is very user friendly, simple. It is not overly complicated.* [Physician #8, female]

Concerns about extra time involved in using DDInteract were not evident among the participants. One physician stated:

*I am trying to imagine having this conversation without the tool. I don't even think the tool would even make it longer; I think what it does is cut down on having to overly explain things, cut down on the feeling that I have to reemphasize things because it was a visual tool, and the patient is seeing what I am seeing, so they don't ask for repeating. So I probably save some questions too. In all I think it would save some time.* [Physician #8, female]

Another physician felt that the tool would encourage more conversations with the patient and help them understand the risks associated with bleeding and anticoagulation treatment:

*I think it was fast to use. I see thrombosis patients, so I do a lot of counseling about warfarin or anticoagulation. I think I tend to assume that people have already been through education, but this would be nice because it will slow me down and help me actually understand what the risk is, which I might just summarize very quickly, but I think this would be really helpful from a patient standpoint to actually, like, seeing is something that explains it a little better.* [Physician #10, female]

When patients were asked about the icon array in DDInteract as an approach to display the risk of bleeding, some comments included:

*I think seeing the graphic portrayal of the different risk levels and how to treat the pain was very helpful. I think it was well done...I thought this tool was much more thorough than the information that I've gotten in the normal clinic visit. I did not know the number of bleeds per 100 patients; that has never been discussed with me during my doctor visit.* [Patient #10, female]

*The icon array definitely makes sense to me.* [Patient #12, male]

*I learn better by what I see.* [Patient #9, female]

*I think the chart with different numbers showing different reactions you get by taking ibuprofen kind of spells it out like black and white for me.* [Patient #8, male]

Table S10 in [Multimedia Appendix 1](#) provides additional comments made by the physicians and patients.

## Discussion

### Principal Results

Results of the formative evaluation suggest that there was a positive opinion of DDInteract as a potentially useful tool to help facilitate SDM concerning concomitant therapies that could result in a DDI. Physicians perceived the tool as intuitive, easy to use, and not increasing the amount of time during an encounter. After the simulated encounter, patients exposed to the DDInteract tool commented that they liked the way the information was presented, including the quantitative estimation of the risk and how risk changed according to patient factors in the user interface.

Physicians that used the DDInteract reported enhanced SDM compared to physicians randomized to UC using a standardized and validated tool (SDM-Q-9) [31,37]. This study demonstrates that DDInteract assisted in facilitating discussion, increasing patient comprehension, providing different treatment options with their respective advantages and disadvantages, and providing clarity around the treatment choices and the final decision. This contrasts with current warnings about potential interactions that are clinician orientated and not designed to be shared with patients.

Due to the ubiquitous use of NSAID pain relievers, including both prescription and over-the-counter products, patients play an important role in preventing harm from DDIs. Through education, patients can reduce harm by avoiding interacting OTC medications and monitoring signs and symptoms of adverse consequences [38]. One of the most important aspects of DDInteract is the risk array, which dynamically changes when risk factors or medications are selected. Usability studies have demonstrated that this approach is appealing and helpful for displaying the risk of harm [25]. Furthermore, physicians using DDInteract during the encounter received no additional

training in SDM but rated SDM aspects higher than physicians randomized to UC.

### Secondary Results

Results from the SUS index suggest that physicians rated DDInteract as more logical, efficient, valuable, easy to use, and helpful/effective than physicians in the UC who rated online compendia and traditional EHR tools. This finding is not surprising given the lack of patient-specific information provided by both EHR systems and online resources for DDIs. Only 1 attribute of the SUS (“learned something”) did not differ significantly between groups ( $P=.06$ ). Patient perceptions of DDInteract and UC tools did not differ on the SUS scale, a finding that might be attributable to patients wanting more information about DDIs [39,40].

DDInteract appeared to be intuitive and easy to use according to participant ratings on the NASA-TLX and UTAUT instruments. These results support those of a previously reported usability study that was conducted when designing the tool [25]. The NASA-TLX has been previously used to evaluate the workload associated with the new Clinical Decision Support Tool similar to DDInteract [41]. Compared to UC, physicians rated DDInteract higher on the UTAUT domains of performance expectancy, effort expectancy, attitude toward the technology, and social influence. Furthermore, DDInteract was associated with lower anxiety than UC resources for DDIs. These findings align with a metaregression of studies that evaluated effective clinical decision support using the UTAUT model and suggested that effort expectancy, facilitating conditions, and performance expectancy had a significant impact on clinician behavior [42]. Numerous studies evaluated medication-related CDS using the UTAUT, including studies examining a comprehensive CDS suite of tools called Sentri7/Quantifi [43], use of a decision aid for psychotic disorders [44], medication management in oncology settings [45], and a tool to help remind HIV positive men to use antiviral therapies when undertaking sexual activities [46]. However, to our knowledge, none have applied the UTAUT model to DDI-specific SDM.

Patient responses to the DCS did not differ between groups, which is unsurprising. Studies using the DCS in the context of nonvalvular atrial fibrillation (NVAF) found mixed results with patient decision aids [47,48]. NVAF decision aids focused on anticoagulant treatments show lower DCS scores or a decrease in the DCS score after using the decision aid, but these differences were not significantly different when compared with groups not using such aids [18,47,48]. In addition, the small sample size may contribute to the lack of significant differences between the groups with respect to the DCS in our study.

The time required to use CDS and SDM applications is a common concern among clinicians. Our study found that physicians did not perceive DDInteract to affect the amount of time or effort spent with patients (PBC scale question on performing SDM without extending the time of consultation favored DDInteract,  $P=.03$ ), though there was a nonsignificant 5-minute difference in the time to conduct the simulated encounter, with DDInteract encounters being longer. During the debrief, 1 physician proposed that DDInteract could actually save time because it visually explains the risk of harm and

various treatment options, as compared to having to explain the risks verbally. This disconnect between observed time to use the DDInteract and perceived time to use the tool is probably due to clinicians, who had no previous experience with the tool, having to learn how to use DDInteract during the simulation. Clinicians found the tool easy to use and intuitive, indicating a reduced time to use the tool in future encounters.

DDInteract is a novel approach to providing information about potential interactions. Currently, online drug compendia and warnings within EHR systems have numerous issues [49]. They do not account for individual patient risk factors or provide quantitative estimates of the risk. Additionally, they lack the ability to explore changes in the risk with therapeutic alternatives and are almost completely text based without any visual display of information. Furthermore, they may contain incorrect information. For example, Drugs.com currently states that the mechanism to control bleeding risk in patients on warfarin and ibuprofen is to closely monitor the international normalized ratio (INR); however, concurrent use of NSAIDs and warfarin interaction is a pharmacodynamic interaction that will not affect the INR [24]. Studies have found patients' INRs to be within range among individuals experiencing bleeding and taking both medications, and a systematic review did not find an effect on INR for ibuprofen or naproxen [50,51]. A review of 7 Chinese apps that support DDI checking and target health care professionals had tremendous variation in information about DDIs, with over 50% lacking accurate information on DDIs [52]. Although DDInteract is currently limited in scope in terms of the interactions it includes, the approach is generalizable to other interactions, and efforts are underway to expand these interactions.

### Limitations

This study has several limitations that should be considered when interpreting the results. We recruited fewer participants than our a priori sample size (16 dyads), but we still observed significant differences between the study groups in terms of the

primary outcome of SDM quality and several secondary outcomes of interest. COVID-19 protocols and restrictions limited our recruitment of subjects because clinic visits have been in part shifted to telehealth, making it more difficult to contact patients, and because physicians were overburdened with demands due to the pandemic [53,54]. In addition, this study focused on a specific DDI involving warfarin, which is being prescribed less because of alternative anticoagulants that require less frequent monitoring. Several physicians commented that they would have liked to see DDInteract incorporate other anticoagulants and other bleeding risks beyond GI. As stated previously, efforts are underway to broaden the interactions incorporated into the tool. Another limitation is that the clinical scenario required dyads to role-play a scenario that may not occur in clinical practice, where a patient is seeking advice from the physician about treating pain while on an anticoagulant. While this scenario may be infrequent in most practice settings, previous studies have found this to occur in up to 25% of patients receiving warfarin [2,3]. Furthermore, many physicians in the study indicated they were aware of the interaction when researchers explained the scenario prior to the simulated visit. Thus, many practitioners may have behaved differently when provided with DDInteract or other DDI information. Nonetheless, after the simulated visit, many practitioners exposed to DDInteract were enthusiastic about the tool. Therefore, we believe DDInteract may be useful in many practice settings, especially as the tool is expanded to support SDM for other drug interactions, such as other oral anticoagulants.

### Conclusions

DDInteract is a novel tool designed to facilitate SDM during encounters concerning potentially harmful DDIs. The tool improved SDM, was not overly burdensome according to the NASA-TLX and UTAUT scales, and was generally well received by both physicians and patients. However, further research is needed to evaluate the impact of DDInteract on exposure to harmful DDIs and other important health outcomes.

### Acknowledgments

This project has been supported by the Agency for Health Care Research and Quality (U18HS027099). The content is solely the responsibility of the authors and does not necessarily represent the official views of the Agency for Health Care Research and Quality. Additionally, we would like to thank project coordinator Maria Wright for her administrative support in all aspects of this study.

### Conflicts of Interest

None declared.

Multimedia Appendix 1

Supplementary Tables S1-S10.

[DOCX File, 29 KB - [formative\\_v6i10e40018\\_app1.docx](#) ]

### References

1. Bleeding risk awareness survey. National Blood Clot Alliance. Philadelphia, PA; 2020 Jun 23. URL: <https://www.stopthecлот.org/news/bleeding-risk-survey/> [accessed 2021-12-14]

2. Malone D, Hutchins D, Hauptert H, Hansten P, Duncan B, Van Bergen RC, et al. Assessment of potential drug-drug interactions with a prescription claims database. *Am J Health Syst Pharm* 2005 Oct 01;62(19):1983-1991. [doi: [10.2146/ajhp040567](https://doi.org/10.2146/ajhp040567)] [Medline: [16174833](https://pubmed.ncbi.nlm.nih.gov/16174833/)]
3. Gabilondo M, Loza J, Pereda A, Caballero O, Zamora N, Gorostiza A, et al. Quality of life in patients with nonvalvular atrial fibrillation treated with oral anticoagulants. *Hematology* 2021 Dec 25;26(1):277-283. [doi: [10.1080/16078454.2021.1892329](https://doi.org/10.1080/16078454.2021.1892329)] [Medline: [33631081](https://pubmed.ncbi.nlm.nih.gov/33631081/)]
4. Delaney JA, Opatrny L, Brophy JM, Suissa S. Drug drug interactions between antithrombotic medications and the risk of gastrointestinal bleeding. *CMAJ* 2007 Aug 14;177(4):347-351 [FREE Full text] [doi: [10.1503/cmaj.070186](https://doi.org/10.1503/cmaj.070186)] [Medline: [17698822](https://pubmed.ncbi.nlm.nih.gov/17698822/)]
5. Villa Zapata L, Hansten PD, Panic J, Horn JR, Boyce RD, Gephart S, et al. Risk of bleeding with exposure to warfarin and nonsteroidal anti-inflammatory drugs: a systematic review and meta-analysis. *Thromb Haemost* 2020 Jul;120(7):1066-1074 [FREE Full text] [doi: [10.1055/s-0040-1710592](https://doi.org/10.1055/s-0040-1710592)] [Medline: [32455439](https://pubmed.ncbi.nlm.nih.gov/32455439/)]
6. Maggioni AP, Dondi L, Andreotti F, Pedrini A, Calabria S, Ronconi G, et al. Four-year trends in oral anticoagulant use and declining rates of ischemic stroke among 194,030 atrial fibrillation patients drawn from a sample of 12 million people. *Am Heart J* 2020 Feb;220:12-19. [doi: [10.1016/j.ahj.2019.10.017](https://doi.org/10.1016/j.ahj.2019.10.017)] [Medline: [31759279](https://pubmed.ncbi.nlm.nih.gov/31759279/)]
7. Shah RU, Rupp AB, Mowery D, Zhang M, Stoddard G, Deshmukh V, et al. Changes in oral anticoagulant treatment rates in atrial fibrillation before and after the introduction of direct oral anticoagulants. *Neuroepidemiology* 2016 Jan 31;47(3-4):201-209. [doi: [10.1159/000455012](https://doi.org/10.1159/000455012)] [Medline: [28135707](https://pubmed.ncbi.nlm.nih.gov/28135707/)]
8. Kohsaka S, Katada J, Saito K, Jenkins A, Li B, Mardekian J, et al. Safety and effectiveness of non-vitamin K oral anticoagulants versus warfarin in real-world patients with non-valvular atrial fibrillation: a retrospective analysis of contemporary Japanese administrative claims data. *Open Heart* 2020 Apr 01;7(1):e001232 [FREE Full text] [doi: [10.1136/openhrt-2019-001232](https://doi.org/10.1136/openhrt-2019-001232)] [Medline: [32341789](https://pubmed.ncbi.nlm.nih.gov/32341789/)]
9. Mullan J, Weston KM, Bonney A, Burns P, Mullan J, Rudd R. Consumer knowledge about over-the-counter NSAIDs: they don't know what they don't know. *Aust N Z J Public Health* 2017 Apr;41(2):210-214. [doi: [10.1111/1753-6405.12589](https://doi.org/10.1111/1753-6405.12589)] [Medline: [27774735](https://pubmed.ncbi.nlm.nih.gov/27774735/)]
10. Cryer B, Barnett MA, Wagner J, Wilcox CM. Overuse and misperceptions of nonsteroidal anti-inflammatory drugs in the United States. *Am J Med Sci* 2016 Nov;352(5):472-480. [doi: [10.1016/j.amjms.2016.08.028](https://doi.org/10.1016/j.amjms.2016.08.028)] [Medline: [27865294](https://pubmed.ncbi.nlm.nih.gov/27865294/)]
11. Elwyn G, Frosch D, Thomson R, Joseph-Williams N, Lloyd A, Kinnersley P, et al. Shared decision making: a model for clinical practice. *J Gen Intern Med* 2012 Oct;27(10):1361-1367 [FREE Full text] [doi: [10.1007/s11606-012-2077-6](https://doi.org/10.1007/s11606-012-2077-6)] [Medline: [22618581](https://pubmed.ncbi.nlm.nih.gov/22618581/)]
12. Shay LA, Lafata JE. Where is the evidence? A systematic review of shared decision making and patient outcomes. *Med Decis Making* 2014 Oct 28;35(1):114-131. [doi: [10.1177/0272989x14551638](https://doi.org/10.1177/0272989x14551638)]
13. Fiorillo A, Barlati S, Bellomo A, Corrivetti G, Nicolò G, Sampogna G, et al. The role of shared decision-making in improving adherence to pharmacological treatments in patients with schizophrenia: a clinical review. *Ann Gen Psychiatry* 2020 Aug 05;19(1):43 [FREE Full text] [doi: [10.1186/s12991-020-00293-4](https://doi.org/10.1186/s12991-020-00293-4)] [Medline: [32774442](https://pubmed.ncbi.nlm.nih.gov/32774442/)]
14. Loewen P. SPARC - stroke prevention in atrial fibrillation risk tool. Vancouver, BC; 2019 Jun 01. URL: <https://www.sparctool.com/> [accessed 2021-11-16]
15. O'Neill ES, Grande SW, Sherman A, Elwyn G, Coylewright M. Availability of patient decision aids for stroke prevention in atrial fibrillation: A systematic review. *Am Heart J* 2017 Sep;191:1-11. [doi: [10.1016/j.ahj.2017.05.014](https://doi.org/10.1016/j.ahj.2017.05.014)] [Medline: [28888264](https://pubmed.ncbi.nlm.nih.gov/28888264/)]
16. Zeballos-Palacios CL, Hargraves IG, Noseworthy PA, Branda ME, Kunneman M, Burnett B, Shared Decision Making for Atrial Fibrillation (SDM4AFib) Trial Investigators. Developing a conversation aid to support shared decision making: reflections on designing anticoagulation choice. *Mayo Clin Proc* 2019 Apr;94(4):686-696 [FREE Full text] [doi: [10.1016/j.mayocp.2018.08.030](https://doi.org/10.1016/j.mayocp.2018.08.030)] [Medline: [30642640](https://pubmed.ncbi.nlm.nih.gov/30642640/)]
17. Anticoagulation Toolkit. 2019 Dec 03. URL: <http://www.anticoagulationtoolkit.org/> [accessed 2021-11-20]
18. Hong C, Kim S, Curnew G, Schulman S, Pullenayegum E, Holbrook A. Validation of a patient decision aid for choosing between dabigatran and warfarin for atrial fibrillation. *J Popul Ther Clin Pharmacol* 2013;20(3):e229-e237. [Medline: [24163147](https://pubmed.ncbi.nlm.nih.gov/24163147/)]
19. Noseworthy P, Brito J, Kunneman M, Hargraves I, Zeballos-Palacios C, Montori VM, et al. Shared decision-making in atrial fibrillation: navigating complex issues in partnership with the patient. *J Interv Card Electrophysiol* 2019 Nov;56(2):159-163 [FREE Full text] [doi: [10.1007/s10840-018-0465-5](https://doi.org/10.1007/s10840-018-0465-5)] [Medline: [30327992](https://pubmed.ncbi.nlm.nih.gov/30327992/)]
20. Chou E, Boyce R, Balkan B, Subbian V, Romero A, Hansten P, et al. Designing and evaluating contextualized drug-drug interaction algorithms. *JAMIA Open* 2021 Jan;4(1):oob023 [FREE Full text] [doi: [10.1093/jamiaopen/oob023](https://doi.org/10.1093/jamiaopen/oob023)] [Medline: [33763631](https://pubmed.ncbi.nlm.nih.gov/33763631/)]
21. Medscape Drug Interaction Checker. URL: <https://reference.medscape.com/drug-interactionchecker> [accessed 2020-12-08]
22. Lexicomp drug interactions. Wolters Kluwer. URL: <https://www.wolterskluwer.com/en/solutions/lexicomp/resources/lexicomp-user-academy/drug-interactions-analysis> [accessed 2020-12-08]
23. RxList Drug Interaction Checker. URL: <https://www.rxlist.com/drug-interaction-checker.htm> [accessed 2020-12-09]
24. Drug interactions checker. Drugs.com. URL: [https://www.drugs.com/drug\\_interactions.html](https://www.drugs.com/drug_interactions.html) [accessed 2020-12-08]

25. Reese TJ, Del Fiol G, Morgan K, Hurwitz JT, Kawamoto K, Gomez-Lumbreras A, et al. A Shared Decision-making Tool for Drug Interactions Between Warfarin and Nonsteroidal Anti-inflammatory Drugs: Design and Usability Study. *JMIR Hum Factors* 2021 Oct 26;8(4):e28618 [FREE Full text] [doi: [10.2196/28618](https://doi.org/10.2196/28618)] [Medline: [34698649](https://pubmed.ncbi.nlm.nih.gov/34698649/)]
26. Sepucha KR, Abhyankar P, Hoffman AS, Bekker HL, LeBlanc A, Levin CA, et al. Standards for universal reporting of patient decision aid evaluation studies: the development of SUNDAE checklist. *BMJ Qual Saf* 2018 May;27(5):380-388 [FREE Full text] [doi: [10.1136/bmjqs-2017-006986](https://doi.org/10.1136/bmjqs-2017-006986)] [Medline: [29269567](https://pubmed.ncbi.nlm.nih.gov/29269567/)]
27. Stetler CB, Legro MW, Wallace CM, Bowman C, Guihan M, Hagedorn H, et al. The role of formative evaluation in implementation research and the QUERI experience. *J Gen Intern Med* 2006 Feb;21(S2):S1-S8. [doi: [10.1111/j.1525-1497.2006.00355.x](https://doi.org/10.1111/j.1525-1497.2006.00355.x)] [Medline: [16637954](https://pubmed.ncbi.nlm.nih.gov/16637954/)]
28. Drug-Drug Interaction Clinical Decision Support. URL: <https://ddi-cds.org/apps/> [accessed 2020-07-07]
29. Mandel JC, Kreda DA, Mandl KD, Kohane IS, Ramoni RB. SMART on FHIR: a standards-based, interoperable apps platform for electronic health records. *J Am Med Inform Assoc* 2016 Feb 17 [FREE Full text] [doi: [10.1093/jamia/ocv189](https://doi.org/10.1093/jamia/ocv189)] [Medline: [26911829](https://pubmed.ncbi.nlm.nih.gov/26911829/)]
30. Thiess H, Del Fiol G, Malone DC, Cornia R, Sibilla M, Rhodes B, et al. Coordinated use of Health Level 7 standards to support clinical decision support: Case study with shared decision making and drug-drug interactions. *Int J Med Inform* 2022 Mar 21;162:104749. [doi: [10.1016/j.ijmedinf.2022.104749](https://doi.org/10.1016/j.ijmedinf.2022.104749)] [Medline: [35358893](https://pubmed.ncbi.nlm.nih.gov/35358893/)]
31. Scholl I, Kriston L, Dirmaier J, Buchholz A, Härter M. Development and psychometric properties of the Shared Decision Making Questionnaire--physician version (SDM-Q-Doc). *Patient Educ Couns* 2012 Aug;88(2):284-290. [doi: [10.1016/j.pec.2012.03.005](https://doi.org/10.1016/j.pec.2012.03.005)] [Medline: [22480628](https://pubmed.ncbi.nlm.nih.gov/22480628/)]
32. Brooke J. SUS: A quickdirty usability scale. In: *Usability Evaluation in Industry*. Oxfordshire, UK: CRC Press; 1996:189-194.
33. Hart SG, Staveland LE. Development of NASA-TLX (Task Load Index): results of empirical and theoretical research. In: Hancock PA, Meshkati N, editors. *Human Mental Workload*. Amsterdam, Netherlands: North-Holland; 1988:139-183.
34. Venkatesh V, Morris MG, Davis GB, Davis FD. 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)]
35. 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)]
36. O'Connor AM. Validation of a decisional conflict scale. *Med Decis Making* 1995;15(1):25-30. [doi: [10.1177/0272989X9501500105](https://doi.org/10.1177/0272989X9501500105)] [Medline: [7898294](https://pubmed.ncbi.nlm.nih.gov/7898294/)]
37. Doherr H, Christalle E, Kriston L, Härter M, Scholl I. Use of the 9-item Shared Decision Making Questionnaire (SDM-Q-9 and SDM-Q-Doc) in intervention studies-A systematic review. *PLoS One* 2017 Mar 30;12(3):e0173904 [FREE Full text] [doi: [10.1371/journal.pone.0173904](https://doi.org/10.1371/journal.pone.0173904)] [Medline: [28358864](https://pubmed.ncbi.nlm.nih.gov/28358864/)]
38. Mutebi A, Warholak TL, Hines LE, Plummer R, Malone DC. Assessing patients' information needs regarding drug-drug interactions. *J Am Pharm Assoc* 2013;53(1):39-45. [doi: [10.1331/JAPhA.2013.12038](https://doi.org/10.1331/JAPhA.2013.12038)] [Medline: [23636154](https://pubmed.ncbi.nlm.nih.gov/23636154/)]
39. Justad H, Askfors Y, Shemeikka T, Andersson ML, Hammar T. Patients' use and perceptions of a drug-drug interaction database: a survey of Janusmed interactions. *Pharmacy (Basel)* 2021 Jan 19;9(1):23 [FREE Full text] [doi: [10.3390/pharmacy9010023](https://doi.org/10.3390/pharmacy9010023)] [Medline: [33478093](https://pubmed.ncbi.nlm.nih.gov/33478093/)]
40. Kusch MK, Haefeli WE, Seidling HM. How to meet patients' individual needs for drug information - a scoping review. *Patient Prefer Adherence* 2018;12:2339-2355 [FREE Full text] [doi: [10.2147/PPA.S173651](https://doi.org/10.2147/PPA.S173651)] [Medline: [30464421](https://pubmed.ncbi.nlm.nih.gov/30464421/)]
41. Richardson KM, Fouquet SD, Kerns E, McCulloh RJ. Impact of mobile device-based clinical decision support tool on guideline adherence and mental workload. *Acad Pediatr* 2019 Sep;19(7):828-834 [FREE Full text] [doi: [10.1016/j.acap.2019.03.001](https://doi.org/10.1016/j.acap.2019.03.001)] [Medline: [30853573](https://pubmed.ncbi.nlm.nih.gov/30853573/)]
42. Liu S, Reese T, Kawamoto K, Del Fiol G, Weir C. A theory-based meta-regression of factors influencing clinical decision support adoption and implementation. *J Am Med Inform Assoc* 2021 Oct 12;28(11):2514-2522 [FREE Full text] [doi: [10.1093/jamia/ocab160](https://doi.org/10.1093/jamia/ocab160)] [Medline: [34387686](https://pubmed.ncbi.nlm.nih.gov/34387686/)]
43. English D, Ankem K, English K. Acceptance of clinical decision support surveillance technology in the clinical pharmacy. *Inform Health Soc Care* 2017 Mar 18;42(2):135-152. [doi: [10.3109/17538157.2015.1113415](https://doi.org/10.3109/17538157.2015.1113415)] [Medline: [26890621](https://pubmed.ncbi.nlm.nih.gov/26890621/)]
44. Roebroek LO, Bruins J, Delespaul P, Boonstra A, Castelein S. Qualitative analysis of clinicians' perspectives on the use of a computerized decision aid in the treatment of psychotic disorders. *BMC Med Inform Decis Mak* 2020 Sep 17;20(1):234 [FREE Full text] [doi: [10.1186/s12911-020-01251-6](https://doi.org/10.1186/s12911-020-01251-6)] [Medline: [32943027](https://pubmed.ncbi.nlm.nih.gov/32943027/)]
45. Dabliz R, Poon SK, Ritchie A, Burke R, Penm J. Usability evaluation of an integrated electronic medication management system implemented in an oncology setting using the unified theory of the acceptance and use of technology. *BMC Med Inform Decis Mak* 2021 Jan 06;21(1):4 [FREE Full text] [doi: [10.1186/s12911-020-01348-y](https://doi.org/10.1186/s12911-020-01348-y)] [Medline: [33407411](https://pubmed.ncbi.nlm.nih.gov/33407411/)]
46. Ma Y, Zhong X, Lin B, He W. Factors influencing the intention of MSM to use the PrEP intelligent reminder system. *RMHP* 2021 Nov; Volume 14:4739-4748. [doi: [10.2147/rmhp.s337287](https://doi.org/10.2147/rmhp.s337287)]
47. Loewen PS, Bansback N, Hicklin J, Andrade JG, Kapanen AI, Kwan L, et al. Evaluating the effect of a patient decision aid for atrial fibrillation stroke prevention therapy. *Ann Pharmacother* 2019 Jul;53(7):665-674. [doi: [10.1177/1060028019828420](https://doi.org/10.1177/1060028019828420)] [Medline: [30724102](https://pubmed.ncbi.nlm.nih.gov/30724102/)]
48. Kunneman M, Branda ME, Hargraves IG, Sivly AL, Lee AT, Gorr H, Shared Decision Making for Atrial Fibrillation (SDM4AFib) Trial Investigators. Assessment of shared decision-making for stroke prevention in patients with atrial

- fibrillation: a randomized clinical trial. *JAMA Intern Med* 2020 Sep 01;180(9):1215-1224 [[FREE Full text](#)] [doi: [10.1001/jamainternmed.2020.2908](https://doi.org/10.1001/jamainternmed.2020.2908)] [Medline: [32897386](#)]
49. Shariff A, Belagodu Sridhar S, Abdullah Basha NF, Bin Taleth Alshemeil SSH, Ahmed Aljallaf Alzaabi NA. Assessing consistency of drug-drug interaction-related information across various drug information resources. *Cureus* 2021 Mar 08;13(3):e13766 [[FREE Full text](#)] [doi: [10.7759/cureus.13766](https://doi.org/10.7759/cureus.13766)] [Medline: [33842142](#)]
  50. Knijff-Dutmer EA, Schut GA, van de Laar MA. Concomitant coumarin–NSAID therapy and risk for bleeding. *Ann Pharmacother* 2016 Jun 28;37(1):12-16. [doi: [10.1345/aph.1c157](https://doi.org/10.1345/aph.1c157)]
  51. Holbrook AM, Pereira JA, Labiris R, McDonald H, Douketis JD, Crowther M, et al. Systematic overview of warfarin and its drug and food interactions. *Arch Intern Med* 2005 May 23;165(10):1095-1106. [doi: [10.1001/archinte.165.10.1095](https://doi.org/10.1001/archinte.165.10.1095)] [Medline: [15911722](#)]
  52. Shen C, Jiang B, Yang Q, Wang C, Lu KZ, Gu M, et al. Mobile apps for drug-drug interaction checks in Chinese app stores: systematic review and content analysis. *JMIR Mhealth Uhealth* 2021 Jun 15;9(6):e26262 [[FREE Full text](#)] [doi: [10.2196/26262](https://doi.org/10.2196/26262)] [Medline: [33962910](#)]
  53. Sathian B, Asim M, Banerjee I, Pizarro AB, Roy B, van Teijlingen ER, et al. Impact of COVID-19 on clinical trials and clinical research: A systematic review. *Nepal J Epidemiol* 2020 Sep 30;10(3):878-887 [[FREE Full text](#)] [doi: [10.3126/nje.v10i3.31622](https://doi.org/10.3126/nje.v10i3.31622)] [Medline: [33042591](#)]
  54. Unger JM, Xiao H, LeBlanc M, Hershman DL, Blanke CD. Cancer clinical trial participation at the 1-year anniversary of the outbreak of the COVID-19 pandemic. *JAMA Netw Open* 2021 Jul 01;4(7):e2118433 [[FREE Full text](#)] [doi: [10.1001/jamanetworkopen.2021.18433](https://doi.org/10.1001/jamanetworkopen.2021.18433)] [Medline: [34323986](#)]

## Abbreviations

- DCS:** Decisional Conflict Scale  
**DDI:** drug-drug interaction  
**EHR:** electronic health record  
**FHIR:** Fast Health Interoperability Resources  
**GI:** gastrointestinal  
**INR:** international normalized ratio  
**NASA-TLX:** National Aeronautics and Space Administration Task Load Index  
**NSAID:** nonsteroidal anti-inflammatory drug  
**NVAF:** nonvalvular atrial fibrillation  
**OTC:** over the counter  
**PBC:** Perceived Behavioral Control  
**SDM:** shared decision-making  
**SDM-Q-9:** 9-item Shared Decision-Making Questionnaire  
**SMART:** Substitutable Medical Applications and Reusable Technologies  
**SUNDAE:** Standards for Universal Reporting of Decision Aid Evaluations  
**SUS:** System Usability Scale  
**UC:** usual care  
**UTAUT:** Unified Theory of Acceptance and Use of Technology

*Edited by A Mavragani; submitted 01.06.22; peer-reviewed by J Walsh, P Dunn, T Zhang; comments to author 15.08.22; revised version received 21.08.22; accepted 22.09.22; published 19.10.22.*

*Please cite as:*

Gomez Lumbreras A, Reese TJ, Del Fiol G, Tan MS, Butler JM, Hurwitz JT, Brown M, Kawamoto K, Thiess H, Wright M, Malone DC

*Shared Decision-Making for Drug-Drug Interactions: Formative Evaluation of an Anticoagulant Drug Interaction*

*JMIR Form Res* 2022;6(10):e40018

URL: <https://formative.jmir.org/2022/10/e40018>

doi: [10.2196/40018](https://doi.org/10.2196/40018)

PMID: [36260377](https://pubmed.ncbi.nlm.nih.gov/36260377/)

©Ainhoa Gomez Lumbreras, Thomas J Reese, Guilherme Del Fiol, Malinda S Tan, Jorie M Butler, Jason T Hurwitz, Mary Brown, Kensaku Kawamoto, Henrik Thiess, Maria Wright, Daniel C Malone. Originally published in *JMIR Formative Research* (<https://formative.jmir.org>), 19.10.2022. 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 Cognitive Behavioral Therapy Chatbot (Otis) for Health Anxiety Management: Mixed Methods Pilot Study

Yenushka Goonesekera<sup>1</sup>, BSc, MHealthPsych; Liesje Donkin<sup>1,2</sup>, MSc, PhD

<sup>1</sup>Department of Psychological Medicine, The University of Auckland, Auckland, New Zealand

<sup>2</sup>Department of Psychology and Neuroscience, Auckland University of Technology, Auckland, New Zealand

**Corresponding Author:**

Liesje Donkin, MSc, PhD

Department of Psychology and Neuroscience

Auckland University of Technology

55 Wellesley Street East

Private Bag 92006

Auckland, 1010

New Zealand

Phone: 64 21847886

Email: [liesje.donkin@aut.ac.nz](mailto:liesje.donkin@aut.ac.nz)

## Abstract

**Background:** An increase in health anxiety was observed during the COVID-19 pandemic. However, due to physical distancing restrictions and a strained mental health system, people were unable to access support to manage health anxiety. Chatbots are emerging as an interactive means to deliver psychological interventions in a scalable manner and provide an opportunity for novel therapy delivery to large groups of people including those who might struggle to access traditional therapies.

**Objective:** The aim of this mixed methods pilot study was to investigate the feasibility, acceptability, engagement, and effectiveness of a cognitive behavioral therapy (CBT)-based chatbot (Otis) as an early health anxiety management intervention for adults in New Zealand during the COVID-19 pandemic.

**Methods:** Users were asked to complete a 14-day program run by Otis, a primarily decision tree-based chatbot on Facebook Messenger. Health anxiety, general anxiety, intolerance of uncertainty, personal well-being, and quality of life were measured pre-intervention, postintervention, and at a 12-week follow-up. Paired samples *t* tests and 1-way ANOVAs were conducted to investigate the associated changes in the outcomes over time. Semistructured interviews and written responses in the self-report questionnaires and Facebook Messenger were thematically analyzed.

**Results:** The trial was completed by 29 participants who provided outcome measures at both postintervention and follow-up. Although an average decrease in health anxiety did not reach significance at postintervention ( $P=.55$ ) or follow-up ( $P=.08$ ), qualitative analysis demonstrated that participants perceived benefiting from the intervention. Significant improvement in general anxiety, personal well-being, and quality of life was associated with the use of Otis at postintervention and follow-up. Anthropomorphism, Otis' appearance, and delivery of content facilitated the use of Otis. Technical difficulties and high performance and effort expectancy were, in contrast, barriers to acceptance and engagement of Otis.

**Conclusions:** Otis may be a feasible, acceptable, and engaging means of delivering CBT to improve anxiety management, quality of life, and personal well-being but might not significantly reduce health anxiety.

(*JMIR Form Res* 2022;6(10):e37877) doi:[10.2196/37877](https://doi.org/10.2196/37877)

## KEYWORDS

health anxiety; conversational agent; illness anxiety disorder; COVID-19; iCBT; user experience; anthropomorphism

## Introduction

Given the importance of physical health to survival, it is natural and adaptive for people to be vigilant and fearful about their health, engage in health-promoting behaviors, and scan for

health threats [1,2]. However, when this fear is associated with functional impairment and distress, it is termed health anxiety [3]. Defined as the persistent worry about illness based on the misinterpretation of bodily symptoms, health anxiety is a criterion of illness anxiety disorder. The tendency of



health-anxious individuals to misinterpret typically benign symptoms as a sign of infection [4,5] such as those of COVID-19 (eg, fever, cough, fatigue) likely explains the reported increase in health anxiety during the COVID-19 pandemic [6].

Health anxiety drives the avoidance of people and environments that an individual perceives as threatening, leading to reduced help-seeking. For example, a previous study reported that, compared with controls, individuals with severe health anxiety were more likely to rate photos of healthy people as being less healthy, demonstrating a bias toward evaluating others as a health threat [7]. Alternatively, or in combination with avoidance behaviors, health-anxious individuals may engage in reassurance-seeking to alleviate anxiety. Both avoidance and reassurance-seeking maintain the anxiety as the individual foregoes their opportunity to become aware of their ability to cope with their perceived fear [8,9], thus preventing long-term cognitive change [10].

In addition to the physical health burden of pandemics, psychological well-being is often vulnerable. Health anxiety in particular is associated with greater levels of distress, thus warranting a greater need for psychological support [11-13]. However, during the COVID-19 pandemic, physical distancing measures used to mitigate the virus limited access to psychological care. Therefore, health care services saw service delivery transition from in-person to telehealth. Despite the shift in service delivery, additional therapy tools such as chatbots were needed to meet the increased demand for support.

Chatbots are conversational agents that hold text, speech, or visual-based conversations with users [14]. The interactions of chatbots are determined by 2 main mechanisms: artificial intelligence (AI) and decision trees. AI-based chatbots use a complex mathematical algorithm to produce specific predefined outputs based on the information input by users. Decision tree-based chatbots instead follow a prewritten script that users interact with by choosing prewritten responses. The mimicking of human conversation within the technology lends well to the use of chatbots in a mental health context. Previous studies have demonstrated that chatbot interventions improve psychological difficulties including major depressive disorder [15], panic disorder [16], posttraumatic stress disorder [17], antipsychotic medication adherence [18], and perceived stress [19].

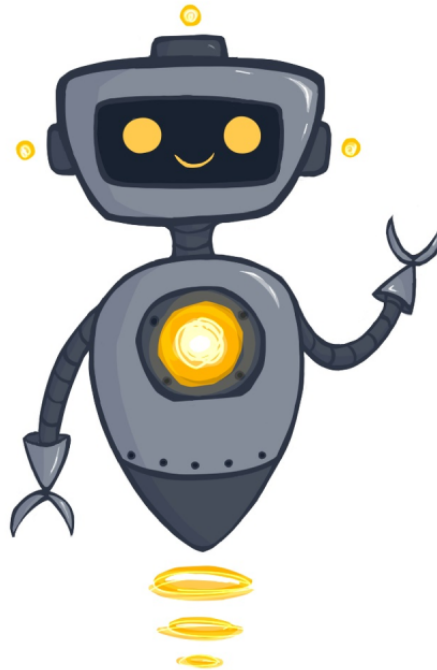
Cognitive behavioral therapy (CBT) has the greatest evidence base for health anxiety management compared with several control conditions, including treatment as usual, medication, placebo, waitlist, support groups, non-CBT-based psychoeducation, and other psychological interventions [20,21]. CBT has been adapted for internet-based use (iCBT) in treating health anxiety and has shown similar effects to face-to-face CBT [22]. iCBT has been delivered using several digital health intervention (DHI) media such as computer programs, websites, emails, videos, mobile applications, and, more recently, chatbots. Although iCBT has been adapted to suit several digital modes of delivery to treat health anxiety, to the best of our knowledge, there has not yet been a chatbot designed to help people manage health anxiety using CBT.

CBT for health anxiety management delivered through chatbot technology may be a suitable DHI, especially in a pandemic context. A chatbot may provide an initial management intervention that reduces the burden of help-seeking individuals who may otherwise overutilize health services. Moreover, it may encourage interaction for avoidant individuals who would otherwise be reluctant to seek care, due to the presence of clinicians and anxiety-provoking settings (eg, hospitals and clinics). Although chatbots have demonstrated improvements in general well-being, few have been used to address specific clinically significant mental health disorders such as health anxiety.

## Methods

### Recruitment

This mixed methods pilot study was conducted using a single group, pretest-posttest intervention framework and semistructured interviews to assess the feasibility, acceptability, engagement, and effectiveness of an automated conversational agent, “Otis” (Figure 1), as a brief health anxiety management program. Participants were recruited via Facebook and Instagram advertising, Twitter promotions by the authors’ and faculty accounts, and email invitations between May 2020 and July 2020. Participants aged 18 years and older who had access to a Facebook Messenger account and internet access that allowed for daily use of Messenger during the 14-day program were included in the study. Participants were excluded if they were not living in New Zealand or unable to understand written English to consent to study participation and complete the program.

**Figure 1.** Avatar of Otis.

### Ethical Considerations

The study was approved by The University of Auckland Human Participants Ethics Committee (UoA Reference: 024655). Once participants clicked on the advertising, a conversation with Otis on Facebook Messenger was opened to invite participants to read a participant information sheet. The participant information sheet outlined the inclusion criteria, study timeline, and author contact details and provided a link to an online consent form.

### Procedures

On completion of the baseline questionnaire, delivered using the web-based survey software Qualtrics [23], participants were provided a randomized user identification number to access Otis on Facebook Messenger. At the conclusion of the 14-day program or abandonment, whichever came first, participants were contacted to complete a postintervention questionnaire. Those who completed the postintervention questionnaire were invited to complete an interview with one of the researchers about their user experience and then contacted again in 12 weeks to complete a follow-up questionnaire. Participants were offered an entry in a prize draw to win one of 5 NZ \$50 (US \$28.36)

retail vouchers per completed questionnaire. Additional entries were offered to those who completed the interview.

### Intervention

Otis was primarily a decision tree-based conversational agent or chatbot that was designed to deliver daily modules of CBT in the form of a 14-day program. Participants chatted with Otis for 5 minutes to 10 minutes per day to learn to manage health anxiety. The chatbot was developed using Chatfuel, a widely used commercial chatbot engine for Facebook Messenger, and accessed through the Facebook Messenger app or the desktop site. Previous effective CBT interventions for health anxiety have been approximately 12 modules [24,25]. Therefore, 12 modules focused on content were also chosen for Otis, with 1 module available to participants per day.

We included 2 additional modules at the start and end of the program, making the program 14 days long in total. The first module focused on introducing the intervention and collecting participant data, and the content was summarized in the final module. The content of the chatbot was derived from previous iCBT interventions and the authors' clinical knowledge. Table 1 provides an overview of the core components that were covered in the 12 modules.

**Table 1.** Core components of the chatbot intervention.

Component	Description
Psychoeducation	Health anxious thoughts and behaviors, importance of managing health anxiety
CBT <sup>a</sup> five-part model	Introduction to different components of the model and the relationships between the components
Cognitive restructuring and thinking errors	Aid identification and naming of unhelpful thinking patterns
Anxiety reduction	Mindfulness, relaxation, delaying worry exercises (eg, worry time)
Exposure	Discuss the importance of having an opportunity to endure anxiety, complete exposure ladder

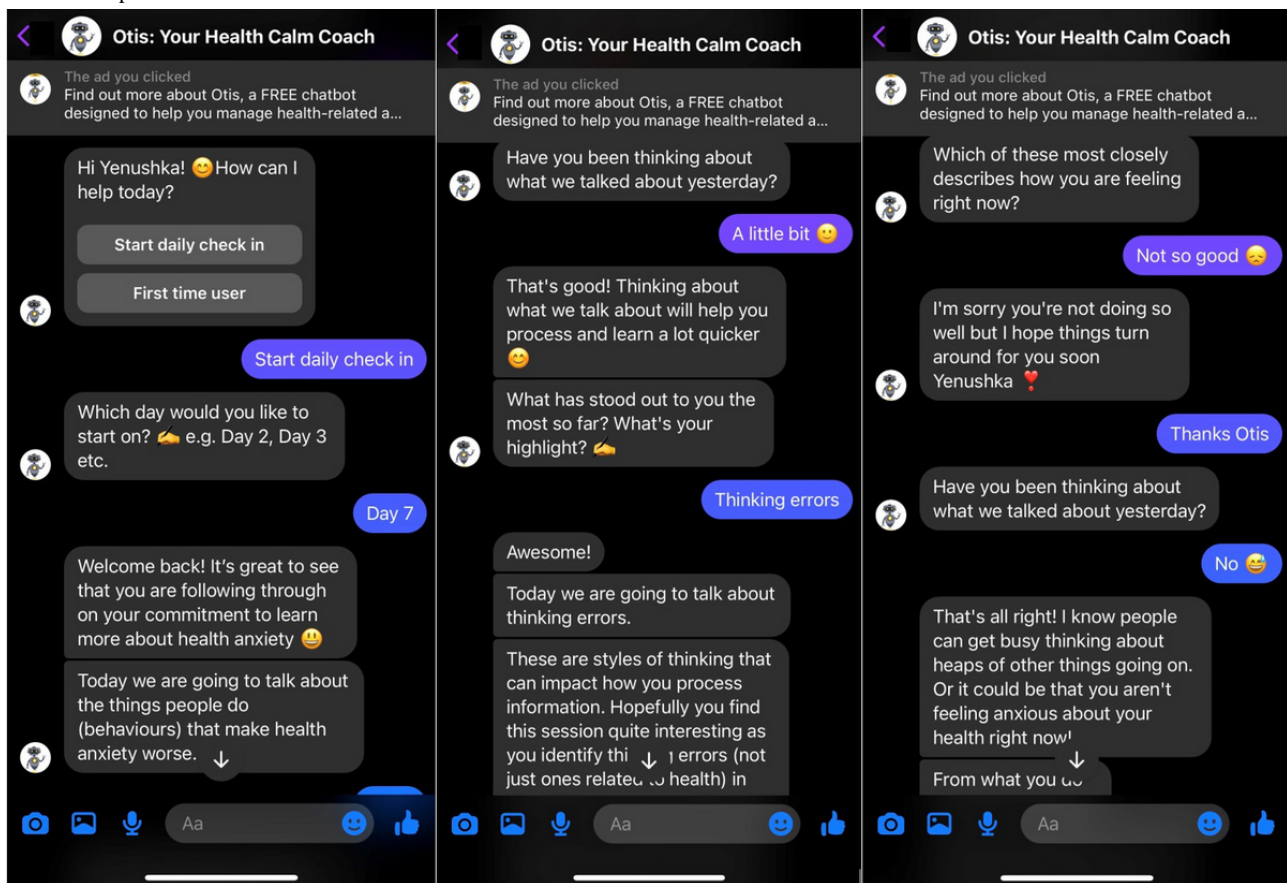
<sup>a</sup>CBT: cognitive behavioral therapy

During the initial development phase of a chatbot, it can produce crude responses that users may perceive as invalidating, robotic, or disconnected from their response [26]. Therefore, a primarily decision tree–based chatbot was chosen as the most suitable mechanism given the possibly sensitive nature of discussing mental health online. Additionally, using a decision tree–based chatbot ensured that participants received a consistent intervention. The conversations participants have within a decision tree–based chatbot are based on a prewritten script so that users follow a set conversation pathway rather than receiving a machine-learned response. Participants interacted with the chatbot by choosing predetermined text and emojis responses called “quick replies” (Figure 2). Quick replies ensured that the conversation did not deviate from the day’s content, as different replies often prompted the same response from Otis. To simulate empathy, some quick replies prompted specific validation that diverged from the pathway to form the “branches” of the tree before eventually converging back to the main script. When a message from Otis was followed by a Writing Hand emoji, this indicated to the user that they could type their own message rather than use a quick reply. Some

basic AI rules were built in to reciprocate user greetings (eg, “hello,” “how are you?” “goodbye”), change notification preferences, and direct participants to crisis helplines if they asked for help or used risk words indicating suicidal ideation.

Participants were allocated 21 days to complete the 14-day program to maximize engagement and to account for fluctuating motivation and availability. On the first day, participants were able to choose a check-in time so that the chatbot triggered a notification reminder at the same time each day. If a notification had not been responded to or a participant had not completed the day’s module in full within 3 days, an additional notification was sent. A similar notification was sent if a participant had not actively used Otis for 7 days. The chatbot sent a link to the postintervention questionnaire when a participant had completed the program or 21 days following the completion of the baseline questionnaire but not the program (abandonment), whichever came first. When the program was abandoned or completed, users could no longer access the chatbot; however, the crisis line feature was available for safety. Otis was not made available to new users at the end of the study period.

Figure 2. Examples of conversations with Otis.



## Measures

### Demographics

At baseline, participant demographics (age, gender, ethnicity, education level, employment status, and household income), technology use, mental health service utilization, and medication use were recorded.

### Outcome Measures

#### Short Health Anxiety Inventory

The Short Health Anxiety Inventory (SHAI-18) [27] is an 18-item self-report questionnaire that assesses the severity of health anxiety over the past 6 months. It has demonstrated comparable validity and reliability (Cronbach  $\alpha=.86$ ) to its initial 64-item counterpart, the Health Anxiety Inventory [2]. Each item of the measure consists of 4 statements and requires

participants to select a statement that best describes their feelings over the previous 6 months. Each statement is scored between 0 and 3 and summed, with higher scores indicating more severe health anxiety in clinical, nonclinical, and medical samples [28]. The cut-off scores of the SHAI-18 vary across studies [29-31]; however, a previous study concluded that a score of 27 and greater differentiated health anxiety from other anxiety disorders [2]. Therefore, a cut-off score of 27 was chosen for the present study.

### Generalized Anxiety Disorder-7

The Generalized Anxiety Disorder 7-item scale (GAD-7) is a brief self-report measure that assesses the severity of anxious thoughts and behaviors over the past 2 weeks. The measure is based on the Diagnostic Manual of Mental Disorders (DSM)-IV criteria for generalized anxiety disorder, in which the severity of anxiety is calculated by summing the participant scores on each item. Scores of 5, 10, and 15 indicate mild, moderate, and severe anxiety, respectively. The GAD-7 has been validated in adult populations, with a Cronbach  $\alpha$  of .92 and high convergent and discriminant validity [32].

### Short Intolerance of Uncertainty Scale 7

As health anxiety is correlated with intolerance of uncertainty [2,33,34], the Short Intolerance of Uncertainty Scale (IUS-12) [35], a 12-item self-report measure that assesses an individual's response to ambiguous situations, uncertainty, and future events, was administered. For each item, participants are asked to rate the degree to which each item is characteristic of them. The ratings are weighted from 1 to 5 and summed, with higher scores indicating greater intolerance of uncertainty. The measure has demonstrated excellent internal consistency (Cronbach  $\alpha$ =.94), good test-retest reliability, and convergent and divergent validity [36].

### Personal Well-being Domain

The Office of National Statistics Personal Well-being Domain (ONS4) is a brief, 4-item self-report assessment that measures potential changes in well-being. The ONS4 is widely used among various patient populations [36-40] and has demonstrated good internal reliability (Cronbach  $\alpha$ =.90) and convergent validity [41].

**Textbox 1.** Acceptability measure items used in the pilot study of Otis.

1. It was helpful.
2. It was easy to use.
3. It was fun.
4. Otis was nice to look at.
5. It worked smoothly (eg, without major technical glitches).
6. I would like to keep it on my device.

### Interviews

Upon completing the postintervention questionnaire, participants were invited to complete an interview with one of the researchers about their user experience. The semistructured interviews collected qualitative data about the feasibility,

### World Health Organization-Five Well-being Index

The World Health Organization-Five Well-being Index (WHO-5) is a 5-item self-report tool that assesses an individual's perceived positive mood, vitality, general interest, and quality of life. The raw scores range from 0 to 25 and are calculated by summing the total scores, with 0 representing the worst and 25 representing the best possible quality of life. A vast set of literature supports the validity and reliability of the measure in several populations [42-45].

### Adherence

Facebook's basic participant usage data were also recorded, including the number of users and rates of completion. The completion rates of the 14-day program were used as a measure of adherence (<4 days = poor adherence; <7 days = low adherence; <10 days = moderate adherence; <14 days = high adherence).

### Reasons for Abandonment

At postintervention, participants were asked to indicate their reasons for abandoning the program if they did not complete all 14 days. From choices of lack of time, technical difficulties, boredom, no longer requiring the chatbot, and being unable to apply the skills they learned through the chatbot, participants were asked to choose the option that best described their reason for abandonment.

### Acceptability Rating

Mixed format questions were used to assess enjoyment and satisfaction with the chatbot. The first measure used consisted of 6 statements relating to the chatbot user experience developed by the researcher (Textbox 1). Participants were required to indicate the extent to which each statement was true in their experience with using the chatbot. The scale, ranging from "not at all" to "definitely," was weighted from 0 to 4. The ratings across items were summed, with higher scores indicating greater satisfaction and acceptability of Otis. For the second measure, participants were asked to provide an overall rating of the chatbot on a scale from 1 to 10. Open-ended questions about how Otis could be improved, what participants found most useful about Otis, and what they liked the most and least about the chatbot were included.

acceptability, and engagement of Otis. The interviewer also asked participants about the content and content delivery of the chatbot (GIFs, images, videos, text). Participants were interviewed primarily over the video calling platform Zoom [46]; however, some participants completed the interview over the phone due to connectivity difficulties in rural areas.

## Statistical Analysis

The quantitative data were analyzed using SPSS Version 26 (IBM Corp, Armonk, NY) [47]. Statistical significance was defined as  $P \leq .05$ . Relationships between demographic characteristics and outcome measures (health anxiety, general anxiety, intolerance of uncertainty, perceived well-being, and quality of life) at a single time point were explored using  $t$  tests or chi-square tests. Adherence was tested by using the average completion rate and the number of participants who persisted through the program. A dose-response relationship between adherence (categorized into poor, low, moderate, and high) and change scores of the outcome measures at postintervention and 12-week follow-up were explored using 1-way analyses of variance (ANOVAs). Average participants' satisfaction ratings were used to partially test acceptability. Finally, the effectiveness of the chatbot was tested using paired sample  $t$  tests of the changes in outcome measures from baseline to postintervention and follow-up. The qualitative data were manually coded by the authors.

Interviews and feedback collected during the program, at postintervention, and 12-week follow-up were included in the qualitative analysis. An inductive thematic analysis was conducted to determine factors relating to participant acceptance of and engagement with Otis. Emerging factors were then organized into themes to answer the questions of acceptability and engagement. Engagement was defined as factors related to participants' interest and persistence with the intervention, while participants' use and perceptions of Otis as a chatbot for health anxiety management were defined as factors related to acceptability.

## Results

### Attrition

Of the 69 participants who completed baseline measures, 35 participants (51%) completed the postintervention questionnaire, and 29 (42%) completed the 12-week follow-up. In addition, 7 participants signaled interest in and completed an interview with a researcher.

**Table 2.** Reasons for abandoning the chatbot program (n=29).

Reasons for abandonment of Otis	Results, n (%)
I didn't have time to use the chatbot.	14 (48)
I had technical problems.	7 (24)
I didn't feel anxious about my health, so I didn't need it anymore.	6 (21)
I tried to put the skills into practice, but it didn't work for me.	1 (3)
It was boring.	1 (3)

There was a significant decrease in general anxiety from baseline (mean 9.69, SD 4.38) to postintervention (mean 7.58, SD 4.38;  $t_{28}=3.30$ ,  $P=.003$ ). The decrease in general anxiety was maintained at the 12-week follow-up (mean 7.00, SD 4.99;  $t_{28}=3.26$ ,  $P=.003$ ). On average, self-reported personal well-being increased from baseline (mean 5.70, SD 1.30) to postintervention (mean 6.31, SD 1.06) and 12-week follow-up (mean 6.50, SD 1.55). These were significant increases at both postintervention

## Demographics

The final sample of 29 participants who completed all follow-up questionnaires was predominantly female (25/29, 86%), ranging in age from 21 years to 80 years old with a mean age of 37.4 (SD 15.1) years. New Zealand Europeans or Pākehā were the most common ethnic groups in the sample (20/29, 69%). Most participants (16/29, 55%) reported they participated because they wanted help to manage their anxiety, followed by a curiosity about chatbot technology (8/29, 28%). Approximately one-half (19/29, 66%) of the sample had not used any DHI for anxiety before the study. Chi-square tests revealed no significant demographic differences between those who completed baseline and those who completed all follow-up questionnaires.

## Evaluation Outcomes

Participants completed an average of 9.8 (SD 4.43) days or 70% of the chatbot intervention before the 21-day period lapsed. Of the 29 participants included in the final analysis, 12 (41%) completed the full 14-day program. The most common reason for abandonment was lack of time (14/29, 35%). See Table 2 for a summary of the reasons cited for abandoning the program. Due to a technical fault in Facebook's data analytics system, data relating to the unique interactions participants had with Otis could not be accurately reported. The mean overall user experience was 18.24 (SD 3.38) out of a possible score of 30. The overall acceptability of the chatbot was rated favorably on a scale of 1 to 10 (mean 8.24, SD 1.80).

On average, there was a reduction in health anxiety from baseline (mean 19.93, SD 8.82) to posttreatment (mean 19.31, SD 9.02) and 12-week follow-up (mean 17.38, SD 10.03). However, this improvement did not reach significance at posttreatment ( $t_{28}=0.61$ ,  $P=.55$ ) or follow-up ( $t_{28}=1.82$ ,  $P=.08$ ). The same pattern was observed in the intolerance of uncertainty of participants at baseline (mean 9.69, SD 4.34), for which the reduction in these scores did not reach significance at posttreatment (mean 7.59, SD 4.38;  $t_{28}=1.82$ ,  $P=.08$ ) or follow-up (mean 7, SD 4.99;  $t_{28}=0.44$ ,  $P=.66$ ).

( $t_{28}=-1.12$ ,  $P=.03$ ) and follow-up ( $t_{28}=-1.47$ ,  $P=.03$ ). There was also significant improvement in self-reported quality of life at postintervention ( $t_{28}=-2.39$ ,  $P=.02$ ) and 12-week follow-up ( $t_{28}=-3.64$ ,  $P<.001$ ; Table 3). Paired sample  $t$  tests of health anxiety, intolerance of uncertainty, general anxiety, personal well-being, and quality of life for all participants (n=35) who completed the postintervention questionnaire revealed no difference in outcomes.

Paired sample *t* tests of health anxiety, intolerance of uncertainty, general anxiety, personal well-being, and quality of life for all participants ( $n=35$ ) who completed the postintervention questionnaire revealed similar results (Table 4). Health anxiety did not decrease significantly from baseline (mean 19.91, SD 8.36) to postintervention (mean 19.34, SD 8.45;  $t_{34}=0.65$ ,  $P=.52$ ) nor did intolerance of uncertainty from baseline (mean 35.63, SD 10.50) to postintervention (mean

33.40, SD 10.5;  $t_{34}=1.88$ ,  $P=.07$ ). The decrease in general anxiety (baseline: mean 10.20, SD 4.80; postintervention: mean 8.26, SD 4.96) as measured by the GAD-7 was significant ( $t_{34}=3.58$ ,  $P<.001$ ). As seen in the final sample, personal well-being significantly increased from baseline (mean 5.66, SD 1.26) to postintervention (mean 6.15, SD 1.05;  $t_{34}=-2.11$ ,  $P=.04$ ), as did quality of life (baseline: mean 9.49, SD 4.35; postintervention: mean 11.09, SD 4.53;  $t_{34}=-2.62$ ,  $P=.01$ ).

**Table 3.** Effect of the intervention on health anxiety, intolerance of uncertainty, general anxiety, personal well-being, and quality of life.

Measure	Baseline (n=29), mean (SD)	Postintervention (n=29), mean (SD)	Follow-up (n=29), mean (SD)	Baseline to postintervention		Baseline to follow-up	
				<i>t</i> (df)	<i>P</i> value	<i>t</i> (df)	<i>P</i> value
Health anxiety (SHAI-18 <sup>a</sup> )	19.93 (8.83)	19.31 (9.03)	17.38 (10.03)	0.61 (28)	.55	-2.28 (28)	.08
Intolerance of uncertainty (IUS-12 <sup>b</sup> )	34.10 (9.70)	31.72 (9.71)	33.38 (10.78)	1.82 (28)	.08	0.44 (28)	.66
General anxiety (GAD-7 <sup>c</sup> )	9.69 (4.34)	7.59 (4.38)	7.00 (4.99)	3.30 (28)	.003	3.26 (28)	.003
Personal well-being (ONS4 <sup>d</sup> )	5.70 (1.30)	6.31 (1.06)	6.50 (1.55)	-2.33 (28)	.03	-2.28 (28)	.03
Quality of life (WHO-5 <sup>e</sup> )	9.62 (4.53)	11.31 (4.90)	12.93 (5.08)	-2.39 (28)	.02	-3.64 (28)	.001

<sup>a</sup>SHAI-18: Short Health Anxiety Scale.

<sup>b</sup>IUS-12: Short Intolerance of Uncertainty Scale.

<sup>c</sup>GAD-7: Generalized Anxiety Disorder 7-item Scale.

<sup>d</sup>ONS4: Office of National Statistics Personal Well-being Domain.

<sup>e</sup>WHO-5: World Health Organization Five Well-being Index.

**Table 4.** Effect of the intervention on health anxiety, intolerance of uncertainty, general anxiety, personal well-being, and quality of life for all participants who completed the postintervention questionnaire.

Measure	Baseline (n=35), mean (SD)	Postintervention, (n=35), mean (SD)	Baseline to postintervention	
			<i>t</i> (df)	<i>P</i> value
Health anxiety (SHAI-18 <sup>a</sup> )	19.91 (8.36)	19.34 (8.45)	0.65 (34)	.52
Intolerance of uncertainty (IUS-12 <sup>b</sup> )	35.63 (10.50)	33.40 (10.50)	1.88 (34)	.07
General anxiety (GAD-7 <sup>c</sup> )	10.20 (4.80)	8.26 (4.96)	3.58 (34)	.001
Personal well-being (ONS4 <sup>d</sup> )	5.66 (1.26)	6.15 (1.05)	-2.11 (34)	.04
Quality of life (WHO-5 <sup>e</sup> )	9.49 (4.35)	11.09 (4.53)	-2.62 (34)	.01

<sup>a</sup>SHAI-18: Short Health Anxiety Scale.

<sup>b</sup>IUS-12: Short Intolerance of Uncertainty Scale.

<sup>c</sup>GAD-7: Generalized Anxiety Disorder 7-item Scale.

<sup>d</sup>ONS4: Office of National Statistics Personal Well-being Domain.

<sup>e</sup>WHO-5: World Health Organization Five Well-being Index.

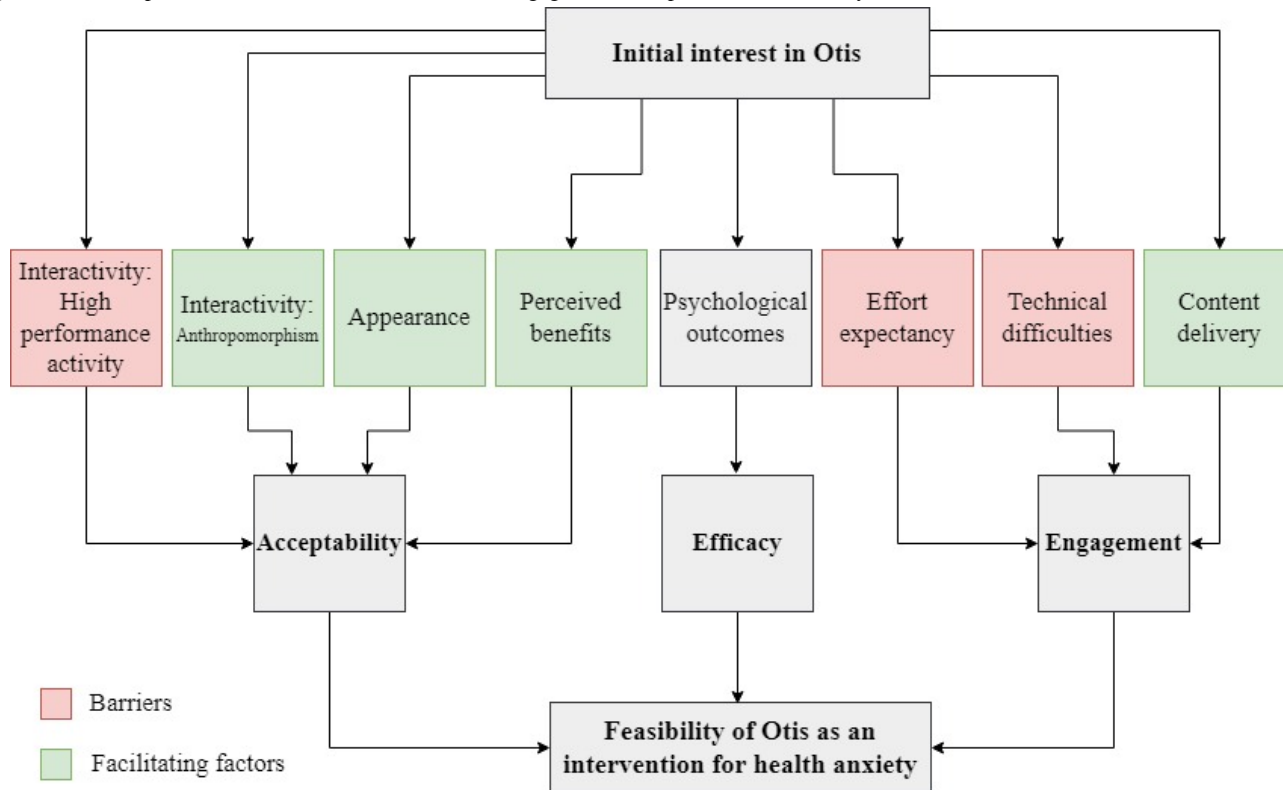
One-way ANOVAs and post hoc tests (Tukey B and Games-Howell) demonstrated no significant difference in IUS-12, GAD-7, ONS4, and WHO-5 change scores among poor, low, moderate, and high adherence groups at postintervention and the 12-week follow-up. However, the SHAI-18 change scores at follow-up were significantly different between participants who moderately and poorly adhered to the program ( $F_{3,25}=3.59$ ,  $P=.02$ ). Scatterplots were used to further investigate the correlation between change scores and the

number of days completed, but no other remarkable relationships were identified.

Across the qualitative data, 3 key factors related to engagement and acceptance emerged (Figure 3). Quotes related to these factors can be found in Multimedia Appendix 1. Participants reported content delivery, technical difficulties, and effort expectancy as factors influencing their engagement with the chatbot, while themes relating to the chatbot's appearance, perceived benefits, and interactivity (how interactive the chatbot was) influenced acceptability. Within the factors of interactivity,

2 subfactors, anthropomorphizing of Otis and high-performance expectancy, were found to distinctly affect acceptability.

**Figure 3.** A conceptualization of the factors related to the engagement, acceptance, and feasibility of Otis.



### Engagement

Regarding effort expectancy, participants reported that, although the “check-ins” were short, conversing with the chatbot daily was tedious and the expectation of participant time was too much. Social engagements, fatigue, and daily tasks were commonly cited as distractions.

Technological difficulties were cited as one of the main reasons for abandonment in the quantitative data. It was reflected in the feedback and interviews, as people discussed the frustration of Facebook and Chatfuel malfunctions and difficulties with understanding how to interact with Otis.

Regarding content delivery, the delivery of information in short “check-ins” and short messages was favored by the participants who reported that, if the information was presented in a long set of text, it would be more tiring to read and less engaging. Additionally, participants noted that the pace of information delivery was manageable and easy to comprehend.

### Acceptance

Interactivity was assessed by the anthropomorphizing of Otis and high-performance expectancy. Regarding the anthropomorphizing of Otis, nearly all participants attributed human-like characteristics to Otis and gave Otis feedback on days 4, 11, and 14 as if it were a person. Despite the robot avatar, participants explained that the content of the conversation, use of emojis, and experience of texting made Otis feel like another “person” they were chatting with online. Regarding high-performance expectancy, although participants attributed human-like characteristics to Otis, they were also

aware of the limitations of a decision tree-based chatbot. Participants described a lack of personalization that could have been achieved with “smart” or AI chatbots. Others described instances where they were unable to discuss certain topics in-depth or address them at all as they were not part of the script.

Otis’ appearance appeared to be important for the acceptance of the chatbot as an intervention for health anxiety. There appeared to be a preference for a chatbot avatar over a human or human-like design.

Regarding perceived benefits, participants reported that they found Otis helpful and made them gain greater insight into health anxiety and develop strategies to challenge their anxious thoughts. It appeared that these benefits saw participants accepting Otis as a valid intervention for health anxiety.

## Discussion

### Principal Findings

Based on a review of the literature and to the best of our knowledge, this is the first study to assess the feasibility, efficacy, acceptability, and engagement of a chatbot (Otis) for health anxiety management among adults. The results suggested that, although there was no significant reduction in health anxiety, the chatbot was accepted as an intervention for health anxiety management and was associated with decreased generalized anxiety and increased perceived anxiety management, personal well-being, and quality of life. The study identified several barriers (high performance expectancy, effort expectancy, technical difficulties) and facilitators (appearance, anthropomorphism, perceived benefits, content delivery) to the

engagement and acceptability of chatbots delivering a psychological intervention.

### Comparison With Prior Work

This study found that approximately 41% of participants completed the study and that, on average, the final sample completed 70% of the intervention, which is higher than seen in most of the literature [48-50]. The high rate of adherence may indicate (1) a highly motivated sample of participants considering that most participants wanted help with managing anxiety (55%), (2) frequent phone users in the sample (>10 times per day = 35%), or (3) that the conversational nature of a chatbot is a more engaging medium than others for delivering online CBT programs.

Although Otis may be engaging, the intervention was not associated with a significant reduction in health anxiety at postintervention or the 12-week follow-up, which may indicate that CBT delivered via chatbot as a treatment for health anxiety is ineffective. This is inconsistent with what has been reported in previous studies in which CBT delivered via digital media was effective in reducing health anxiety in adult populations [22]. Instead, the results were consistent with a systematic review and meta-analysis of mental health chatbots demonstrating weak evidence for chatbots producing clinically significant psychological outcomes [26].

Failure to reach significance may also be explained by the low statistical power of the study due to a small sample size. Additionally, the results may demonstrate a floor effect during the COVID-19 pandemic, during which it was adaptive to be vigilant about health, in which case the SHAI-18 may have been measuring atypical behavior seen and encouraged during a pandemic. Therefore, although the chatbot was associated with a decrease in health anxiety, perhaps in the context of a pandemic, Otis could have no further effect in lowering health anxiety.

The qualitative results revealed that participants perceived Otis as useful in reducing their health anxiety. The inconsistency between the quantitative and qualitative results may indicate that the program was associated with improved anxiety management but that health anxiety persisted as an adaptive response to COVID-19 and New Zealand's several lockdowns. Furthermore, only 5 participants in the study met the cut-off score (25) for health anxiety [2], which further supports the supposition of a floor effect in the remainder of the sample. Additionally, these results may reflect the perception of increased usefulness beyond the actual functionality of a chatbot that is associated with anthropomorphism [51].

To the best of our knowledge, no study has shown significantly improved anxiety compared with a control group using a CBT-based chatbot. For example, a study into Woebot's effect on anxiety found that the information-only condition was as effective in reducing anxiety as the chatbot user group [52]. As exposure and habituation are key components in the treatment of anxiety, chatbots may not be a suitable medium, as the technology is unable to actively incorporate and encourage these components of treatment in the same way a therapist could. Otis was also associated with significant improvements in users'

personal well-being and quality of life at both postintervention and the 12-week follow-up. These results are congruent with a large body of literature demonstrating improved anxiety is linked to a better quality of life and well-being. Lockdowns have a significant impact on mental well-being and other aspects related to quality of life. Therefore, an alternative explanation for the improvement across these measures may be a result of New Zealand moving into less restrictive lockdown alert levels during the course of the study.

A systematic review of barriers to facilitators of engagement with mobile health (mHealth) interventions reported that a user's ability to integrate the intervention into their lives impacted engagement [53]. The importance of integration is reflected in this study in which the perceived high effort expectancy of using Otis competed with other activities leading to forgetfulness or deprioritizing of use, as seen in other studies [54,55]. The themes of technical difficulties and content delivery affecting engagement with Otis reflected the results in the systematic review by Borghouts and colleagues [53]. Technical difficulties are a key issue to consider in mHealth interventions as they are often unavoidable. For example, technology is constantly updated to make improvements to the user experience. However, these updates can cause systems to crash and slow, causing disruptions to users. Disruption can significantly impact a user's trust in the system and motivation to engage especially when in a discussion about their mental health. Finally, participants found that Otis' short and varied content delivery (text, videos, audio, and GIFs) was engaging, which is congruent with other studies assessing the effects of content delivery on overall user experience [56]. Despite relatively high engagement and reported acceptability, participants did not consider Otis a substitute for a health professional, in line with previous research that showed a preference for CBT and medication in the treatment of health anxiety over iCBT [57].

An unexpected finding in this study was participants' preference for a nonhuman chatbot avatar. Though previous work suggests that anthropomorphism is amplified by human-looking avatars [58], even in the absence of a human avatar, users responded socially to Otis. A recent study reported that users' perceived trustworthiness and affinity were lower when a realistic digital avatar was used as compared with a human being but remained unaffected by the knowledge of whether a chatbot was controlled by a person or AI [59]. Although the aforementioned study did not compare human and nonhuman avatars, its findings indicate that appearance makes a significant difference to the user experience. Participants in this study reported feelings of mistrust and worries of judgment as they would have with a real human if Otis were not a robot figure, thus being a barrier to acceptance as an intervention for health anxiety. These findings are in direct conflict with previous research suggesting that, if users were to apply the same social rules to their interactions with a chatbot, it may improve a user's perception of social presence [60].

A study that appears to be in partial support of the findings of this study reported that too many anthropomorphic characteristics in a chatbot avatar set up high expectations, leading to user disappointment [61], thus supporting the idea that high expectations of functionality can lead to feelings of



disappointment and frustration for participants. Upgrading Otis and other chatbots to AI-based chatbots that can understand and better process user input may meet the expectations of users in the future. The results presented in this study and the existing literature should be considered in work with “digital humans,” which aim to make digital interactions more personalized and lifelike. Consideration should be given for which contexts and populations the highly anthropomorphic digital human characters would be suitable.

### Limitations

The observational design of this study limited the ability to imply causation. Additionally, the results may have been confounded by unidentifiable variables such as psychological input outside of the intervention and changes in COVID-19 restrictions. Although the plethora of pilot and feasibility studies within the field of mHealth is a common criticism, the use of randomized controlled trials only provides evidence of the efficacy of a digital intervention at a certain point in time. Pilot and feasibility studies instead grow understanding of factors involved in digital interactions to be implemented in evolving technology. The sample was mostly New Zealand European and highly educated, thus limiting the generalizability of the

findings across the population. Additionally, bias within the interpretation of the qualitative results, a common critique of qualitative research [62], is another limitation of the study.

### Conclusions

In conclusion, Otis, a CBT-based chatbot, may be a feasible intervention for the management but not treatment of health anxiety among adults. The study found that, although there was no significant improvement in health anxiety, participants reported benefiting from the intervention, which is further evidenced by improvements in general anxiety, personal well-being, and quality of life. The results suggest that mental health chatbots are a feasible supplementary treatment that can alleviate strain on psychological services. Therefore, future studies should continue to evaluate the feasibility of chatbots in this space as the technology develops. This pilot study makes a unique contribution to the understanding of digital interactions, health anxiety, and the potential of chatbots in mental health care. Future studies in the area should continue to investigate the “therapeutic relationship” between people and technology to aid the development of mHealth, thereby facilitating the democratization of mental health services.

### Conflicts of Interest

None disclosed.

### Multimedia Appendix 1

Participant quotes regarding engagement and acceptance of Otis as an intervention for health anxiety management. [[DOCX File, 18 KB - formative\\_v6i10e37877\\_app1.docx](#)]

### References

1. Looper KJ, Kirmayer LJ. Hypochondriacal concerns in a community population. *Psychol Med* 2001 May 17;31(4):577-584. [doi: [10.1017/s0033291701003737](https://doi.org/10.1017/s0033291701003737)] [Medline: [11352360](https://pubmed.ncbi.nlm.nih.gov/11352360/)]
2. Abramowitz JS, Deacon BJ, Valentiner DP. The Short Health Anxiety Inventory: psychometric properties and construct validity in a non-clinical sample. *Cognit Ther Res* 2007 Feb 15;31(6):871-883 [FREE Full text] [doi: [10.1007/s10608-006-9058-1](https://doi.org/10.1007/s10608-006-9058-1)] [Medline: [32214558](https://pubmed.ncbi.nlm.nih.gov/32214558/)]
3. Mertens G, Gerritsen L, Duijndam S, Salemink E, Engelhard IM. Fear of the coronavirus (COVID-19): Predictors in an online study conducted in March 2020. *J Anxiety Disord* 2020 Aug;74:102258 [FREE Full text] [doi: [10.1016/j.janxdis.2020.102258](https://doi.org/10.1016/j.janxdis.2020.102258)] [Medline: [32569905](https://pubmed.ncbi.nlm.nih.gov/32569905/)]
4. Taylor S, Asmundson GJG. *Treating Health Anxiety: A Cognitive-Behavioral Approach*. New York, NY: The Guilford Press; 2004.
5. Wheaton MG, Abramowitz JS, Berman NC, Fabricant LE, Olatunji BO. Psychological predictors of anxiety in response to the H1N1 (swine flu) pandemic. *Cogn Ther Res* 2011 Feb 22;36(3):210-218. [doi: [10.1007/s10608-011-9353-3](https://doi.org/10.1007/s10608-011-9353-3)]
6. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*, 4th edition. Washington, DC: American Psychiatric Association; 2000.
7. Hedman E, Lekander M, Karshikoff B, Ljótsson B, Axelsson E, Axelsson J. Health anxiety in a disease-avoidance framework: Investigation of anxiety, disgust and disease perception in response to sickness cues. *J Abnorm Psychol* 2016 Oct;125(7):868-878. [doi: [10.1037/abn0000195](https://doi.org/10.1037/abn0000195)] [Medline: [27732023](https://pubmed.ncbi.nlm.nih.gov/27732023/)]
8. Olatunji BO, Etzel EN, Tomarken AJ, Ciesielski BG, Deacon B. The effects of safety behaviors on health anxiety: an experimental investigation. *Behav Res Ther* 2011 Nov;49(11):719-728. [doi: [10.1016/j.brat.2011.07.008](https://doi.org/10.1016/j.brat.2011.07.008)] [Medline: [21839987](https://pubmed.ncbi.nlm.nih.gov/21839987/)]
9. Salkovskis PM. The importance of behaviour in the maintenance of anxiety and panic: a cognitive account. *Behav Psychother* 2009 Jun 16;19(1):6-19. [doi: [10.1017/S0141347300011472](https://doi.org/10.1017/S0141347300011472)]
10. Helbig-Lang S, Petermann F. Tolerate or eliminate? A systematic review on the effects of safety behavior across anxiety disorders. *Clinical Psychology: Science and Practice* 2010 Sep;17(3):218-233. [doi: [10.1111/j.1468-2850.2010.01213.x](https://doi.org/10.1111/j.1468-2850.2010.01213.x)]

11. Ersoy K, Altin B, Bayram Sarıkaya B, Güngörmüş Özkardaş O. The comparison of impact of health anxiety on dispositional hope and psychological well-being of mothers who have children diagnosed with autism and mothers who have normal children, in Covid-19 pandemic. *Social Sciences Research Journal* 2020;9(2):117-126 [[FREE Full text](#)]
12. Lee S, Creed FH, Ma Y, Leung CM. Somatic symptom burden and health anxiety in the population and their correlates. *J Psychosom Res* 2015 Jan;78(1):71-76. [doi: [10.1016/j.jpsychores.2014.11.012](#)] [Medline: [25466323](#)]
13. Sunderland M, Newby JM, Andrews G. Health anxiety in Australia: prevalence, comorbidity, disability and service use. *Br J Psychiatry* 2013 Jan 02;202(1):56-61. [doi: [10.1192/bjp.bp.111.103960](#)] [Medline: [22500013](#)]
14. Abd-Alrazaq AA, Alajlani M, Alalwan AA, Bewick BM, Gardner P, Househ M. An overview of the features of chatbots in mental health: A scoping review. *Int J Med Inform* 2019 Dec;132:103978. [doi: [10.1016/j.ijmedinf.2019.103978](#)] [Medline: [31622850](#)]
15. Bickmore TW, Mitchell SE, Jack BW, Paasche-Orlow MK, Pfeifer LM, Odonnell J. Response to a relational agent by hospital patients with depressive symptoms. *Interact Comput* 2010 Jul 01;22(4):289-298 [[FREE Full text](#)] [doi: [10.1016/j.intcom.2009.12.001](#)] [Medline: [20628581](#)]
16. Oh J, Jang S, Kim H, Kim J. Efficacy of mobile app-based interactive cognitive behavioral therapy using a chatbot for panic disorder. *Int J Med Inform* 2020 Aug;140:104171. [doi: [10.1016/j.ijmedinf.2020.104171](#)] [Medline: [32446158](#)]
17. Tielman ML, Neerinx MA, Bidarra R, Kybartas B, Brinkman W. A therapy system for post-traumatic stress disorder using a virtual agent and virtual storytelling to reconstruct traumatic memories. *J Med Syst* 2017 Aug 11;41(8):125 [[FREE Full text](#)] [doi: [10.1007/s10916-017-0771-y](#)] [Medline: [28699083](#)]
18. Bickmore TW, Puskar K, Schlenk EA, Pfeifer LM, Sereika SM. Maintaining reality: Relational agents for antipsychotic medication adherence. *Interacting with Computers* 2010 Jul;22(4):276-288. [doi: [10.1016/j.intcom.2010.02.001](#)]
19. Ly KH, Ly A, Andersson G. A fully automated conversational agent for promoting mental well-being: A pilot RCT using mixed methods. *Internet Interv* 2017 Dec;10:39-46 [[FREE Full text](#)] [doi: [10.1016/j.invent.2017.10.002](#)] [Medline: [30135751](#)]
20. Cooper K, Gregory JD, Walker I, Lambe S, Salkovskis PM. Cognitive behaviour therapy for health anxiety: a systematic review and meta analysis. *Behav. Cogn. Psychother* 2017 Feb 23;45(2):110-123. [doi: [10.1017/s1352465816000527](#)]
21. Olatunji BO, Kauffman BY, Meltzer S, Davis ML, Smits JA, Powers MB. Cognitive-behavioral therapy for hypochondriasis/health anxiety: a meta-analysis of treatment outcome and moderators. *Behav Res Ther* 2014 Jul;58:65-74. [doi: [10.1016/j.brat.2014.05.002](#)] [Medline: [24954212](#)]
22. Axelsson E, Hedman-Lagerlöf E. Cognitive behavior therapy for health anxiety: systematic review and meta-analysis of clinical efficacy and health economic outcomes. *Expert Rev Pharmacoecon Outcomes Res* 2019 Dec;19(6):663-676. [doi: [10.1080/14737167.2019.1703182](#)] [Medline: [31859542](#)]
23. Qualtrics XM. URL: <https://www.qualtrics.com/> [accessed 2022-10-01]
24. Craske MG, Roy-Byrne PP, Stein MB, Sullivan G, Sherbourne C, Bystritsky A. Treatment for anxiety disorders: Efficacy to effectiveness to implementation. *Behav Res Ther* 2009 Nov;47(11):931-937 [[FREE Full text](#)] [doi: [10.1016/j.brat.2009.07.012](#)] [Medline: [19632667](#)]
25. Hedman E, Andersson E, Ljótsson B, Andersson G, Rück C, Lindfors N. Cost-effectiveness of Internet-based cognitive behavior therapy vs. cognitive behavioral group therapy for social anxiety disorder: results from a randomized controlled trial. *Behav Res Ther* 2011 Nov;49(11):729-736. [doi: [10.1016/j.brat.2011.07.009](#)] [Medline: [21851929](#)]
26. Abd-Alrazaq AA, Alajlani M, Ali N, Denecke K, Bewick BM, Househ M. Perceptions and opinions of patients about mental health chatbots: scoping review. *J Med Internet Res* 2021 Jan 13;23(1):e17828 [[FREE Full text](#)] [doi: [10.2196/17828](#)] [Medline: [33439133](#)]
27. Salkovskis PM, Rimes KA, Warwick HMC, Clark DM. The Health Anxiety Inventory: development and validation of scales for the measurement of health anxiety and hypochondriasis. *Psychol Med* 2002 Jul 15;32(5):843-853. [doi: [10.1017/s0033291702005822](#)] [Medline: [12171378](#)]
28. Alberts NM, Hadjistavropoulos HD, Jones SL, Sharpe D. The Short Health Anxiety Inventory: a systematic review and meta-analysis. *J Anxiety Disord* 2013 Jan;27(1):68-78. [doi: [10.1016/j.janxdis.2012.10.009](#)] [Medline: [23247202](#)]
29. Rode S, Salkovskis P, Dowd H, Hanna M. Health anxiety levels in chronic pain clinic attenders. *J Psychosom Res* 2006 Feb;60(2):155-161. [doi: [10.1016/j.jpsychores.2005.07.005](#)] [Medline: [16439268](#)]
30. Seivewright H, Salkovskis P, Green J, Mullan N, Behr G, Carlin E, et al. Prevalence and service implications of health anxiety in genitourinary medicine clinics. *Int J STD AIDS* 2004 Aug;15(8):519-522. [doi: [10.1258/0956462041558122](#)] [Medline: [15307961](#)]
31. Tang NKY, Wright KJ, Salkovskis PM. Prevalence and correlates of clinical insomnia co-occurring with chronic back pain. *J Sleep Res* 2007 Mar;16(1):85-95 [[FREE Full text](#)] [doi: [10.1111/j.1365-2869.2007.00571.x](#)] [Medline: [17309767](#)]
32. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch Intern Med* 2006 May 22;166(10):1092-1097. [doi: [10.1001/archinte.166.10.1092](#)] [Medline: [16717171](#)]
33. Fergus TA, Valentiner DP. Intolerance of uncertainty moderates the relationship between catastrophic health appraisals and health anxiety. *Cogn Ther Res* 2011 Sep 9;35(6):560-565. [doi: [10.1007/s10608-011-9392-9](#)]
34. Norr AM, Albanese BJ, Oglesby ME, Allan NP, Schmidt NB. Anxiety sensitivity and intolerance of uncertainty as potential risk factors for cyberchondria. *J Affect Disord* 2015 Mar 15;174:64-69. [doi: [10.1016/j.jad.2014.11.023](#)] [Medline: [25486275](#)]

35. Carleton RN, Norton MPJ, Asmundson GJ. Fearing the unknown: a short version of the Intolerance of Uncertainty Scale. *J Anxiety Disord* 2007 Jan;21(1):105-117. [doi: [10.1016/j.janxdis.2006.03.014](https://doi.org/10.1016/j.janxdis.2006.03.014)] [Medline: [16647833](https://pubmed.ncbi.nlm.nih.gov/16647833/)]
36. Buhr K, Dugas M. The intolerance of uncertainty scale: psychometric properties of the English version. *Behaviour Research and Therapy* 2002 Aug;40(8):931-945. [doi: [10.1016/s0005-7967\(01\)00092-4](https://doi.org/10.1016/s0005-7967(01)00092-4)]
37. Carson J, Prescott J, Allen R, McHugh S. Winter is coming: age and early psychological concomitants of the Covid-19 pandemic in England. *JPMH* 2020 Jul 27;19(3):221-230. [doi: [10.1108/jpmh-06-2020-0062](https://doi.org/10.1108/jpmh-06-2020-0062)]
38. de Albornoz SC, Chen G. Relationship between health-related quality of life and subjective wellbeing in asthma. *J Psychosom Res* 2021 Mar;142:110356. [doi: [10.1016/j.jpsychores.2021.110356](https://doi.org/10.1016/j.jpsychores.2021.110356)] [Medline: [33454566](https://pubmed.ncbi.nlm.nih.gov/33454566/)]
39. Eichstaedt JC, Yaden DB, Ribeiro F, Adler A, Kern ML. Lifestyle and wellbeing: Exploring behavioral and demographic covariates in a large US sample. *Intl. J. Wellbeing* 2020 Sep 30;10(4):87-112. [doi: [10.5502/ijw.v10i4.831](https://doi.org/10.5502/ijw.v10i4.831)]
40. Gao L, Moodie M, Chen G. Measuring subjective wellbeing in patients with heart disease: relationship and comparison between health-related quality of life instruments. *Qual Life Res* 2019 Apr 2;28(4):1017-1028. [doi: [10.1007/s11136-018-2094-y](https://doi.org/10.1007/s11136-018-2094-y)] [Medline: [30604342](https://pubmed.ncbi.nlm.nih.gov/30604342/)]
41. Benson T, Sladen J, Liles A, Potts HWW. Personal Wellbeing Score (PWS)—a short version of ONS4: development and validation in social prescribing. *BMJ Open Qual* 2019 Apr 11;8(2):e000394. [doi: [10.1136/bmjoq-2018-000394](https://doi.org/10.1136/bmjoq-2018-000394)]
42. Bech P, Olsen LR, Kjoller M, Rasmussen NK. Measuring well-being rather than the absence of distress symptoms: a comparison of the SF-36 Mental Health subscale and the WHO-Five Well-Being Scale. *Int J Methods Psychiatr Res* 2003 Jun;12(2):85-91 [FREE Full text] [doi: [10.1002/mpr.145](https://doi.org/10.1002/mpr.145)] [Medline: [12830302](https://pubmed.ncbi.nlm.nih.gov/12830302/)]
43. McDowell I. Measures of self-perceived well-being. *J Psychosom Res* 2010 Jul;69(1):69-79. [doi: [10.1016/j.jpsychores.2009.07.002](https://doi.org/10.1016/j.jpsychores.2009.07.002)] [Medline: [20630265](https://pubmed.ncbi.nlm.nih.gov/20630265/)]
44. Sischka PE, Costa AP, Steffgen G, Schmidt AF. The WHO-5 well-being index – validation based on item response theory and the analysis of measurement invariance across 35 countries. *Journal of Affective Disorders Reports* 2020 Dec;1:100020. [doi: [10.1016/j.jadr.2020.100020](https://doi.org/10.1016/j.jadr.2020.100020)]
45. 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/)]
46. Zoom. URL: <https://zoom.us/> [accessed 2022-10-01]
47. IBM SPSS software. IBM. URL: <https://www.ibm.com/spss> [accessed 2022-10-01]
48. Christensen H, Griffiths KM, Farrer L. Adherence in internet interventions for anxiety and depression. *J Med Internet Res* 2009 Apr 24;11(2):e13 [FREE Full text] [doi: [10.2196/jmir.1194](https://doi.org/10.2196/jmir.1194)] [Medline: [19403466](https://pubmed.ncbi.nlm.nih.gov/19403466/)]
49. 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/)]
50. van Ballegooijen W, Cuijpers P, van Straten A, Karyotaki E, Andersson G, Smit JH, et al. Adherence to Internet-based and face-to-face cognitive behavioural therapy for depression: a meta-analysis. *PLoS One* 2014;9(7):e100674 [FREE Full text] [doi: [10.1371/journal.pone.0100674](https://doi.org/10.1371/journal.pone.0100674)] [Medline: [25029507](https://pubmed.ncbi.nlm.nih.gov/25029507/)]
51. Rietz T, Benke I, Maedche A. The impact of anthropomorphic and functional chatbot design features in enterprise collaboration systems on user acceptance. 2019 Presented at: 14th International Conference on Wirtschaftsinformatik; February 24-27, 2019; Siegen, Germany.
52. 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/)]
53. Borghouts J, Eikley E, Mark G, De Leon C, Schueller SM, Schneider M, et al. Barriers to and facilitators of user engagement with digital mental health interventions: systematic review. *J Med Internet Res* 2021 Mar 24;23(3):e24387 [FREE Full text] [doi: [10.2196/24387](https://doi.org/10.2196/24387)] [Medline: [33759801](https://pubmed.ncbi.nlm.nih.gov/33759801/)]
54. Donkin L, Glozier N. Motivators and motivations to persist with online psychological interventions: a qualitative study of treatment completers. *J Med Internet Res* 2012 Jun 22;14(3):e91 [FREE Full text] [doi: [10.2196/jmir.2100](https://doi.org/10.2196/jmir.2100)] [Medline: [22743581](https://pubmed.ncbi.nlm.nih.gov/22743581/)]
55. Simblett SK, Bruno E, Siddi S, Matcham F, Giuliano L, López JH, RADAR-CNS Consortium. Patient perspectives on the acceptability of mHealth technology for remote measurement and management of epilepsy: A qualitative analysis. *Epilepsy Behav* 2019 Aug;97:123-129. [doi: [10.1016/j.yebeh.2019.05.035](https://doi.org/10.1016/j.yebeh.2019.05.035)] [Medline: [31247523](https://pubmed.ncbi.nlm.nih.gov/31247523/)]
56. Sinha Deb K, Tuli A, Sood M, Chadda R, Verma R, Kumar S, et al. Is India ready for mental health apps (MHApps)? A quantitative-qualitative exploration of caregivers' perspective on smartphone-based solutions for managing severe mental illnesses in low resource settings. *PLoS One* 2018 Sep 19;13(9):e0203353 [FREE Full text] [doi: [10.1371/journal.pone.0203353](https://doi.org/10.1371/journal.pone.0203353)] [Medline: [30231056](https://pubmed.ncbi.nlm.nih.gov/30231056/)]
57. Soucy JN, Hadjistavropoulos HD. Treatment acceptability and preferences for managing severe health anxiety: Perceptions of internet-delivered cognitive behaviour therapy among primary care patients. *J Behav Ther Exp Psychiatry* 2017 Dec;57:14-24. [doi: [10.1016/j.jbtep.2017.02.002](https://doi.org/10.1016/j.jbtep.2017.02.002)] [Medline: [28242411](https://pubmed.ncbi.nlm.nih.gov/28242411/)]
58. Nass C, Kim EY, Lee FJ. When my face is the interface: an experimental comparison of interacting with one's own face or someone else's face. 1998 Presented at: SIGCHI Conference on Human Factors in Computing Systems; April 18-23, 1998; Los Angeles, CA. [doi: [10.1145/274644.274667](https://doi.org/10.1145/274644.274667)]

59. Seymour M, Yuan L, Dennis AR, Riemer K. Facing the Artificial: Understanding Affinity, Trustworthiness, and Preference for More Realistic Digital Humans. 2020 Presented at: 53rd Hawaii International Conference on System Sciences; January 7-10, 2020; Maui, HI. [doi: [10.24251/hicss.2020.574](https://doi.org/10.24251/hicss.2020.574)]
60. Qiu L, Benbasat I. Evaluating anthropomorphic product recommendation agents: a social relationship perspective to designing information systems. *Journal of Management Information Systems* 2014 Dec 08;25(4):145-182. [doi: [10.2753/MIS0742-1222250405](https://doi.org/10.2753/MIS0742-1222250405)]
61. Ben Mimoun MS, Poncin I, Garnier M. Case study—Embodied virtual agents: An analysis on reasons for failure. *Journal of Retailing and Consumer Services* 2012 Nov;19(6):605-612. [doi: [10.1016/j.jretconser.2012.07.006](https://doi.org/10.1016/j.jretconser.2012.07.006)]
62. Mays N, Pope C. Rigour and qualitative research. *BMJ* 1995 Jul 08;311(6997):109-112 [FREE Full text] [doi: [10.1136/bmj.311.6997.109](https://doi.org/10.1136/bmj.311.6997.109)] [Medline: [7613363](https://pubmed.ncbi.nlm.nih.gov/7613363/)]

## Abbreviations

- AI:** artificial intelligence  
**ANOVA:** analysis of variance  
**CBT:** cognitive behavioral therapy  
**DHI:** digital health intervention  
**DSM:** Diagnostic Manual of Mental Disorders  
**GAD-7:** Generalized Anxiety Disorder 7-item Scale  
**iCBT:** internet-based cognitive behavioral therapy  
**IUS-12:** Intolerance of Uncertainty Scale  
**mHealth:** mobile health  
**ONS4:** Office of National Statistics Personal Well-being Domain  
**SHAI-18:** Short Health Anxiety Scale  
**WHO-5:** World Health Organization Five Well-being Index

*Edited by A Mavragani; submitted 10.03.22; peer-reviewed by SK Choi, P Shah, A Joseph, K Uludag, S El kefi; comments to author 07.08.22; revised version received 01.09.22; accepted 20.09.22; published 20.10.22.*

*Please cite as:*

*Goonesekera Y, Donkin L*

*A Cognitive Behavioral Therapy Chatbot (Otis) for Health Anxiety Management: Mixed Methods Pilot Study*

*JMIR Form Res* 2022;6(10):e37877

URL: <https://formative.jmir.org/2022/10/e37877>

doi: [10.2196/37877](https://doi.org/10.2196/37877)

PMID: [36150049](https://pubmed.ncbi.nlm.nih.gov/36150049/)

©Yenushka Goonesekera, Liesje Donkin. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 20.10.2022. 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

# Social Media Communication and Network Correlates of HIV Infection and Transmission Risks Among Black Sexual Minority Men: Cross-sectional Digital Epidemiology Study

Lindsay E Young<sup>1</sup>, PhD; Jack Lipei Tang<sup>1</sup>, MPhil; Yuanfeixue Nan<sup>1</sup>, MA

Annenberg School for Communication and Journalism, University of Southern California, Los Angeles, CA, United States

**Corresponding Author:**

Lindsay E Young, PhD

Annenberg School for Communication and Journalism

University of Southern California

3502 Watt Way

Los Angeles, CA, 90089

United States

Phone: 1 610 331 1930

Email: [lindsay.young@usc.edu](mailto:lindsay.young@usc.edu)

## Abstract

**Background:** In the United States, HIV disproportionately affects Black cisgender sexual minority men (BSMM). Although epidemiological and behavioral surveillance are integral to identifying BSMM at risk of HIV infection and transmission, overreliance on self-reported data, inability to observe social contexts, and neglect of populations with limited engagement in health care systems limits their effectiveness. Digital epidemiological approaches drawing on social media data offer an opportunity to overcome these limitations by passively observing in organic settings activities, beliefs, behaviors, and moods that indicate health risks but are otherwise challenging to capture.

**Objective:** The primary aim of this study was to determine whether features of Facebook communication and networks were associated with biological, behavioral, and psychological indicators of HIV infection and transmission risk.

**Methods:** Facebook and survey data were collected from BSMM aged 18 to 35 years living in Chicago (N=310). Participants' Facebook posts were characterized using 4 culturally tailored topic dictionaries related to aspects of HIV protection and risk among BSMM (sexual health; substance use; sex behavior; and ballroom culture, a salient subculture in lesbian, gay, bisexual, transgender, and queer communities of color). Social network methods were used to capture structural features of BSMM's Facebook friendships (centrality, brokerage, and local clustering) and Facebook group affiliations. Multivariable regressions revealed relationships between these Facebook features and 5 ground truth indicators of HIV infection and transmission risk (sexually transmitted infection incidence, condomless sex, sex drug use, biomedical prevention, and depression).

**Results:** Although analysis of participants' Facebook posts revealed that HIV-related topics occupied a small portion of the total messages posted by each participant, significant associations were found between the following HIV risk indicators and Facebook features: *Condomless sex*, including communication about sexual health (odds ratio [OR] 1.58, 95% CI 1.09-2.29), ballroom culture (OR 0.76, 95% CI 0.63-0.93), and friendship centrality (OR 0.69, 95% CI 0.52-0.92); *Sex drug use*, including communication about substance use (OR 1.81, 95% CI 1.17-2.79) and friendship centrality (OR 0.73, 95% CI 0.55-0.96) and brokerage (OR 0.71, 95% CI 0.51-0.99); *Biomedical prevention*, including communication about ballroom culture (OR 0.06, 95% CI 0.01-0.71); and *Depression*, including communication about sexual health ( $\beta=-0.72$ , 95% CI  $-1.42$  to  $-0.02$ ), ballroom culture ( $\beta=.80$ , 95% CI 0.27-1.34), friendship centrality ( $\beta=-0.90$ , 95% CI  $-1.60$  to  $-0.21$ ), and Facebook group affiliations ( $\beta=.84$ , 95% CI 0.25-1.43). Facebook features provided no significant explanatory value for sexually transmitted infection incidence.

**Conclusions:** Finding innovative strategies to detect BSMM at risk of contracting or transmitting HIV is critical to eliminating HIV disparities in this community. The findings suggest that social media data enable passive observance of social and communicative contexts that would otherwise go undetected using traditional HIV surveillance methods. As such, social media data are promising complements to more traditional data sources.

(JMIR Form Res 2022;6(10):e37982) doi:[10.2196/37982](https://doi.org/10.2196/37982)

**KEYWORDS**

HIV surveillance; HIV prevention; digital epidemiology; social media; social networks; sexual minority men; men who have sex with men

## Introduction

### Background

Despite significant strides in preventing and treating HIV in the United States, the burden of HIV remains disproportionately concentrated among cisgender gay, bisexual, same gender-loving, and other sexual minority men (SMM), of which Black or African American SMM (BSMM) demonstrate particular vulnerability. Of the estimated 34,800 new HIV infections in 2019, a total of 66% (23,100) were among SMM, with BSMM accounting for the plurality (37%) of those cases [1] despite comprising only approximately 9% of the SMM population [2]. With greater prevalence of HIV among BSMM comes greater opportunity for its transmission within BSMM networks as BSMM tend to select sexual partners of the same race [3]. Opportunities for transmission are further compounded by the fact that many BSMM are unaware of their HIV infection [4], in part because of underestimations of personal risk and limited access to high-quality HIV prevention services [5]. People who do not know that they have HIV do not receive timely medical care and unknowingly infect others.

At the patient level, the risks of both HIV infection and transmission are bound in a web of intersecting and mutually exacerbating biological, behavioral, and psychosocial factors. Biologically, HIV risk cannot be discussed in isolation from other sexually transmitted infections (STIs), which are known to increase susceptibility to and transmission of HIV given their shared pathways of transmission [6]. Behaviorally, engagement in unprotected anal intercourse (hereafter, condomless sex) and other sex behaviors that can compromise condom use decisions, for example, use of drugs during sex (hereafter, sex drug use), account for most new infections and transmission instances [7]. Alternatively, protective biomedical behaviors such as adherence to daily pre-exposure prophylaxis (PrEP) for HIV-negative individuals (ie, PrEP as prevention) and antiretroviral medications for people living with HIV (ie, treatment as prevention) are viewed as highly efficacious HIV prevention practices [8]. However, linkage and retention in both regimens of HIV prevention are lower among BSMM relative to their White counterparts, in large part because of systemic vulnerabilities (eg, poverty, housing instability, medical mistrust, and structural racism) that create barriers to medication adherence [9-12].

Finally, adverse mental health conditions such as depression and low self-esteem are also correlates of HIV risk through their associations with high-risk behaviors such as condomless sex, substance use, and multiple concurrent partnerships [13-15]. Depression may increase the likelihood of engaging in these behaviors to mitigate negative feelings brought on by HIV stigma, homophobia, and discrimination [16] and may compromise motivations to change these behaviors [14]. Moreover, depression prevalence among people living with HIV can compromise their ability to remain adherent to HIV

medications and, therefore, achieve sustained viral suppression [17,18].

### Traditional HIV Risk Surveillance and Its Limitations

Key to efforts to attenuate the HIV epidemic among SMM of all races and ethnicities is the ability to actually observe the aforementioned indicators of HIV risk in individuals before moments of infection and transmission such that prevention can actually occur [19,20]. Traditionally, researchers and public health professionals have relied on surveillance data that suffer from considerable lag time in reporting as they are often captured through large-scale behavioral surveys or clinical case visits, which require extensive resources for collection and analysis [21,22]. Perhaps more importantly, with their reliance on surveys and clinical visits, traditional surveillance strategies inevitably neglect populations that have limited engagement with the health care system because of issues such as low perceptions of personal risk, experiences with stigma and discrimination, and socioeconomic disadvantage. With studies suggesting limited access to health care services for BSMM, including HIV care [23], data predicated on an individual's use of health care services may not represent those at greatest risk of contracting or unknowingly transmitting HIV or those who are most vulnerable to falling out of care once diagnosed with HIV. Finally, with their emphasis on individual risk behaviors, survey- and interview-based surveillance methods are not well suited to observe the social contexts that nurture behavioral risk, such as social networks [24], social and behavioral norms [25], and peer-group dynamics [26].

### Using Social Media to Surveil Social and Communicative Contexts Associated With HIV Risks

More recently, the mainstreaming of social media and other digital communication technologies has given way to the emerging field of *digital epidemiology* [22,27]. Digital epidemiology rests on the idea that digital communications such as social media posts are made in organic environments and that what people post about reflects activities, behaviors, and dispositions that are relevant to various health outcomes. Demonstration studies have put these assumptions to the test and showed that what people talk about on social media platforms can be used to predict health outcomes such as depression [28], sexual risk [29], and HIV and STI *hotspots* [30].

Moreover, digital epidemiology is poised to take advantage of the networked (or relational) infrastructure that undergirds social media to detect features of interpersonal networks through which HIV risks are conferred [31,32]. For example, previous research focusing on offline networks has shown that the cohesion of an individual's network [33,34], an individual's position vis-à-vis other network members [24], and their exposure to other risky individuals [35-38] affect disease spread and risk susceptibility. In addition, looser affiliations formed around broader groups or organizational structures have also been shown to nurture and confer HIV risks [31], for example, physical-world social

venues (eg, bars, clubs, and bathhouses) [39-44] and their digital analogs (eg, dating apps and lesbian, gay, bisexual, transgender, and queer social networking groups) [45,46].

## Objectives

Given the need to engage more high-risk BSMM in the HIV prevention and treatment continuum of care and known limitations of traditional approaches to HIV risk surveillance, it is essential to understand whether alternative, more organic sources of social and behavioral information can be used to improve surveillance of HIV risk among BSMM. To this end, we took an initial step to evaluate the feasibility of using digital traces of BSMM's Facebook communication and network connections as an additional source of informal health information typically not reported to medical officials or health departments [47]. Specifically, we sought to ascertain (1) the extent to which BSMM posted on Facebook about topics known to be associated with HIV prevention and risk in their community and (2) the extent to which those topics and the features of their Facebook friendship connections were associated with biological, behavioral, and psychological indicators of HIV infection and transmission risk.

## Methods

### Participants

This study draws on parent study data collected from 2016 to 2018 from participants enrolled in a social network PrEP for prevention intervention for BSMM living in Chicago, Illinois, United States. As described elsewhere [48,49], participants were eligible to take part in the parent study if they (1) were aged 18 to 35 years; (2) identified as Black or African American; (3) were assigned male at birth; (4) had had sex with a man in the past 12 months; and, because of our interest in social media use, (5) had an active Facebook profile. Participants were recruited using a probability-based variant of snowball sampling called respondent-driven sampling (RDS) [50]. RDS referral chains began with an initial set of "seeds" meeting study eligibility who were then enlisted to recruit up to 6 of their peers ("sprouts") who also met study eligibility. The process continued until the recruitment target was reached (N=423).

The analyses featured in this study draw on data collected at the parent study's 12-month assessment as this was the only assessment that included questions about sexual risk behavior. In total, 82% (347/423) of the parent study participants were retained at 12 months, and only those who had complete data on all variables of interest (310/347, 89.3%) were included in our analyses. Bivariate analyses (not shown) comparing participants included in the analytic sample (310/347, 89.3%) with filtered cases (37/347, 10.7%) showed no significant differences in age, education, sexual identity, or HIV status.

### Data Sources and Measures

#### Overview

All study measures were derived from 3 sources of data obtained with consent from participants: (1) a computer-assisted self-administered survey, which included modules on HIV prevention engagement, sex behaviors, substance use, and

demographics; (2) biomedical testing, which determined participants' HIV status and viral load if they were HIV positive; and (3) a manual download of participants' Facebook friend lists, group memberships, and timeline data, which enabled the construction of their web-based friendship and group affiliation networks and analysis of their public communication content, respectively.

#### Outcomes

The outcomes included 5 indicators of HIV infection and transmission risk that spanned biological, behavioral, and psychosocial domains. As a biological indicator of risk, we included a measure of recent *STI incidence*, which was operationalized dichotomously as whether a participant self-reported having been diagnosed with at least one STI other than HIV (eg, syphilis, chlamydia, or human papillomavirus) in the past 12 months. Behavioral indicators included measures of condomless sex, sex drug use, and biomedical prevention engagement. Engagement in *condomless sex* (ie, sex without a condom) and *sex drug use* (ie, using drugs or alcohol to enhance the sexual experience or make sex easier) were defined dichotomously as inconsistent condom use (ie, not always using a condom) and having ever engaged in sex drug use in the past 6 months. *Biomedical prevention engagement* was operationalized as either being on PrEP to prevent contracting HIV or achieving viral suppression to prevent transmitting HIV [6]. As such, the measure of biomedical prevention engagement is intentionally status neutral, and not being engaged in this form of prevention is taken to be the indication of risk. *Depression* was included as a psychosocial indicator of risk and was measured using the Revised Center for Epidemiologic Studies Depression Scale [51], which consists of 10 items (Cronbach  $\alpha=.81$ ), each investigating the frequency of a depressive feeling within the last week (0="rarely or none of the time" to 3="all of the time"). Item scores were aggregated to create a composite depression score.

#### Facebook Communication Features

To capture the degree to which participants posted about topics that have been linked to HIV protections and vulnerabilities, we constructed a series of topical dictionaries to enable an automated content analysis of their timeline posts. To develop each dictionary, an iterative, mixed methods approach was used, described in greater detail in [Multimedia Appendix 1](#) [52-54], that drew on extant literature, the expertise of BSMM themselves, internet sources, and endogenous patterns in the data themselves. Topics featured in this study include *sexual health*; *substance use*; *sex behavior*; and *house/ballroom culture*, a subaltern system of chosen family (also known as *Houses*) and identity-affirming competitions (also known as *Balls*) that feature prominently in the Black gay community and that are thought to provide sources of support that help buffer BSMM members from HIV vulnerabilities such as financial hardship, housing instability, and stigma and discrimination [52-54]. Example terms and posts associated with each topic dictionary are shown in [Table 1](#). Using the flagged posts, we then created 4 participant-level topic variables representing the number of posts a participant made with a given topic orientation. In addition, we accounted for the *positive affect* of posts, which

has been linked to HIV-adjacent outcomes such as depression [28] and substance use [55]. Positive affect was measured using the Linguistic Inquiry and Word Count (Pennebaker Conglomerates, Inc) program [56].

**Table 1.** Topic dictionaries.

Topics	Example keywords	Sample post	Posts, n (%)
Sexual health	<i>HIV, AIDS, condom, PrEP, PEP, sexually transmitted, STD, get tested, and poz</i>	“condoms have become my best friend kinda sorta”	420 (0.18)
Sex behavior	<i>Sex, bareback, masturbate, blow job, give head, cock, horny, and thirst trap</i>	“I’m bored horny with nothing and nobody to do lmaooooo this isn’t happening to me right now lol”	3452 (1.5)
Substance use	<i>Drugs, meth, poppers, weed, marijuana, liquor, alcohol, boozie, turnt, and getting loud</i>	“i went out last night turnt up fell asleep woke up still turnt”	3754 (1.63)
Ballroom culture	House names: <i>Balenciaga, Mizrahi, and Cartier</i> ; status and roles: <i>legends, icons, and house mother</i> ; competition categories: <i>realness, sex siren, and vogue</i>	“rules of twist you must walk your realness category in order to get respected as a twister if you don’t walk realness please don’t think im gone let you sit next to me in the twist line you cn vogue wit the vogue fems”	4510 (1.96)

### Facebook Network Features

In line with previous work showing associations between how SMM are positioned vis-à-vis one another in their social networks and their prevention and risk engagement [24,57], we also accounted for 4 measures of an individual’s web-based structural embeddedness in a Facebook friendship network among the other BSMM in the study. *Eigenvector centrality* measures the degree to which an individual is connected with other well-connected network members and, hence, their relative closeness to others in the network [58,59]. *Network brokerage* represents the degree to which an individual connects disparate subcommunities within the friendship network and was measured using the brokerage measure by Everett and Valente [60]. A measure of *local clustering* [61] was also included to represent the degree to which actors in a network create social cliques. Finally, *group network size* accounts for an individual’s group-mediated embeddedness and was operationalized as the number of Facebook groups to which an individual belonged.

### Controls

The models were adjusted for several control measures, including (1) a dichotomous measure of HIV status (1=HIV positive; 0=HIV negative) defined based on blood tests for those who consented to blood tests or self-reports for those who opted out, (2) a measure of the total number of posts made in the past 12 months to control for the varying volume of posts across participants, and (3) the other 4 indicators of HIV infection and transmission risk.

### Statistical Analysis

Descriptive statistics and multivariable logistic (for dichotomous outcomes) and linear (for numeric outcomes) regression models were estimated using RStudio [62] (version 1.4.1717; R Foundation for Statistical Computing). Each outcome was regressed on the same set of factors, including the Facebook communication and network features and the other HIV prevention and risk outcomes. In addition, all numeric (interval or ratio) covariates were standardized for ease of interpretation. The effects in the logistic regression models are reported as odds ratios (ORs), whereas the coefficients are reported in the linear regression model of depression. RDS sampling weights

were not included in the regressions as this would amount to assuming heteroskedasticity, where respondents with high weights would be assumed to provide the most accurate information. There is no reason for such an assumption.

### Ethical Considerations

All study procedures were approved by the Institutional Review Board of the University of Chicago (IRB15-1250). Informed consent was obtained from all the participants. For all nonparticipant Facebook friends of study participants, a waiver of consent was obtained from the Institutional Review Board given the minimal risk to these individuals. The parent randomized trial is registered at ClinicalTrials.gov (identifier NCT02896699).

## Results

### Descriptives

Participants in the analytic sample were, on average, aged 25.8 (SD 4.21) years. Most had earned a high school-level education (193/310, 62.3%) and identified as gay (188/310, 60.6%), whereas a quarter identified as bisexual (80/310, 25.8%). Summary statistics of key variables in our analysis are provided in Table 2. Of the 310 BSMM in the analytic sample, 147 (47.4%) were living with HIV, 69 (22.3%) reported an STI diagnosis in the past 12 months, 103 (33.2%) were either on PrEP or were virally suppressed (ie, biomedical prevention engagement), 193 (62.3%) reported condomless sex, and 149 (48.1%) reported sex drug use. A Revised Center for Epidemiologic Studies Depression Scale score of 10 is considered depressed; on average, participants were slightly below that threshold (mean 9.30, SD 5.94; range 0-28).

Analysis of participants’ Facebook posts revealed that HIV-related topics accounted for a small portion of the total messages posted by participants (mean 730.63, SD 836.79). Overall, participants posted a median of 4 messages (IQR 1-14) about substance use, a median of 3 messages (IQR 0-12) about sex behavior, a median of 2 messages (IQR 0-11) about ballroom culture, and a median of 0 messages (IQR 0-2) about sexual health (Table 2).



**Table 2.** Summary statistics of HIV infection and transmission risk indicators and Facebook communication and network features of Black sexual minority men (N=310).

Characteristic	Values
<b>HIV infection and transmission risk indicators</b>	
STI <sup>a</sup> incidence, n (%)	69 (22.3)
Condomless sex, n (%)	193 (62.3)
Sex drug use, n (%)	149 (48.1)
Biomedical prevention engagement, n (%)	103 (33.2)
Depression score, mean (SD)	9.30 (5.94)
<b>Facebook communication features</b>	
Sexual health posts, median (IQR)	0 (0-2)
Substance use posts, median (IQR)	4 (1-14)
Sex behavior posts, median (IQR)	3 (0-12)
Ballroom culture posts, median (IQR)	2 (0-11)
Positive affect, mean (SD)	4.84 (2.42)
<b>Facebook network features</b>	
Eigenvector centrality, mean (SD)	0.17 (0.16)
Brokerage, mean (SD)	28.67 (36.41)
Local clustering, mean (SD)	0.16 (0.10)
Facebook group count, mean (SD)	58.24 (64.98)
HIV-positive status, n (%)	147 (47.4)
Total number of posts, mean (SD)	730.63 (836.79)

<sup>a</sup>STI: sexually transmitted infection.

### Associations Between Facebook Features and HIV Infection and Transmission Risks

The multivariable regressions of the HIV infection and transmission risk indicators are shown in Table 3. First, the logistic regression of STI incidence revealed that, once the positive and significant effects of engagement in biomedical prevention and condomless sex were accounted for, Facebook network and communication features offered no additional explanatory power.

Regarding condomless sex, individuals who posted about sexual health had greater odds of reporting condomless sex in the past 6 months (OR 1.58, 95% CI 1.09-2.29), whereas posting about aspects of ballroom culture decreased those odds (OR 0.76, 95% CI 0.63-0.93). Facebook network features were also revealing. Individuals who were connected with other well-connected BSMM (ie, eigenvector centrality) had decreased odds of engaging in condomless sex (OR 0.69, 95% CI 0.52-0.92). A negative trend was also observed for network brokerage (ie, connecting disparate subcommunities of BSMM on Facebook), although the relationship was not statistically significant ( $P=.07$ ). Among the other HIV risk indicators, only STI incidence was positively associated with condomless sex (Table 3).

The model of sex drug use revealed several pronounced effects of Facebook network and communication features. Among the communication topics, individuals who posted about substances

were more likely to report using drugs during sex (OR 1.81, 95% CI 1.17-2.79), whereas adopting a more positive affective tone in posts decreased those odds (OR 0.73, 95% CI 0.55-0.96). With respect to network features, individuals who were Facebook friends with other well-connected BSMM (ie, eigenvector centrality; OR 0.73, 95% CI 0.55-0.96) and those who bridged disparate subcommunities of BSMM on Facebook (ie, brokerage; OR 0.71, 95% CI 0.51-0.99) were less likely to report sex drug use. Among the other HIV risk indicators, depression was positively associated with sex drug use (Table 3).

With respect to biomedical prevention engagement, after controlling for the positive and significant effects of HIV status and STI incidence, the results showed that individuals who posted more often about aspects of ballroom culture were less likely to be on PrEP or achieve viral suppression through antiretroviral therapy adherence (OR 0.06, 95% CI 0.01-0.71). Meanwhile, individuals who posted about sexual health were more likely to adhere to biomedical prevention ( $P=.05$ ), although this trend did not meet the criterion of significance.

Finally, regarding depression, individuals who posted about sexual health tended to be less depressed ( $\beta=-0.72$ , 95% CI -1.42 to -0.02), whereas those who posted about ballroom culture tended to be more depressed ( $\beta=.74$ , 95% CI 0.11-0.37). Furthermore, posting about substance use ( $P=.07$ ) and sex behavior ( $P=.06$ ) showed negative and positive trends, respectively, albeit nonsignificant ones. Of the Facebook

network features, being connected with other well-connected BSMM on Facebook (ie, eigenvector centrality;  $\beta=-0.90$ , 95% CI  $-1.60$  to  $-0.21$ ) and being affiliated with more Facebook groups ( $\beta=.84$ , 95% CI  $0.25$ - $1.43$ ) were negatively and positively associated with depression, respectively. Sex drug use was also positively associated with depression (Table 3).

**Table 3.** Multivariable logistic and linear regression models to assess associations between Facebook communication and network features and HIV infection and transmission risk indicators (N=310).

Independent variables	STI <sup>a</sup> incidence <sup>b</sup> , OR <sup>c</sup> (95% CI)	Condomless sex <sup>b</sup> , OR (95% CI)	Sex drug use <sup>b</sup> , OR (95% CI)	Biomedical prevention <sup>b</sup> , OR (95% CI)	Depression <sup>d</sup> , $\beta$ (95% CI)
<b>Facebook communication features</b>					
Sexual health content (std <sup>e</sup> )	0.99 (0.73 to 1.33)	1.58 (1.09 to 2.29) <sup>f</sup>	1.22 (0.89 to 1.67)	1.48 (1.00 to 2.20) <sup>g</sup>	-0.72 (-1.42 to -0.02) <sup>f</sup>
Substance use content (std)	1.06 (0.72 to 1.56)	1.13 (0.81 to 1.58)	1.81 (1.17 to 2.79) <sup>h</sup>	0.89 (0.58 to 1.36)	-0.80 (-1.65 to 0.05) <sup>g</sup>
Sex behavior content (std)	0.88 (0.66 to 1.18)	1.02 (0.72 to 1.43)	1.02 (0.73 to 1.43)	1.20 (0.82 to 1.77)	.63 (-0.03 to 1.04) <sup>g</sup>
Ballroom culture content (std)	1.24 (0.92 to 1.66)	0.76 (0.63 to 0.93) <sup>h</sup>	0.94 (0.76 to 1.18)	0.06 (0.01 to 0.71) <sup>f</sup>	.80 (0.27 to 1.34) <sup>h</sup>
Positive affect (std)	0.81 (0.56 to 1.15)	1.08 (0.83 to 1.42)	0.73 (0.55 to 0.96) <sup>f</sup>	0.85 (0.66 to 1.08)	.18 (-0.45 to 0.81)
<b>Facebook network features</b>					
Eigenvector centrality (std)	1.22 (0.88 to 1.70)	0.69 (0.52 to 0.92) <sup>f</sup>	0.73 (0.55 to 0.96) <sup>f</sup>	0.97 (0.69 to 1.36)	-0.90 (-1.60 to -0.21) <sup>f</sup>
Brokerage (std)	0.97 (0.61 to 1.54)	0.74 (0.53 to 1.03) <sup>g</sup>	0.71 (0.51 to 0.99) <sup>f</sup>	0.78 (0.44 to 1.37)	-0.44 (-1.29 to 0.42)
Local clustering coefficient (std)	0.96 (0.61 to 1.51)	0.78 (0.53 to 1.15)	0.89 (0.60 to 1.30)	1.13 (0.76 to 1.66)	.01 (-1.03 to 1.04)
Facebook group count (std)	0.82 (0.60 to 1.13)	1.00 (0.78 to 1.30)	1.01 (0.79 to 1.29)	0.85 (0.66 to 1.10)	.84 (0.25 to 1.43) <sup>h</sup>
<b>Controls</b>					
HIV status (positive)	1.84 (0.97 to 3.47) <sup>g</sup>	0.90 (0.53 to 1.52)	0.85 (0.51 to 1.40)	3.20 (1.85 to 5.53) <sup>i</sup>	.30 (-1.02 to 1.62)
Total number of posts (std)	1.03 (0.62 to 1.72)	0.75 (0.50 to 1.11)	0.64 (0.40 to 1.03) <sup>g</sup>	1.12 (0.68 to 1.84)	.65 (-0.60 to 1.90)
STI incidence	— <sup>j</sup>	2.66 (1.40 to 5.04) <sup>h</sup>	1.00 (0.54 to 1.87)	2.65 (1.43 to 4.89) <sup>h</sup>	.27 (-1.34 to 1.87)
Condomless sex	2.71 (1.44 to 5.55) <sup>h</sup>	—	1.19 (0.70 to 2.01)	0.90 (0.52 to 1.57)	-0.12 (-1.57 to 1.32)
Sex drug use	1.00 (0.51 to 1.72)	1.18 (0.70 to 2.00)	—	1.02 (0.59 to 1.76)	2.69 (1.27 to 4.11) <sup>i</sup>
Biomedical prevention	2.59 (1.42 to 4.73) <sup>h</sup>	0.85 (0.49 to 1.48)	0.92 (0.53 to 1.58)	—	-0.34 (-1.70 to 1.02)
Depression (std)	1.05 (0.77 to 1.44)	0.98 (0.76 to 1.27)	1.64 (1.27 to 2.11) <sup>i</sup>	0.90 (0.69 to 1.17)	—
Intercept	0.07 (0.03 to 0.15) <sup>i</sup>	1.46 (0.91 to 2.33)	0.96 (0.59 to 1.57)	0.17 (0.09 to 0.35) <sup>i</sup>	8.05 (6.58 to 9.52) <sup>i</sup>

<sup>a</sup>STI: sexually transmitted infection.

<sup>b</sup>Logistic regression model.

<sup>c</sup>OR: odds ratio.

<sup>d</sup>Linear regression model.

<sup>e</sup>Std: SD unit change.

<sup>f</sup> $P<.05$ .

<sup>g</sup> $P<.10$ .

<sup>h</sup> $P<.01$ .

<sup>i</sup> $P<.001$ .

<sup>j</sup>Outcome variable; excluded as a covariate.

## HIV Status Subgroup Analysis of Engagement in Biomedical Prevention

In the primary analysis, we prioritized a status-neutral measure of biomedical prevention engagement in light of evidence that, similar to PrEP, viral suppression through antiretroviral therapy adherence is itself an effective form of biomedical HIV prevention (ie, treatment as prevention) [63,64]. Despite this and calls to prioritize a status-neutral continuum of care, we acknowledge that the circumstances that enable medication initiation and adherence may be different for people living with HIV than for people at risk of HIV.

For this reason, we performed a stratified subgroup analysis to determine whether Facebook features were differently associated with each subgroup's likelihood of being engaged in biomedical forms of HIV health care (Multimedia Appendix 2). As shown, the negative association between posting about ballroom culture and biomedical engagement that was observed in the primary analysis was only observed among BSMM living with HIV in the stratified analysis. As such, BSMM living with HIV who showed signs of identifying with the ballroom community were

less likely to be virally suppressed. In addition, the positive association between STI incidence and biomedical engagement that we observed in the unstratified model was shown to be significant only among HIV-negative BSMM.

## Improvements in Model Fit

We performed likelihood ratio tests for all binary outcomes and hierarchical regression for the continuous outcome of depression to determine whether the addition of Facebook features to models with the HIV risk indicators only improved model fit (Table 4). Variables that met a  $P < .10$  Cronbach  $\alpha$  criterion in the primary analysis (Table 3) were included in this portion of the analysis. For each outcome, eligible HIV risk indicators were included in model 1, with eligible Facebook features added to that in model 2. As shown, with the exception of STI incidence (not shown), for which Facebook features provided no additional explanatory power in the primary analysis, the addition of Facebook features significantly improved model fit over models with HIV risk indicators alone. For condomless sex, biomedical prevention engagement, and depression, model improvements were significant at the  $P < .01$  level. For sex drug use, model improvement was significant at the  $P < .001$  level.

**Table 4.** A comparison of model fits: models with indicators of HIV infection and transmission only (model 1) versus models with Facebook features added (model 2).

Dependent variable and model	LR <sup>a</sup> chi-square ( <i>df</i> )	<i>F</i> test ( <i>df</i> )	Model 1 vs Model 2 LR chi-square ( <i>df</i> )	Model 1 vs Model 2 <i>F</i> change ( <i>df</i> )	<i>P</i> value
<b>Condomless sex</b>		N/A <sup>c</sup>	18.2 (4) <sup>d</sup>		.001
Model 1 <sup>b</sup>	6.8 (1)				
Model 2 <sup>e</sup>	25.0 (5)				
<b>Sex drug use</b>		N/A	21.4 (5) <sup>d</sup>		<.001
Model 1	16.5 (1)				
Model 2	37.9 (6)				
<b>Biomedical prevention</b>		N/A	11.5 (2) <sup>d</sup>		.003
Model 1	31.6 (2)				
Model 2	43.2 (4)				
<b>Depression<sup>f</sup></b>	N/A		N/A	3.14 (6, 302) <sup>d</sup>	.005
Model 1		16.88 (1, 308)			
Model 2		5.21 (7, 302)			

<sup>a</sup>LR: likelihood ratio.

<sup>b</sup>For each outcome, model 1 includes all HIV infection and transmission risk indicators that met a Cronbach  $\alpha = .10$  criterion in previous models.

<sup>c</sup>N/A: not applicable.

<sup>d</sup>The statistical significance of the comparison of the two models.

<sup>e</sup>For each outcome, model 2 adds to model 1 the Facebook network and communication variables that met a Cronbach  $\alpha = .10$  criterion in previous models.

<sup>f</sup>Depression is a numeric measure, so an *F* test and *F* change are reported instead of LR chi-square tests.

## Discussion

### Principal Findings

In this study, we argue that digital epidemiology—the practice of drawing on digital traces of web-based communication and social interaction to detect individuals with certain health

risks—offers a critical step forward in closing the gap on HIV disparities as it may yield informal health information not otherwise found in more formal medical records and offers a way to detect at-risk individuals whose medical records are sparse in the first place. To test these assumptions, we demonstrated the potential of using Facebook communication and network data collected from a cohort of young BSMM, a

population experiencing a disproportionate burden of new HIV incidences in the United States, to identify individuals who reveal key biological, behavioral, and psychological indicators of HIV infection and transmission risk. The fact that we were able to detect significant relationships between Facebook communication and network features and multiple HIV risk indicators while also adjusting for correlates of those outcomes that were directly observed using surveys and blood tests suggests that the effects of these social media variables are relatively robust.

Our demonstration yielded several insights about contingencies of the digital epidemiology approach in general and, more specifically, the relationship between various social media features and HIV risk indicators in this particular cohort of BSMM. Regarding the contingencies of the approach, although we intentionally examined 5 risk indicators that represented different aspects of HIV risk (ie, biological, behavioral, and psychosocial), our results clearly showed that Facebook communication and network features were not universally predictive of these outcomes. For example, features of network embeddedness, specifically attenuated social integration with other BSMM, were strong indicators of sex behaviors that can place BSMM at risk of HIV (ie, condomless sex and sex drug use) and mental health risks such as depression. The same was not true for biological risks such as STI incidence or for engagement in biomedical forms of HIV care, for which the more directly observed epidemiological and behavioral variables were far better predictors.

Regarding our statistical findings, several relationships stand out that can bring additional focus to these contingencies. A first set of noteworthy findings pertains to the effects of being embedded in web-based friendships with other BSMM. We learned that web-based friendships with other BSMM, both in the form of being connected with well-connected others and being a network bridge, were associated with decreased likelihoods of engaging in sex drug use and condomless sex. This suggests that those who engage in these sex behaviors may be at risk of becoming disenfranchised from the larger BSMM community, a finding we have found support for elsewhere in work currently under review. This raises concern as the isolation of already at-risk behavioral subgroups may lead to the adoption of additional compounding risk behaviors [65,66]. Therefore, public health outreach to these behavioral communities may need to incorporate ways to integrate them into the larger social fabric of BSMM, which will increase their access to social support from other members of the BSMM community and expose them to alternative behavioral norms and a more diverse range of behavioral choices. Social isolation among BSMM who engage in sex behaviors that place them at risk of HIV infection and transmission is also concerning given its obvious implications for facilitating or exacerbating depression [67], which itself had an unambiguous direct effect on the likelihood of using sex drugs.

The impact of social embeddedness also surfaced in the model of depression. Confirming our previous assertions about the relationship between social isolation (vis-à-vis other BSMM) and depression, we found that the BSMM in our sample who had fewer Facebook friendships with other well-connected

BSMM (ie, had lower eigenvector centrality) tended to score higher on depression. Furthermore, and somewhat surprisingly, we also found that BSMM who belonged to more Facebook groups also tended to be more depressed. We surmise that this relationship is either the result of compensation behavior, whereby BSMM who are more depressed tend to seek social connection through group settings to help them reduce their depression—perhaps to compensate for their attenuated integration with other BSMM via friendships—or the result of negative interactions that occur in Facebook groups that may exacerbate negative mindsets. That our analysis was only cross-sectional and not longitudinal means that the directionality of this and all other relationships discussed cannot be adequately ascertained.

Second, there were significant relationships between posting about substance use and using drugs to enhance the sexual experience (ie, sex drug use) and a trending association between posting about sexual health and being engaged in sexual health care (ie, biomedical prevention). This supports the general intuition that most people have about communication on general-purpose platforms such as Facebook—that there tends to be a strong correlation between the behaviors that one talks about in digital spaces and the behaviors that one engages in real-life spaces. In many ways, the naturalistic settings in which posts on social media platforms are made (eg, at home, in school, and socializing with friends) mean that what one communicates is likely to reflect behaviors that are relevant to their routine activities and thinking [28]. That said, observing HIV risk behaviors and activities can be resource-intensive and can suffer from serious lags relative to when the moment of risk occurred when retrospective self-reports are used. Thus, knowing that digitally archived social media communication could be leveraged to detect at-risk individuals in place of behavioral observations or self-reports is an exciting public health advancement that could make detecting individuals at risk of HIV incidence or transmission and other comorbidities easier and more responsive to real-time moments of risk.

Finally, our results also suggest that communication on social media can be used to detect BSMM subcommunities such as the ballroom community, which tends to attract individuals who are already structurally vulnerable because of limited access to medical care, socioeconomic disadvantages, housing instability, experiences with rejection from biological families, and stigma and discrimination [54]. These factors may then induce additional HIV risk behaviors such as sex drug use [68] and survival sex [52], which can further increase their HIV risk. Our analysis confirmed this in 2 ways. Posting about aspects of ballroom culture was negatively associated with the likelihood of being on PrEP if HIV negative or virally suppressed if living with HIV and positively associated with heightened depression. Thus, although the ballroom community undoubtedly provides its members with kinship and a sense of sexual and gender affirmation, those drawn to it may be more likely to demonstrate pre-existing hardships and vulnerabilities that continue to affect the degree of HIV risk and prevention that they engage in or are exposed to within the ballroom community. As such, social media communication does more than reflect behaviors such as substance use. Its ability to detect members of sociocultural

milieus with other known HIV vulnerabilities is also a critical advancement in HIV risk-reduction strategies.

### Clinical Implications

Our results suggest that social media data could be leveraged to improve surveillance and modeling of certain types of HIV infection and transmission risk among BSMM while also bringing our attention to new digital “fingerprints” that could serve as early warning signs of an individual’s HIV risk potential. Although the utility of digital epidemiological approaches is often discussed in the context of developing automated public health tracking systems at scale with large amounts of publicly available social media data, we argue for a more person-focused application of social media–assisted surveillance that can lead to more personalized forms of intervention and care. For instance, with knowledge of important digital fingerprints of potential risk, social media archives from consenting clients could help frontline health and social welfare staff (eg, HIV counselors, community health workers, and case managers) profile their clients in terms of their risk of future HIV-related outcomes, thereby serving as a barometer to guide their decisions on screening, treatment, and other service recommendations. Furthermore, having the ability to monitor extreme changes in a client’s communication and relational dynamics (eg, changes in mood or suddenly dissolved friendships), which have been linked in previous work to sexual risk engagement [69], can alert frontline staff to potential crises that warrant impromptu “soft touch” interventions in the form of supportive check-ins. As such, we see social media’s epidemiological promise in its ability to enable more responsive care and the ability to intervene in risky social contexts as they unfold in near real time.

### Limitations

This study represents an important first step in determining the efficacy of using digital epidemiological approaches to identify individuals at risk of HIV incidence or transmission. However, it is indeed just that, a first step. As such, its current limitations correspond to obvious next steps. Foremost of those limitations is that the data we used in this study are cross-sectional, which limits us to making correlational as opposed to causal inferences. A next step in our research agenda is to build predictive models that draw on all 3 waves of data from all 3 sources (ie, Facebook, surveys, and laboratory tests) using logistic and machine learning approaches.

Second, although the featured analyses account for the effects of both self-reported and passively observed social media indicators of HIV infection and transmission risk potential, it remains to be seen whether social media indicators on their own would be more helpful than self-reported health and behavioral

data. Our intuition suggests that this would not be the case. Our results showed that passively observed social media indicators of an individual’s risk potential improved the predictive performance of models that included self-reported data alone, but we remain dubious about the prospects of using social media data as replacements for self-reported data. Rather, they are more appropriately seen as complements of one another.

Third, dictionary-based approaches to digital content analysis are limiting in that they only identify topically relevant keywords (as opposed to deeper meanings) and necessarily rely on the researchers’ ability to build a robust dictionary. Furthermore, as themes are defined by the researcher before performing the content analysis, more emergent themes and topics that may be related to HIV risk go unidentified and unexplored. Latent space topic modeling such as latent Dirichlet allocation would be an appropriate modeling approach to those ends. However, we are cautious about expressing too much enthusiasm for this approach, particularly when applied to Facebook content. Unlike Twitter, which often engenders a more intentional and focused tone, Facebook posts tend to be more diffuse and stream of consciousness, which we surmise will make detecting meaningful latent topics more challenging in these models.

### Conclusions

To end the HIV epidemic in the United States, ambitious goals have been set to reduce the number of incident infections by 90% by 2030, with prioritized intervention among BSMM [8]. Meeting this goal demands an innovative and multipronged strategy to identify individuals who are at high risk of HIV infection and transmission and engage them in prevention or treatment care continuums. Although well-funded epidemiological and behavioral surveillance programs are and should remain the primary engine of this work, they do not come without limitations. This study established that social media offers a complementary informal source of health information that can be used to sharpen our ability to detect individuals at risk of HIV and reach people who may otherwise be missed by surveillance that privileges those who engage more regularly with the mainstream health care system. Indeed, our analysis showed that social media communication and network features are correlates of several indicators of HIV infection and transmission risk among BSMM in our sample, although not uniformly. Moreover, the inclusion of social media variables seemed to capture protective and risky features of BSMM’s social lives that were not being captured in the self-reported data. Further research is needed to verify the acceptability and feasibility of incorporating social media data collection into established surveillance and prevention and treatment practice and identify ways to leverage insights from those efforts into near–real-time interventions.

### Acknowledgments

This work was supported by the National Institute of Child Health and Human Development (R00HD094648). Parent study data collection was supported by the National Institute of Allergy and Infectious Diseases (R01AI20700). The contents of this publication are solely the responsibility of the authors and do not necessarily represent the official views of the National Institutes of Health or the Department of Health and Human Services.

## Conflicts of Interest

None declared.

### Multimedia Appendix 1

Construction of topic dictionaries.

[[DOCX File , 87 KB - formative\\_v6i10e37982\\_app1.docx](#) ]

### Multimedia Appendix 2

Multivariable logistic regression models stratified by HIV status.

[[DOCX File , 18 KB - formative\\_v6i10e37982\\_app2.docx](#) ]

## References

- Centers for Disease Control and Prevention. Diagnoses of HIV infection in the United States and dependent areas, 2018. Centers for Disease Control and Prevention. 2018. URL: <https://www.cabdirect.org/globalhealth/abstract/20203225669> [accessed 2021-10-22]
- Lieb S, Fallon SJ, Friedman SR, Thompson DR, Gates GJ, Liberti TM, et al. Statewide estimation of racial/ethnic populations of men who have sex with men in the U.S. *Public Health Rep* 2011;126(1):60-72 [FREE Full text] [doi: [10.1177/003335491112600110](https://doi.org/10.1177/003335491112600110)] [Medline: [21351396](https://pubmed.ncbi.nlm.nih.gov/21351396/)]
- Adimora A, Schoenbach V. Social context, sexual networks, and racial disparities in rates of sexually transmitted infections. *J Infect Dis* 2005 Mar 01;191 Suppl 1(s1):S115-S122. [doi: [10.1086/425280](https://doi.org/10.1086/425280)] [Medline: [15627221](https://pubmed.ncbi.nlm.nih.gov/15627221/)]
- Wejnert C, Hess KL, Rose CE, Balaji A, Smith JC, Paz-Bailey G, NHBS Study Group. Age-specific race and ethnicity disparities in HIV infection and awareness among men who have sex with men--20 US cities, 2008-2014. *J Infect Dis* 2016 Mar 01;213(5):776-783 [FREE Full text] [doi: [10.1093/infdis/jiv500](https://doi.org/10.1093/infdis/jiv500)] [Medline: [26486637](https://pubmed.ncbi.nlm.nih.gov/26486637/)]
- Centers for Disease ControlPrevention (CDC). Prevalence and awareness of HIV infection among men who have sex with men --- 21 cities, United States, 2008. *MMWR Morb Mortal Wkly Rep* 2010 Sep 24;59(37):1201-1207 [FREE Full text] [Medline: [20864920](https://pubmed.ncbi.nlm.nih.gov/20864920/)]
- Cohen MS, Council OD, Chen JS. Sexually transmitted infections and HIV in the era of antiretroviral treatment and prevention: the biologic basis for epidemiologic synergy. *J Int AIDS Soc* 2019 Aug 30;22 Suppl 6(S6):e25355 [FREE Full text] [doi: [10.1002/jia2.25355](https://doi.org/10.1002/jia2.25355)] [Medline: [31468737](https://pubmed.ncbi.nlm.nih.gov/31468737/)]
- HIV risk behaviors. Centers for Disease Control and Prevention. URL: <https://www.cdc.gov/hiv/risk/estimates/riskbehaviors.html> [accessed 2021-10-22]
- Fauci AS, Redfield RR, Sigounas G, Weahkee MD, Giroir BP. Ending the HIV epidemic: a plan for the United States. *JAMA* 2019 Mar 05;321(9):844-845. [doi: [10.1001/jama.2019.1343](https://doi.org/10.1001/jama.2019.1343)] [Medline: [30730529](https://pubmed.ncbi.nlm.nih.gov/30730529/)]
- Holloway I, Beltran R, Shah SV, Cordero L, Garth G, Smith T, et al. Structural syndemics and antiretroviral medication adherence among Black sexual minority men living with HIV. *J Acquir Immune Defic Syndr* 2021 Dec 15;88(S1):S12-S19 [FREE Full text] [doi: [10.1097/QAI.0000000000002806](https://doi.org/10.1097/QAI.0000000000002806)] [Medline: [34757988](https://pubmed.ncbi.nlm.nih.gov/34757988/)]
- Maragh-Bass AC, Gamble T, El-Sadr WM, Hanscom B, Tolley EE. Exploring individual-level barriers to HIV medication adherence among men who have sex with men in the HIV Prevention Trials Network (HPTN 065) study. *AIDS Care* 2021 Nov 07;33(11):1404-1413. [doi: [10.1080/09540121.2020.1828799](https://doi.org/10.1080/09540121.2020.1828799)] [Medline: [33025791](https://pubmed.ncbi.nlm.nih.gov/33025791/)]
- Russ S, Zhang C, Liu Y. Pre-exposure prophylaxis care continuum, barriers, and facilitators among black men who have sex with men in the United States: a systematic review and meta-analysis. *AIDS Behav* 2021 Jul 12;25(7):2278-2288. [doi: [10.1007/s10461-020-03156-x](https://doi.org/10.1007/s10461-020-03156-x)] [Medline: [33438151](https://pubmed.ncbi.nlm.nih.gov/33438151/)]
- Bowleg L, Malekzadeh AN, Mbaba M, Boone CA. Ending the HIV epidemic for all, not just some: structural racism as a fundamental but overlooked social-structural determinant of the US HIV epidemic. *Curr Opin HIV AIDS* 2022 Mar 01;17(2):40-45. [doi: [10.1097/COH.0000000000000724](https://doi.org/10.1097/COH.0000000000000724)] [Medline: [35102051](https://pubmed.ncbi.nlm.nih.gov/35102051/)]
- Card KG, Lachowsky NJ, Armstrong HL, Cui Z, Wang L, Sereda P, et al. The additive effects of depressive symptoms and polysubstance use on HIV risk among gay, bisexual, and other men who have sex with men. *Addict Behav* 2018 Jul;82:158-165 [FREE Full text] [doi: [10.1016/j.addbeh.2018.03.005](https://doi.org/10.1016/j.addbeh.2018.03.005)] [Medline: [29533847](https://pubmed.ncbi.nlm.nih.gov/29533847/)]
- Hutton HE, Lyketos CG, Zenilman JM, Thompson RE, Erbeling EJ. Depression and HIV risk behaviors among patients in a sexually transmitted disease clinic. *Am J Psychiatry* 2004 May;161(5):912-914. [doi: [10.1176/appi.ajp.161.5.912](https://doi.org/10.1176/appi.ajp.161.5.912)] [Medline: [15121659](https://pubmed.ncbi.nlm.nih.gov/15121659/)]
- Wilson PA, Stadler G, Boone MR, Bolger N. Fluctuations in depression and well-being are associated with sexual risk episodes among HIV-positive men. *Health Psychol* 2014 Jul;33(7):681-685 [FREE Full text] [doi: [10.1037/a0035405](https://doi.org/10.1037/a0035405)] [Medline: [24467260](https://pubmed.ncbi.nlm.nih.gov/24467260/)]
- Alvy LM, McKirnan DJ, Mansergh G, Koblin B, Colfax GN, Flores SA, Project MIX Study Group. Depression is associated with sexual risk among men who have sex with men, but is mediated by cognitive escape and self-efficacy. *AIDS Behav* 2011 Aug 9;15(6):1171-1179. [doi: [10.1007/s10461-010-9678-z](https://doi.org/10.1007/s10461-010-9678-z)] [Medline: [20217471](https://pubmed.ncbi.nlm.nih.gov/20217471/)]

17. Gonzalez J, Batchelder AW, Psaros C, Safren SA. Depression and HIV/AIDS treatment nonadherence: a review and meta-analysis. *J Acquir Immune Defic Syndr* 2011 Oct 01;58(2):181-187 [FREE Full text] [doi: [10.1097/QAI.0b013e31822d490a](https://doi.org/10.1097/QAI.0b013e31822d490a)] [Medline: [21857529](https://pubmed.ncbi.nlm.nih.gov/21857529/)]
18. Holloway I, Tan D, Dunlap S, Palmer L, Beougher S, Cederbaum J. Network support, technology use, depression, and ART adherence among HIV-positive MSM of color. *AIDS Care* 2017 Sep 10;29(9):1153-1161 [FREE Full text] [doi: [10.1080/09540121.2017.1325435](https://doi.org/10.1080/09540121.2017.1325435)] [Medline: [28488886](https://pubmed.ncbi.nlm.nih.gov/28488886/)]
19. McDavid K, McKenna MT. HIV/AIDS risk factor ascertainment: a critical challenge. *AIDS Patient Care STDS* 2006 Apr;20(4):285-292. [doi: [10.1089/apc.2006.20.285](https://doi.org/10.1089/apc.2006.20.285)] [Medline: [16623627](https://pubmed.ncbi.nlm.nih.gov/16623627/)]
20. Gallagher KM, Sullivan PS, Lansky A, Onorato IM. Behavioral surveillance among people at risk for HIV infection in the U.S.: the national HIV behavioral surveillance system. *Public Health Rep* 2016 Aug 02;122(1\_suppl):32-38. [doi: [10.1177/00333549071220s106](https://doi.org/10.1177/00333549071220s106)]
21. Garrett R, Young SD. Digital public health surveillance tools for alcohol use and HIV risk behaviors. *AIDS Behav* 2021 Dec 17;25(Suppl 3):333-338 [FREE Full text] [doi: [10.1007/s10461-021-03221-z](https://doi.org/10.1007/s10461-021-03221-z)] [Medline: [33730254](https://pubmed.ncbi.nlm.nih.gov/33730254/)]
22. Salathé M, Bengtsson L, Bodnar TJ, Brewer DD, Brownstein JS, Buckee C, et al. Digital epidemiology. *PLoS Comput Biol* 2012 Jul 26;8(7):e1002616 [FREE Full text] [doi: [10.1371/journal.pcbi.1002616](https://doi.org/10.1371/journal.pcbi.1002616)] [Medline: [22844241](https://pubmed.ncbi.nlm.nih.gov/22844241/)]
23. Pierce SJ, Miller RL, Morales MM, Forney J. Identifying HIV prevention service needs of African American men who have sex with men: an application of spatial analysis techniques to service planning. *J Public Health Manag Pract* 2007 Jan;Suppl(Supplement):S72-S79. [doi: [10.1097/00124784-200701001-00012](https://doi.org/10.1097/00124784-200701001-00012)] [Medline: [17159472](https://pubmed.ncbi.nlm.nih.gov/17159472/)]
24. Shah NS, Iveniuk J, Muth SQ, Michaels S, Jose J, Laumann EO, et al. Structural bridging network position is associated with HIV status in a younger Black men who have sex with men epidemic. *AIDS Behav* 2014 Mar 12;18(2):335-345 [FREE Full text] [doi: [10.1007/s10461-013-0677-8](https://doi.org/10.1007/s10461-013-0677-8)] [Medline: [24337699](https://pubmed.ncbi.nlm.nih.gov/24337699/)]
25. Latkin CA, Forman V, Knowlton A, Sherman S. Norms, social networks, and HIV-related risk behaviors among urban disadvantaged drug users. *Social Sci Med* 2003 Feb;56(3):465-476. [doi: [10.1016/s0277-9536\(02\)00047-3](https://doi.org/10.1016/s0277-9536(02)00047-3)]
26. Kelly JA, St Lawrence JS, Diaz YE, Stevenson LY, Hauth AC, Brasfield TL, et al. HIV risk behavior reduction following intervention with key opinion leaders of population: an experimental analysis. *Am J Public Health* 1991 Mar;81(2):168-171. [doi: [10.2105/ajph.81.2.168](https://doi.org/10.2105/ajph.81.2.168)] [Medline: [1990853](https://pubmed.ncbi.nlm.nih.gov/1990853/)]
27. Eysenbach G. Infodemiology and infoveillance: framework for an emerging set of public health informatics methods to analyze search, communication and publication behavior on the Internet. *J Med Internet Res* 2009 Mar 27;11(1):e11 [FREE Full text] [doi: [10.2196/jmir.1157](https://doi.org/10.2196/jmir.1157)] [Medline: [19329408](https://pubmed.ncbi.nlm.nih.gov/19329408/)]
28. De Choudhury M, Gamon M, Counts S, Horvitz E. Predicting depression via social media. In: Proceedings of the International AAAI Conference on Web and Social Media. 2013 Presented at: International AAAI Conference on Web and Social Media; Jul 8-11, 2013; Cambridge, Massachusetts USA.
29. Razi A, Kim S, Alsoubai A, Stringhini G, Solorio T, De Choudhury M, et al. A human-centered systematic literature review of the computational approaches for online sexual risk detection. *Proc ACM Human Comput Interact* 2021 Oct 13;5(CSCW2):1-38. [doi: [10.1145/3479609](https://doi.org/10.1145/3479609)]
30. Young SD, Rivers C, Lewis B. Methods of using real-time social media technologies for detection and remote monitoring of HIV outcomes. *Prev Med* 2014 Jun;63:112-115 [FREE Full text] [doi: [10.1016/j.ypmed.2014.01.024](https://doi.org/10.1016/j.ypmed.2014.01.024)] [Medline: [24513169](https://pubmed.ncbi.nlm.nih.gov/24513169/)]
31. Schneider J. Sociostructural 2-mode network analysis: critical connections for HIV transmission elimination. *Sex Transm Dis* 2013 Jun;40(6):459-461. [doi: [10.1097/01.olq.0000430672.69321.0a](https://doi.org/10.1097/01.olq.0000430672.69321.0a)] [Medline: [23677020](https://pubmed.ncbi.nlm.nih.gov/23677020/)]
32. The Sexual Organization of the City. Chicago: University of Chicago Press; 2005.
33. Potterat JJ, Rothenberg RB, Muth SQ. Network structural dynamics and infectious disease propagation. *Int J STD AIDS* 1999 Mar 25;10(3):182-185. [doi: [10.1258/0956462991913853](https://doi.org/10.1258/0956462991913853)] [Medline: [10340199](https://pubmed.ncbi.nlm.nih.gov/10340199/)]
34. Rothenberg R, Baldwin J, Trotter R, Muth S. The risk environment for HIV transmission: results from the Atlanta and Flagstaff network studies. *J Urban Health* 2001 Sep;78(3):419-432 [FREE Full text] [doi: [10.1093/jurban/78.3.419](https://doi.org/10.1093/jurban/78.3.419)] [Medline: [11564846](https://pubmed.ncbi.nlm.nih.gov/11564846/)]
35. Schneider JA, Cornwell B, Ostrow D, Michaels S, Schumm P, Laumann EO, et al. Network mixing and network influences most linked to HIV infection and risk behavior in the HIV epidemic among Black men who have sex with men. *Am J Public Health* 2013 Jan;103(1):e28-e36. [doi: [10.2105/ajph.2012.301003](https://doi.org/10.2105/ajph.2012.301003)]
36. Morris M, Goodreau S, Moody J. Sexual networks, concurrency, and STD/HIV. In: Sexually Transmitted Diseases. 4th edition. New York, United States: McGraw-Hill International Book; 2007.
37. Rice E, Milburn NG, Rotheram-Borus MJ. Pro-social and problematic social network influences on HIV/AIDS risk behaviours among newly homeless youth in Los Angeles. *AIDS Care* 2007 May 25;19(5):697-704 [FREE Full text] [doi: [10.1080/09540120601087038](https://doi.org/10.1080/09540120601087038)] [Medline: [17505933](https://pubmed.ncbi.nlm.nih.gov/17505933/)]
38. Rice E. The positive role of social networks and social networking technology in the condom-using behaviors of homeless young people. *Public Health Rep* 2010 Jul 01;125(4):588-595. [doi: [10.1177/003335491012500414](https://doi.org/10.1177/003335491012500414)] [Medline: [20597459](https://pubmed.ncbi.nlm.nih.gov/20597459/)]
39. Grov C. HIV risk and substance use in men who have sex with men surveyed in bathhouses, bars/clubs, and on Craigslist.org: venue of recruitment matters. *AIDS Behav* 2012 May 12;16(4):807-817 [FREE Full text] [doi: [10.1007/s10461-011-9999-6](https://doi.org/10.1007/s10461-011-9999-6)] [Medline: [21748276](https://pubmed.ncbi.nlm.nih.gov/21748276/)]

40. Grov C, Parsons JT, Bimbi DS. Sexual risk behavior and venues for meeting sex partners: an intercept survey of gay and bisexual men in LA and NYC. *AIDS Behav* 2007 Nov 8;11(6):915-926. [doi: [10.1007/s10461-006-9199-y](https://doi.org/10.1007/s10461-006-9199-y)] [Medline: [17206536](https://pubmed.ncbi.nlm.nih.gov/17206536/)]
41. Elwood WN, Greene K, Carter KK. Gentlemen don't speak: communication norms and condom use in bathhouses. *J Applied Commun Res* 2003 Nov;31(4):277-297. [doi: [10.1080/1369681032000132564](https://doi.org/10.1080/1369681032000132564)]
42. Fujimoto K, Williams ML, Ross MW. Venue-based affiliation networks and HIV risk-taking behavior among male sex workers. *Sex Transm Dis* 2013 Jun;40(6):453-458 [FREE Full text] [doi: [10.1097/OLQ.0b013e31829186e5](https://doi.org/10.1097/OLQ.0b013e31829186e5)] [Medline: [23677019](https://pubmed.ncbi.nlm.nih.gov/23677019/)]
43. Binson D, Woods WJ, Pollack L, Paul J, Stall R, Catania JA. Differential HIV risk in bathhouses and public cruising areas. *Am J Public Health* 2001 Sep;91(9):1482-1486. [doi: [10.2105/ajph.91.9.1482](https://doi.org/10.2105/ajph.91.9.1482)] [Medline: [11527785](https://pubmed.ncbi.nlm.nih.gov/11527785/)]
44. Grov C, Hirshfield S, Remien RH, Humberstone M, Chiasson MA. Exploring the venue's role in risky sexual behavior among gay and bisexual men: an event-level analysis from a national online survey in the U.S. *Arch Sex Behav* 2013 Mar 20;42(2):291-302 [FREE Full text] [doi: [10.1007/s10508-011-9854-x](https://doi.org/10.1007/s10508-011-9854-x)] [Medline: [22012413](https://pubmed.ncbi.nlm.nih.gov/22012413/)]
45. Holloway IW, Dunlap S, Del Pino HE, Hermanstynne K, Pulsipher C, Landovitz RJ. Online social networking, sexual risk and protective behaviors: considerations for clinicians and researchers. *Curr Addict Rep* 2014 Sep 19;1(3):220-228 [FREE Full text] [doi: [10.1007/s40429-014-0029-4](https://doi.org/10.1007/s40429-014-0029-4)] [Medline: [25642408](https://pubmed.ncbi.nlm.nih.gov/25642408/)]
46. Young LE, Fujimoto K, Schneider JA. HIV prevention and sex behaviors as organizing mechanisms in a Facebook group affiliation network among young Black men who have sex with men. *AIDS Behav* 2018 Oct 13;22(10):3324-3334 [FREE Full text] [doi: [10.1007/s10461-018-2087-4](https://doi.org/10.1007/s10461-018-2087-4)] [Medline: [29536285](https://pubmed.ncbi.nlm.nih.gov/29536285/)]
47. Charles-Smith LE, Reynolds TL, Cameron MA, Conway M, Lau EH, Olsen JM, et al. Using social media for actionable disease surveillance and outbreak management: a systematic literature review. *PLoS One* 2015 Oct 5;10(10):e0139701 [FREE Full text] [doi: [10.1371/journal.pone.0139701](https://doi.org/10.1371/journal.pone.0139701)] [Medline: [26437454](https://pubmed.ncbi.nlm.nih.gov/26437454/)]
48. Schneider J, Young L, Ramachandran A, Michaels S, Cohen H, Robinson I, PrEP Chicago Network Intervention Team. A pragmatic randomized controlled trial to increase PrEP uptake for HIV prevention: 55-week results from PrEPChicago. *J Acquir Immune Defic Syndr* 2021 Jan 01;86(1):31-37 [FREE Full text] [doi: [10.1097/QAI.0000000000002518](https://doi.org/10.1097/QAI.0000000000002518)] [Medline: [33306562](https://pubmed.ncbi.nlm.nih.gov/33306562/)]
49. Young LE, Schumm P, Alon L, Bouris A, Ferreira M, Hill B, et al. PrEP Chicago: a randomized controlled peer change agent intervention to promote the adoption of pre-exposure prophylaxis for HIV prevention among young Black men who have sex with men. *Clin Trials* 2018 Feb 11;15(1):44-52 [FREE Full text] [doi: [10.1177/1740774517730012](https://doi.org/10.1177/1740774517730012)] [Medline: [28862483](https://pubmed.ncbi.nlm.nih.gov/28862483/)]
50. Heckathorn DD. Respondent-driven sampling: a new approach to the study of hidden populations. *Social Problems* 1997 May;44(2):174-199. [doi: [10.1525/sp.1997.44.2.03x0221m](https://doi.org/10.1525/sp.1997.44.2.03x0221m)]
51. Eaton WW, Smith C, Ybarra M, Muntaner C, Tien A. Center for epidemiologic studies depression scale: review and revision (CESD and CESD-R). In: *The Use of Psychological Testing for Treatment Planning and Outcomes Assessment: Instruments for Adults*. Hillsdale, NJ: Lawrence Erlbaum Associates Publishers; 2004.
52. Arnold EA, Bailey MM. Constructing home and family: how the ballroom community supports African American GLBTQ youth in the face of HIV/AIDS. *J Gay Lesbian Soc Serv* 2009 Jan 01;21(2-3):171-188 [FREE Full text] [doi: [10.1080/10538720902772006](https://doi.org/10.1080/10538720902772006)] [Medline: [23136464](https://pubmed.ncbi.nlm.nih.gov/23136464/)]
53. Bailey MM. Performance as intravention: ballroom culture and the politics of HIV/AIDS in Detroit. *Souls* 2009 Sep 08;11(3):253-274. [doi: [10.1080/10999940903088226](https://doi.org/10.1080/10999940903088226)]
54. Kubicek K, McNeeley M, Holloway IW, Weiss G, Kipke MD. "It's like our own little world": resilience as a factor in participating in the Ballroom community subculture. *AIDS Behav* 2013 May 23;17(4):1524-1539 [FREE Full text] [doi: [10.1007/s10461-012-0205-2](https://doi.org/10.1007/s10461-012-0205-2)] [Medline: [22618891](https://pubmed.ncbi.nlm.nih.gov/22618891/)]
55. Ding T, Hasan F, Bickel WK, Pan S. Interpreting social media-based substance use prediction models with knowledge distillation. In: *Proceedings of the 2018 IEEE 30th International Conference on Tools with Artificial Intelligence (ICTAI)*. 2018 Presented at: 2018 IEEE 30th International Conference on Tools with Artificial Intelligence (ICTAI); Nov 05-07, 2018; Volos, Greece. [doi: [10.1109/ictai.2018.00100](https://doi.org/10.1109/ictai.2018.00100)]
56. Pennebaker J, Booth RJ, Boyd RL, Francis ME. *Linguistic Inquiry and Word Count: LIWC2015*. Austin, TX: Pennebaker Conglomerates; 2015.
57. Schneider J, Michaels S, Bouris A. Family network proportion and HIV risk among black men who have sex with men. *J Acquir Immune Defic Syndr* 2012 Dec 15;61(5):627-635 [FREE Full text] [doi: [10.1097/QAI.0b013e318270d3cb](https://doi.org/10.1097/QAI.0b013e318270d3cb)] [Medline: [23011395](https://pubmed.ncbi.nlm.nih.gov/23011395/)]
58. Bonacich P. Factoring and weighting approaches to status scores and clique identification. *J Mathematical Sociol* 1972 Jan;2(1):113-120. [doi: [10.1080/0022250x.1972.9989806](https://doi.org/10.1080/0022250x.1972.9989806)]
59. Hanneman R, Riddle M. *Introduction to Social Network Methods*. Riverside, CA: University of California; 2005.
60. Everett MG, Valente TW. Bridging, brokerage and betweenness. *Soc Networks* 2016 Jan;44:202-208 [FREE Full text] [doi: [10.1016/j.socnet.2015.09.001](https://doi.org/10.1016/j.socnet.2015.09.001)] [Medline: [28239229](https://pubmed.ncbi.nlm.nih.gov/28239229/)]
61. Watts DJ, Strogatz SH. Collective dynamics of 'small-world' networks. *Nature* 1998 Jun;393(6684):440-442. [doi: [10.1038/30918](https://doi.org/10.1038/30918)]



62. RStudio: Integrated Development Environment for R. Boston, MA: RStudio; 2022. URL: <http://www.rstudio.com/> [accessed 2022-10-13]
63. Myers J, Braunstein SL, Xia Q, Scanlin K, Edelstein Z, Harriman G, et al. Redefining prevention and care: a status-neutral approach to HIV. *Open Forum Infect Dis* 2018 Jun;5(6):ofy097 [FREE Full text] [doi: [10.1093/ofid/ofy097](https://doi.org/10.1093/ofid/ofy097)] [Medline: [29977957](https://pubmed.ncbi.nlm.nih.gov/29977957/)]
64. Pagkas-Bather J, Brewer R, Bouris A. Status-neutral interventions to support health equity for Black sexual minority men. *Curr HIV/AIDS Rep* 2022 Aug 06;19(4):265-280. [doi: [10.1007/s11904-022-00610-4](https://doi.org/10.1007/s11904-022-00610-4)] [Medline: [35794447](https://pubmed.ncbi.nlm.nih.gov/35794447/)]
65. Prestage G, Hammoud M, Jin F, Degenhardt L, Bourne A, Maher L. Mental health, drug use and sexual risk behavior among gay and bisexual men. *Int J Drug Policy* 2018 May;55:169-179. [doi: [10.1016/j.drugpo.2018.01.020](https://doi.org/10.1016/j.drugpo.2018.01.020)] [Medline: [29429865](https://pubmed.ncbi.nlm.nih.gov/29429865/)]
66. Ryan C, Huebner D, Diaz RM, Sanchez J. Family rejection as a predictor of negative health outcomes in white and Latino lesbian, gay, and bisexual young adults. *Pediatrics* 2009 Jan;123(1):346-352. [doi: [10.1542/peds.2007-3524](https://doi.org/10.1542/peds.2007-3524)] [Medline: [19117902](https://pubmed.ncbi.nlm.nih.gov/19117902/)]
67. Matthews T, Danese A, Wertz J, Odgers CL, Ambler A, Moffitt TE, et al. Social isolation, loneliness and depression in young adulthood: a behavioural genetic analysis. *Soc Psychiatry Psychiatr Epidemiol* 2016 Mar 3;51(3):339-348 [FREE Full text] [doi: [10.1007/s00127-016-1178-7](https://doi.org/10.1007/s00127-016-1178-7)] [Medline: [26843197](https://pubmed.ncbi.nlm.nih.gov/26843197/)]
68. Young LE, Jonas AB, Michaels S, Jackson JD, Pierce ML, Schneider JA, uConnect Study Team. Social-structural properties and HIV prevention among young men who have sex with men in the ballroom house and independent gay family communities. *Soc Sci Med* 2017 Feb;174:26-34 [FREE Full text] [doi: [10.1016/j.socscimed.2016.12.009](https://doi.org/10.1016/j.socscimed.2016.12.009)] [Medline: [27987435](https://pubmed.ncbi.nlm.nih.gov/27987435/)]
69. Schneider J, Cornwell B, Jonas A, Lancki N, Behler R, Skaathun B, et al. Network dynamics of HIV risk and prevention in a population-based cohort of young Black men who have sex with men. *Net Sci* 2017 Feb 01;5(3):381-409. [doi: [10.1017/nws.2016.27](https://doi.org/10.1017/nws.2016.27)]

## Abbreviations

**BSMM:** Black sexual minority men

**OR:** odds ratio

**PrEP:** pre-exposure prophylaxis

**RDS:** respondent-driven sampling

**SMM:** sexual minority men

**STI:** sexually transmitted infection

*Edited by A Mavragani; submitted 14.03.22; peer-reviewed by C Copen, A Sharma; comments to author 25.07.22; revised version received 27.08.22; accepted 06.09.22; published 20.10.22.*

*Please cite as:*

*Young LE, Tang JL, Nan Y*

*Social Media Communication and Network Correlates of HIV Infection and Transmission Risks Among Black Sexual Minority Men: Cross-sectional Digital Epidemiology Study*

*JMIR Form Res* 2022;6(10):e37982

URL: <https://formative.jmir.org/2022/10/e37982>

doi: [10.2196/37982](https://doi.org/10.2196/37982)

PMID: [36264617](https://pubmed.ncbi.nlm.nih.gov/36264617/)

©Lindsay E Young, Jack Lipei Tang, Yuanfeixue Nan. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 20.10.2022. 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

# Mental Illness Concordance Between Hospital Clinical Records and Mentions in Domestic Violence Police Narratives: Data Linkage Study

George Karystianis<sup>1</sup>, BSc, MSc, PhD; Rina Carines Cabral<sup>1</sup>, BSc, MDS; Armita Adily<sup>1</sup>, BMedSc, MPH, PhD; Wilson Lukmanjaya<sup>2</sup>, BSc; Peter Schofield<sup>3</sup>, BS, BSc (Hons), MEd, MSc, MD; Iain Buchan<sup>4</sup>, BSc (Hons), MB ChB, MD, DPH; Goran Nenadic<sup>5</sup>, BSc, MSc, PhD; Tony Butler<sup>1</sup>, BSc, MSc, PhD

<sup>1</sup>School of Population Health, University of New South Wales, Sydney, Australia

<sup>2</sup>School of Computer Science, University of Technology, Sydney, Australia

<sup>3</sup>Hunter Medical Research Institute, Newcastle, Australia

<sup>4</sup>Institute of Population Health, University of Liverpool, Liverpool, United Kingdom

<sup>5</sup>School of Computer Science, University of Manchester, Manchester, United Kingdom

**Corresponding Author:**

George Karystianis, BSc, MSc, PhD

School of Population Health

University of New South Wales

Level 3, Samuels Building, Gate 11, Botany Street

UNSW Kensington Campus

Sydney, 2052

Australia

Phone: 61 93852517

Email: [g.karystianis@unsw.edu.au](mailto:g.karystianis@unsw.edu.au)

## Abstract

**Background:** To better understand domestic violence, data sources from multiple sectors such as police, justice, health, and welfare are needed. Linking police data to data collections from other agencies could provide unique insights and promote an all-of-government response to domestic violence. The New South Wales Police Force attends domestic violence events and records information in the form of both structured data and a free-text narrative, with the latter shown to be a rich source of information on the mental health status of persons of interest (POIs) and victims, abuse types, and sustained injuries.

**Objective:** This study aims to examine the concordance (ie, matching) between mental illness mentions extracted from the police's event narratives and mental health diagnoses from hospital and emergency department records.

**Methods:** We applied a rule-based text mining method on 416,441 domestic violence police event narratives between December 2005 and January 2016 to identify mental illness mentions for POIs and victims. Using different window periods (1, 3, 6, and 12 months) before and after a domestic violence event, we linked the extracted mental illness mentions of victims and POIs to clinical records from the Emergency Department Data Collection and the Admitted Patient Data Collection in New South Wales, Australia using a unique identifier for each individual in the same cohort.

**Results:** Using a 2-year window period (ie, 12 months before and after the domestic violence event), less than 1% (3020/416,441, 0.73%) of events had a mental illness mention and also a corresponding hospital record. About 16% of domestic violence events for both POIs (382/2395, 15.95%) and victims (101/631, 16.01%) had an agreement between hospital records and police narrative mentions of mental illness. A total of 51,025/416,441 (12.25%) events for POIs and 14,802/416,441 (3.55%) events for victims had mental illness mentions in their narratives but no hospital record. Only 841 events for POIs and 919 events for victims had a documented hospital record within 48 hours of the domestic violence event.

**Conclusions:** Our findings suggest that current surveillance systems used to report on domestic violence may be enhanced by accessing rich information (ie, mental illness) contained in police text narratives, made available for both POIs and victims through the application of text mining. Additional insights can be gained by linkage to other health and welfare data collections.

(JMIR Form Res 2022;6(10):e39373) doi:[10.2196/39373](https://doi.org/10.2196/39373)

**KEYWORDS**

data linkage; mental health; domestic violence; police records; hospital records; text mining

## Introduction

Domestic violence is defined as “any incident of threatening behavior, violence or (psychological, physical, sexual, financial, emotional) abuse between adults who are or have been an intimate partner or family member, regardless of gender or sexuality” [1], but it can also occur in other relationships including caregivers, guardians, parents, and a dependent person or those living together in a household such as flatmates [2]. Domestic violence is a global public health problem resulting in a significant economic and health burden on the community. In Australia, domestic violence is the leading cause of morbidity and mortality for women, surpassing risk factors such as obesity and smoking [3]. In 2018, 1 in 6 women and 1 in 16 men experienced physical or sexual violence by their current/former partner [4]. Research also shows that children exposed to domestic violence experience long-term effects on their development with an increased risk of poor mental health, learning difficulties, and behavioral problems [5].

Domestic violence has been linked to deaths, physical injuries ranging from minor traumas to those requiring hospitalization, depression, substance use, risky sexual behaviors, eating disorders, posttraumatic stress disorder, suicidal ideation and attempts, acts of self-harm, and exacerbation of psychotic symptoms [4,6-11]. Associations have been found between mental health conditions (eg, bipolar disorder, schizophrenia) and the perpetration of violence [12-18]. Evidence also suggests that people with mental illnesses are at a greater risk of victimization compared with those without [8,10,13,19-22]. Men and women with severe mental illness (eg, psychotic disorders) are 2-8 times more likely to experience any form of domestic abuse and have poor health outcomes (eg, suicide attempt, substance abuse) than the general population [19,23]. Thus, knowledge of the mental health of those involved in domestic violence may enable better prevention and intervention measures to be developed and implemented.

Administrative data collections represent a significant public resource [24,25]. Data linkage, the process that brings together individual-level data from different data collections, can provide valuable research insights and community benefits, although accessing them can be challenging as identified by the Australian Productivity Commission, and thus their use is often discouraged [26-29]. Data linkage offers a powerful, relatively inexpensive, population-wide, and accurate source of information to explore risk and protective factors. Linked administrative data have potentially very large sample sizes, access to the entire population information served by an agency, data on hard-to-reach populations, minimal loss to follow-up for certain outcomes, and a high level of external validity necessary for policy making [28,30]. These data are free from many of the measurement issues and biases that have been associated with qualitative research approaches (eg, in-depth victim interviews) that feature occasionally small sample sizes, are time consuming, and can have selection bias and high cost [31]. However, tapping into existing, routinely collected data using novel approaches

such as text mining with subsequent linkage to different sectors’ data can complement these primary data collection methods in the area of domestic violence.

A common source of domestic violence data around the world are police records. In the Australian state of New South Wales (NSW), the NSW Police Force (NSWPF) attends and subsequently records details on thousands of domestic violence events each year. In NSW, a domestic violence event is defined as an incident of domestic dispute that involves any form of violence or abuse between a person of interest (POI)—an individual accused of perpetrating any form of violence or abuse toward another individual—and a victim. Information related to domestic violence events is recorded both as structured data (fixed fields, eg, covering demographic information such as name, date of birth, Aboriginal status) for the POI and the victim, and as a free-text narrative that describes details of the event (eg, cause, mental health status, threats of subsequent violence) based on the police officers’ observations and testimonies from the involved parties or the presence of witnesses (eg, neighbors, roommates, friends, family members).

Although the narratives may be used as an aide-mémoire for police officers and lawyers should the case proceed to court, to date they have not been used systematically for research and monitoring purposes due to their voluminous nature and the time taken to inspect and glean relevant information. We recently demonstrated that these event narratives contain rich information on perpetrators and victims of domestic violence such as mental illness [32], victim injuries, and abuse types [33]. We also demonstrated that such information could be extracted automatically, and that it could be used for population-wide domestic violence monitoring, surveillance, and prediction of domestic violence-related offences [34-36]. While other sources of domestic violence data might offer complementary information that will help to understand domestic violence, barriers to fully display its incidence and prevalence exist [37,38], leading to a potentially underrepresentation of this public health problem.

As mental illness is an important factor in domestic violence and we had previously identified a significant number of police events that contained mental illness mentions for both victims and POIs, the aim of this study is to examine the concordance between these mentions and corresponding diagnoses recorded by tertiary health services. This approach enables an assessment of the extent to which these 2 collections overlap and highlight their utility for domestic violence surveillance. This also allows an appraisal of relative merits, among which collection has more value to policy development in the domestic violence area.

## Methods

### Domestic Violence Police Events

We initially examined 492,393 police-recorded domestic violence events from January 2005 to December 2016 that were flagged in the fixed fields with 1 of the following tags:

“domestic” as the offence type; “domestic violence related” as the associated factor of the police event; “spouse/partner (including ex-spouse/ex-partner),” “boy/girlfriend (including ex-boy/ex-girlfriend),” “parent/guardian (including step/foster),” “child (including step/foster),” “sibling,” “other member of family (including kin),” or “carer” as the relationship status between the victim and the POI. We note that domestic violence events can contain events that the police attended but no crime was committed.

### Text Mining Method and Normalization of Mental Illness Mentions

We previously developed a text mining methodology which we applied to domestic violence police event narratives that extracted mental illness mentions for POIs and victims [32]. The approach was implemented through the General Architecture for Text Engineering [39], which was chosen for its support of rule-based methods and easy manipulation of unstructured data. Rules were created after observing common lexical patterns that suggest a mental illness (Multimedia Appendix 1) for a POI (eg, “the POI is suffering from dementia”) or a victim (eg, “the victim was diagnosed with paranoid schizophrenia”), with semantic anchors relating the mention to either a POI (eg, “POI”, “person of interest”, “defendant”) or a victim (eg, “vic”, “PINOP—person in need of protection”, “victim”) in a sample of 200 police-recorded domestic violence events. This also included cases where:

- an unspecified mental illness was recorded (eg, “the defendant has mental health issues”);
- psychotropic drugs were used (eg, “the victim takes Valium”, “accused takes a number of antidepressants”), which might indicate a mental illness; these were categorized into 4 groups (antianxiety, antidepressants, neuroleptics, and antipsychotics); and
- individuals had traumatic brain injury or a drug prescription abuse (unspecified in the text regarding the medication), substance abuse (unspecified in the text regarding the substance), and drug-induced disorder (unspecified in the text regarding the drug).

The rules were combined with dictionaries that contained mental illness terms taken from the World Health Organization’s International Classification of Diseases, tenth revision (ICD-10) for Mental and Behavioral Disorders categories [40,41], including common abbreviations and synonyms. The methodology was fully evaluated against the manual annotations of mental illness mentions for POIs and victims by 2 experts (in domestic violence and neuropsychiatry, respectively, with an interannotator agreement of 90% [42]) in a random sample of 100 police-recorded domestic violence events, returning an average  $F_1$ -score [43] of 84% (87.0% for POIs and 81.0% for victims); a detailed description of the methodology including its design, error analysis, and limitations has been published elsewhere [32].

As the extracted mentions ranged from specific (eg, “oppositional defiance disorder”) to general descriptions (eg, “behavioral problems”), these were mapped to ICD-10’s Mental and Behavioral Disorders categories using 3 levels (Multimedia

Appendix 2) to conduct analysis of the results. Ambiguous case mapping was resolved by using the neuropsychiatry expertise of the fifth author (PS). A detailed analysis of the normalization process has been published elsewhere [32,34].

Although domestic violence events can have more than 1 POI or victim, the previous text mining methodology was unable to associate the extracted mental health “mention” with a specific POI or victim if more than 1 POI or victim was present in the same event. Thus, in this study we focused only on those events that included a single POI and a single victim. This resulted in 416,441 domestic violence events involving 214,148 unique POIs and 244,218 unique victims. By processing the associated narratives, 64,587 domestic violence events had at least one (unique) mention of mental illness for POIs and victims [32,34].

### NSW Health Records

Two sources of hospital records were used: the Admitted Patient Data Collection (APDC) and the Emergency Department Data Collection (EDDC). The APDC includes records for all hospital admissions from all NSW public and private hospitals and day procedure centers, whereas the EDDC provides information on presentations to emergency rooms in public hospitals in NSW. Each record in both collections includes a mental health diagnosis, the age of the patient at the time of diagnosis, and the start/end date of the admission. The EDDC contains diagnostic codes in the ICD-10 (eg, F10.0 stands for “mental and behavioral disorders due to use of alcohol, acute intoxication”), International Classification of Diseases, ninth revision (ICD-9; eg, 295.0 stands for “simple schizophrenia type”), and the Systematized Nomenclature of Medicine - Clinical Terms (SNOMED-CT; eg, 42344001 stands for “alcohol-induced psychosis”) formats, whereas the APDC has only ICD-10 codes. To ensure consistency with the data extracted from the police reports, the SNOMED-CT and ICD-9 diagnostic codes were mapped to the equivalent ICD-10 codes. For the conversion of the SNOMED-CT codes, we used a schema developed by the National eHealth Transition Authority, Australian Digital Health Service, where each SNOMED-CT code was assigned a corresponding ICD-10 code. The ICD-9 codes were converted through an online automatic converter, a process that was supervised by the fifth author (PS) [41]. Codes that could have several mappings to ICD-10 were assigned to 1 code only based on the expertise of the fifth author (PS).

The respective window of these 2 collections involved records from July 2001 to September 2018 (APDC) and from January 2005 to September 2018 (EDDC), which covered the period of the domestic violence event data set (January 2005 to December 2016).

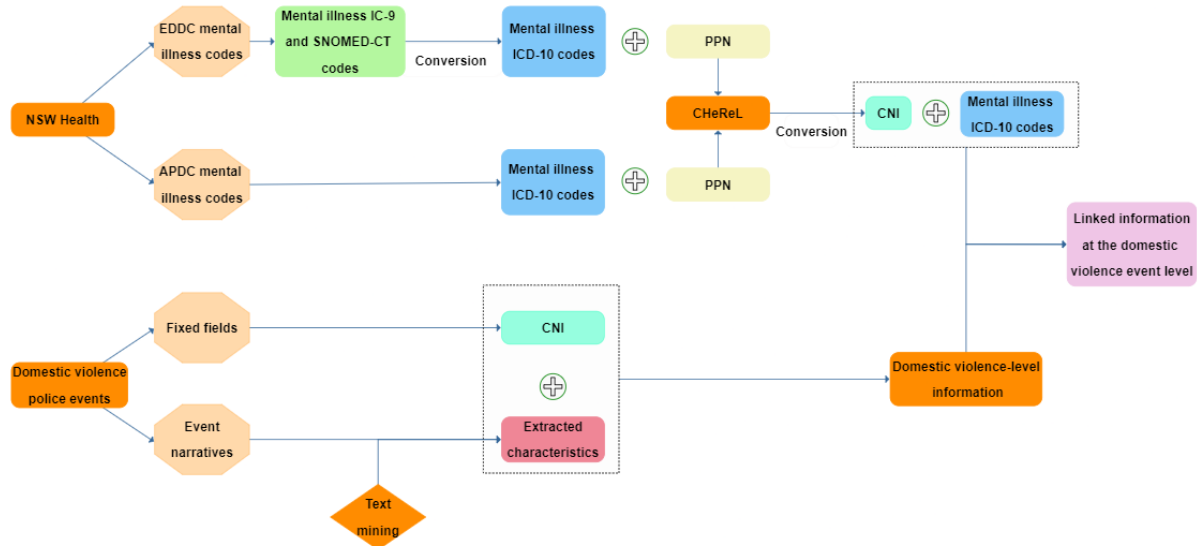
### Linkage Between Police Records and NSW Health Records

The police’s unique Criminal Number Index (CNI) for each individual involved in a domestic violence event was converted into a unique Project Person Number (PPN) by the Centre for Health Record Linkage to link these individuals with the EDDC and APDC data using probabilistic record linkage [44]. An overview of the data linkage process including the application of the text mining methodology on the domestic violence police

narratives is shown in [Figure 1](#). A detailed description of all the information (ie, text mined, fixed fields, and external mental

health diagnoses) used to describe a domestic violence event is provided in [Multimedia Appendix 3](#).

**Figure 1.** An overview of the data used in the linkage process. CNI is the police's unique number for each individual involved in a domestic violence event. PPN is the converted CNI that is required to link the individuals of a police record with the EDDC and APDC data collections. APDC: Admitted Patient Data Collection; CHeReL: Centre for Health Record Linkage; CNI: Criminal Number Index; EDDC: Emergency Department Data Collection; ICD-9: International Classification of Disease, ninth Revision; ICD-10: International Classification of Disease, tenth Revision; NSW: New South Wales; PPN: Project Person Number; SNOMED-CT: Systematized Nomenclature of Medicine - Clinical Terms.



## Concordance Analysis

To examine concordance between the police and secondary care systems, we used descriptive statistics. It should be noted that an individual might have more than 1 diagnosis or police mental illness mentions. Thus, at an event level, concordance was defined as an exact agreement on a mental illness diagnosis code between a police mention and a hospital record for an individual (ie, how many events have the same code in police mentions and hospital records irrelevant of how many other diagnosis codes might be associated for this particular individual at this event). Concordance for a specific mental illness followed the same approach (ie, how many events had an agreement on a specific mental illness code between police mentions and hospital records). We used only the first level of ICD-10 to inspect the concordance. This means that more specific mentions from either side (eg, a depression mention from police narratives or a bipolar disorder diagnosis from a hospital admission matched at the first level of ICD-10, which was “mood [affective] disorders”) were aggregated to the first level for reporting purposes.

Given that the dates for hospital admissions and emergency department presentations were unlikely to align in a temporal sense with the domestic violence event, we used window periods of varying lengths (ie, 1, 3, 6, and 12 months) before and after the event to examine the concordance between the extracted (police mentions) and diagnostic information (hospital records). This approach was adopted to examine existing mental health contacts with secondary and tertiary health care systems. Those conditions seen by primary care providers or parts of the health care system other than hospitals would not have been captured by our data linkage.

## Ethics Approval

Permission to access the domestic violence events was granted by the NSWPF following ethics approval from the University of NSW Human Research Ethics Committee (approval number HC16558). Approval was also granted by the NSW Ministry of Health (approval number HC16558) to access information contained in the APDC and EDDC.

## Results

### Overview

Using no window period for the linkage between the hospital and police records for 416,441 domestic violence police events, the total number of events with a hospital record of mental illness was 142,324/416,441 (34.18%) and the number of events with mental illness police mentions was 64,587/416,441 (15.51%; [Table 1](#)).

Using a 2-year window period (ie, 12 months before and after the domestic violence event), the number of events with a corresponding hospital record with any mental illness was reduced to 34,337/416,441 (8.25%; [Table 2](#)). Predictably, the larger the window period used, the more overlap there was between the police and hospital records related to mental illness. The least overlap was seen when applying the shortest (ie, 1 month before and after the domestic violence event) window period, where only 699 police and hospital systems recorded a mental illness for either the POI or victim ([Figure 2](#)).

Using a 2-year window, a total of 5560 domestic violence events had both a hospital record and a police mention of mental illness for either the POI or victim ([Table 2](#)). Regardless of the diagnosis, in 3020/5560 (54.32%) events the mental illnesses recorded by the police and hospital systems agreed in terms of

having a diagnosis present for the POI (2395/3020, 79.30%) and victim (631/3020, 20.89%; Figure 3). However, 2540/5560 (45.68%) events reported mental health information for different

individuals (ie, the police recorded a mental illness mention in an event for a POI, whereas the linked hospital record of mental illness for the same event was attributed to the victim).

**Table 1.** Number of domestic violence events with or without a recorded mental illness in the APDC<sup>a</sup>/EDDC<sup>b</sup> collections or a reported (unique) mention from the police narratives—no window period applied (N=416,441)<sup>c</sup>.

APDC/EDDC diagnosis code	Police mental illness mentions, n	
	Yes (n=64,587)	No (n=351,854)
Yes (n=142,324)	22,190	120,134
No (n=274,117)	42,397	231,720

<sup>a</sup>APDC: Admitted Patient Data Collection.

<sup>b</sup>EDDC: Emergency Department Data Collection.

<sup>c</sup>“Yes” refers to the number of domestic violence events/hospital records that have a mental illness mention/diagnosis.

**Table 2.** Number of domestic violence events with or without a recorded mental illness in the APDC<sup>a</sup>/EDDC<sup>b</sup> collections or a reported mention from the police narratives; 2-year window period applied (ie, 12 months before and after an event occurred; N=416,441)<sup>c</sup>.

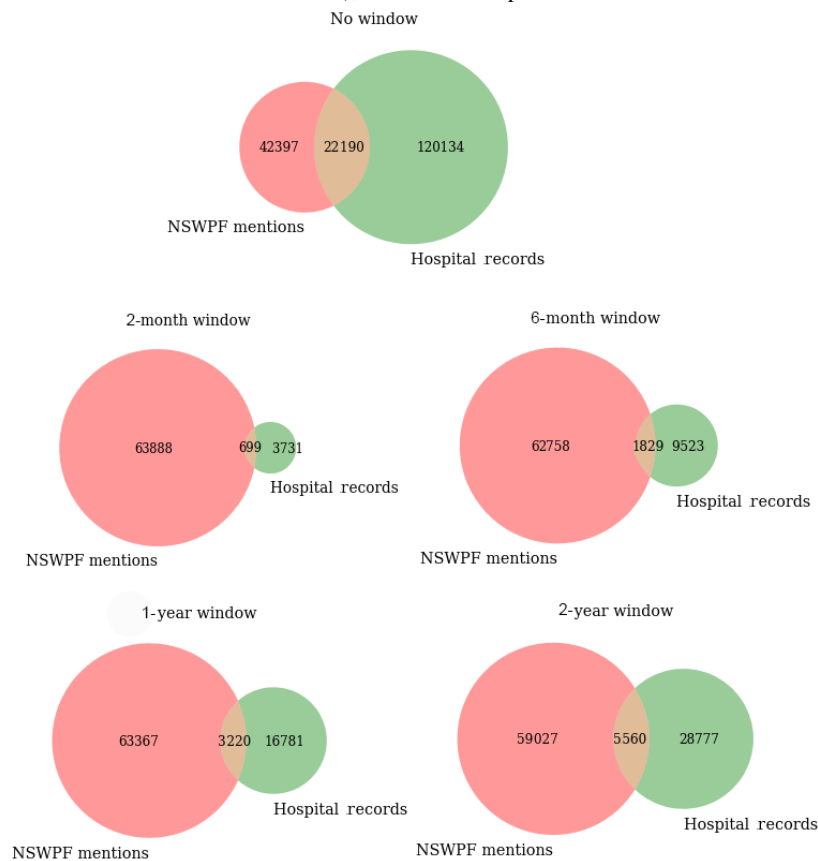
APDC/EDDC diagnosis code	Police mental illness mentions, n	
	Yes (n=64,587)	No (n=351,854)
Yes (n=34,337)	5560	28,777
No (n=382,104)	59,027	323,077

<sup>a</sup>APDC: Admitted Patient Data Collection.

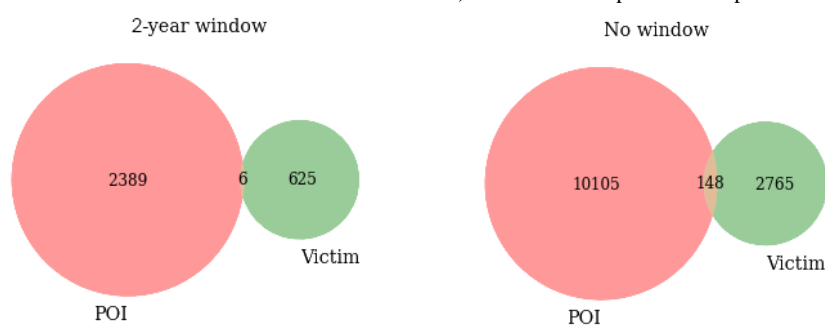
<sup>b</sup>EDDC: Emergency Department Data Collection.

<sup>c</sup>“Yes” refers to domestic violence events/hospital records that have a mental illness mention/diagnosis.

**Figure 2.** Number of domestic violence police events with a police mention and a hospital record of mental illness using various window periods (ie, 1, 3, 6, and 12 months before and after the domestic violence event), and no window period. NSWPF: New South Wales Police Force.



**Figure 3.** Number of domestic violence police events with a police mention and hospital record of mental illness for a POI and a victim using a 2-year window (ie, 12 months before and after the domestic violence event occurred) and no window period. POI: person of interest.



### Concordance

Using a 2-year window period, the concordance between hospital and police records for POIs was in only 382 events (382/2395, 15.95%), whereas for victims it was in 101 (101/631, 16.01%) events. The concordance rate was highest for mental and behavioral disorders due to psychoactive substance use when taking into consideration the total number of police events with a respective mention (POIs: 151 events out of 5616, 2.69%; victims: 26 events out of 1098, 2.37%; [Table 3](#)).

When no window period was applied, the concordance rate between police and hospital records increased: for POIs it was 19.56% (2005/10,253), whereas for victims it was 18.50% (539/2913). Despite mood disorders having a higher number of events that saw concordance between police and hospital records, their concordance rate was lower in both POIs (756/11,985, 6.31%) and victims (233/4108, 5.67%), whereas mental and behavioral disorders due to psychoactive substance use had the highest concordance rate (692/5616, 12.32% for POIs and 115/1098, 10.47% for victims; [Table 4](#)).

**Table 3.** Number of domestic violence police events with the same mental illness diagnosis (presented as ICD-10<sup>a</sup> mental illness groups) in the police and hospital systems for POI<sup>b</sup> and victims—2-year window period applied.

Mental illness group (ICD-10)	Total <sup>c</sup> (POI)	Number of events with a concordance <sup>d</sup> (POI)	Concordance rate, %	Total <sup>c</sup> (victim)	Number of events with a concordance <sup>d</sup> (victim)	Concordance rate, %
Mental and behavioral disorders due to psychoactive substance use	5616	151	2.7	1098	26	2.4
Mood (affective) disorders	11,985	110	0.9	4108	35	0.9
Schizophrenia, schizotypal, delusional, and other nonmood psychotic disorders	4529	50	1.1	878	3	0.3
Anxiety, dissociative, stress related, somatoform, and other nonpsychotic mental disorders	2855	29	1.0	1832	21	1.1
Unspecified mental disorder	21,839	28	0.1	4198	10	0.2
Behavioral and emotional disorders with onset usually occurring in childhood and adolescence	7097	9	0.1	1721	2	0.1
Intentional self-harm	2694	6	0.2	821	2	0.2
Disorders of adult personality and behavior	1067	3	0.3	355	1	0.3
Pervasive and specific developmental disorders	1338	1	0.1	401	1	0.2
Intellectual disability	1196	0	0.0	782	1	0.1

<sup>a</sup>ICD-10: International Classification of Disease, tenth revision.

<sup>b</sup>POI: person of interest.

<sup>c</sup>Total number of events with a text-mined mention of mental illness

<sup>d</sup>Number of events with an agreement between a text mining (police) mental illness mention and a mental health (hospital) diagnosis code.

**Table 4.** Number of domestic violence police events with the same mental illness diagnosis (presented in ICD-10<sup>a</sup> mental illness groups) in the police and hospital systems for POI<sup>b</sup> and victims—no window period applied.

Mental illness group (ICD-10)	Total <sup>c</sup> (POI)	Number of events with a concordance <sup>d</sup> (POI)	Concordance rate, %	Total <sup>c</sup> (victim)	Number of events with a concordance <sup>d</sup> (victim)	Concordance rate, %
Mood (affective) disorders	11,985	756	6.31	4108	233	5.7
Mental and behavioral disorders due to psychoactive substance use	5616	692	12.3	1098	115	10.5
Unspecified mental disorder	21,839	187	0.9	4198	44	1.0
Anxiety, dissociative, stress related, somatoform, and other nonpsychotic mental disorders	2855	181	6.3	1832	101	5.5
Schizophrenia, schizotypal, delusional, and other nonmood psychotic disorders	4529	133	2.9	878	32	3.6
Behavioral and emotional disorders with onset usually occurring in childhood and adolescence	7097	85	1.2	1721	13	0.8
Intentional self-harm	2694	31	1.2	821	10	1.2
Disorders of adult personality and behavior	1067	21	2.0	355	4	1.1
Intellectual disability	1196	5	0.4	782	2	0.3
Pervasive and specific developmental disorders	1338	3	0.2	401	3	0.7
Mental disorders due to known physiological conditions	496	2	0.4	579	1	0.2
Symptoms and signs involving cognition, perception, emotional state, and behavior	160	1	0.6	72	1	1.4

<sup>a</sup>ICD-10: International Classification of Disease, tenth revision.

<sup>b</sup>POI: person of interest.

<sup>c</sup>Total number of events with a text-mined mention of mental illness.

<sup>d</sup>Number of events with an agreement between a text mining (police) mental illness mention and a mental health (hospital) diagnosis code.

### Events With a Police Mention and No Hospital Record of Mental Illness

In a 2-year window period, 59,027/416,441 (14.17%) domestic violence events had a police mention but no hospital record of mental illness (Table 2), with 51,025/59,027 (86.44%) event mentions related to POIs and 14,802/59,027 (25.08%) to victims. Unspecified mental disorders (generic mentions of mental illness

in the narrative text such as mental disorder, some form of mental illness) were the most common in POI (20,890/51,025, 40.94%) and victims (4043/14,802, 27.31%), followed by mood disorders (11,427/51,025, 22.39%, in POIs and 3938/14,802, 26.60%, in victims). Anxiety disorders in victims (1752/14,802, 11.84%) had more than double the rate when compared with that of the POIs (2731/51,025, 5.35%; Table 5).



**Table 5.** Number of domestic violence police events with a mention of mental illness for POIs<sup>a</sup> and victims without a respective hospital record—2-year window period applied (n=51,025)<sup>b</sup>.

Mental illness group (ICD-10 <sup>c</sup> )	Events (POI; n=51,025), n (%)	Number of events (victim; n=14,802), n (%)
Unspecified mental disorder	20,890 (40.94)	4043 (27.31)
Mood (affective) disorders	11,427 (22.39)	3938 (26.60)
Behavioral and emotional disorders with onset usually occurring in childhood and adolescence	6791 (13.31)	1645 (11.11)
Mental and behavioral disorders due to psychoactive substance use	5360 (10.50)	1064 (7.19)
Schizophrenia, schizotypal, delusional, and other nonmood psychotic disorders	4306 (8.44)	840 (5.67)
Anxiety, dissociative, stress related, somatoform, and other nonpsychotic mental disorders	2731 (5.35)	1752 (11.84)
Intentional self-harm	2577 (5.05)	783 (5.29)
Substance abuse	2177 (4.27)	302 (2.04)
Pervasive and specific developmental disorders	1278 (2.50)	378 (2.55)
Intellectual disability	1143 (2.24)	737 (4.98)
Disorders of adult personality and behavior	1021 (2.0)	339 (2.29)
Injury of unspecified body region	658 (1.29)	208 (1.41)
Traumatic brain injury	531 (1.04)	197 (1.33)
Mental disorders due to known physiological conditions	479 (0.94)	557 (3.76)
Medications: antidepressants	311 (0.61)	105 (0.71)
Symptoms and signs involving cognition, perception, emotional state, and behavior	154 (0.30)	68 (0.46)
Medications (antipsychotics)	104 (0.20)	13 (0.09)
Medications for anxiety	73 (0.14)	20 (0.14)
Other degenerative diseases of the nervous system	51 (0.10)	43 (0.29)
Chromosomal abnormalities, not elsewhere classified	46 (0.09)	30 (0.20)
Unspecified drug-induced disorders	43 (0.08)	0 (0.0)
Behavioral syndromes associated with physiological disturbances and physical factors	22 (0.04)	17 (0.1)
Systemic atrophies primarily affecting the central nervous system	10 (0.02)	6 (0.04)
Diseases of the nervous system	3 (0.01)	2 (0.01)
Drug prescription abuse	3 (0.01)	0 (0.0)
Medications (neuroleptics)	1 (0.0)	0 (0.0)

<sup>a</sup>POI: person of interest.

<sup>b</sup>A single event could contain multiple mentions of a mental illness group (ie, number of mentions).

<sup>c</sup>ICD-10: International Classification of Disease, tenth revision.

### Events With Hospital Records and No Police Mention of Mental Illness

In the 2-year window, 28,777/416,441 (6.91%) domestic violence events had a hospital record but no police mention of mental illness (Table 6), with 15,340/28,777 (53.31%) and 16,793/28,777 (58.36%) events having a hospital record for

POIs and victims, respectively. The most prevalent mental illness was mental and behavioral disorders due to psychoactive substance use (9112/15,340, 59.40%, for POIs and 10,017/16,793, 59.65%, for victims), followed by mood disorders (3892/15,340, 25.37%, for POIs and 4070/16,793, 24.24%, for victims; Table 6).

**Table 6.** Number of domestic violence police events with a mention of mental illness for POIs<sup>a</sup> and victims without a respective hospital record—2-year window period applied<sup>b</sup>.

Mental illness group (ICD-10 <sup>c</sup> )	Events (POI; n=15,340)	Events (victim; n=16,793)
Mental and behavioral disorders due to psychoactive substance use	9112 (59.40)	10,017 (59.65)
Mood (affective) disorders	3892 (25.37)	4070 (24.24)
Anxiety, dissociative, stress related, somatoform, and other nonpsychotic mental disorders	3604 (23.49)	3976 (23.68)
Schizophrenia, schizotypal, delusional, and other nonmood psychotic disorders	2627 (17.13)	2789 (16.61)
Disorders of adult personality and behavior	1710 (11.15)	1853 (11.03)
Behavioral and emotional disorders with onset usually occurring in childhood and adolescence	676 (4.41)	708 (4.22)
Unspecified mental disorder	644 (4.20)	641 (3.82)
Intentional self-harm	636 (4.15)	629 (3.75)
Mental disorders due to known physiological conditions	615 (4.01)	724 (4.31)
Symptoms and signs involving cognition, perception, emotional state, and behavior	486 (3.17)	580 (3.45)
Intellectual disability	312 (2.03)	345 (2.05)
Pervasive and specific developmental disorders	239 (1.56)	243 (1.45)
Behavioral syndromes associated with physiological disturbances and physical factors	172 (1.12)	203 (1.21)
Other degenerative diseases of the nervous system	65 (0.42)	58 (0.35)
Systematic atrophies primarily affecting the central nervous system	15 (0.10)	12 (0.07)
Chromosomal abnormalities, not elsewhere classified	12 (0.08)	14 (0.08)
Injury of unspecified body region	1 (0.01)	1 (0.01)

<sup>a</sup>POI: person of interest.

<sup>b</sup>A single event could have several mental illness groups.

<sup>c</sup>ICD-10: International Classification of Disease, tenth revision.

## Discussion

### Principal Findings

These findings build on our previous work, showing that when police attend domestic violence events, they report on a significant number of victims and perpetrators with mental illness [29,31]. Almost 16% (64,587/416,441, 15.51%) of domestic violence events attended by the police had a mention of mental illness [34], which is high considering that this is not proactively probed by the attending officers.

In this study, we explored links between mentions of mental illness in police reports and hospital data. We linked almost half a million police-attended domestic violence events to records from emergency departments (EDDC) and hospital admissions (APDC). Only 3020/416,441 (0.73%) domestic violence events with extracted mental illness information from the police narrative had a corresponding hospital record using a 2-year window period (ie, 12 months before and after the event). This reinforces the view that surveillance systems based solely on hospital presentations are very limited in terms of coverage. We found that concordance (ie, the same mental illness between a police mention and a hospital record for an individual) between these 2 data sources was 15.95% (382/2395) for POI and 16.01% (101/631) for victims. Yet, this lack of

concordance between the police and the hospital system is likely explained by the severity of the mental illness, whereby relatively few less severe mental disorders, which may be seen by primary care providers, result in a hospital admission or emergency room attendance.

This does not necessarily reflect poor reporting quality by the NSWPF for several reasons. First, police officers in NSW only undergo a short (1-4 days) period of mental health training as part of their initial training (Chief Inspector Matthew McCarthy, NSWPF, personal communication, July 2021). Second, their primary focus while attending domestic violence events is on victim safety, and thus, recording mental health information is not a priority. Third, recording mental health status is likely to be opportunistic with different motivations in reporting this to the police by victims, perpetrators, and witnesses. The number of police-attended domestic violence events with no mention of mental illness but with a respective hospital record remained roughly the same before (8626 for POIs; 9447 for victims) and after (8157 for POIs; 9804 for victims) the date of the event when a 2-year window period was applied.

Most domestic violence police events with a mental illness mention had no recorded hospital diagnosis within a 2-year window period (59,027/64,587, 91.39%), whereas 83.81% (28,777/34,337) of the hospital admissions that involved a POI

or victim had no corresponding mention of mental illness in the police data. Events in which the NSWPF did not record a mental illness mention but had a hospital record could be explained from the 2-year window. It is possible that when police officers attended the event, the individual may not have been diagnosed with a mental illness at that point in time. This can be observed in 28,777/416,441 (6.91%) domestic violence events with a hospital record for POIs and victims, with most of the diagnoses recorded related to substance use (ie, mental and behavioral disorders due to psychoactive substance use), mood, anxiety, and schizophrenic disorders within the 2-year window.

### Concordance With Specific Mental Illnesses

The highest concordance rate between police and health data occurred in events that reported mental and behavioral disorders due to psychoactive substance use for POIs (151 events out of 5616, 2.69%) and victims (26 events out of 1098, 2.37%), with the concordance rate increasing 10-fold when no window period was applied (692/5616, 12.32%, for POIs and 115/1098, 10.47%, for victims). Admissions to residential drug treatment programs within the health system likely explain the higher concordance in terms of diagnoses related to psychoactive substance use. Recent research has demonstrated that substance use (including alcohol consumption) increases both the risk for and impact of domestic violence, highlighting the need for policy that advocates interventions that can address both drug use and violence in combination [45]. These findings have implications about the effect that alcohol and drug abuse might have on the perpetration of domestic violence [46]. Interestingly, unspecified mental disorders were the main recorded group within the police narratives, which is consistent with police's role not being focused on recording specific details on mental health status.

### Future Implications

Overall, the police appear to have substantial visibility on mental illness in domestic violence, covering the scope and the severity of conditions. Our findings indicated that the NSWPF encounter myriad mental illnesses in the context of domestic violence with more than 120 types identified [32,34]. This further demonstrates the usefulness of employing police narrative data for surveillance and monitoring purposes as we have previously suggested [36].

Partnerships between the police, public health, and welfare sectors with regard to data sharing could help improve domestic violence, health, and justice outcomes, but it often does not occur. For example, data (or algorithm) sharing arrangements could assist in more focused policies and initiatives to improve front-line responses when mental illness is identified, for example, through crisis intervention teams [47-49]. Combining information extracted from domestic violence police records with diagnoses from the health system in privacy preserving ways could help to further explain behaviors and motivations and improve domestic violence prevention. Primary health care providers such as general practitioners (GPs) are also likely to have good visibility on domestic violence for injuries not requiring hospitalization, and less severe forms of mental health and should be considered for linkage as part of a comprehensive data capture system.

No single data collection is likely to have complete coverage of domestic violence, but we have demonstrated that 2 of the big government agencies, the police and secondary and tertiary health care systems (ie, hospitals), can be effectively linked. To improve coverage and identify repeat perpetrators or victims, multiple data sources (eg, GP victim visits vs police attendance) will be required. We identified that only 0.73% (3020/416,441) of domestic violence events had a mental illness recorded in both police records and hospital records, suggesting that other agencies' data are needed for comprehensive domestic violence reporting.

Interestingly, within the 12-month period, only 841 events for POIs and 919 events for victims (out of 416,441) had a hospital record of mental illness within 48 hours since the domestic violence event occurred. Further, apart from those codes indicating presentations due to psychoactive substance abuse, the recorded codes referred to medical reasons for hospitalization that were unlikely related to domestic violence. This supports the idea that hospitals only see a small fraction of conditions arising from domestic violence situations or do not screen admitted individuals for domestic violence [37,38], which are recognized as limitations of the current domestic violence surveillance [36].

Notwithstanding these factors, we believe that first contact with the police as a consequence of a domestic violence event represents an important opportunity to inquire about the mental health status and enable referral pathways for those requiring treatment and support. This would not be without its challenges but may be important in preventing and minimizing future recurrences of domestic violence. It is possible that these cases of mental illness in domestic violence could be captured from additional data collections that are more appropriate for less severe forms of mental illnesses such as GP visits and psychological services, both of which are covered by Australia's Medicare Benefits Schedule. However, this requires ongoing dialog and discussion between agencies holding different data collections as well as sophisticated agreements to ensure privacy is protected.

### Limitations

Linking police-recorded domestic violence events with health records from hospital admissions and emergency department presentations has several limitations. There are unique challenges in the appropriate research use of linked administrative data, for example, with respect to bias from linkage errors where records cannot be linked or are linked together incorrectly [28,50]. When linking data sets, a proportion of cases match and a proportion will remain unmatched [50]. The window period of 1 year before and after the date of each event might have limited the number of linked events. Thus, there might have been cases with individuals that were diagnosed after the date when the domestic violence event occurred (eg, after 2 years). Therefore, an expansion of the time frame might return a larger number of linked events (Figure 2) and could potentially assist in providing more population-based information for this type of linkage research and more nuanced conclusions.

Common mental disorders are seen mostly in primary care, while serious mental illness are seen in specialist secondary and tertiary care services. Therefore, common mental disorders are underrepresented in our data. Obtaining and linking other types of clinical data such as GP visits and psychotropic medication prescriptions would embellish the picture of mental health and domestic violence. Police-attended domestic violence events may not result in or be preceded by GP attendances either, and thus remain undetected by the health system. This is a further argument for the benefit of using police data for domestic violence surveillance and employing text mining to extract salient information such as mental illness mentions.

Finally, it is unclear whether any of these extracted mentions of mental illness from police-attended domestic violence events are accurate. These data are not collected for research or reporting purposes and therefore researchers do not get to decide which variables should be recorded [31]. Concordance between police mentions and hospital records with a 2-year window yielded 15.95% (382/2395) for POIs and 16.01% (101/631) for victims, highlighting the need to have a variety of other data sources that can be linked to police records to increase the visibility on the needs of those involved in domestic violence situations, covering law enforcement, social care, health care, welfare.

## Conclusions

Domestic violence is a significant public health problem with its myriad facets being described and recorded in various data collections that cover the health, housing, welfare, justice, child neglect, and finance sectors. However, these different data sources often do not readily share information with each other, preventing the development of a comprehensive picture of domestic violence that can be used for monitoring and prevention purposes. Using a 2-year period to link half a million domestic violence police events with emergency department and hospital admissions, the concordance between the identified mental health mentions from the police narratives and those recorded by the hospital system was 15.95% (382/2395 events) for POIs and 16.0% (101/631 events) for victims. Thus, our findings reinforce the view that surveillance systems based solely on hospital presentations are limited in terms of coverage, with hospitals seeing only a fraction of mental illness conditions arising from domestic violence situations. This demonstrates that police have good visibility of mental illness in domestic violence covering both the scope and the severity of conditions. Despite the set of limitations that the police data come with, the coverage of mental health within the domestic violence police narratives highlights the need to be able to link various data sources for large-scale surveillance and reporting that could prove beneficial for at-risk populations.

---

## Conflicts of Interest

IB is the Chief Data Scientist Advisor for Astra Zeneca.

---

### Multimedia Appendix 1

Mental illnesses listed in the ICD-10 including the eight new categories targeted for extraction in domestic violence events with examples as they appeared in the police event narratives.

[[DOCX File , 30 KB - formative\\_v6i10e39373\\_app1.docx](#) ]

---

### Multimedia Appendix 2

The ICD-10 Mental and Behavioural Disorders schema used to map the extracted mental illness mentions containing three levels (first, second and third).

[[DOCX File , 41 KB - formative\\_v6i10e39373\\_app2.docx](#) ]

---

### Multimedia Appendix 3

Fields that describe a police recorded domestic violence event. This information is a combination of the outputs from the application of the text mining methodology on the domestic violence event narratives, the demographic and spatiotemporal information included in the respective fixed fields and the linked mental health diagnoses obtained from the Admitted Patient Data Collection (APDC) and Emergency Department Data Collection (EDDC) datasets. POI refers to the person of interest.

[[DOCX File , 47 KB - formative\\_v6i10e39373\\_app3.docx](#) ]

---

## References

1. Home Office. Crime in England and Wales: Findings From the British Crime Survey and Police Recorded Crime. In: Crime in England and Wales: Findings From the British Crime Survey and Police Recorded Crime. London, UK: Home Office; 2009.
2. Briodi A. Domestic Violence is a crime electronic resource: police and community booklet. Sydney, NSW: Sydney City Council and NSWPF; 2010. URL: <http://www.cityofsydney.nsw.gov.au/Community/Safety/DomesticViolenceIsACrime.asp> [accessed 2022-03-13]
3. VicHealth. The Health Costs of Violence: Measuring the Burden of Diseases Caused by Intimate Partner Violence. 2005. URL: <https://www.vichealth.vic.gov.au/media-and-resources/publications/the-health-costs-of-violence> [accessed 2022-01-18]

4. Australian Institute of Health and Welfare (AIHW). Family, domestic and sexual violence in Australia. 2018. URL: <https://www.aihw.gov.au/reports/domestic-violence/family-domestic-sexual-violence-in-australia-2018/contents/summary> [accessed 2022-02-10]
5. Campo M. Children's exposure to domestic and family violence: Key issues and responses. *Journal of the Home Economics Institute of Australia* 2015;22(3):33.
6. Sheridan DJ, Nash KR. Acute injury patterns of intimate partner violence victims. *Trauma Violence Abuse* 2007 Jul;8(3):281-289. [doi: [10.1177/1524838007303504](https://doi.org/10.1177/1524838007303504)] [Medline: [17596345](https://pubmed.ncbi.nlm.nih.gov/17596345/)]
7. Robinson L, Spilsbury K. Systematic review of the perceptions and experiences of accessing health services by adult victims of domestic violence. *Health Soc Care Community* 2008 Jan;16(1):16-30. [doi: [10.1111/j.1365-2524.2007.00721.x](https://doi.org/10.1111/j.1365-2524.2007.00721.x)] [Medline: [18181812](https://pubmed.ncbi.nlm.nih.gov/18181812/)]
8. Howard LM, Trevillion K, Khalifeh H, Woodall A, Agnew-Davies R, Feder G. Domestic violence and severe psychiatric disorders: prevalence and interventions. *Psychol Med* 2010 Jun;40(6):881-893. [doi: [10.1017/S0033291709991589](https://doi.org/10.1017/S0033291709991589)] [Medline: [19891808](https://pubmed.ncbi.nlm.nih.gov/19891808/)]
9. Trevillion K, Oram S, Feder G, Howard LM. Experiences of domestic violence and mental disorders: a systematic review and meta-analysis. *PLoS One* 2012;7(12):e51740 [FREE Full text] [doi: [10.1371/journal.pone.0051740](https://doi.org/10.1371/journal.pone.0051740)] [Medline: [23300562](https://pubmed.ncbi.nlm.nih.gov/23300562/)]
10. Khalifeh H, Moran P, Borschmann R, Dean K, Hart C, Hogg J, et al. Domestic and sexual violence against patients with severe mental illness. *Psychol Med* 2015 Mar;45(4):875-886 [FREE Full text] [doi: [10.1017/S0033291714001962](https://doi.org/10.1017/S0033291714001962)] [Medline: [25180908](https://pubmed.ncbi.nlm.nih.gov/25180908/)]
11. Van Deirse TB, Wilson AB, Macy RJ, Cuddeback GS. Intimate Partner Violence and Women with Severe Mental Illnesses: Needs and Challenges from the Perspectives of Behavioral Health and Domestic Violence Service Providers. *J Behav Health Serv Res* 2019 Apr;46(2):283-293. [doi: [10.1007/s11414-018-9624-9](https://doi.org/10.1007/s11414-018-9624-9)] [Medline: [29790039](https://pubmed.ncbi.nlm.nih.gov/29790039/)]
12. Junginger J, McGuire L. Psychotic motivation and the paradox of current research on serious mental illness and rates of violence. *Schizophr Bull* 2004;30(1):21-30. [doi: [10.1093/oxfordjournals.schbul.a007064](https://doi.org/10.1093/oxfordjournals.schbul.a007064)] [Medline: [15176759](https://pubmed.ncbi.nlm.nih.gov/15176759/)]
13. Choe JY, Teplin LA, Abram KM. Perpetration of violence, violent victimization, and severe mental illness: balancing public health concerns. *Psychiatr Serv* 2008 Feb;59(2):153-164 [FREE Full text] [doi: [10.1176/ps.2008.59.2.153](https://doi.org/10.1176/ps.2008.59.2.153)] [Medline: [18245157](https://pubmed.ncbi.nlm.nih.gov/18245157/)]
14. Fazel S, Gulati G, Linsell L, Geddes JR, Grann M. Schizophrenia and violence: systematic review and meta-analysis. *PLoS Med* 2009 Aug;6(8):e1000120 [FREE Full text] [doi: [10.1371/journal.pmed.1000120](https://doi.org/10.1371/journal.pmed.1000120)] [Medline: [19668362](https://pubmed.ncbi.nlm.nih.gov/19668362/)]
15. Fazel S, Lichtenstein P, Grann M, Goodwin GM, Långström N. Bipolar disorder and violent crime: new evidence from population-based longitudinal studies and systematic review. *Arch Gen Psychiatry* 2010 Sep;67(9):931-938. [doi: [10.1001/archgenpsychiatry.2010.97](https://doi.org/10.1001/archgenpsychiatry.2010.97)] [Medline: [20819987](https://pubmed.ncbi.nlm.nih.gov/20819987/)]
16. Van Dorn R, Volavka J, Johnson N. Mental disorder and violence: is there a relationship beyond substance use? *Soc Psychiatry Psychiatr Epidemiol* 2012 Mar;47(3):487-503. [doi: [10.1007/s00127-011-0356-x](https://doi.org/10.1007/s00127-011-0356-x)] [Medline: [21359532](https://pubmed.ncbi.nlm.nih.gov/21359532/)]
17. Oram S, Flynn SM, Shaw J, Appleby L, Howard LM. Mental illness and domestic homicide: a population-based descriptive study. *Psychiatr Serv* 2013 Oct;64(10):1006-1011. [doi: [10.1176/appi.ps.201200484](https://doi.org/10.1176/appi.ps.201200484)] [Medline: [23820784](https://pubmed.ncbi.nlm.nih.gov/23820784/)]
18. Labrum T, Solomon P, Marcus S. Victimization and Perpetration of Violence Involving Persons With Mood and Other Psychiatric Disorders and Their Relatives. *Psychiatr Serv* 2020 May 01;71(5):498-501. [doi: [10.1176/appi.ps.201900384](https://doi.org/10.1176/appi.ps.201900384)] [Medline: [32041513](https://pubmed.ncbi.nlm.nih.gov/32041513/)]
19. Khalifeh H, Dean K. Gender and violence against people with severe mental illness. *Int Rev Psychiatry* 2010;22(5):535-546. [doi: [10.3109/09540261.2010.506185](https://doi.org/10.3109/09540261.2010.506185)] [Medline: [21047165](https://pubmed.ncbi.nlm.nih.gov/21047165/)]
20. Khalifeh H, Moran P, Borschmann R, Dean K, Hart C, Hogg J, et al. Domestic and sexual violence against patients with severe mental illness. *Psychol Med* 2014 Sep 04;45(4):875-886. [doi: [10.1017/s0033291714001962](https://doi.org/10.1017/s0033291714001962)]
21. Bhavsar V, Dean K, Hatch SL, MacCabe JH, Hotopf M. Psychiatric symptoms and risk of victimisation: a population-based study from Southeast London. *Epidemiol Psychiatr Sci* 2019 Apr;28(2):168-178 [FREE Full text] [doi: [10.1017/S2045796018000537](https://doi.org/10.1017/S2045796018000537)] [Medline: [30196801](https://pubmed.ncbi.nlm.nih.gov/30196801/)]
22. Sariaslan A, Arseneault L, Larsson H, Lichtenstein P, Fazel S. Risk of Subjection to Violence and Perpetration of Violence in Persons With Psychiatric Disorders in Sweden. *JAMA Psychiatry* 2020 Apr 01;77(4):359-367 [FREE Full text] [doi: [10.1001/jamapsychiatry.2019.4275](https://doi.org/10.1001/jamapsychiatry.2019.4275)] [Medline: [31940015](https://pubmed.ncbi.nlm.nih.gov/31940015/)]
23. Neria Y, Bromet EJ, Carlson GA, Naz B. Assaultive trauma and illness course in psychotic bipolar disorder: findings from the Suffolk county mental health project. *Acta Psychiatr Scand* 2005 May;111(5):380-383. [doi: [10.1111/j.1600-0447.2005.00530.x](https://doi.org/10.1111/j.1600-0447.2005.00530.x)] [Medline: [15819732](https://pubmed.ncbi.nlm.nih.gov/15819732/)]
24. Boyd JH, Ferrante AM, O'Keefe CM, Bass AJ, Randall SM, Semmens JB. Data linkage infrastructure for cross-jurisdictional health-related research in Australia. *BMC Health Serv Res* 2012 Dec 29;12:480 [FREE Full text] [doi: [10.1186/1472-6963-12-480](https://doi.org/10.1186/1472-6963-12-480)] [Medline: [23272652](https://pubmed.ncbi.nlm.nih.gov/23272652/)]
25. Productivity Commission, Canberra. Annual Report 2012-13, Annual Report Series. 2013. URL: <https://www.pc.gov.au/about/governance/annual-reports/2012-13/annual-report-2012-13.pdf> [accessed 2022-02-10]
26. Black C, Roos L. Linking and combining data to develop statistics for understanding the population's health. In: *Health Statistics: Shaping Policy and Practice to Improve the Population's Health*. Oxford, UK: Oxford University Press; 2005:214-240.

27. Ainsworth J, Buchan I. Combining Health Data Uses to Ignite Health System Learning. *Methods Inf Med* 2015;54(6):479-487. [doi: [10.3414/ME15-01-0064](https://doi.org/10.3414/ME15-01-0064)] [Medline: [26395036](https://pubmed.ncbi.nlm.nih.gov/26395036/)]
28. Harron K, Dibben C, Boyd J, Hjern A, Azimae M, Barreto ML, et al. Challenges in administrative data linkage for research. *Big Data Soc* 2017 Dec 05;4(2):2053951717745678 [FREE Full text] [doi: [10.1177/2053951717745678](https://doi.org/10.1177/2053951717745678)] [Medline: [30381794](https://pubmed.ncbi.nlm.nih.gov/30381794/)]
29. Productivity Commission, Canberra C. Data Availability and Use: Overview & Recommendations, Report No 82. 2017. URL: <https://www.pc.gov.au/inquiries/completed/data-access/report> [accessed 2022-03-08]
30. Holman CDJ, Bass AJ, Rosman DL, Smith MB, Semmens JB, Glasson EJ, et al. A decade of data linkage in Western Australia: strategic design, applications and benefits of the WA data linkage system. *Aust Health Rev* 2008 Nov;32(4):766-777. [doi: [10.1071/ah080766](https://doi.org/10.1071/ah080766)] [Medline: [18980573](https://pubmed.ncbi.nlm.nih.gov/18980573/)]
31. Brownell MD, Jutte DP. Administrative data linkage as a tool for child maltreatment research. *Child Abuse Negl* 2013;37(2-3):120-124. [doi: [10.1016/j.chiabu.2012.09.013](https://doi.org/10.1016/j.chiabu.2012.09.013)] [Medline: [23260116](https://pubmed.ncbi.nlm.nih.gov/23260116/)]
32. Karystianis G, Adily A, Schofield P, Knight L, Galdon C, Greenberg D, et al. Automatic Extraction of Mental Health Disorders From Domestic Violence Police Narratives: Text Mining Study. *J Med Internet Res* 2018 Sep 13;20(9):e11548 [FREE Full text] [doi: [10.2196/11548](https://doi.org/10.2196/11548)] [Medline: [30213778](https://pubmed.ncbi.nlm.nih.gov/30213778/)]
33. Karystianis G, Adily A, Schofield PW, Greenberg D, Jorm L, Nenadic G, et al. Automated Analysis of Domestic Violence Police Reports to Explore Abuse Types and Victim Injuries: Text Mining Study. *J Med Internet Res* 2019 Mar 12;21(3):e13067 [FREE Full text] [doi: [10.2196/13067](https://doi.org/10.2196/13067)] [Medline: [30860490](https://pubmed.ncbi.nlm.nih.gov/30860490/)]
34. Karystianis G, Simpson A, Adily A, Schofield P, Greenberg D, Wand H, et al. Prevalence of Mental Illnesses in Domestic Violence Police Records: Text Mining Study. *J Med Internet Res* 2020 Dec 24;22(12):e23725 [FREE Full text] [doi: [10.2196/23725](https://doi.org/10.2196/23725)] [Medline: [33361056](https://pubmed.ncbi.nlm.nih.gov/33361056/)]
35. Karystianis G, Cabral RC, Han SC, Poon J, Butler T. Utilizing Text Mining, Data Linkage and Deep Learning in Police and Health Records to Predict Future Offenses in Family and Domestic Violence. *Front Digit Health* 2021;3:602683 [FREE Full text] [doi: [10.3389/fdgh.2021.602683](https://doi.org/10.3389/fdgh.2021.602683)] [Medline: [34713088](https://pubmed.ncbi.nlm.nih.gov/34713088/)]
36. Karystianis G, Adily A, Schofield PW, Wand H, Lukmanjaya W, Buchan I, et al. Surveillance of Domestic Violence Using Text Mining Outputs From Australian Police Records. *Front Psychiatry* 2021;12:787792 [FREE Full text] [doi: [10.3389/fpsy.2021.787792](https://doi.org/10.3389/fpsy.2021.787792)] [Medline: [35222105](https://pubmed.ncbi.nlm.nih.gov/35222105/)]
37. Waalen J, Goodwin MM, Spitz AM, Petersen R, Saltzman LE. Screening for intimate partner violence by health care providers. Barriers and interventions. *Am J Prev Med* 2000 Nov;19(4):230-237. [doi: [10.1016/s0749-3797\(00\)00229-4](https://doi.org/10.1016/s0749-3797(00)00229-4)] [Medline: [11064226](https://pubmed.ncbi.nlm.nih.gov/11064226/)]
38. Daugherty JD, Houry DE. Intimate partner violence screening in the emergency department. *J Postgrad Med* 2008;54(4):301-305 [FREE Full text] [doi: [10.4103/0022-3859.43513](https://doi.org/10.4103/0022-3859.43513)] [Medline: [18953150](https://pubmed.ncbi.nlm.nih.gov/18953150/)]
39. Cunningham H, Tablan V, Roberts A, Bontcheva K. Getting more out of biomedical documents with GATE's full lifecycle open source text analytics. *PLoS Comput Biol* 2013 Feb 7;9(2):e1002854 [FREE Full text] [doi: [10.1371/journal.pcbi.1002854](https://doi.org/10.1371/journal.pcbi.1002854)] [Medline: [23408875](https://pubmed.ncbi.nlm.nih.gov/23408875/)]
40. World Health Organization. The International Classification of Disease 10 of Mental and Behavioural Disorders. 2017. URL: [https://www.who.int/substance\\_abuse/terminology/icd\\_10/en/](https://www.who.int/substance_abuse/terminology/icd_10/en/) [accessed 2022-01-16]
41. ICD-10-CM Codes: mental, behavioral and neurodevelopmental disorders. 2017. URL: <https://www.icd10data.com/ICD10CM/Codes/F01-F99> [accessed 2021-11-17]
42. Ananiadou S, McNaught J. Text Mining for Biology and Biomedicine. Boston, MA: Artec House Publishers; 2006.
43. Ananiadou S, Kell DB, Tsujii J. Text mining and its potential applications in systems biology. *Trends Biotechnol* 2006 Dec;24(12):571-579. [doi: [10.1016/j.tibtech.2006.10.002](https://doi.org/10.1016/j.tibtech.2006.10.002)] [Medline: [17045684](https://pubmed.ncbi.nlm.nih.gov/17045684/)]
44. Borthwick A, Buechi M, Goldberg A. Key concepts in the choicemaker 2 record matching system. 2003 Presented at: InProceedings of the KDD Workshop on Data Cleaning, Record Linkage, and Object Consolidation; 2003; Washington DC.
45. Coomber K, Mayschak R, Liknaitzky P, Curtis A, Walker A, Hyder S, et al. The Role of Illicit Drug Use in Family and Domestic Violence in Australia. *J Interpers Violence* 2021 Aug;36(15-16):NP8247-NP8267. [doi: [10.1177/0886260519843288](https://doi.org/10.1177/0886260519843288)] [Medline: [30973034](https://pubmed.ncbi.nlm.nih.gov/30973034/)]
46. Yates S. "An exercise in careful diplomacy": talking about alcohol, drugs and family violence. *Policy Design and Practice* 2019 Jul 11;2(3):258-274 [FREE Full text] [doi: [10.1080/25741292.2019.1638697](https://doi.org/10.1080/25741292.2019.1638697)]
47. Shepherd JP, Sumner SA. Policing and Public Health-Strategies for Collaboration. *JAMA* 2017 Apr 18;317(15):1525-1526 [FREE Full text] [doi: [10.1001/jama.2017.1854](https://doi.org/10.1001/jama.2017.1854)] [Medline: [28358946](https://pubmed.ncbi.nlm.nih.gov/28358946/)]
48. Jacoby SF, Kollar LMM, Ridgeway G, Sumner SA. Health system and law enforcement synergies for injury surveillance, control and prevention: a scoping review. *Inj Prev* 2018 Aug;24(4):305-311 [FREE Full text] [doi: [10.1136/injuryprev-2017-042416](https://doi.org/10.1136/injuryprev-2017-042416)] [Medline: [28971857](https://pubmed.ncbi.nlm.nih.gov/28971857/)]
49. Marshall WA, Egger ME, Pike A, Bozeman MC, Franklin GA, Nash NA, et al. Recidivism rates following firearm injury as determined by a collaborative hospital and law enforcement database. *J Trauma Acute Care Surg* 2020 Aug;89(2):371-376. [doi: [10.1097/TA.0000000000002746](https://doi.org/10.1097/TA.0000000000002746)] [Medline: [32345906](https://pubmed.ncbi.nlm.nih.gov/32345906/)]

50. Bohensky MA, Jolley D, Sundararajan V, Evans S, Pilcher DV, Scott I, et al. Data Linkage: A powerful research tool with potential problems. *BMC Health Serv Res* 2010 Dec 22;10(1):346. [doi: [10.1186/1472-6963-10-346](https://doi.org/10.1186/1472-6963-10-346)] [Medline: [21176171](https://pubmed.ncbi.nlm.nih.gov/21176171/)]

## Abbreviations

**APDC:** Admitted Patient Data Collection  
**CHeReL:** Centre for Health Record Linkage  
**CNI:** Criminal Number Index  
**EDDC:** Emergency Department Data Collection  
**ICD-9:** International Classification of Disease, ninth revision  
**ICD-10:** International Classification of Disease, tenth revision  
**NSW:** New South Wales  
**NSWPF:** New South Wales Police Force  
**POI:** person of interest  
**PPN:** Project Person Number  
**SNOMED-CT:** Systematized Nomenclature of Medicine - Clinical Terms

*Edited by A Mavragani; submitted 08.05.22; peer-reviewed by V Bhavsar, L Shang, A Gutiérrez-Sacristán; comments to author 01.06.22; revised version received 24.06.22; accepted 12.09.22; published 20.10.22.*

*Please cite as:*

*Karystianis G, Cabral RC, Adily A, Lukmanjaya W, Schofield P, Buchan I, Nenadic G, Butler T*

*Mental Illness Concordance Between Hospital Clinical Records and Mentions in Domestic Violence Police Narratives: Data Linkage Study*

*JMIR Form Res* 2022;6(10):e39373

URL: <https://formative.jmir.org/2022/10/e39373>

doi: [10.2196/39373](https://doi.org/10.2196/39373)

PMID: [36264613](https://pubmed.ncbi.nlm.nih.gov/36264613/)

©George Karystianis, Rina Carines Cabral, Armita Adily, Wilson Lukmanjaya, Peter Schofield, Iain Buchan, Goran Nenadic, Tony Butler. Originally published in *JMIR Formative Research* (<https://formative.jmir.org>), 20.10.2022. 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

# Internet-Delivered Self-help for Adults With ADHD (MyADHD): Usability Study

Robin Maria Francisca Kenter<sup>1,2</sup>, PhD; Adrian Schønning<sup>1</sup>, MSc; Yavuz Inal<sup>3</sup>, PhD

<sup>1</sup>Department of Clinical Psychology, Faculty of Psychology, University of Bergen, Bergen, Norway

<sup>2</sup>Section for Research-Driven Innovation, Division of Psychiatry, Haukeland University Hospital, Bergen, Norway

<sup>3</sup>Department of Design, Faculty of Architecture and Design, Norwegian University of Science and Technology, Gjøvik, Norway

**Corresponding Author:**

Robin Maria Francisca Kenter, PhD

Department of Clinical Psychology

Faculty of Psychology

University of Bergen

Christies gate 12

Bergen, 5015

Norway

Phone: 47 48464314

Email: [robken@helse-bergen.no](mailto:robken@helse-bergen.no)

## Abstract

**Background:** Although effective pharmacological treatment exists, many adults with attention-deficit/hyperactivity disorder (ADHD) prefer a nonpharmacological option for managing their symptoms. Internet-delivered self-help interventions have the potential to address this unmet supportive care need reported by adults with ADHD, at relatively low costs. However, if the intervention does not offer optimal functions, content, and layout, it could decrease adherence and engagement and potentially compromise the effectiveness of such interventions. Thus, there is a need for examining the usability and factors that enhance and impair the usability of internet-delivered self-help interventions.

**Objective:** This study evaluates the usability of an internet-delivered self-help intervention for adults with ADHD (MyADHD). The main goals were to (1) collect qualitative and quantitative data on usability and (2) identify usability problems.

**Methods:** Individual think-aloud interviews and staged usability testing (N=5) were conducted to evaluate the usability of the MyADHD intervention in terms of function, content, and design. MyADHD end users provided iterative feedback to maximize engagement and usability. They performed tasks involved in operating the intervention and provided “think-aloud” commentary and postsession usability ratings. The interviews were recorded, transcribed verbatim, and analyzed.

**Results:** Participants were satisfied with the overall usability of the program. The average perceived usability score out of 100 was 70 for the first round of testing and improved to 77.5 after applying modifications, with a mean score of 75.5 (SD 5.9) for all rounds of usability testing. The analysis of the interviews revealed 3 central themes: functionality, content, and layout.

**Conclusions:** Optimizing the usability of internet-delivered self-guided interventions is a critical step in the design and development process. The usability testing in this study provided valuable information from users’ perspectives on the content and platform of the intervention. Analysis revealed the need for intervention enhancement with regard to design, functionality, and content from the perspective of potential end users. Overall, participants saw value in the MyADHD intervention and were confident that they could use it for the self-management of symptoms and expressed the desire to use the entire intervention when it becomes available. Through this development process, we produced an intervention that is likely to be used successfully and is ready for deployment in a randomized controlled trial.

**Trial Registration:** ClinicalTrials.gov NCT04511169; <https://clinicaltrials.gov/ct2/show/NCT04511169>

(*JMIR Form Res* 2022;6(10):e37137) doi:[10.2196/37137](https://doi.org/10.2196/37137)

**KEYWORDS**

usability testing; user evaluation; attention-deficit/hyperactivity disorder; self-guided intervention; internet-delivered; self-help; adults with attention-deficit/hyperactivity disorder; intervention; usability study; care needs; usability



## Introduction

### Background

Digital technology for the treatment and management of mental health issues has become increasingly available in recent years [1]. For common mental health disorders like depression and anxiety, internet-delivered interventions are part of official treatment guidelines [2]. Internet-delivered interventions either can be delivered with some form of guidance or can be entirely self-guided [3]. Effective internet-delivered self-guided interventions can dramatically increase treatment accessibility at relatively low costs, as they can be distributed over the internet directly to the user's computer, without the need for therapist guidance [4]. However, in self-guided internet-delivered interventions, a commonly described issue is low adherence and engagement [5].

### Internet-Delivered Interventions for Adults With ADHD

For people with attention-deficit/hyperactivity disorder (ADHD), few evidence-based internet-delivered interventions for adults exist today. ADHD in adults is characterized by pervasive symptoms of inattention, hyperactivity/impulsiveness, or both that persist across different situations [6]. With an estimated prevalence of 2%-3%, ADHD can be counted as a common mental disorder in adulthood [7-9]. Although effective pharmacological treatment exists [10], many adults with ADHD prefer a nonpharmacological option for managing their symptoms [11]. In this case, self-guided psychological interventions that can be delivered over the internet might have the potential to meet the treatment preferences of adults with ADHD. These types of interventions can help manage symptoms of ADHD by providing general information, tailored advice, support, and skill training via technology (eg, smartphones, tablets, or websites). The few studies that examine internet-delivered psychological treatments for adults with ADHD show good clinical outcomes [12-14]. Self-guided internet-delivered interventions for adults with ADHD have the potential to offer evidence-based interventions to many individuals, with 24/7 access to information, in the comfort of their own homes, where they can read and reread materials.

However, self-guided interventions are also associated with higher dropout rates than guided interventions. Since adherence is related to outcome, it is important to focus on all factors that could improve adherence to self-guided internet-delivered interventions. This starts with designing good systems that address the needs of the end users and have good usability. The primary reasons for low engagement with internet-delivered interventions has been a lack of user-centered design and poor usability [15]. Interventions with poor usability may make it difficult to interact with the intervention, which could lower the engagement, which then can lead to suboptimal clinical outcomes or rejection of the intervention by the user. This might be especially a problem for adults with ADHD, where the cardinal symptoms often are inattention and impulsiveness. It is, therefore, important to test the usability of a new internet-delivered intervention for ADHD during the development process, so that usability flaws are known at an

early stage and can be fixed before the intervention goes to full trial.

### Usability

Usability is defined by the International Organization for Standardization [16] as "the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use." It also refers to the quality of a user's experience when interacting with systems, including websites and mobile apps. The usability of an internet-delivered intervention has a direct influence on user satisfaction. When an internet-delivered intervention is not well designed, it can prevent users from effectively and efficiently using the program. Usability evaluation is a very valuable resource to understand how easy it is to use a developed system for new users. It can be summative (collecting evidence that the intervention is usable) or formative (informing the redesign and improvement of the intervention) [17].

This study describes the results from a formative usability evaluation of the MyADHD intervention [18] that gathered feedback from end users and improved the design as part of an iterative design process.

### Formative Usability Evaluation of MyADHD

In this study, we evaluated the usability of a first prototype of an internet-based intervention for adults with ADHD. The main goals were to (1) collect qualitative and quantitative data on usability and (2) identify usability problems. An iterative development process was used to promote the further development of the content, visual design, and interaction design. Once the first round of usability testing was completed, we evaluated, determined improvements, implemented the improvements, and retested the updated prototype. We included 5 rounds with usability testing with 5 potential end users, as 5 users are often enough to identify 80% of all usability problems [19]. The goal was to identify understandability, ease of learning the intervention's platform, and appropriateness of the intervention. Through the testing, we were provided with feedback about what does and does not work in the intervention for end users and determined whether the features of the intervention were acceptable and feasible for users.

### Objectives

This study aims to investigate the usability of a self-guided internet-delivered intervention (MyADHD), which targets symptoms of ADHD among adults.

## Methods

### Procedures

The usability evaluation was performed in 2 parts: (1) the expert evaluation and (2) the user evaluation. This study reports on findings from the first stage of the formative research process: usability testing. The aim is to investigate the usability of a self-guided internet-delivered intervention in terms of function, content, and design.

### Part 1: Expert Evaluation

The first and last author of the paper (RMFK and YI) conducted an expert evaluation prior to the user evaluation. Through this evaluation method, we detected usability problems with the interface early in the process [20]. The expert evaluation involved an evaluation on how well a design supported its target audience in achieving their specified goals with effectiveness, efficiency, and satisfaction. The usability expert (YI) conducted a review according to a set of generally accepted usability guidelines, as well as their personal knowledge of the design's domain. The evaluation was based on Nielsen's heuristics. Problems detected by the evaluators were addressed

#### Textbox 1. Tasks.

1. Go to the program (adhd.youwell.no) and log in
2. Start module 1, read the material, and say out loud what you think
3. Start module 2 and try one of the breathing exercises
4. Log your experience of the breathing exercise in the exercise log
5. Complete module 2
6. Go to the home page
7. Fill out the questionnaires of module 3
8. Go to my diary and fill out an entry
9. Go to my calendar page and see this week's homework assignments
10. Try another breathing exercise and log your experience again

First, the participants provided demographic information and reported on their experience with personal computers and the use of internet. Next, they were asked to perform a series of tasks in the platform and intervention. Two facilitators were present in each usability session. One led the session and 1 observed. Participants were asked to "think aloud" (ie, provide continuous commentary) in accordance with Concurrent Think-aloud Method (CTA) while operating the system [21]. The goal with CTA was to encourage participants to keep a running stream of consciousness as they worked with the tasks. This helped to understand participants' thoughts as they tried to work through the intervention and elicit real-time feedback and emotional responses to the intervention. This is a common approach to usability testing that enables evaluation of the ease with which a system is learned [22]. Participants provided feedback on the MyADHD intervention (ie, layout, color, text, readability, and videos), and the content and provided information on their preferences regarding the use of MyADHD. If the participant asked a question, the facilitator remained neutral and replied with "What do you think?" or "What would you do?"

During the CTA, the observer took notes and used an audio recorder to capture all that the participant did and said. At the end of the session, the System Usability Scale [23] was administered as an online survey in the platform. The information retrieved from the think-aloud interviews was used to further optimize the MyADHD intervention.

immediately, so they did not influence the second part: the user evaluation.

### Part 2: User Evaluation

Five individuals diagnosed with ADHD were recruited via the Norwegian patient association (ADHD Norge) to participate in a 1-hour laboratory-based usability test. The inclusion criteria for the usability tests were (1) a self-reported ADHD diagnosis, (2) age >18 years, and (3) willing to participate and able to meet at the laboratory. In the laboratory, participants used a laptop and were asked to use the prototype of MyADHD intervention. The test entailed the completion of 10 goal-oriented tasks (see [Textbox 1](#)).

### Ethics Approval

This study was reviewed and approved by the Norwegian Regional Committee for Medical and Health Research Ethics, REC South East #203804. The participants provided written informed consent to participate in this study. As a reimbursement for their time, participants received gift vouchers worth NOK 200 (US \$19).

### Questionnaires

#### The System Usability Scale

The System Usability Scale (SUS) [23] is a widely used source for the assessment of the perceived usability of an evaluated system [24]. The SUS has sufficient reliability (coefficient  $\alpha$  .9). It contains 10 items with a response from 1 (strongly disagree) to 5 (strongly agree), with 5 positive statements and 5 negative statements. The items are scored on a 5-point Likert scale. The sum of the items leads to a general measure of perceived usability with the total score varying between 0 and 100. Based on different studies in the literature, an overall usability score of below 70 shows poor usability of the evaluated system [25,26]. Higher overall scores represent a high quality of usability. The scale allows item-based analysis by calculating the mean score and standard deviation of each item. Minor wording changes do not appear to affect SUS scores [25,27]. In our research, we tailored the scale by replacing the term "system" with "the intervention."

Additionally, the participants were asked questions on sociodemographic and internet use characteristics.

## The MyADHD Intervention

MyADHD development was theory based and person based [28]. The planning phase consisted of qualitative focus groups, the development of guiding principles, and a literature review. The first prototype of the intervention was delivered via an online secure portal, which is accessible on laptops and personal computers. The intervention is a structured self-guided intervention with modified elements from cognitive behavioral therapy, dialectical behavioral therapy, and goal management training to target specific challenges experienced by adults with ADHD. Kenter et al [18] described the development of the content of the intervention. The intervention consists of 7 training modules. In this study, we used the first 3 modules of the intervention (see [Multimedia Appendix 1](#)).

**Table 1.** Overview of the first 3 modules of MyADHD.

Module	Rationale and content	Exercises and videos
1. Start module	Goal setting and practical information about how to use the internet-delivered intervention	<ul style="list-style-type: none"> <li>• One goal setting exercise</li> <li>• Describe how ADHD<sup>a</sup> affects your life exercise</li> <li>• One lived-experience video</li> </ul>
2. Mindful awareness	Inattention is a core symptom of ADHD. In this module, participants are given information about different aspects of attention and concentration and how to cope with impairment. In this module, the participants start training on mindful awareness (“being here and now”) by focusing on their breathing (based on dialectical behavioral therapy).	<ul style="list-style-type: none"> <li>• Three different types of breathing exercises</li> <li>• Two lived-experience videos</li> </ul>
3. Inhibition training	Impulsivity and loss of impulse control are common among adults with ADHD. This module consists of exercises focusing on impulse control and goal-oriented/goal-directed behavior (stop, observe, proceed, and check; based on goal management training).	<ul style="list-style-type: none"> <li>• Two STOP exercises</li> <li>• Two lived-experience videos</li> </ul>

<sup>a</sup>ADHD: attention-deficit/hyperactivity disorder.

## Data Collection and Analysis

### Tasks and Questionnaires

In each individual usability testing session, following informed consent, the participants were first asked to think aloud while using the program and doing the tasks (see [Textbox 1](#)). The procedure was pretested with an individual who was not involved in the study. A facilitator guided the participant through the usability sessions by presenting the tasks and reminding the participants to think aloud by questions like “Tell me what you’re thinking,” “what are you looking for?,” and “What do you think now?” The sessions were audiotaped and transcribed verbatim. We coded and analyzed the transcripts using NVIVO [29].

### Data Analyses

The usability tests were recorded and transcribed verbatim by the second author. Transcripts were analyzed thematically using

The main goals of the intervention are to help participants obtain improved functioning in daily life activities and offer strategies that aim to reduce stress, reduce inattention, and improve quality of life. Each module includes psychoeducation alongside the text, audio, and video material instructing participants in the use of specific techniques. Further, modules include case vignettes and lived-experience videos that clarify the written content and help participants make connections between the material and their own experiences. Since all modules have the same structure and ingredients, we only used the first 3 modules in our test. [Table 1](#) shows an overview of the intervention modules that were used in this study.

an iterative coding procedure. The focus of the analysis was on the features of the intervention that needed to be redesigned or improved. The categories were identified using an iterative process of reading and rereading the transcripts. Usability issues were coded into categories.

## Results

### Participants

The participants had a mean age of 38.4 (SD 16.3) years. Three were men and 4 were highly educated (high school or higher). One participant reported difficulty using computers and the internet, while the others had reported good computer skills, see [Table 2](#). All participants had ADHD diagnoses, and none of them had previously used an internet-delivered intervention before.

**Table 2.** Demographics.

	Value
<b>Demographics</b>	
<b>Age (years)</b>	
Mean (SD)	38.4 (16.3)
Range	25-62
<b>Gender, n</b>	
Male	3
Female	2
<b>Educational level, n</b>	
Middle	1
High	4
<b>Internet and computer skills, n</b>	
Self-reported good computer skills	4
Self-reported good internet skills	4

## Usability Issues

The transcripts revealed 3 main categories of barriers that limited usability: (1) functionality, (2) content, and (3) layout.

### Functionality

This category referred to the need for the intervention to be easy to use, navigate, and have a logical flow. Analysis showed that usability was limited when navigation was difficult. The participants experienced problems related to, for example, unclear navigation (“I can’t move back to the homepage,” “where do I go to see my past entries?”), unclear functionality (“Does the calendar synchronize with my own phone?”), and it was unclear whether filled-out questionnaires were saved and submitted (“Did I save this now? I am not sure, can I go back? How do I check this?”). Furthermore, textboxes were too small to type in larger amounts of text.

In preparation of new rounds of usability testing, back and forward buttons were added, more explanation was offered with the different elements such as calendar, automatic feedback that answers on questionnaires were sent in, and textboxes were made larger. At the end of usability testing, the participants found the calendar still confusing to use. Due to this, the calendar was removed from the intervention, and the participants were encouraged to use their own private calendar, for example, their mobile phone calendar or paper calendar. After this adjustment, think-aloud comments to the functionality were more neutral (“OK, now I click here to go to the next page”), positive (“Nice! I did not remember what I was supposed to do, but I can go back to read the previous page”), or focused on the specific intervention content.

### Content

Overall, the participants had a general positive impression of the content in the program. The program felt “positive,” “light,” and “useful.” The exercises were perceived as helpful and useful (“This really speaks to me,” “this makes perfect sense”). Usability testing revealed that several participants (n=4) had

difficulty understanding the text when it was too lengthy, with large blocks of text, making the information difficult to process, which limited the usability.

A point of improvement involved the wording in the content, whereby language was perceived as too negative (“I do not recognize myself in having problems, I might have challenges, not problems”) and too complicated (“I have to read the most sentences twice”). There were also inconsistencies in wording that made it difficult to use the intervention (“first you call it dashboard, and now I have to go back to the homepage,” “it is called a program or an intervention? It seems now that these are two different things”).

This resulted in changes in the wording; more positive wording was used, which resulted in more positive comments about the wording (“It is nice with empathy about that it can be difficult to be part of this ADHD intervention”). Wording was checked for inconsistencies, spelling errors were removed, and text was shortened or chunked up in smaller sections so that it became easier to read. We also added a short summary at the end of each module. All participants liked the videos, exercises, and psychoeducational text and viewed them as helpful.

### Layout

All participants commented that they saw value in the images, visual aids, and videos. They liked the layout and described it as “calm,” “beautiful, yeah it is esthetic,” and “friendly” with “nice colors and pretty images.” Regarding the questionnaires, the participants found it hard to fill out the questionnaires because of the large amount of text and answering options.

### Satisfaction With the Intervention

All participants felt confident they would be able to use the platform (n=3 strongly agreed, n=2 agreed). The participants were satisfied with how easy it was to use and viewed the intervention as helpful. All participants were able to perform the tasks and learn to use the intervention and the platform on their own.

## Perceived Usability Outcomes

The results for perceived usability (as measured by the SUS) are presented in Table 3. The analysis of the data identified that

the intervention was rated with an average score of 75.5 (SD 5.9). Of the 5 participants, 1 rated the MyADHD intervention as “excellent,” while 4 others rated it as “good.”

**Table 3.** Result of the System Usability Scale.

Items	P1 <sup>a</sup>	P2	P3	P4	P5	Mean
1. I think that I would like to use this intervention	4	3	3	4	3	3.4
2. I found the intervention unnecessarily complex	2	2	2	1	1	1.6
3. I found the intervention easy to use	4	4	4	5	5	4.4
4. I think I would need support from a technical person to use this intervention	1	2	1	1	1	1.2
5. I found the various functions in this intervention were well integrated	3	4	4	3	3	3.4
6. I thought there was too much inconsistency in this intervention	3	2	2	3	3	2.6
7. I would imagine that most people would learn to use this intervention quickly	3	3	4	5	4	3.8
8. I found the intervention very cumbersome to use	2	2	2	1	2	1.8
9. I felt very confident using the intervention	3	5	3	4	5	4
10. I needed to learn a lot of things before I could start using this intervention	1	2	2	1	2	1.6
SUS <sup>b,c</sup> score (0-100 scale)	70	72.5	72.5	85	77.5	75.5

<sup>a</sup>P1-5: participant 1-5.

<sup>b</sup>SUS: System Usability Scale.

<sup>c</sup>Total score reversed to a 0-100 scale.

## Discussion

### Principal Results

Usability testing demonstrated that all individuals could perform the desired tasks, and that they learned to use the intervention quickly. Results from the SUS revealed that the program obtained high scores, indicating that the intervention was considered very useful. Staged iterative usability testing was essential for discovering intervention enhancement needs (eg, more visual aids, more buttons) and for resolving design, functionalities (eg, more feedback on actions and explanation to functionalities), and content inadequacy (eg, inconsistencies, difficult wording, lengthy text). Overall, the participants saw value in the MyADHD intervention and were confident they could use it for self-management of symptoms and expressed the desire to use the entire intervention when it becomes available.

Optimizing usability early in the process is a critical step in the development process of self-guided internet-delivered interventions [28], especially with self-guided interventions, where uptake and adherence are often low [4]. While there is a large number of studies examining the feasibility and efficacy of internet-delivered interventions for common mental health disorders like depression and anxiety [1], there are only a few studies that examine internet-delivered interventions for adults with ADHD [12,13,30]. These studies reported low adherence to the interventions. This is problematic as adherence is related to outcome [31,32]. Additionally, low usability could have an impact on the effectiveness of such interventions [33]. A study on usability is therefore important and relevant, as it may help to ensure that the interventions are well designed and therefore increase the interest and number of people who can benefit from

it. End users have high expectations of digital products, and because of difficulties with inattention, adults with ADHD specifically might drop out if the intervention does not have optimal usability. We did not find other studies examining the usability of internet-delivered interventions for adults with ADHD.

The overall SUS scores found in this study are comparable to other usability studies of internet-delivered interventions. A study of transdiagnostic internet-delivered treatment [34] found a mean SUS score of 81.89; a study on blended care [35] found a mean SUS score of 76.3 for the internet-delivered treatment, while a study on web-based support for people with mild intellectual disabilities [36] found a low mean SUS score of 56.4 and 51.1.

### Limitations and Strengths

For usability tests, 5 users are often enough to identify 80% of all usability problems [19]. The optimal number of participants required is not clear. This depends on the problems raised by the participants and whether the participants give an adequate reflection of potential end users. Despite the small sample, participants with different profiles and backgrounds were included, and within this sample, theme saturation was achieved. We believe that our sample reflects the population of potential end users of MyADHD.

The staged iterative usability tests provided knowledge regarding whether specific tasks could be performed in the intervention and gave direct input on how potential end users would use the intervention. The study was not conducted to identify every single usability problem but rather to show how usability testing with a small sample could identify usability problems that experts did not recognize prior, which allowed us to make

significant improvements to the intervention before going full trial. The methods used were effective in identifying elements that needed modification.

## Conclusions

Innovative technologies can play an important role in helping adults with ADHD manage their symptoms better. For such

interventions, delivered over the internet without clinician support, to be viable, they need to be developed with the needs, characteristics, and preferences of their intended end users in mind. At the end of a user-centered development process, with usability evaluations, the MyADHD intervention was deemed ready for testing in real-world conditions.

## Acknowledgments

This publication is part of the Introducing Mental Health Through Adaptive Technology (INTROMAT) project. The Research Council of Norway (NFR# 259293) funded the study. The funding body played no role in the design of the study, the collection, analysis, and interpretation of data; or in writing the manuscript.

## Authors' Contributions

RMFK is the chief investigator in this study, led the proposal and study protocol development, contributed to the analysis, and drafted the manuscript. AS contributed to the execution of the study, the analysis, and the manuscript. YI contributed to the study design, the development of the platform interface, and the drafting of the manuscript. All authors critically reviewed and approved the final version of the manuscript.

## Conflicts of Interest

None declared.

## Multimedia Appendix 1

Screenshot of the intervention.

[[DOCX File, 359 KB - formative\\_v6i10e37137\\_app1.docx](#)]

## References

1. Andersson G, Titov N, Dear BF, Rozental A, Carlbring P. Internet-delivered psychological treatments: from innovation to implementation. *World Psychiatry* 2019;18(1):20-28 [FREE Full text] [doi: [10.1002/wps.20610](https://doi.org/10.1002/wps.20610)] [Medline: [30600624](https://pubmed.ncbi.nlm.nih.gov/30600624/)]
2. Folker AP, Mathiasen K, Lauridsen SM, Stenderup E, Dozeman E, Folker MP. Implementing internet-delivered cognitive behavior therapy for common mental health disorders: a comparative case study of implementation challenges perceived by therapists and managers in five European internet services. *Internet Interv* 2018 Mar;11:60-70. [doi: [10.1016/j.invent.2018.02.001](https://doi.org/10.1016/j.invent.2018.02.001)]
3. 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)]
4. 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;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/)]
5. 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 Apr;45(13):2717-2726. [doi: [10.1017/s0033291715000665](https://doi.org/10.1017/s0033291715000665)]
6. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. Washington, DC: American Psychiatric Association; 2013:591-643.
7. Faraone SV, Biederman J, Mick E. The age-dependent decline of attention deficit hyperactivity disorder: a meta-analysis of follow-up studies. *Psychol Med* 2006 Feb;36(2):159-165. [doi: [10.1017/S003329170500471X](https://doi.org/10.1017/S003329170500471X)] [Medline: [16420712](https://pubmed.ncbi.nlm.nih.gov/16420712/)]
8. Franke B, Michelini G, Asherson P, Banaschewski T, Bilbow A, Buitelaar JK, et al. Live fast, die young? A review on the developmental trajectories of ADHD across the lifespan. *Eur Neuropsychopharmacol* 2018 Oct;28(10):1059-1088. [doi: [10.1016/j.euroneuro.2018.08.001](https://doi.org/10.1016/j.euroneuro.2018.08.001)]
9. Kessler RC, Adler LA, Barkley R, Biederman J, Conners CK, Faraone SV, et al. Patterns and predictors of attention-deficit/hyperactivity disorder persistence into adulthood: results from the National Comorbidity Survey Replication. *Biol Psychiatry* 2005 Jun;57(11):1442-1451. [doi: [10.1016/j.biopsych.2005.04.001](https://doi.org/10.1016/j.biopsych.2005.04.001)]
10. Spencer T, Biederman J, Wilens T. Pharmacotherapy of attention deficit hyperactivity disorder. *Child Adolesc Psychiatr Clin N Am* 2000 Jan;9(1):77-97. [Medline: [10674191](https://pubmed.ncbi.nlm.nih.gov/10674191/)]
11. Solberg BS, Haavik J, Halmøy A. Health care services for adults with ADHD: patient satisfaction and the role of psycho-education. *J Atten Disord* 2015 Jun;23(1):99-108. [doi: [10.1177/1087054715587941](https://doi.org/10.1177/1087054715587941)]

12. Moëll B, Kollberg L, Nasri B, Lindefors N, Kaldo V. Living SMART — a randomized controlled trial of a guided online course teaching adults with ADHD or sub-clinical ADHD to use smartphones to structure their everyday life. *Internet Interv* 2015 Mar;2(1):24-31. [doi: [10.1016/j.invent.2014.11.004](https://doi.org/10.1016/j.invent.2014.11.004)]
13. Nasri B. *New Approach to the Treatment of ADHD Through Internet and Smartphones: From Impairment to Improvement*. Sweden: Karolinska Institutet; 2017.
14. Nordby ES, Kenter RM, Lundervold AJ, Nordgreen T. A self-guided internet-delivered intervention for adults with ADHD: a feasibility study. *Internet Interv* 2021 Sep;25:100416. [doi: [10.1016/j.invent.2021.100416](https://doi.org/10.1016/j.invent.2021.100416)]
15. Torous J, Nicholas J, Larsen ME, Firth J, Christensen H. Clinical review of user engagement with mental health smartphone apps: evidence, theory and improvements. *Evid Based Ment Health* 2018 Aug;21(3):116-119. [doi: [10.1136/eb-2018-102891](https://doi.org/10.1136/eb-2018-102891)] [Medline: [29871870](https://pubmed.ncbi.nlm.nih.gov/29871870/)]
16. ISO 9241-11:2018: Ergonomics of Human-System Interaction - Part 11: Usability: Definitions and Concepts. International Organization for Standardization. 2018. URL: <https://www.iso.org/standard/63500.html> [accessed 2022-09-12]
17. Inal Y, Wake JD, Guribye F, Nordgreen T. Usability evaluations of mobile mental health technologies: systematic review. *J Med Internet Res* 2020 Jan;22(1):e15337 [FREE Full text] [doi: [10.2196/15337](https://doi.org/10.2196/15337)] [Medline: [31904579](https://pubmed.ncbi.nlm.nih.gov/31904579/)]
18. Kenter RMF, Lundervold AJ, Nordgreen T. A self-guided internet-delivered intervention for adults with ADHD: a protocol for a randomized controlled trial. *Internet Interv* 2021 Dec;26:100485 [FREE Full text] [doi: [10.1016/j.invent.2021.100485](https://doi.org/10.1016/j.invent.2021.100485)] [Medline: [34877262](https://pubmed.ncbi.nlm.nih.gov/34877262/)]
19. Nielsen J, Landauer KT. A mathematical model of the finding of usability problems. In: *Proceedings of the INTERACT'93 and CHI'93 Conference on Human Factors in Computing Systems*. 1993 Presented at: CHI '93; April 24-29, 1993; Amsterdam, The Netherlands. [doi: [10.1145/169059.169166](https://doi.org/10.1145/169059.169166)]
20. Nielsen J. Finding usability problems through heuristic evaluation. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 1992 Presented at: CHI'92; May 3-7, 1992; Monterey, CA. [doi: [10.1145/142750.142834](https://doi.org/10.1145/142750.142834)]
21. Fan M, Lin J, Chung C, Truong KN. Concurrent think-aloud verbalizations and usability problems. *ACM Trans Comput-Hum Interact* 2019;26(5):1-35. [doi: [10.1145/3325281](https://doi.org/10.1145/3325281)]
22. Brinck T, Gergle D, Wood SD. *Usability for the Web: Designing Web Sites That Work*. San Francisco, CA: Elsevier; 2001.
23. Brooke J. Sus: a quick and dirty usability. *Usability Eval Ind* 1996;189(3):4-7. [doi: [10.1201/9781498710411-35](https://doi.org/10.1201/9781498710411-35)]
24. Lewis JR. The System Usability Scale: past, present, and future. *Int J Human-Comput Interact* 2018 Mar;34(7):577-590. [doi: [10.1080/10447318.2018.1455307](https://doi.org/10.1080/10447318.2018.1455307)]
25. Bangor A, Kortum P, Miller J. Determining what individual SUS scores mean: adding an adjective rating scale. *J Usability Stud* 2009;4(3):114-123 [FREE Full text]
26. McLellan S, Muddimer A, Camille Peres S. The effect of experience on system usability scale ratings. *J Usability Stud* 2012;7(2):56-67 [FREE Full text]
27. Lewis JR, Sauro J. The factor structure of the system usability scale. In: *International Conference on Human Centered Design*. Springer; 2009 Presented at: HCI International 2009; July 19-24, 2009; San Diego, CA. [doi: [10.1007/978-3-642-02806-9\\_12](https://doi.org/10.1007/978-3-642-02806-9_12)]
28. Yardley L, Ainsworth B, Arden-Close E, Muller I. The person-based approach to enhancing the acceptability and feasibility of interventions. *Pilot Feasibility Stud* 2015;1:37 [FREE Full text] [doi: [10.1186/s40814-015-0033-z](https://doi.org/10.1186/s40814-015-0033-z)] [Medline: [27965815](https://pubmed.ncbi.nlm.nih.gov/27965815/)]
29. Castleberry A. NVivo 10 [software program]. Version 10. QSR International; 2012. *Am J Pharmaceutical Education* 2014 Feb;78(1):25-27. [doi: [10.5688/ajpe78125](https://doi.org/10.5688/ajpe78125)]
30. Pettersson R, Söderström S, Edlund-Söderström K, Nilsson KW. Internet-based cognitive behavioral therapy for adults With ADHD in outpatient psychiatric care. *J Atten Disord* 2016 Jul;21(6):508-521. [doi: [10.1177/1087054714539998](https://doi.org/10.1177/1087054714539998)]
31. Titov N, Dear B, Nielsens 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/)]
32. Simpson HB, Marcus SM, Zuckoff A, Franklin M, Foa EB. Patient adherence to cognitive-behavioral therapy predicts long-term outcome in obsessive-compulsive disorder. *J Clin Psychiatry* 2012 Sep;73(09):1265-1266. [doi: [10.4088/jcp.12107879](https://doi.org/10.4088/jcp.12107879)]
33. Zapata BC, Fernández-Alemán JL, Idri A, Toval A. Empirical studies on usability of mHealth apps: a systematic literature review. *J Med Syst* 2015 Jan;39(2):1. [doi: [10.1007/s10916-014-0182-2](https://doi.org/10.1007/s10916-014-0182-2)]
34. Díaz-García A, González-Robles A, Fernández-Álvarez J, Castilla D, Mira A, Bretón J. Usability of a transdiagnostic internet-delivered protocol for anxiety and depression in community patients. In: Ciproso P, Serino S, Villani D, editors. *Pervasive Computing Paradigms for Mental Health*. MindCare 2019. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering. Cham: Springer; 2019.
35. Etzelmueller A, Radkovsky A, Hannig W, Berking M, Ebert DD. Patient's experience with blended video- and internet based cognitive behavioural therapy service in routine care. *Internet Interv* 2018 Jun;12:165-175. [doi: [10.1016/j.invent.2018.01.003](https://doi.org/10.1016/j.invent.2018.01.003)]
36. de Wit J, Dozeman E, Ruwaard J, Alblas J, Riper H. Web-based support for daily functioning of people with mild intellectual disabilities or chronic psychiatric disorders: A feasibility study in routine practice. *Internet Interv* 2015 May;2(2):161-168. [doi: [10.1016/j.invent.2015.02.007](https://doi.org/10.1016/j.invent.2015.02.007)]

---

**Abbreviations**

**ADHD:** attention-deficit/hyperactivity disorder

**SUS:** System Usability Scale

---

*Edited by A Mavragani; submitted 08.02.22; peer-reviewed by H Mehdizadeh, V Prasad; comments to author 04.07.22; revised version received 30.08.22; accepted 30.08.22; published 21.10.22.*

*Please cite as:*

*Kenter RMF, Schønning A, Inal Y*

*Internet-Delivered Self-help for Adults With ADHD (MyADHD): Usability Study*

*JMIR Form Res 2022;6(10):e37137*

*URL: <https://formative.jmir.org/2022/10/e37137>*

*doi: [10.2196/37137](https://doi.org/10.2196/37137)*

*PMID: [36269662](https://pubmed.ncbi.nlm.nih.gov/36269662/)*

©Robin Maria Francisca Kenter, Adrian Schønning, Yavuz Inal. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 21.10.2022. 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 Development and Evaluation of “Life Age”—a Primary Prevention and Population-Focused Risk Communication Tool: Feasibility Study With a Single-Arm Repeated Measures Design

Adeogo Akinwale Olusan<sup>1,2\*</sup>, MD, MSc; Suzanne Barr<sup>2\*</sup>, RD, MSc, MEd, PhD; Mark Cobain<sup>2,3\*</sup>, BSc, PhD; Holly Whelan<sup>2,3\*</sup>, BSc, PhD

<sup>1</sup>Department of Cardiology, Belfast Health and Social Care (HSC) Trust, Belfast, United Kingdom

<sup>2</sup>Faculty of Medicine, National Heart and Lung Institute, Imperial College London, London, United Kingdom

<sup>3</sup>Younger Lives Limited, London, United Kingdom

\*all authors contributed equally

**Corresponding Author:**

Adeogo Akinwale Olusan, MD, MSc

Department of Cardiology

Belfast Health and Social Care (HSC) Trust

274 Grosvenor Road,

Belfast, BT12 6BA

United Kingdom

Phone: 44 7423161441

Email: [aakinwale@yahoo.co.uk](mailto:aakinwale@yahoo.co.uk)

## Abstract

**Background:** Communicating cardiovascular risk to the general population requires forms of communication that can enhance risk perception and stimulate lifestyle changes associated with reduced cardiovascular risk.

**Objective:** The aim of this study was to evaluate the motivational potential of a novel lifestyle risk assessment (“Life Age”) based on factors predictive of both premature mortality and psychosocial well-being.

**Methods:** A feasibility study with a single-arm repeated measures design was conducted to evaluate the potential efficacy of Life Age on motivating lifestyle changes. Participants were recruited via social media, completed a web-based version of the Life Age questionnaire at baseline and at follow-up (8 weeks), and received 23 e-newsletters based on their Life Age results along with a mobile tracker. Participants’ estimated Life Age scores were analyzed for evidence of lifestyle changes made. Quantitative feedback of participants was also assessed.

**Results:** In total, 18 of 27 participants completed the two Life Age tests. The median baseline Life Age was 1 year older than chronological age, which was reduced to –1.9 years at follow-up, representing an improvement of 2.9 years ( $P=.02$ ). There were also accompanying improvements in Mediterranean diet score ( $P=.001$ ), life satisfaction ( $P=.003$ ), and sleep ( $P=.05$ ). Quantitative feedback assessment indicated that the Life Age tool was easy to understand, helpful, and motivating.

**Conclusions:** This study demonstrated the potential benefit of a novel Life Age tool in generating a broad set of lifestyle changes known to be associated with clinical risk factors, similar to “Heart Age.” This was achieved without the recourse to expensive biomarker tests. However, the results from this study suggest that the motivated lifestyle changes improved both healthy lifestyle risks and psychosocial well-being, consistent with the approach of Life Age in merging the importance of a healthy lifestyle and psychosocial well-being. Further evaluation using a larger randomized controlled trial is required to fully evaluate the impact of the Life Age tool on lifestyle changes, cardiovascular disease prevention, and overall psychosocial well-being.

(*JMIR Form Res* 2022;6(10):e37385) doi:[10.2196/37385](https://doi.org/10.2196/37385)

**KEYWORDS**

Life Age; Heart Age; cardiovascular risk; lifestyle change; psychosocial well-being; population focused risk communication tool; health promotion; risk perception; premature mortality; risk communication

## Introduction

### Background

Age-based approaches to risk communication have become popular in recent years. The concept of “Heart Age” was previously developed in 2007 [1] to enhance the perception of personal cardiovascular risk in those who would benefit from lifestyle change [2], thereby leading to risk factor reduction [3]. Heart Age has subsequently been adopted by national health organizations [4,5], and its novelty is in being rooted in validated risk models but with an output format that is engaging for users. However, the requirement for biometric factors (eg, blood cholesterol, blood glucose, and blood pressure assessments) in risk models can prevent widespread adoption and potentially lead to misclassification of risk [6].

### Development of the “Life Age” concept

Using lifestyle factors, as opposed to clinical risk factors, in age-based tests could be more useful in promoting cardiovascular health than clinical risk factors in the wider population. In addition to communicating disease risks, communicating on psychosocial well-being factors that constitute the broader lifestyle (life satisfaction, perceived stress, insomnia, and positive and negative mood) may be of equal importance.

Several large observational studies across diverse populations have demonstrated the additive impact of a common set of lifestyle factors on all-cause mortality [7]. Thus, we developed a lifestyle-based assessment of mortality risk by combining published relative risks of all-cause mortality for body weight, physical activity, alcohol, adherence to a Mediterranean diet, and smoking into an overall lifestyle risk score. Relative risks were converted to “effective age” scores for each lifestyle factor using the method devised by Spiegelhalter et al [8] and represented the modifiable mortality risk component of “Life Age.” The range of years that could feasibly be added or removed from a person’s age ranged from –6 years (BMI<23, nonsmoker, no or low alcohol, optimal diet, and high physical activity) to +28 years (BMI>40, smoker, binge drinker, poor diet, and sedentary lifestyle). However, among previous web-based users (n=2000, unpublished), 95% of data fell within a range of –16 to +16 years.

To create the psychosocial well-being component of Life Age, questionnaire scores for life satisfaction, positive and negative mood, hours of sleep, and perceived stress were converted into years (ranging from –2 to +2 years) and based on the distance of score from population norms. This ensured that Life Age scores for psychosocial well-being equalled the impact of healthy lifestyle factors.

### Objective

The aim of this study was to evaluate the motivational potential of a novel lifestyle risk assessment (Life Age) based on factors predictive of both premature mortality and psychosocial well-being.

## Methods

### Study Design

A pilot nonrandomized, single-arm, repeated measures feasibility study (unpublished) on postgraduate students at Imperial College London (London, United Kingdom) examined the impact of taking the web-based version of the Life Age questionnaire without any intervention. We observed a 1.3-year improvement in participants’ Life Age after 8 weeks ( $P=.006$ ). To further evaluate the impact of Life Age on lifestyle change, we conducted a feasibility study using a nonrandomized interventional, single-arm, repeated measures design.

### Ethics Approval

Ethics approval was provided by the Imperial College Research Ethics Committee on April 25, 2018 (18IC4516). The study was carried out in accordance with the tenets of the Declaration of Helsinki. All participants provided consent for data follow-up and publication.

### Participant Recruitment, Newsletters, Mobile Tracker, and Web-Based Feedback Assessment

Participants were recruited via social media. Interested participants completed a web-based version of the Life Age questionnaire at baseline and at follow-up (8 weeks apart). They received 23 e-newsletters based on their Life Age results through email (between weeks 1 and 7) along with a downloadable mobile tracker. The newsletters were developed by the “Younger Lives” expert team, and content was standardized, validated, and centered around the following: (1) setting up participants’ personalized targets on the basis of their initial Life Age score and how to achieve these targets; (2) setting up their environment; that is, setting up the kitchen, particularly a Mediterranean style diet also while eating out, <14 units of weekly alcohol intake, exercising by doing at least 150 minutes of moderate-intensity activity or at least 75 minutes of vigorous-intensity activity weekly [9], self-monitoring using bathroom scales and tape measures for weight and waist measurements, respectively, fitness tracker pedometers on mobile phones or wrist watches, and setting up the bedroom for a good sleep routine; (3) working on daily step counts, diet, being appreciative of good things in life, and understanding and protecting one’s emotional well-being; (4) advice on daily tracking of activities to create lifelong habits; (5) advice on the importance of positive thinking; (6) advice on tips to managing stress at nights; and (7) advice on maximizing health and happiness. Based on their Life Age, the newsletters were formulated to stimulate lifestyle changes based on the distance from recommended lifestyle behaviors in conjunction with their psychosocial well-being; these were provided to encourage them to “get younger” through lifestyle change, and the mobile tracker helped to monitor progress.

The downloadable mobile tracker helped to simplify overall progress tracking. It sets targets on the basis of participants’ Life Age results and maintains a daily log of participants’ body weight, step counts, activity, and sleep. However, participants had to complete a quick 1-minute check-in at the end of each day and also record their weight and waist circumference once

a week. The application also helps them to understand if their overall lifestyle has been aging or helping to keep them “young.”

Participants’ estimated Life Age scores were analyzed at baseline and at follow-up for lifestyle change and whether change in health was related to change in psychosocial factors. A quantitative feedback assessment using a web-based questionnaire on ease of use and understanding, motivation, and self-reported lifestyle changes was also conducted.

### Sample Size and Statistical Analysis

The sample size calculation used for the previous pilot study was based on the assumption that pre- and post-2-month intervention might change the mean Life Age by approximately 0.5 years, and the range of change from the participant who decreased his/her Life Age the most and the least is 3 years. Based on these estimates, the sample size was calculated using an SD of 0.75 with a significance level of 5% and power of 90%, and a sample size of 30 participants was agreed on after correcting for a 20% estimated dropout rate. The same sample size was thus used in this feasibility study.

Statistical analysis was performed using Stata software (StataCorp). Parametric and nonparametric data were analyzed using a paired *t* test and Wilcoxon signed-rank test, respectively. A paired *t* test was used to assess differences among Mediterranean diet score, life satisfaction, and combined mood score, whereas the Wilcoxon signed-rank test was used to assess the differences in Life Age, BMI, body weight, physical activity, perceived stress level, and sleep.

The Mediterranean diet score is based on a 14-item score, with a low adherence score being <7, and a higher adherence score being ≥8 according to the Prevención con Dieta Mediterránea (PREDIMED) trial [10]. With regard to mood, higher the

positive mood score, better the outcome, and vice versa for negative mood score. The positive mood scores were based on a scale from 1 to 5 points, and its components include being proud of oneself, alert, inspired, determined, attentive, active, interested, excited, strong, and enthusiastic. The negative mood scores were also based on a scale from 1 to 5 points, and its components include being irritable, ashamed, nervous, jittery, afraid, distressed, upset, guilty, scared, and hostile. However, owing to the need for a combined mood variable within our statistical analysis, we proposed the following formula: combined mood = positive mood + (negative mood × -1).

Throughout the data analysis, a *P* value of ≤.05 was considered statistically significant.

### Inclusion Criteria

Participants aged between 30 and 60 years, who signed the consent form, completed the web-based Life Age test, provided self-reported body measurements, have a good understanding of the English language, live within the United Kingdom, and use an iPhone owing to mobile tracker compatibility were included in the study.

## Results

### Sample Characteristics

Between April and May 2018, a total of 27 eligible individuals were enrolled in the study at baseline. At baseline, their average chronological age was 37 years, Life Age was 1 year older than the chronological age, BMI was 24.2 kg/m<sup>2</sup>, waist circumference was 81.3 cm, body weight was 68 kg, amount of physical activity per week was 13.3 metabolic equivalents of task (METs) per hour, and 17 (63%) of them were female. Other baseline characteristics are shown in [Table 1](#).

**Table 1.** Characteristics of participants between baseline and follow-up.

Characteristics	Baseline (n=27)	Follow up (n=18)	P value
Female sex, n (%)	17 (63)	13 (72)	N/A <sup>a</sup>
Age (years), median (IQR)	37 (30 to 56)	37 (30 to 57)	N/A
BMI (kg/m <sup>2</sup> ), median (IQR)	24.2 (19.4 to 34.7)	22.8 (19.5 to 31.8)	.21
BMI<25 kg/m <sup>2</sup> , n (%)	15 (56)	12 (67)	N/A
BMI=25-29.9 kg/m <sup>2</sup> , n (%)	7 (26)	4 (22)	N/A
BMI>30 kg/m <sup>2</sup> , n (%)	5 (18)	2 (11)	N/A
Waist circumference (cm), median (IQR)	81.3 (67 to 111)	77.5 (50 to 101.6)	.20
Weight (kg), median (IQR)	68 (51 to 116)	65 (50 to 93)	.18
<b>Smoking status, n (%)</b>			
Never smoked	20 (74)	15 (83)	N/A
Ex-smoker	7 (26)	3 (17)	N/A
Current smoker	0 (0)	0 (0)	N/A
Nutrition (units), mean (SD)	7.1 (1.8)	8.5 (1.8)	.001
Weekly physical activity, median (IQR)	13.3 (0 to 43.3)	16.7 (0 to 38)	.12
Life age (years), median (IQR)	1 (-5 to 12.75)	-1.9 (-6.3 to 11.5)	.02
Life satisfaction (units), mean (SD)	30.6 (6.3)	34.2 (5.2)	.003
Stress level, median (IQR)	15 (5 to 27)	14 (9 to 31)	.06
Sleep (hours), median (IQR)	6.5 (4.5 to 8)	7 (5 to 8)	.05
<b>Mood score</b>			
Positive mood, mean (SD)	35.7 (6.4)	37.5 (6.0)	.22
Negative mood, median (IQR)	17 (10 to 36)	15.5 (10 to 27)	.21

## Main Findings

In total, 18 of 27 participants completed both Life Age tests, which were separated by an 8-week interval. At follow-up, the median Life Age was -1.9 years, which was younger than the chronological age, representing a 2.9-year reduction in Life Age over 8 weeks ( $P=.02$ ). An analysis of individual risk factor components revealed an improvement of 1.4 units in the Mediterranean diet score ( $P=.001$ ), 3.6-unit increase in life satisfaction ( $P=.003$ ), and a 0.5-hour increase in sleep ( $P=.05$ ), whereas perceived stress levels improved to a degree approaching statistical significance ( $P=.06$ ; Table 1). Clinically relevant improvements were observed in BMI (1.4 kg/m<sup>2</sup>), waist circumference (3.8 cm), body weight (3 kg), physical activity (3.4 METS per hour), and mood (3.3 units); however, these failed to reach the preassigned level of statistical significance. There was a high dropout rate, which resulted in a smaller sample at the end of the study, which affected some secondary outcomes including BMI, weight, physical activity, and mood.

## Web-Based Feedback Assessment Findings

Web-based feedback assessment from 16 of 18 participants who completed the study revealed that 10 of 16 (63%) participants felt that the lifestyle recommendations provided were helpful, 9 (56%) would recommend the Life Age tool to their friends and family, 11 (69%) found the mobile tracker easy to understand and user friendly, 9 (56%) found the newsletters

informative and user friendly, 11 (69%) would recommend the use of the mobile tracker and newsletter to their friends and family, 11 (69%) were motivated to change their lifestyle via the use of newsletters, 12 (75%) found the dietary advice within the newsletters most useful, and 11 (69%) felt confident to continue with their lifestyle changes after the study ended.

## Discussion

### Principal Findings and Comparison With Previous Work

This feasibility study explored the development and impact of a novel lifestyle risk assessment tool, Life Age, on lifestyle changes. We observed a significant improvement in the overall Life Age metric, adherence to Mediterranean diet, life satisfaction, and sleep. Furthermore, there were clinically relevant improvements in perceived stress level, BMI, waist circumference, body weight, physical activity, and mood, but these failed to reach the preassigned level of statistical significance, in part owing to high variability at baseline and lack of power in the study. The improvement in Life Age was in agreement with the findings of the initial pilot study in 2017 (unpublished), which reported a median improvement of 1.3 years ( $P=.006$ ).

### Comparison Between Different Age-Based Approaches: Life Age and Heart Age

Although Heart Age was developed to enhance cardiovascular risk perception with the aim of facilitating lifestyle changes and subsequent risk factor reduction, it has been adopted by national health organizations [1-5]; however, a recently published systematic review of 16 studies assessed the effects of Heart Age intervention and reported that absolute risk could be reduced but with minimal evidence that it motivates lifestyle behavior. When combined with behavioral change strategies, there is evidence that it can improve clinical outcomes [11]. In comparison, this feasibility study on Life Age reported significant lifestyle behavior changes, particularly in the adherence to a Mediterranean diet and improvement in some psychosocial factors such as life satisfaction and sleep facilitated by the use of e-newsletters and a mobile tracker.

### Comparison With Known Evidence on Mediterranean Diet, Life Satisfaction, and Sleep

Adherence to a Mediterranean diet was encouraged throughout this study, and there was an improvement in the Mediterranean diet score by 1.4 units from 7.1 to 8.5 units ( $P=.001$ ). A score of  $\geq 8$  indicates a higher level of adherence to the Mediterranean diet [10]. Evidence from the PREDIMED trial and the Health, Alcohol and Psychosocial factors In Eastern Europe (HAPIEE) study have demonstrated that increased adherence to this diet reduces major adverse cardiovascular events [12,13]. Similarly, there was an improvement in life satisfaction score by 3.6 units from 30.6 to 34.2 units ( $P=.003$ ) at the end of this study. Few studies have shown an association between low life satisfaction and an increased risk of cardiovascular disease [14,15]. Furthermore, there was an improvement in average sleep duration from 6.5 hours to 7 hours at the end of this study ( $P=.05$ ). Evidence from systematic reviews has shown that both short and long sleep durations ( $<7$  hours and  $>7$  hours, respectively) are associated with an increased risk of cardiovascular events and all-cause mortality [16,17].

### Limitations and Strengths

A limitation in this study was the potential for under- or overestimation of self-reported body measurements. Attempts to minimize the impact of measurement bias was addressed by asking participants to provide measurements of recent visits to general practitioners, surgeons, or the gymnasium. A further limitation was a higher dropout rate and a smaller sample size

at the end of the trial, which affected some secondary outcomes such as BMI, weight, physical activity, and mood. However, as this was a feasibility study, these limitations are likely to be addressed in a larger randomized trial. Finally, the total sample is small, and participants reported health parameters that were healthier than the average population, potentially limiting the generalizability of this study. Nevertheless, this study has several strengths: first, minimal resources were used in this feasibility study, thus making it cost-effective at this stage of development. Second, the use of lifestyle assessments instead of clinical risk factors reduced the burden on participants in this feasibility study and no harm was encountered. Third, although this was the second study evaluating the impact of the use of this novel Life Age tool, it is the first to combine the Life Age assessment with follow-up material such as newsletters and mobile tracker. This enabled us to evaluate the longer-term impact on participant behavior. Finally, feedback assessment showed that a significant proportion of the participants found the intervention to be useful and user friendly.

### Future Direction

Based on the findings of this feasibility study, we propose a larger randomized controlled trial to fully evaluate the longer-term impact of the Life Age tool on lifestyle changes and risk factors, in addition to a head-to-head comparison with risk factor-based tests such as Heart Age to understand whether lifestyle factors or clinical risk factors are modified equally or differently by different approaches.

### Conclusions

This study demonstrated the potential benefit of a novel Life Age tool in generating a broad set of lifestyle changes known to be associated with clinical risk factors similar to Heart Age. This was achieved without the recourse to expensive biomarker tests. However, results from this study suggested that the motivated lifestyle changes improved both healthy lifestyle risks and psychosocial well-being, consistent with the approach of Life Age in merging the importance of a healthy lifestyle and psychosocial well-being. Further evaluation using a larger randomized controlled trial is required to fully evaluate the impact and relative merit of the Life Age tool on lifestyle changes, cardiovascular risk factors, and overall psychosocial well-being. Comparison of this assessment versus commonly used risk assessments that include biomarkers can help identify the value associated with the noninvasive approach to risk assessment.

### Acknowledgments

AAO acknowledges the support and input of SB, MC, and HW in ensuring that this project was successful. The abstract of this paper was presented at the 2019 EuroPrevent conference.

### Data Availability

The data sets generated or analyzed during this study are available on request from the corresponding author, these data are not publicly available owing to privacy and ethical restrictions.

## Authors' Contributions

AAO was the lead in study conceptualization; formal analysis; and writing, reviewing, and editing of this manuscript, as this was part of his dissertation work for the completion of an MSc degree in preventive cardiology at the Imperial College in 2018. SB, MC, and HW were supporting cosupervisors during this period.

## Conflicts of Interest

MC is one of the developers of Heart Age. MC and HW work for Younger Lives Limited, which provides health and well-being services, including services that include the Life Age test. AAO and SB report no conflicts of interest.

## References

1. D'Agostino RB, Vasan RS, Pencina MJ, Wolf PA, Cobain M, Massaro JM, et al. General cardiovascular risk profile for use in primary care. *Circulation* 2008 Feb 12;117(6):743-753. [doi: [10.1161/circulationaha.107.699579](https://doi.org/10.1161/circulationaha.107.699579)]
2. Soureti A, Hurling R, Murray P, van Mechelen W, Cobain M. Evaluation of a cardiovascular disease risk assessment tool for the promotion of healthier lifestyles. *Eur J Cardiovasc Prev Rehabil* 2010 Oct;17(5):519-523. [doi: [10.1097/hjr.0b013e328337ccd3](https://doi.org/10.1097/hjr.0b013e328337ccd3)]
3. Lopez-Gonzalez AA, Aguilo A, Frontera M, Bennasar-Veny M, Campos I, Vicente-Herrero T, et al. Effectiveness of the Heart Age tool for improving modifiable cardiovascular risk factors in a Southern European population: a randomized trial. *Eur J Prev Cardiol* 2015 Mar 03;22(3):389-396. [doi: [10.1177/2047487313518479](https://doi.org/10.1177/2047487313518479)] [Medline: [24491403](https://pubmed.ncbi.nlm.nih.gov/24491403/)]
4. Patel RS, Lagord C, Waterall J, Moth M, Knapton M, Deanfield JE. Online self-assessment of cardiovascular risk using the Joint British Societies (JBS3)-derived heart age tool: a descriptive study. *BMJ Open* 2016 Sep 28;6(9):e011511 [FREE Full text] [doi: [10.1136/bmjopen-2016-011511](https://doi.org/10.1136/bmjopen-2016-011511)] [Medline: [27683512](https://pubmed.ncbi.nlm.nih.gov/27683512/)]
5. Yang Q, Zhong Y, Ritchey M, Cobain M, Gillespie C, Merritt R, et al. Vital signs: predicted heart age and racial disparities in heart age among U.S. adults at the state level. *MMWR Morb Mortal Wkly Rep* 2015 Sep 04;64(34):950-958 [FREE Full text] [doi: [10.15585/mmwr.mm6434a6](https://doi.org/10.15585/mmwr.mm6434a6)] [Medline: [26335037](https://pubmed.ncbi.nlm.nih.gov/26335037/)]
6. Neufingerl N, Cobain MR, Newson RS. Web-based self-assessment health tools: who are the users and what is the impact of missing input information? *J Med Internet Res* 2014 Sep 26;16(9):e215 [FREE Full text] [doi: [10.2196/jmir.3146](https://doi.org/10.2196/jmir.3146)] [Medline: [25261155](https://pubmed.ncbi.nlm.nih.gov/25261155/)]
7. Khaw K, Wareham N, Bingham S, Welch A, Luben R, Day N. Correction: combined impact of health behaviours and mortality in men and women: the EPIC-Norfolk prospective population study. *PLoS Med* 2008 Mar 25;5(3):e70. [doi: [10.1371/journal.pmed.0050070](https://doi.org/10.1371/journal.pmed.0050070)]
8. Spiegelhalter D. How old are you, really? Communicating chronic risk through 'effective age' of your body and organs. *BMC Med Inform Decis Mak* 2016 Aug 05;16(1):104 [FREE Full text] [doi: [10.1186/s12911-016-0342-z](https://doi.org/10.1186/s12911-016-0342-z)] [Medline: [27496144](https://pubmed.ncbi.nlm.nih.gov/27496144/)]
9. Piepoli MF, Hoes AW, Agewall S, Albus C, Brotons C, Catapano AL, ESC Scientific Document Group. *Eur Heart J* 2016 Aug 01;37(29):2315-2381 [FREE Full text] [doi: [10.1093/eurheartj/ehw106](https://doi.org/10.1093/eurheartj/ehw106)] [Medline: [27222591](https://pubmed.ncbi.nlm.nih.gov/27222591/)]
10. Martínez-González MA, García-Arellano A, Toledo E, Salas-Salvadó J, Buil-Cosiales P, Corella D, PREDIMED Study Investigators. A 14-item Mediterranean diet assessment tool and obesity indexes among high-risk subjects: the PREDIMED trial. *PLoS One* 2012 Aug 14;7(8):e43134 [FREE Full text] [doi: [10.1371/journal.pone.0043134](https://doi.org/10.1371/journal.pone.0043134)] [Medline: [22905215](https://pubmed.ncbi.nlm.nih.gov/22905215/)]
11. Bonner C, Batcup C, Cornell S, Fajardo MA, Hawkes AL, Trevena L, et al. Interventions using heart age for cardiovascular disease risk communication: systematic review of psychological, behavioral, and clinical effects. *JMIR Cardio* 2021 Nov 05;5(2):e31056 [FREE Full text] [doi: [10.2196/31056](https://doi.org/10.2196/31056)] [Medline: [34738908](https://pubmed.ncbi.nlm.nih.gov/34738908/)]
12. Estruch R, Ros E, Salas-Salvadó J, Covas M, Corella D, Arós F, et al. Primary prevention of cardiovascular disease with a Mediterranean diet supplemented with extra-virgin olive oil or nuts. *N Engl J Med* 2018 Jun 21;378(25):e34. [doi: [10.1056/nejmoa1800389](https://doi.org/10.1056/nejmoa1800389)]
13. Stefler D, Malyutina S, Kubinova R, Pajak A, Peasey A, Pikhart H, et al. Mediterranean diet score and total and cardiovascular mortality in Eastern Europe: the HAPIEE study. *Eur J Nutr* 2017 Feb 17;56(1):421-429 [FREE Full text] [doi: [10.1007/s00394-015-1092-x](https://doi.org/10.1007/s00394-015-1092-x)] [Medline: [26578528](https://pubmed.ncbi.nlm.nih.gov/26578528/)]
14. Baumann M, Tchicaya A, Lorentz N, Le Bihan E. Life satisfaction and longitudinal changes in physical activity, diabetes and obesity among patients with cardiovascular diseases. *BMC Public Health* 2017 Dec 02;17(1):925 [FREE Full text] [doi: [10.1186/s12889-017-4925-0](https://doi.org/10.1186/s12889-017-4925-0)] [Medline: [29197375](https://pubmed.ncbi.nlm.nih.gov/29197375/)]
15. Baumann M, Tchicaya A, Vanderpool K, Lorentz N, Le Bihan E. Life satisfaction, cardiovascular risk factors, unhealthy behaviours and socioeconomic inequality, 5 years after coronary angiography. *BMC Public Health* 2015 Jul 15;15:668 [FREE Full text] [doi: [10.1186/s12889-015-2047-0](https://doi.org/10.1186/s12889-015-2047-0)] [Medline: [26174092](https://pubmed.ncbi.nlm.nih.gov/26174092/)]
16. Yin J, Jin X, Shan Z, Li S, Huang H, Li P, et al. Relationship of sleep duration with all-cause mortality and cardiovascular events: a systematic review and dose-response meta-analysis of prospective cohort studies. *JAMA* 2017 Sep 22;6(9). [doi: [10.1161/jaha.117.005947](https://doi.org/10.1161/jaha.117.005947)]

17. Cappuccio F, Cooper D, D'Elia L, Strazzullo P, Miller M. Sleep duration predicts cardiovascular outcomes: a systematic review and meta-analysis of prospective studies. *Eur Heart J* 2011 Jun;32(12):1484-1492. [doi: [10.1093/eurheartj/ehr007](https://doi.org/10.1093/eurheartj/ehr007)] [Medline: [21300732](https://pubmed.ncbi.nlm.nih.gov/21300732/)]

## Abbreviations

**HAPIEE:** Health, Alcohol and Psychosocial factors In Eastern Europe

**METS:** metabolic equivalents of task

**PREDIMED:** Prevención con Dieta Mediterránea

*Edited by T Leung; submitted 18.02.22; peer-reviewed by N Maglaveras, C Bonner, M Vidal dos Santos, P Hwang, A King; comments to author 24.07.22; revised version received 15.09.22; accepted 13.10.22; published 24.10.22.*

*Please cite as:*

*Olusan AA, Barr S, Cobain M, Whelan H*

*The Development and Evaluation of “Life Age”—a Primary Prevention and Population-Focused Risk Communication Tool: Feasibility Study With a Single-Arm Repeated Measures Design*

*JMIR Form Res* 2022;6(10):e37385

URL: <https://formative.jmir.org/2022/10/e37385>

doi: [10.2196/37385](https://doi.org/10.2196/37385)

PMID: [36279163](https://pubmed.ncbi.nlm.nih.gov/36279163/)

©Adeogo Akinwale Olusan, Suzanne Barr, Mark Cobain, Holly Whelan. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 24.10.2022. 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

# An Automated Surveillance System (SurCeGGID) for the French Sexually Transmitted Infection Clinics: Epidemiological Monitoring Study

Ndeindo Ndeikoundam Ngangro<sup>1</sup>, MPH, MD, PhD; Corinne Pioche<sup>1</sup>, MSc; Sophie Vaux<sup>1</sup>, PharmD, MPH; Delphine Viriot<sup>1</sup>, MPH; Julien Durand<sup>1</sup>, MSc; Bénédicte Berat<sup>1</sup>, MSc; Mohammed Hamdaoui<sup>1</sup>, MSc; Florence Lot<sup>1</sup>, MPH, MD

Infectious Diseases Department, Sante publique France, Saint-Maurice, France

**Corresponding Author:**

Ndeindo Ndeikoundam Ngangro, MPH, MD, PhD

Infectious Diseases Department

Sante publique France

12 Rue du Val d'Osne

Saint-Maurice, 94410

France

Phone: 33 (0)141796804

Email: [ndeindo.ndeikoundam@santepubliquefrance.fr](mailto:ndeindo.ndeikoundam@santepubliquefrance.fr)

## Abstract

**Background:** Viral and bacterial sexually transmitted infections (STIs) are public health concerns worldwide, but surveillance systems are not comprehensive enough to design and monitor accurately STI control strategies in most countries. In 2016, 320 STI clinics (CeGIDDs in French) were implemented in France, primarily targeting most exposed populations, although access is free of charge for anybody.

**Objective:** This article describes the mandatory surveillance system (SurCeGIDD) based on CeGIDD's individual data aiming to better guide STI prevention.

**Methods:** A decree ensured the use of software to manage consultations in CeGIDDs and to transfer surveillance data. A web service was implemented to secure data transfer from CeGIDDs' software to a centralized database. CeGIDDs can also transfer data in CSV format via a secured data-sharing platform. Then, data are automatically checked before integration. Data on sociodemographic variables, sexual exposure, blood exposure, symptoms, STI tests, STI diagnoses, and sexual health services delivery were collected for the previous year (n-1). Preliminary and descriptive analyses of 2017-2018 data transmitted in 2018 and 2019, respectively, were performed using numbers and proportions for qualitative variables.

**Results:** In 2017, 54/320 (16.9%) CeGIDDs transmitted their data. In 2018, this number of participants increased to 143/320 (44.7%) CeGIDDs. The corresponding volume of records increased from 2414 in 2017 to 382,890 in 2018. In 2018, most attendances were hospital based (263,480/382,890, 68.81%). In 2018, attendees were mostly men 227,326/379,921 (59.84%), while 151,963/379,921 (40%) were women 632/379,921 (0.17%) transgenders. The median age was 27 years for men, 23 years for women, and 30 years for transgender. Half of the attendees (81,964/174,932, 46.85%) were heterosexual men, 69,016/174,932 (39.45%) heterosexual women, 20,764/174,932 (11.87%) men who have sex with men, and 3188/174,932 (1.82%) women who have sex with women. A majority of them were born in France (227,698/286,289, 79.53%) and unemployed 115,913/211,707 (54.75%). The positivity rates were 0.37% for 205,348 HIV serologies, 1.31% for 131,551 hepatitis B virus serologies, 7.16% for 161,241 *Chlamydia trachomatis* PCR, 2.83% for 146,649 gonorrhea PCR, 1.04% for the syphilis combination of treponema and nontreponema serologies, and 5.96% for 13,313 *Mycoplasma genitalium* PCR.

**Conclusions:** Despite challenges, the effectiveness of the SurCeGIDD surveillance based on routine patients' records was demonstrated. The wide range of information, including socioeconomic determinants, might help to better guide and evaluate the prevention policies and services delivery. However, the growing volumes of information will require adapted tools and algorithms for the data management and analyses.

(JMIR Form Res 2022;6(10):e31136) doi:[10.2196/31136](https://doi.org/10.2196/31136)

**KEYWORDS**

HIV; hepatitis B; hepatitis C; STI; sexual health; epidemiology; surveillance



## Introduction

Viral and bacterial sexually transmitted infections (STIs) are public health concerns worldwide because of their high frequency, transmissibility, morbidity, potential complications (infertility, pelvic inflammatory diseases, cancers due to the human papillomavirus [HPV], acquired immunodeficiency due to the HIV), attributable mortality, the multidrug-resistance threat, the risk of congenital transmission, and the impact on the sexual health in the absence of an effective treatment or prevention [1-15].

In 2020, 37.7 million people were infected with HIV and 680,000 deaths were attributable to this infection globally [16,17]. There were 570,000 new cervical cancer cases and about 311,000 deaths attributable to the HPV in 2020 [16,18]. In 2019, 296 million people were infected with chronic hepatitis B, with 1.5 million newly acquired infections and 500,000 liver cancer cases associated with hepatitis B virus (HBV) and hepatitis C virus (HCV) [16,18]. The number of new cases of STI was estimated at 374 million per year and more than 1 million per day [16]. In 2020, 2.3 million deaths related to HIV, hepatitis, and STI worldwide have been reported [16].

In response, the public health strategies aiming to reach the 2030 Sustainable Development Goals (SDG) by eliminating or controlling these epidemics rely on evidence-based interventions [16]. Therefore, the most relevant epidemiological information is required to guide, implement, and evaluate these responses, monitoring progress toward the SDG. Nevertheless, the representativeness of surveillance data remains a major concern in most countries, limiting their use to draw evidence-based control strategies at national and subnational levels [12].

In France, the national sexual health strategy was launched in 2017, aiming to protect sexual health, and to reduce significantly STI incidence by 2030 [19]. In 2016, 320 new STI clinics (Centres gratuits d'information, de dépistage et de diagnostic [CeGIDDs]) were implemented across the country to replace the centers for HIV and viral hepatitis testing and the former STI clinics that were exclusively dedicated to the bacterial STI testing [8,20]. Compared with the former clinics, a wider range of sexual health services were implemented, including STI prevention, testing and treatment, consultations for sexual dysfunctions, sexual violence, sexual education, contraception, and reference to other specialists (eg, proctologist, oncologist) [8,20,21]. Although the CeGIDDs primarily target the most vulnerable populations (men who have sex with men, migrants from endemic countries), the delivery of services is free of charge for the whole population [8,20,21].

Before the implementation of the CeGIDDs, the surveillance systems were based on the sentinel networks for bacterial STI and the mandatory notification for HIV and HBV infections [22-24]. However, their representativeness, sensitivity, timeliness, and flexibility remained challenges even after several decades [22-26]. Indeed, these surveillance systems provided a reliable information at the national level but these indicators could neither describe accurately the epidemics at a lower level, nor give a relevant insight about the CeGIDD users by subregions [12,27].

Since 2016, the CeGIDDs submit a mandatory activity report to the regional health authority each year [16]. As the aggregated data of this report are insufficient to guide the policies [8], the decree of July 2015 ensures that each CeGIDD uses a software “to manage the patients’ records and facilitate the mandatory transmission of case-based data” to Santé publique France (SpF) [21], to establish a comprehensive surveillance of the CeGIDDs.

This article describes SurCeGIDD, the French automated surveillance system of several pathologies, designed to collect case-based data from a wide range of sexual health services, to better guide prevention policies at the national and subnational levels.

## Methods

### Principles of the SurCeGIDD Surveillance

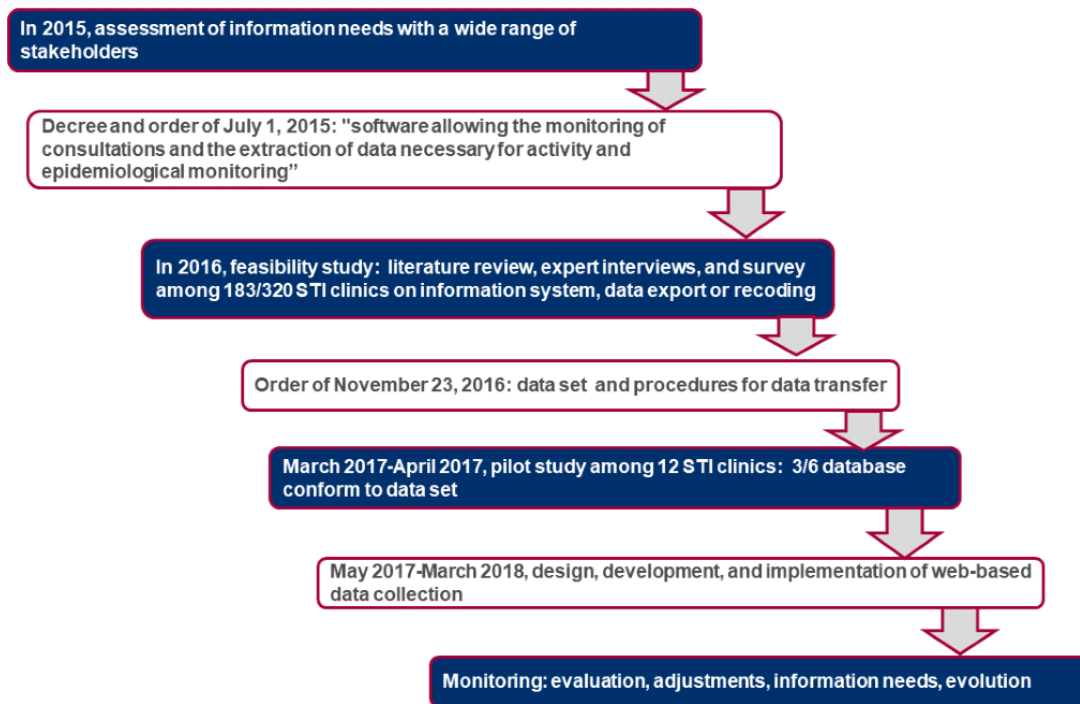
The SurCeGIDD surveillance was developed through a collaborative and iterative approach launched in 2015 (Figure 1). First, the stakeholders engaged in STI clinics, infectious diseases, sexual health promotion, epidemiological surveillance, regional and national health administration, and population representatives were identified through a national concertation meeting held by the French Ministry of Health. The software companies were not engaged in the working groups to prevent any conflict of interest, assuming that the CeGIDDs will contract the implementation of the specifications for SurCeGIDD afterward.

Among the stakeholders, volunteers—regional health agencies (n=3), former STI clinics and HIV-hepatitis testing centers (n=4), charities (n=2), regional response coordination consortium for HIV infection (n=1), SpF, the Health Care and Health Directorates of the Ministry of Health—defined the core principles for the SurCeGIDD surveillance:

- to produce the most relevant epidemiological indicators to guide, implement, and monitor the national sexual health strategy at national, subnational, and local levels;
- to better understand the drivers of these epidemics in France;
- to cover STIs that are not targeted by the classical surveillance in France (genital warts, herpes, *Mycoplasma genitalium*);
- to collect the tests and diagnoses to monitor the positivity rates of STI;
- to cover all sexual health services provided by the CeGIDDs (eg, contraception, HIV preexposure prophylaxis [PrEP], HIV postexposure prophylaxis, vaccination);
- to use patients’ records routinely collected for consultation purposes;
- to use computer-based protocols to facilitate a timely transmission of the required data (no paper form, no online form) and to maximize the coverage of the surveillance;
- to conform to the SurCeGIDD metadata set to enable the comparisons of CeGIDDs, subnational regions, and exposed populations;
- to guarantee patients’ confidentiality by collecting and sharing anonymous data; and
- to secure the data transfer and repository.

Subsequently, the stakeholders mapped the information needs at the national, subnational, and local levels (Textbox 1 [28]). Finally, SpF used this mapping exercise to propose a detailed metadata set that was discussed, amended, and agreed upon by all the stakeholders.

**Figure 1.** Iterative stages of the SurCeGIDD surveillance project. STI: sexually transmitted infection.



**Textbox 1.** Variables collected by the SurCeGIDD surveillance, France from the specification for the surveillance of STI clinics.

#### Mandatory Variables

*Unique identifier:* Clinic, patient, and consultation.

*Sociodemographic variable:* Sex, age, country of birth, date of arrival for migrants, region and department of residence, health insurance coverage.

*Context of consultation:* Anonymity of consultations, outreach testing, use of rapid diagnosis test (point of care).

*History of sexually transmitted infection(s):* Diagnosis of HIV, hepatitis B virus (HBV), hepatitis C virus (HCV), bacterial sexually transmitted infection (STI).

*Sexual exposure:* Sex of partner(s), number of sexual partners in the last 12 months, condom use by sexual practice with steady or casual partner(s) or both.

*Blood exposure:* Injection drug use, health care in a foreign country, blood transfusion before 1992, contact with person(s) infected with HBV, HCV, or HIV.

*Symptoms:* Presence/absence of symptoms (HIV or STI).

*Tests, diagnoses:* Tests for HIV, HBV, HCV, syphilis, *Neisseria gonorrhoeae*, lymphogranuloma venereum, *Mycoplasma genitalium*. Diagnoses of HIV, HBV, HCV, syphilis, *N. gonorrhoeae*, *Chlamydia trachomatis*, *Mycoplasma genitalium*, genital warts, herpes, antimicrobial resistance of *N. gonorrhoeae*.

*Prescription:* Treatments, preexposure prophylaxis, postexposure treatment for HIV, and vaccination (HBV, hepatitis A virus, human papillomavirus).

#### Optional Variables

*Sociodemographic variable:* Employment status, occupation.

*Context of consultation and history of STI:* Reason for initial consultation in this STI clinic, date of last HIV test, diagnosis of STI in the last 12 months.

*Sexual exposures:* Sex for money, service, drug, house.

*Sexual health services:* Emergency contraception, standard contraception, patient transfer inside the clinic or outside for HIV treatment, HBV and HCV treatment, gynecology, obstetrics, proctology, welfare/social service, sexology, postexposure treatment, preexposure prophylaxis, vaccination, other services.

## Feasibility Assessment

Second, to assess the applicability of the SurCeGIDD principles, SpF conducted a nonsystematic review of scientific articles and the gray literature using the keywords (surveillance or individual or cases-based data extraction or data export or data workflow or data flow or interoperability and France or Europe or the names of comparable industrialized countries [eg, United States of America, United Kingdom, England, Germany]) on PubMed, Embase, Scopus, and Google Scholar. The websites of the national public health agencies, references of interesting articles, and the documents shared by the European networks for surveillances were also checked. After scanning the titles and abstracts for the most relevant ones, articles were analyzed using a form to gather information of interest (eg, framework, metadata set, stakeholder, data collection, data sharing, data integration). The use of health care reimbursement data (no clinical details), the mandatory notification system for HIV (electronic or paper form, such as e-DO in France), and the syndromic surveillance system based on algorithms analyzing the global activity of health care services (no biological details) were excluded from this document review, considering the principle of using clinical and biological records routinely collected during the consultations [23,29-32].

In France, the literature review shows that the surveillance of the colorectal cancer relies on a validated data set of clinical and biological information [33]. This surveillance covers 50 screening centers using 3 software to manage the patients' records. In the absence of interoperability, an application was developed to recode, extract, and transfer anonymized data from the screening centers to SpF through a secured data-sharing platform [33]. An intermittent data workflow is generated yearly by the screening centers, by activating a specific command on their computer [33]. Then, incoming data are checked and merged using SAS (SAS Institute Inc.) [33]. Another approach is an automated data extraction from 3 multisite laboratories (3Labos surveillance) including tests and diagnoses for 11 pathogens [34]. The 3Labos surveillance relies on a daily/monthly transmission (depending on pathogens) of individual data (sex, age, residence, test, and diagnosis) to SpF through a web service (a protocol enabling the communication between 2 electronic devices via internet) [34]. The incoming data are checked and merged using a Stata (StataCorp) script [34].

The literature review also showed the advantages and limitations of surveillance systems based on data extraction in other countries as well as the development stages of these surveillance systems: GUMCAD (the Genitourinary Medicine Clinic Activity Dataset) in England; CIDR (Computerised Infectious Disease Reporting) in Ireland; SSuN (STD Surveillance Network) in the United States; MIBA (the Danish Microbiology Database) in Denmark; ACCESS (Australian Collaboration for Coordinated Enhanced Sentinel Surveillance) in Australia; eCR (Electronic Case Reporting) in Utah, USA; SOAP (Dutch STI Database) in the Netherlands; and MAVEN (Massachusetts Virtual Epidemiologic Network) in the United States [35-44]. The main points in the development of these electronic and automated surveillance systems were the precise data

specifications, the stakeholder's involvement, the effective information systems to manage patients' records, and the development and the maintenance of an information platform by the surveillance programs to collect and integrate the data [35-39,44,45]. A long-term strategic framework for the sustainability and the evolution of the surveillance system is also an asset.

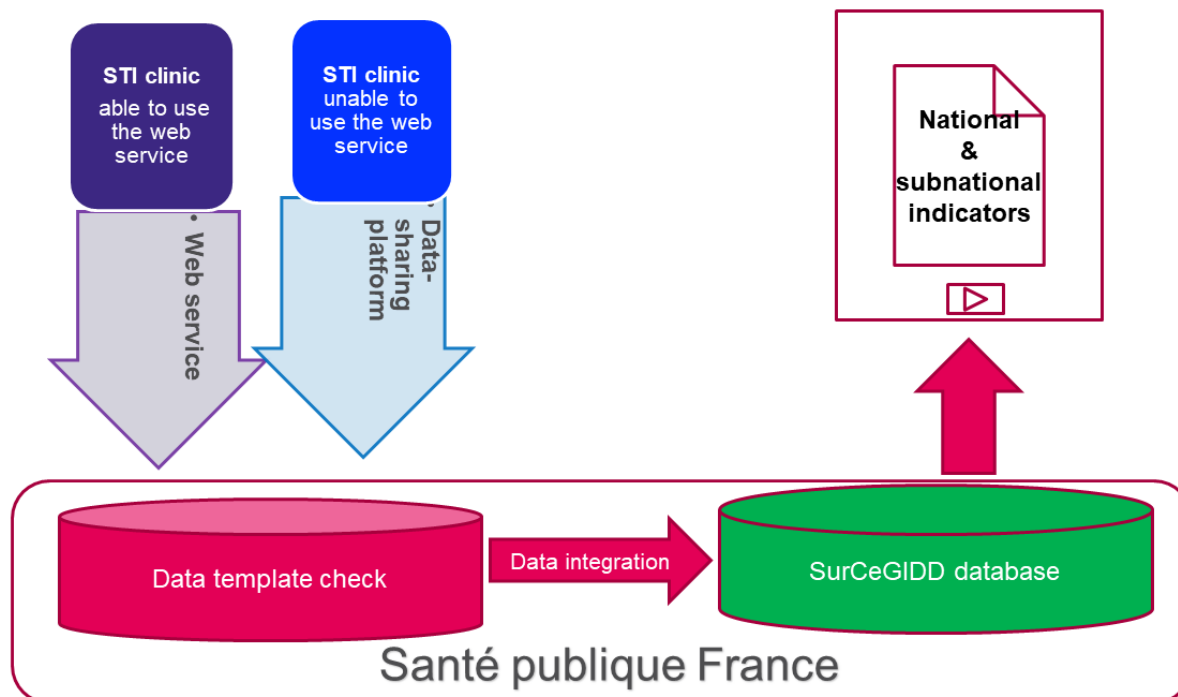
Third, to better interpret the information retrieved from the literature, SpF conducted a series of in-depth interviews with French experts (clinicians, biologists, software specialists, data managers, and epidemiologists), using an interview guide to complete the information on the French context. Then, SpF organized a concertation meeting with several clinicians, biologists, and public health specialists to share the findings of the literature review and the interviews, and to discuss the informatics capabilities of the CeGIDDs as well as the levers and obstacles that could impact the implementation of the SurCeGIDD principles in each subnational region.

Fourth, in May 2016, SpF conducted a feasibility survey, by sending an electronic questionnaire to all CeGIDDs. This questionnaire was used to collect information on the computerization of the CeGIDDs, the specifications of their software for the management of the patients records, and their capability to export or recode and transfer a data set in a predefined CSV format. The response rate was 57.2% (183/320). Among the respondents, most CeGIDDs (147/183, 80.3%) used informatic tools to manage their patients/users. A majority of them could export individual data (117/183, 63.9%). Less than half of the CeGIDDs (79/183, 43.2%) were capable of recoding a data set but only 21.9% (40/183) reported being able to recode and transfer data according to a template. The CeGIDDs used mainly 6 software and numerous "homemade" applications. Interestingly, 2 major software companies decided to apply the SurCeGIDD principles at this stage, and communicated their decision to use the surveillance specifications as soon as they would be published. Few months after the feasibility study, one of these companies covered half of the CeGIDDs.

## Surveillance Architecture

Fifth, the findings of the feasibility assessment were used to adjust the SurCeGIDD metadata set, taking into account the day-to-day practices in the CeGIDDs and their informatic limitation. Then, the order of November 23, 2016, stated the mandatory format for the SurCeGIDD data as well as the procedures for data transfer [46]. The specifications for the surveillance were subsequently published on the SpF website [46].

Considering the SurCeGIDD specifications, the literature review, and the results of the feasibility survey, a web service was designed for the data transmission [42,47]. If any software provider developed an access to the web service using the specifications shared by SpF, the CeGIDD can automatically transfer its individual database to SpF (Figure 2). Otherwise, the CeGIDDs can export and recode their data in the expected CSV format before their transmission to SpF via a secured data-sharing platform.

**Figure 2.** Workflow of sexual health and sexually transmitted infection (STI) data, SurCeGIDD surveillance, France.

Then, the structure of the data is automatically checked and validated before its integration in the SurCeGIDD repository or its return to the CeGIDDs if this control stage fails. During this stage, empty variables are accepted. Warning messages pop-up if the data specifications are not applied. In such a case, SpF helps the CeGIDDs to correct its data set.

A pilot study conducted by SpF among 12 voluntary CeGIDDs helped to test, correct, redesign, and implement the surveillance system between March and April 2017. During the pilot development, 6 CeGIDDs failed to extract their data. Among the 6 data sets that were transmitted to SpF, only 3 were extracted and transferred according to the SurCeGIDD specifications. This highlighted CeGIDDs' difficulties in performing effectively the expected data management, and the need for friendly software requiring no manual tasks to prepare the data.

### Preliminary Statistical Analyses

Preliminary and descriptive analyses of 2017-2018 data were performed using numbers and proportions to assess the data quality and estimate for the first time the positivity rates for STI in the CeGIDDs (number of positive tests divided by the total of tests), giving an overview of the SurCeGIDD surveillance system after its deployment. Stata 14.2 was used for statistical analyses.

### Ethical Consideration

The protocol was reviewed and approved by the data protection officer of SpF. Then, ethical approval was granted by the French Personal Data Protection Authority (CNIL authorization no. 2049450). Data are collected with patients' consent. Only anonymized data are transmitted through the SurCeGIDD system. Subsequently, any potential identifying information is aggregated or suppressed after the elimination of duplicates to

ensure patients confidentiality. Access to database is strictly restricted to the authorized agents of SpF.

## Results

The number of CeGIDD respondents increased from 54/320 (16.9%) in 2017 to 143/320 (44.7%) in 2018. In 2018, 65.7% (94/143) of respondent CeGIDDs transmitted their data using the web service versus 41% (22/54) in 2017. The volume of corresponding records increased from 2414 in 2017 to 382,890 in 2018.

Assessment of the 2018 data quality showed a high level of completeness for some variables: 93%-99% for sex (382,733/382,890, 99.96%), age (366,501/382,890, 95.72%), HIV (378,499/382,890, 98.85%), HBV (370,562/382,890, 96.78%), HCV (365,591/382,890, 95.48%), syphilis (372,002/382,890, 97.16%), chlamydia (368,930/382,890, 96.35%), and gonorrhea tests (367,622/382,890, 6.01%). Nevertheless, the proportion of missing information was greater than 50% for the number and sex of sexual partners in last 12 months (311,703/382,890, 81.41%), PrEP prescription (296,399/382,890, 77.41%), HBV vaccine (325,215/382,890, 84.94%), hepatitis A virus (HAV) vaccine (347,429/382,890, 90.74%), HPV vaccine (338,165/382,890, 88.32%), diagnoses of genital warts (269,492/382,890, 70.38%), diagnoses of herpes (269,489/382,890, 70.38%), and antimicrobial resistance of *Neisseria gonorrhoeae*. Most respondent CeGIDDs reported ongoing efforts to complete next data sets. Numerous CeGIDDs reported different identifiers for the consultations of the same person, whereas others attributed a unique identifier to the same person regardless of the number of his/her visits. Consequently, it was not possible to link repeated consultations of the same person in SurCeGIDD. An advisory committee for SurCeGIDD was set up to better guide the construction of algorithms

enabling effective data management and analyses. Accurate analyses will be performed afterward.

Nevertheless, the first analyses showed that attendees were mostly men 227,326/379,921 (59.84%), whereas 151,963/379,921 (40%) were women and 632/379,921 (0.17%) transgenders in 2018. The median age was 27 years for men, 23 years for women, and 30 years for transgender. Half of the attendees (81,964/174,932, 46.85%) were heterosexual men, 69,016/174,932 (39.45%) heterosexual women, 20,764/174,932 (11.87%) men who have sex with men, and 3188/174,932 (1.82%) women who have sex with women. A majority of them were born in France (227,698/286,289, 79.53%) and unemployed (115,913/211,707, 54.75%). The positivity rates were 0.37% for 205,348 HIV serologies, 1.31% for 131,551 HBV serologies, 7.16% for 161,241 *Chlamydia trachomatis* PCR, 2.83% for 146,649 gonorrhea PCR, 1.04% for the syphilis combination of treponema and nontreponema serologies, and 5.96% for 13,313 *M. genitalium* PCR.

## Discussion

### Principal Findings

Despite numerous challenges, SurCeGIDD surveillance was successfully implemented. The effectiveness of this system to collect automatically important volumes of data was demonstrated by 143/320 (44.7%) CeGIDDs that reported simultaneously the STI tests, diagnoses, clinical details, and patients' socioeconomic characteristics for the first time in France. Although these data need further consolidation, the preliminary analysis confirmed that CeGIDDs' attendees are particularly exposed to STI with high positivity rates: 0.37% for HIV, 1.31% for HBV, 7.16% for Ct, 2.83% for gonorrhea, 1.04% for syphilis, and 5.96% for *M. genitalium*. These results also gave the first insight into *M. genitalium* testing in the CeGIDDs. Further analyses will better describe their situation.

However, a selection of CeGIDDs with greater human, informatic, or financial resources compared with the nonparticipating structures cannot be ruled out. But the participation rate of 143/320 (44.7%) will probably increase in the next years, considering the growing interest of the CeGIDDs and software providers for SurCeGIDD data. Experiences in STI surveillance in Denmark, England, Ireland, Australia, and United States seem to confirm this expectation [35-38,44]. Moreover, the increasing computerization of the CeGIDDs and the adapted support provided by SpF in the use of the specifications might contribute to a better representativeness of the data in the medium term.

This support is also an opportunity for SpF to correct the metadata set, taking into account the operational feedback from the structures. Nevertheless, the time to develop or modify the informatic platforms was underestimated by SpF and stakeholders, delaying the use of the updated metadata set in data collection. This was also observed in other contexts, and it must be better anticipated in informatic development [35-38,44]. Moreover, difficulties to extract or recode data by CeGIDDs according to the specifications hampered the surveillance implementation. Indeed, some CeGIDDs selected

a range of variables considering feasibility burden. Consequently, nonprioritized variables were mostly missing.

The proportion of missing information was also greater than 50% for important variables such as the characteristics of sexual partners (311,703/382,890, 81.41%), PrEP use (296,399/382,890, 77.41%), HBV vaccine (325,215/382,890, 84.94%), HAV (347,429/382,890, 90.74%), HPV vaccine (338,165/382,890, 88.32%), diagnoses of genital warts (269,492/382,890, 70.38%), diagnoses of herpes (269,489/382,890, 70.38%), and antimicrobial resistance of *N. gonorrhoeae*. Some patient's records could not be recoded despite the CeGIDDs' willingness to conform to the surveillance specifications. For example, some CeGIDDs only recorded trademarks of vaccines instead of the names and doses that enable the monitoring of vaccination coverage. Some adjustment might be necessary in the data format to better take CeGIDDs' needs into account. This also underlines the need for appropriate human and financial resources to support CeGIDDs' adherence to the specifications. Nevertheless, data transmission will require no extra labor burden once the specifications have been implemented, in contrast to the use of paper or electronic forms [35,41,42].

Because of the legal context in France, it was not possible to engage software companies in the SurCeGIDD working groups. In England, software providers contributed to the early stage of the surveillance project, applying GUMCAD specification in the development of their tools [36]. Moreover, the lack of specific funding for information systems in CeGIDDs might limit SurCeGIDD's expansion and evolution because each structure has to pay for the implementation of the specification changes. Conversely, incentives and additional funding contributed to ensure the largest and long-term coverage in other contexts [35,37,41,42]. Mandatory use of national specifications by all software companies might be a solution. It might not only enable database transmission by all CeGIDDs, but can also facilitate the implementation of any change simultaneously in all CeGIDDs (additional information) [35,36].

Regarding SurCeGIDD coverage, the list of CeGIDDs including emails and telephone numbers was not updated since 2016, despite the creation of new structures and closures of other ones by regional health agencies [8]. Therefore, it was not possible to engage with noninformed CeGIDDs in the surveillance. SpF advocates for a systematic transmission of updated phone books by regional public health agencies each year, to ensure the best representativeness of the surveillance data.

Considering the volumes of data shared, SpF needs to upgrade its tools and procedures, using algorithms to automatize data monitoring, quality checking, and epidemiological reports at national and subnational levels. There is also a need to formalize rules and process for data sharing and analyses [35,48,49].

Despite its limitations, SurCeGIDD surveillance creates new opportunities for research collaboration on sexual health in France [35,41,50]. Availability of a wide range of individual characteristics (health insurance, occupation) in this database will be an opportunity to better understand the impact of social health inequalities on STI epidemics in France. SurCeGIDD will also help to fill the gap in our knowledge of *M. genitalium*,

genital warts, herpes, contraception, and vaccination in France [8]. Generating a continuous data workflow instead of a yearly transmission of data might be the next improvement of SurCeGIDD, to timely detect and investigate STI events (multidrug resistance, uncommon anatomic location, upsurge, clusters) [41,44,51-53]. An anonymous linkage of patients within the SurCeGIDD database might improve the epidemiological surveillance [35,36,42]. The almost real-time transmission of routinely collected data to monitor COVID-19 diagnoses from laboratories in France demonstrated the feasibility of such evolutions [48]. Moreover, the expansion of this laboratory surveillance toward STI diagnoses might help to complete the picture of STI epidemics in France.

## Conclusions

The automated SurCeGIDD surveillance is functional despite the challenges that hamper its coverage expansion. A long-term vision relying on sustained funding, including informatic developments for STI clinics, and the mandatory use of the specifications by software providers might help to improve the repressiveness of surveillance data to better guide and evaluate prevention strategies and services delivery. Indeed, the timeliness and efficacy of the COVID-19 monitoring systems demonstrate how electronic medical records could effectively replace passive surveillance systems, particularly sentinel surveillance systems commonly overwhelmed by important volumes of information. Moreover, exhaustive medical and laboratory records could help to compare more accurately countries for STI trends, drivers of epidemics, and sexual health services delivery.

## Conflicts of Interest

None declared.

## References

- Rowley J, Vander Hoorn S, Korenromp E, Low N, Unemo M, Abu-Raddad LJ, et al. Chlamydia, gonorrhoea, trichomoniasis and syphilis: global prevalence and incidence estimates, 2016. *Bull. World Health Organ* 2019 Jun 06;97(8):548-562P. [doi: [10.2471/blt.18.228486](https://doi.org/10.2471/blt.18.228486)]
- Unemo M, Bradshaw CS, Hocking JS, de Vries HJC, Francis SC, Mabey D, et al. Sexually transmitted infections: challenges ahead. *The Lancet Infectious Diseases* 2017 Aug;17(8):e235-e279. [doi: [10.1016/s1473-3099\(17\)30310-9](https://doi.org/10.1016/s1473-3099(17)30310-9)]
- Poncin T, Fouere S, Braille A, Camelena F, Agsous M, Bebear C, et al. Multidrug-resistant failing treatment with ceftriaxone and doxycycline in France, November 2017. *Euro Surveill* 2018 May;23(21):1800264 [FREE Full text] [doi: [10.2807/1560-7917.ES.2018.23.21.1800264](https://doi.org/10.2807/1560-7917.ES.2018.23.21.1800264)] [Medline: [29845928](https://pubmed.ncbi.nlm.nih.gov/29845928/)]
- Annual Epidemiological Report 2016: Syphilis. European Centre for Disease Prevention and Control (ECDC). Stockholm, Sweden: ECDC; 2018. URL: <https://www.ecdc.europa.eu/en/publications-data/syphilis-annual-epidemiological-report-2016> [accessed 2022-08-31]
- Annual Epidemiological Report 2016: Chlamydia. European Centre for Disease Prevention and Control (ECDC). Stockholm, Sweden: ECDC; 2018. URL: <https://www.ecdc.europa.eu/en/publications-data/chlamydia-infection-annual-epidemiological-report-2016> [accessed 2022-08-31]
- Annual Epidemiological Report 2016: Gonorrhoea. European Centre for Disease Prevention and Control (ECDC). Stockholm, Sweden: ECDC; 2018. URL: <https://www.ecdc.europa.eu/en/publications-data/gonorrhoea-annual-epidemiological-report-2016> [accessed 2022-08-31]
- Myers KO, Ahmed NU. The Role of HIV in the Progression through the Stages of the Human Papillomavirus to Cervical Cancer Pathway. *AIDS Rev* 2018;20(2):94-1043. [doi: [10.24875/AIDSRev.M18000021](https://doi.org/10.24875/AIDSRev.M18000021)] [Medline: [29938703](https://pubmed.ncbi.nlm.nih.gov/29938703/)]
- Pioche C, Ndeikoundam N, Sarr A, Cazein F, Bruyand M, Viriot D. Activité de dépistage et diagnostic du VIH, des hépatites B et C et des autres IST en CeGIDD, France. *Bull Epidémiol Hebd* 2019;31-32:625-633 [FREE Full text]
- Falla AM, Hofstraat SHI, Duffell E, Hahné SJM, Tavoschi L, Veldhuijzen IK. Hepatitis B/C in the countries of the EU/EEA: a systematic review of the prevalence among at-risk groups. *BMC Infect Dis* 2018 Feb 12;18(1):79 [FREE Full text] [doi: [10.1186/s12879-018-2988-x](https://doi.org/10.1186/s12879-018-2988-x)] [Medline: [29433454](https://pubmed.ncbi.nlm.nih.gov/29433454/)]
- Cassini A, Colzani E, Pini A, Mangen M, Plass D, McDonald S, On Behalf Of The BCoDE Consortium. Impact of infectious diseases on population health using incidence-based disability-adjusted life years (DALYs): results from the Burden of Communicable Diseases in Europe study, European Union and European Economic Area countries, 2009 to 2013. *Euro Surveill* 2018 Apr;23(16):17-00454 [FREE Full text] [doi: [10.2807/1560-7917.ES.2018.23.16.17-00454](https://doi.org/10.2807/1560-7917.ES.2018.23.16.17-00454)] [Medline: [29692315](https://pubmed.ncbi.nlm.nih.gov/29692315/)]
- HIV infection and AIDS. European Centre for Disease Prevention and Control (ECDC). Stockholm, Sweden: ECDC; 2019. URL: <https://www.ecdc.europa.eu/sites/default/files/documents/hiv-surveillance-report-2019.pdf> [accessed 2022-08-31]
- Taylor MM, Wi TE. Transforming and integrating STI surveillance to enhance global advocacy and investment in STI control. *J Int AIDS Soc* 2019 Aug;22 Suppl 6:e25361 [FREE Full text] [doi: [10.1002/jia2.25361](https://doi.org/10.1002/jia2.25361)] [Medline: [31468726](https://pubmed.ncbi.nlm.nih.gov/31468726/)]
- Taylor M, Newman L, Ishikawa N, Laverty M, Hayashi C, Ghidinelli M, et al. Elimination of mother-to-child transmission of HIV and Syphilis (EMTCT): Process, progress, and program integration. *PLoS Med* 2017 Jun;14(6):e1002329 [FREE Full text] [doi: [10.1371/journal.pmed.1002329](https://doi.org/10.1371/journal.pmed.1002329)] [Medline: [28654643](https://pubmed.ncbi.nlm.nih.gov/28654643/)]

14. Govender RD, Hashim MJ, Khan MA, Mustafa H, Khan G. Global Epidemiology of HIV/AIDS: A Resurgence in North America and Europe. *J Epidemiol Glob Health* 2021 Sep;11(3):296-301 [FREE Full text] [doi: [10.2991/jegh.k.210621.001](https://doi.org/10.2991/jegh.k.210621.001)] [Medline: [34270183](https://pubmed.ncbi.nlm.nih.gov/34270183/)]
15. Eisinger RW, Fauci AS. Ending the HIV/AIDS Pandemic. *Emerg Infect Dis* 2018 Mar;24(3):413-416 [FREE Full text] [doi: [10.3201/eid2403.171797](https://doi.org/10.3201/eid2403.171797)] [Medline: [29460740](https://pubmed.ncbi.nlm.nih.gov/29460740/)]
16. World Health Organization (WHO). Global progress report on HIV, viral hepatitissexually transmitted infections, 2021. Accountability for the global health sector strategies 2016-2021: actions for impact. In: World Health Organization. Geneva, Switzerland: World Health Organization; Jan 01, 2021.
17. Global HIV Programme. World Health Organization. URL: <https://www.who.int/teams/global-hiv-hepatitis-and-stis-programmes/hiv/strategic-information/hiv-data-and-statistics> [accessed 2022-08-31]
18. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin* 2021 May;71(3):209-249 [FREE Full text] [doi: [10.3322/caac.21660](https://doi.org/10.3322/caac.21660)] [Medline: [33538338](https://pubmed.ncbi.nlm.nih.gov/33538338/)]
19. National Sexual Health Strategy-2030 Agenda. Ministère des solidarités et de la santé. Paris, France: Ministère des solidarités et de la santé; 2017. URL: [https://solidarites-sante.gouv.fr/IMG/pdf/strategie\\_sante\\_sexuelle\\_def\\_ang.pdf](https://solidarites-sante.gouv.fr/IMG/pdf/strategie_sante_sexuelle_def_ang.pdf) [accessed 2022-08-31]
20. Lailier G, Comboroure J, Florence S, Troussier T, Pilorgé F, Sarr A. Premier bilan sur l'activité des CeGIDD, France, 2016. *Bull Epidémiol Heb* 2018;40-41:818-826 [FREE Full text]
21. Décret n° 2015-796 du 1er juillet 2015 relatif aux centres gratuits d'information, de dépistage et de diagnostic des infections par les virus de l'immunodéficience humaine et des hépatites virales et des infections sexuellement transmissibles. Ministère des solidarités et de la santé. 2015. URL: <https://www.legifrance.gouv.fr/loda/id/JORFTEXT000030824374/> [accessed 2022-08-31]
22. Laporal S, Leon L, Pioche C, Vaux S, Pillonel J, Lot F. Estimation du nombre de diagnostics d'hépatite B aiguë et de l'exhaustivité de la déclaration obligatoire en France en 2016. *Bull Epidémiol Hebd* 2019;24-25:496-501 [FREE Full text]
23. Cazein F, Sommen C, Pillonel J, Bruyan M, Ramus C, Pichon P. Activité de dépistage du VIH et circonstances de découverte de l'infection à VIH France 2018. *Bull Epidémiol Hebd* 2019;32:31 [FREE Full text]
24. Ndeikoundam Ngangro N, Viriot D, Fournet N, Pioche C, De Barbeyrac B, Goubard A. Referents for the regional offices of the French national public health agency. Bacterial sexually transmitted infections in France: recent trends and patients' characteristics in 2016. *Euro Surveill* 2019 Jan;24(5):1800038 [FREE Full text] [doi: [10.2807/1560-7917.ES.2019.24.5.1800038](https://doi.org/10.2807/1560-7917.ES.2019.24.5.1800038)] [Medline: [30722812](https://pubmed.ncbi.nlm.nih.gov/30722812/)]
25. German R, Lee L, Horan J, Milstein R, Pertowski C, Waller M, Guidelines Working Group Centers for Disease ControlPrevention (CDC). Updated guidelines for evaluating public health surveillance systems: recommendations from the Guidelines Working Group. *MMWR Recomm Rep* 2001 Jul 27;50(RR-13):1-35; quiz CE1-CE7. [Medline: [18634202](https://pubmed.ncbi.nlm.nih.gov/18634202/)]
26. Garcia M, Garrett N, Singletary V, Brown S, Hennessy-Burt T, Haney G, et al. An Assessment of Information Exchange Practices, Challenges, and Opportunities to Support US Disease Surveillance in 3 States. *J Public Health Manag Pract* 2018;24(6):546-553 [FREE Full text] [doi: [10.1097/PHH.0000000000000625](https://doi.org/10.1097/PHH.0000000000000625)] [Medline: [29227421](https://pubmed.ncbi.nlm.nih.gov/29227421/)]
27. Dunbar JK, Hughes G, Fenton K. In a challenging environment, intelligent use of surveillance data can help guide sexual health commissioners' choices to maximise public health benefit. *Sex Transm Infect* 2017 May;93(3):151-152. [doi: [10.1136/sextrans-2016-052696](https://doi.org/10.1136/sextrans-2016-052696)] [Medline: [27879291](https://pubmed.ncbi.nlm.nih.gov/27879291/)]
28. Surveillance épidémiologique au sein des CEGIDD. Santé publique France. 2017 Mar 2. URL: <https://www.santepubliquefrance.fr/maladies-et-traumatismes/hepatites-virales/hepatite-c/articles/surveillance-epidemiologique-au-sein-des-cegidd> [accessed 2022-09-06]
29. Semaille C, Barin F, Cazein F, Pillonel J, Lot F, Brand D, et al. Monitoring the dynamics of the HIV epidemic using assays for recent infection and serotyping among new HIV diagnoses: experience after 2 years in France. *J Infect Dis* 2007 Aug 01;196(3):377-383. [doi: [10.1086/519387](https://doi.org/10.1086/519387)] [Medline: [17597452](https://pubmed.ncbi.nlm.nih.gov/17597452/)]
30. Tuppin P, Rudant J, Constantinou P, Gastaldi-Ménager C, Rachas A, de Roquefeuil L, et al. Value of a national administrative database to guide public decisions: From the système national d'information interrégimes de l'Assurance Maladie (SNIIRAM) to the système national des données de santé (SNDS) in France. *Rev Epidemiol Sante Publique* 2017 Oct;65 Suppl 4:S149-S167. [doi: [10.1016/j.respe.2017.05.004](https://doi.org/10.1016/j.respe.2017.05.004)] [Medline: [28756037](https://pubmed.ncbi.nlm.nih.gov/28756037/)]
31. Ndeikoundam Ngangro N, Viriot D, Lucas E, Boussac-Zarebska M, Lot F, Dupin N, et al. Relevance of healthcare reimbursement data to monitor syphilis epidemic: an alternative surveillance through the national health insurance database in France, 2011-2013. *BMJ Open* 2018 Jul 23;8(7):e020336 [FREE Full text] [doi: [10.1136/bmjopen-2017-020336](https://doi.org/10.1136/bmjopen-2017-020336)] [Medline: [30037863](https://pubmed.ncbi.nlm.nih.gov/30037863/)]
32. Caserio-Schönemann CB, Fouillet A, Henry V. Le système de surveillance syndromique SurSaUD®. *Bull Epidémiol Hebd* 2014;38-44:3-4 [FREE Full text]
33. Jezewski-Serra DL, Goulard H, Salines E, Bloch J, D'Acremont M. Dépistage organisé du cancer colorectal: Guide du format des données et définitions des indicateurs de l'évaluation du programme national. Santé publique France. Saint Maurice, France: Santé publique France; 2017. URL: <https://www.santepubliquefrance.fr/recherche/#search=D%C3%A9pistage%20organisé%20du%20cancer%20colorectal%20:%20Guide%20du%20format%20>

- [des%20donn%C3%A9es%20et%20d%C3%A9finitions%20des%20indicateurs%20de%20l%C3%A9valuation%20du%20programme%20national](#) [accessed 2022-08-31]
34. Durand J, Noël H, Septfonds A, Fournet N, Georges S, Félix M, et al. La surveillance des maladies infectieuses à partir des données de laboratoire : projet 3labos. *Médecine et Maladies Infectieuses* 2017 Jun;47(4):S67. [doi: [10.1016/j.medmal.2017.03.165](https://doi.org/10.1016/j.medmal.2017.03.165)]
  35. Callander D, Moreira C, El-Hayek C, Asselin J, van Gemert C, Watchirs Smith L, et al. Monitoring the Control of Sexually Transmissible Infections and Blood-Borne Viruses: Protocol for the Australian Collaboration for Coordinated Enhanced Sentinel Surveillance (ACCESS). *JMIR Res Protoc* 2018 Nov 20;7(11):e11028 [FREE Full text] [doi: [10.2196/11028](https://doi.org/10.2196/11028)] [Medline: [30459142](https://pubmed.ncbi.nlm.nih.gov/30459142/)]
  36. Savage E, Mohammed H, Leong G, Duffell S, Hughes G. Improving surveillance of sexually transmitted infections using mandatory electronic clinical reporting: the genitourinary medicine clinic activity dataset, England, 2009 to 2013. *Euro Surveill* 2014 Dec 04;19(48):20981 [FREE Full text] [doi: [10.2807/1560-7917.es2014.19.48.20981](https://doi.org/10.2807/1560-7917.es2014.19.48.20981)] [Medline: [25496573](https://pubmed.ncbi.nlm.nih.gov/25496573/)]
  37. Rietmeijer CA, Donnelly J, Bernstein KT, Bisette JM, Martins S, Pathela P, et al. Here Comes the SSuN: Early Experiences with the STD Surveillance Network. *Public Health Rep* 2009 Nov 01;124(2\_suppl):72-77. [doi: [10.1177/00333549091240s211](https://doi.org/10.1177/00333549091240s211)]
  38. Voldstedlund M, Haar M, Mølbak K, MiBa Board of Representatives. The Danish Microbiology Database (MiBa) 2010 to 2013. *Euro Surveill* 2014 Jan 09;19(1):20667 [FREE Full text] [doi: [10.2807/1560-7917.es2014.19.1.20667](https://doi.org/10.2807/1560-7917.es2014.19.1.20667)] [Medline: [24434175](https://pubmed.ncbi.nlm.nih.gov/24434175/)]
  39. Whipple A, Jackson J, Ridderhoff J, Nakashima AK. Piloting Electronic Case Reporting for Improved Surveillance of Sexually Transmitted Diseases in Utah. *Online J Public Health Inform* 2019;11(2):e7 [FREE Full text] [doi: [10.5210/ojphi.v11i2.9733](https://doi.org/10.5210/ojphi.v11i2.9733)] [Medline: [31632601](https://pubmed.ncbi.nlm.nih.gov/31632601/)]
  40. Götz HM, van Oeffelen LA, Hoebe CJP, van Benthem BH. Regional differences in chlamydia and gonorrhoeae positivity rate among heterosexual STI clinic visitors in the Netherlands: contribution of client and regional characteristics as assessed by cross-sectional surveillance data. *BMJ Open* 2019 Jan 21;9(1):e022793 [FREE Full text] [doi: [10.1136/bmjopen-2018-022793](https://doi.org/10.1136/bmjopen-2018-022793)] [Medline: [30670509](https://pubmed.ncbi.nlm.nih.gov/30670509/)]
  41. Birkhead GS, Klompas M, Shah NR. Uses of electronic health records for public health surveillance to advance public health. *Annu Rev Public Health* 2015 Mar 18;36:345-359. [doi: [10.1146/annurev-publhealth-031914-122747](https://doi.org/10.1146/annurev-publhealth-031914-122747)] [Medline: [25581157](https://pubmed.ncbi.nlm.nih.gov/25581157/)]
  42. Troppy S, Haney G, Cocoros N, Cranston K, DeMaria A. Infectious disease surveillance in the 21st century: an integrated web-based surveillance and case management system. *Public Health Rep* 2014;129(2):132-138 [FREE Full text] [doi: [10.1177/003335491412900206](https://doi.org/10.1177/003335491412900206)] [Medline: [24587547](https://pubmed.ncbi.nlm.nih.gov/24587547/)]
  43. Slurink IAL, Op de Coul ELM, Heijne JCM, van Wees DA, Hoenderboom BM, Visser M. Sexually transmitted infections in the Netherlands in 2018. In: National Institute for Public Health and the Environment (RIVM). Bilthoven, Netherlands: National Institute for Public Health and the Environment (RIVM); 2019:2018.
  44. Chaintarli K, Jackson S, Cotter S, O'Donnell J. Evaluation and comparison of the National Tuberculosis (TB) Surveillance System in Ireland before and after the introduction of the Computerised Electronic Reporting System (CIDR). *Epidemiol. Infect* 2018 Jul 06;146(14):1756-1762. [doi: [10.1017/s0950268818001796](https://doi.org/10.1017/s0950268818001796)]
  45. Dudareva-Vizule S, Haar K, Sailer A, Jansen K, Hamouda O, Wisplinghoff H. Chlamydia trachomatis laboratory sentinel team. Establishment of a voluntary electronic Chlamydia trachomatis laboratory surveillance system in Germany, 2008 to 2014. *Euro Surveill* 2017 Feb 09;22(6):30459 [FREE Full text] [doi: [10.2807/1560-7917.ES.2017.22.6.30459](https://doi.org/10.2807/1560-7917.ES.2017.22.6.30459)] [Medline: [28205505](https://pubmed.ncbi.nlm.nih.gov/28205505/)]
  46. Arrêté du 23 novembre 2016 fixant le modèle de rapport d'activité et de performance devant être fourni au directeur général de l'agence régionale de santé et à l'Agence nationale de santé publique par les centres gratuits d'information, de dépistage et de diagnostic des infections par les virus de l'immunodéficience humaine et des hépatites virales et des infections sexuellement transmissibles (CeGIDD). Ministère de la santé et des solidarités. Paris, France: Ministère de la santé et des solidarités; 2016. URL: [https://www.legifrance.gouv.fr/jorf/article\\_jo/JORFARTI000033507489](https://www.legifrance.gouv.fr/jorf/article_jo/JORFARTI000033507489) [accessed 2022-08-31]
  47. Romano P, Marra D. SWS: accessing SRS sites contents through Web Services. *BMC Bioinformatics* 2008 Mar 26;9(S2):1-7. [doi: [10.1186/1471-2105-9-s2-s15](https://doi.org/10.1186/1471-2105-9-s2-s15)]
  48. GUMCAD STI Surveillance System Data Specification and Technical Guidance. In: Public Health England. London, UK: Public Health England; 2019.
  49. Data quality monitoring and surveillance system evaluation - A handbook of methods and applications. European Centre for Disease Prevention and Control (ECDC). Stockholm, Sweden: ECDC; 2014. URL: <https://www.ecdc.europa.eu/sites/default/files/media/en/publications/Publications/Data-quality-monitoring-surveillance-system-evaluation-Sept-2014.pdf> [accessed 2022-08-31]
  50. Chen MY, Langan S, Benchimol EI. Routinely collected electronic health data and STI research: RECORD extension to the STROBE guidelines. *Sex Transm Infect* 2016 Feb;92(1):2-3. [doi: [10.1136/sextrans-2015-052360](https://doi.org/10.1136/sextrans-2015-052360)] [Medline: [26668088](https://pubmed.ncbi.nlm.nih.gov/26668088/)]
  51. Faensen D, Claus H, Benzler J, Ammon A, Pfoch T, Breuer T, et al. SurvNet@RKI - a multistate electronic reporting system for communicable diseases. *Euro Surveill* 2006 Apr;11(4):7-8. [doi: [10.2807/esm.11.04.00614-en](https://doi.org/10.2807/esm.11.04.00614-en)] [Medline: [29208145](https://pubmed.ncbi.nlm.nih.gov/29208145/)]



52. Olesen B, Anhøj J, Rasmussen KP, Mølbak K, Voldstedlund M. MiBAAlert-a new information tool to fight multidrug-resistant bacteria in the hospital setting. *Int J Med Inform* 2016 Nov;95:43-48. [doi: [10.1016/j.ijmedinf.2016.09.001](https://doi.org/10.1016/j.ijmedinf.2016.09.001)] [Medline: [27697231](https://pubmed.ncbi.nlm.nih.gov/27697231/)]
53. Pullano G, Di Domenico L, Sabbatini CE, Valdano E, Turbelin C, Debin M, et al. Underdetection of cases of COVID-19 in France threatens epidemic control. *Nature* 2021 Feb;590(7844):134-139. [doi: [10.1038/s41586-020-03095-6](https://doi.org/10.1038/s41586-020-03095-6)] [Medline: [33348340](https://pubmed.ncbi.nlm.nih.gov/33348340/)]

## Abbreviations

**ACCESS:** Australian Collaboration for Coordinated Enhanced Sentinel Surveillance

**CeGIDD:** Centre gratuits d'information, de dépistage et de diagnostic

**CIDR:** Computerised Infectious Disease Reporting

**eCR:** Electronic Case Reporting

**GUMCAD:** Genitourinary Medicine Clinic Activity Dataset

**HAV:** hepatitis A virus

**HBV:** hepatitis B virus

**HCV:** hepatitis C virus

**HPV:** human papillomavirus

**MAVEN:** Massachusetts Virtual Epidemiologic Network

**MIBA:** Danish Microbiology Database

**PrEP:** preexposure prophylaxis

**SDG:** Sustainable Development Goals

**SOAP:** Dutch STI Database

**SpF:** Santé publique France

**SSuN:** STD Surveillance Network

**STI:** sexually transmitted infection

*Edited by Y Khader; submitted 10.06.21; peer-reviewed by R Poluru, P Nguyen, G Umeh; comments to author 11.12.21; revised version received 05.05.22; accepted 23.05.22; published 25.10.22.*

*Please cite as:*

*Ndeikoundam Ngangro N, Pioche C, Vaux S, Viriot D, Durand J, Berat B, Hamdaoui M, Lot F*

*An Automated Surveillance System (SurCeGGID) for the French Sexually Transmitted Infection Clinics: Epidemiological Monitoring Study*

*JMIR Form Res* 2022;6(10):e31136

URL: <https://formative.jmir.org/2022/10/e31136>

doi: [10.2196/31136](https://doi.org/10.2196/31136)

PMID: [36282557](https://pubmed.ncbi.nlm.nih.gov/36282557/)

©Ndeindo Ndeikoundam Ngangro, Corinne Pioche, Sophie Vaux, Delphine Viriot, Julien Durand, Bénédicte Berat, Mohammed Hamdaoui, Florence Lot. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 25.10.2022. 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

# User Behavior of a Publicly Available, Free-to-Use, Self-guided mHealth App for Depression: Observational Study in a Global Sample

Langting Su<sup>1,2</sup>, MA; Page Lyn Anderson<sup>1</sup>, PhD, ABPP

<sup>1</sup>Department of Psychology, Georgia State University, Atlanta, GA, United States

<sup>2</sup>Inquiry Health, LLC, Sheridan, WY, United States

**Corresponding Author:**

Page Lyn Anderson, PhD, ABPP

Department of Psychology

Georgia State University

Psychology Department, Urban Life Building

140 Decatur Street

Atlanta, GA, 30033

United States

Phone: 1 404 822 7009

Email: [panderson@gsu.edu](mailto:panderson@gsu.edu)

## Abstract

**Background:** Reducing the burden of depression is a global health concern. Self-guided mobile health (mHealth) apps are one approach to address this problem. However, there is little research on self-guided mHealth apps in a global sample or on how they are used in the real world. These gaps in our knowledge must be addressed to bring the promise of mHealth apps for reducing the global burden of depression closer to reality.

**Objective:** The purpose of this study is to examine the naturalistic user behavior of MoodTools, a publicly available, free-to-use, self-guided mHealth app designed to improve symptoms of depression, in a global community sample.

**Methods:** Mobile analytics data were collected from all unique downloads of the Android version of MoodTools between March 1, 2016, and February 28, 2018. Due to the deidentification and data aggregation process, no demographic or personal identifying information was tied to individual user data. All information was stored in aggregated, anonymized data files on Google Analytics' storage database. Google's software development kit was used to securely capture data about the number of downloads, location of downloads, number of app sessions, frequency and duration of app sessions, time between app sessions, and user retention, allowing for examination of which app's tools were viewed and for how long, including *Information* (psychoeducation), *Test* (self-monitoring using the Patient Health Questionnaire [PHQ-9]), *Thought Diary* (targeting negative cognitions), *Activities* (behavioral activation), *Videos* (curated from YouTube), and *Safety Plan* (safety plan development and links to quickly access crisis management resources).

**Results:** MoodTools was used by 158,930 people from 198 countries, including countries where English was not the primary language and in low- and middle-income countries. After the initial download, 51.14% (n=81,277) of users returned to the app after the initial download, and retention rates decreased with each subsequent app session. The typical person used the app for 3 sessions for a total of 12 minutes over 90 days. The most frequently visited tools were *Test* and *Thought Diary* (n=393,549, 24.32%). On average, users completed and reviewed the results of the PHQ-9 for 49 seconds and 53 seconds, respectively, and spent 3 minutes and 5 seconds on *Thought Diary*.

**Conclusions:** Self-guided mHealth apps could be one approach (among the many needed) to reduce the burden of depression. Observational data collected in this study show a global interest in MoodTools, including in low- and middle-income countries and countries where English is not the primary language. Future research is needed to determine whether people who use self-guided apps experience improvement in depressive symptoms, and if so, what "dosage" provides a meaningful benefit.

(JMIR Form Res 2022;6(10):e35538) doi:[10.2196/35538](https://doi.org/10.2196/35538)

**KEYWORDS**

mHealth; depression; smartphone; mobile app; digital health; global mental health; MoodTools; mobile health; mental health; mobile phone; low- and middle-income countries

## *Introduction*

Mobile health (mHealth) is any medical or health practice supported by mobile devices such as smartphones [1]. An estimated 5.2 billion people have mobile devices [2], so mHealth interventions have the potential to reach almost two-thirds of the world's population. Depression is a leading cause of global disability [3] that is associated not only with personal suffering but also unemployment, poor physical health, poor social function, and suicide [4,5]. Pharmacotherapy and psychotherapy improve depressive symptoms, but there are widespread shortages of trained mental health professionals to deliver these therapies that are not expected to improve in the near future [6,7], which limits the reach of these efficacious interventions. Self-guided internet-based interventions can be accessed through the web and used without support from mental health professionals to deliver cost-effective behavioral health services worldwide [8,9]. Specifically, internet-based interventions for depression began gaining traction in the 2000s with Beating the Blues [10,11] and MoodGYM [12], both of which use a fixed sequence navigation format, in which each session is built on the preceding one until the intervention is completed. Although internet-based interventions are efficacious for the treatment of depression [13], nearly half of the participants in clinical trials do not complete a full course of these interventions [14], and the dissemination and implementation of internet-based interventions for depression remain limited [15,16]. With the advent of smartphones in the late 2000s, mobile technology and interventions delivered via downloadable apps became more accessible. Smartphone-based interventions for depression have been shown to improve depressive symptoms [17] and have the potential to address depression worldwide, as the rate of smartphone adoption globally is 65% and is on the rise in countries with emerging economies [18,19].

There is a gap between how people use mHealth interventions in the real world and how researchers evaluate them. Although randomized controlled and feasibility trials find that mobile and web-based interventions are efficacious, the findings do not hold in practical settings [16,20,21]. Many studies use research-only versions of interventions that are not available to the public afterward. Out of 18 apps that were recently reviewed in a meta-analysis of randomized clinical trials evaluating

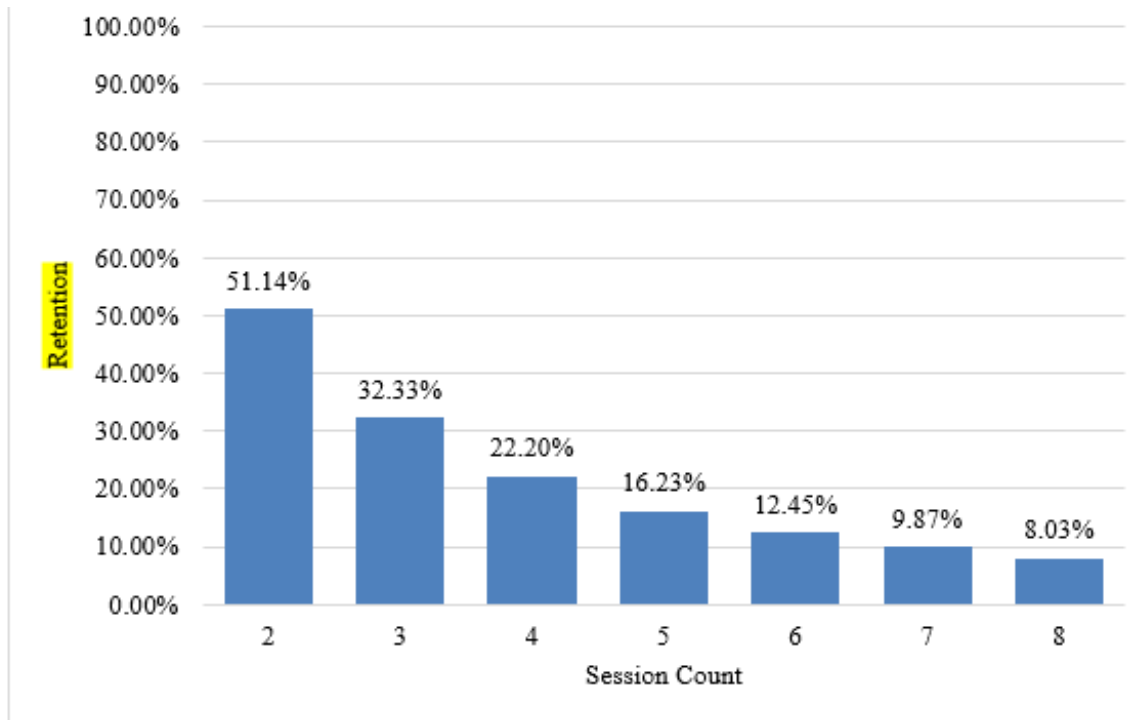
smartphone interventions for depression and anxiety, only 5 are currently available for public download [17]. Conversely, some of the more popular behavioral health apps have tens of millions of downloads on the app marketplace [22], but most are unevaluated [23]. The research gap on mHealth interventions and real-world usage must be addressed to bring the promise of mHealth for reducing the global burden of depression closer to reality.

There is very little ecologically valid research on the use of publicly available self-guided mHealth interventions for mental health [24]. To date, researchers have described the naturalistic behavior of users of apps targeting psychological disorders in only two studies [25,26]. IntelliCare is a suite of 13 smartphone apps designed to improve symptoms of depression and anxiety. Each self-guided app is accessed via a central hub and targets a single function (eg, cultivating gratitude). PTSD Coach is a free, publicly available self-guided app created by the United States Department of Veterans Affairs and the Department of Defense for managing acute distress related to posttraumatic stress disorder [27]. Return-visit users of PTSD Coach reported higher momentary distress levels compared to first-time users, suggesting that the app is being used in moments of need.

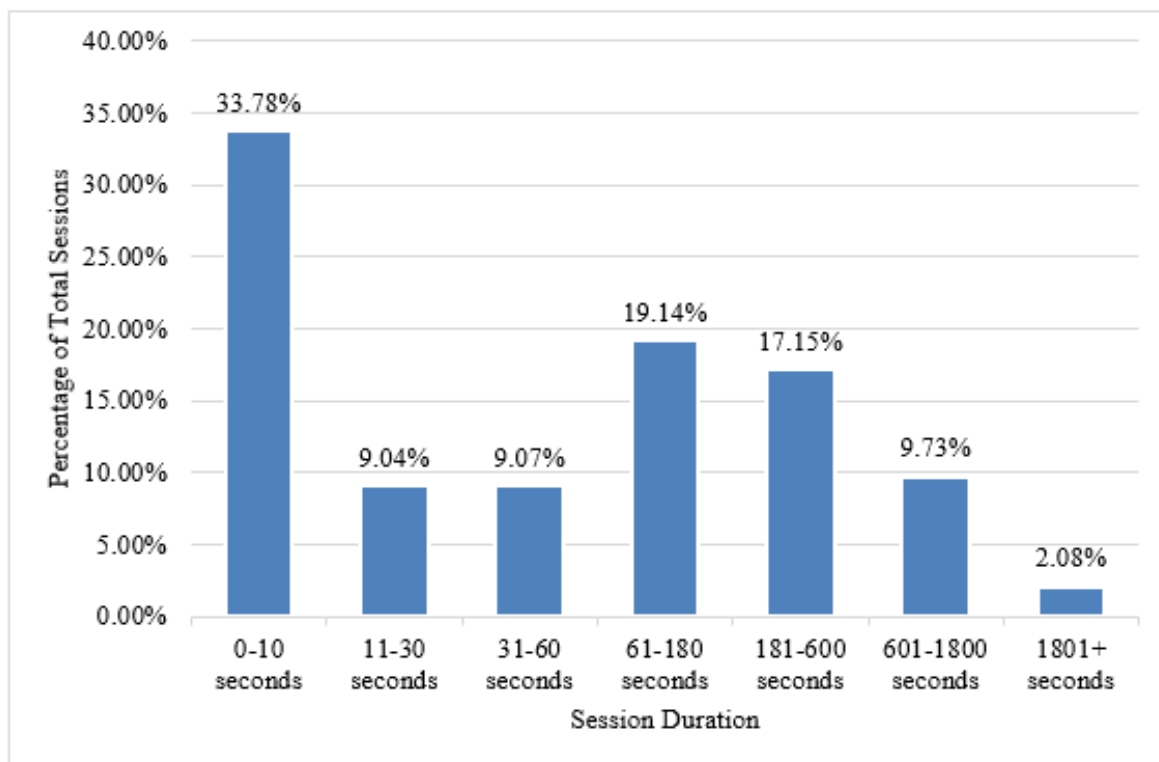
There are very few studies on self-guided apps for mental health and well-being within a global sample. PTSD Coach was downloaded in 86 countries, with non-US downloads making up 12% of total downloads [25]. A preliminary study of Wysa, a self-guided artificial intelligence–based chatbot app designed to promote mental well-being using a text-based conversational interface, reported that among the global community of users sampled, participants in the study downloaded the app in the United States, Europe, and Asia [28]. A randomized controlled trial of Headspace, a self-guided app that teaches mindfulness practice through guided meditations, found that its participants represented 11 countries [29]. These studies suggest that there is a global interest in using self-guided apps to improve mental health.

This study aims to describe how MoodTools, a publicly available, free-to-use, self-guided mental health app for depressive symptoms, is used “in the wild” among a global sample.

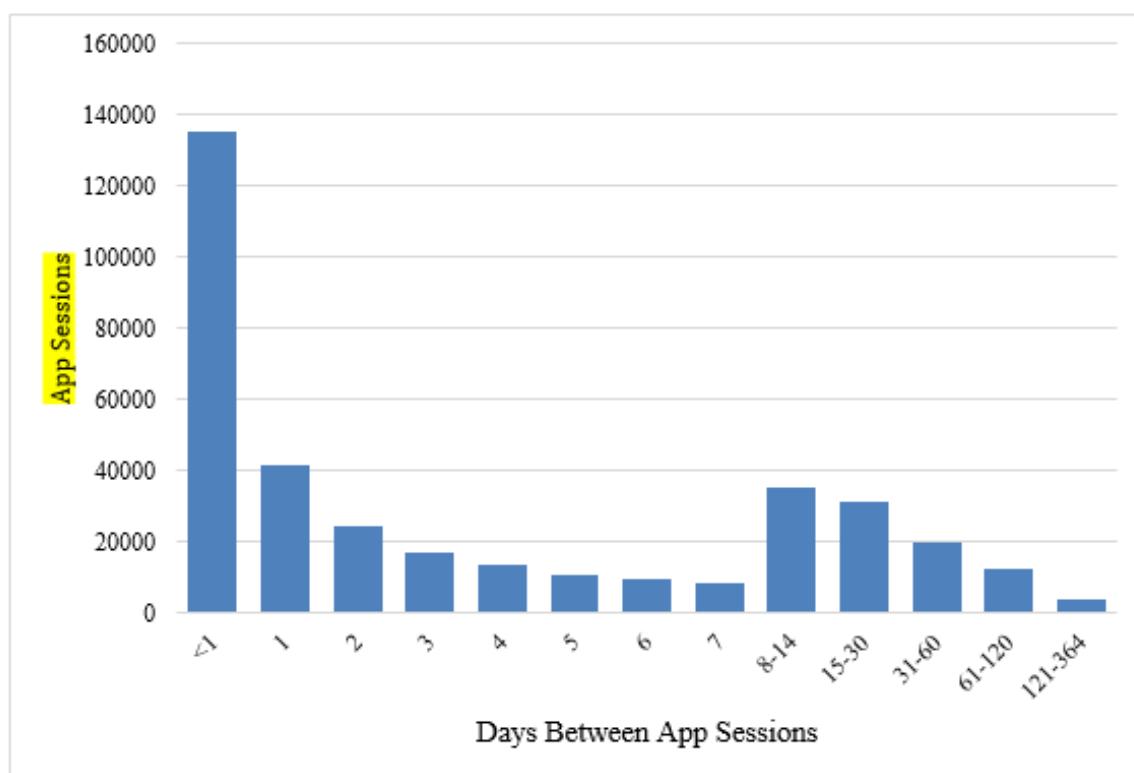
**Figure 1.** User retention across sessions.



**Figure 2.** Duration of MoodTools sessions.



**Figure 3.** The amount of time between downloading MoodTools and returning to the app.



## Methods

### Overview of MoodTools

MoodTools was published on Google Play for Android devices in June 2014 and on Apple App Store for iOS devices in 2015. Since its release in 2014, it has been downloaded on iOS and Android devices over 500,000 times. MoodTools is a fully automated, self-guided smartphone app for iOS (ie, iPhone and iPad) and Android devices. All content is self-contained, with the exception of external links that take the users out of the app; there is no human interaction. The app is exclusively in English. MoodTools contains 6 features called tools. The *Information* tool contains psychoeducation about depression. The *Test* tool contains the mobile form of the Patient Health Questionnaire (PHQ-9), a 9-item depression screening questionnaire that has been validated in the paper form [30] and in the mobile form [31]. If a user responds with “more than half the days” or “nearly every day” to item 9 of the PHQ-9 (ie, “Over the last 2 weeks, how often have you been bothered by thoughts that you would be better off dead or of hurting yourself in some way”), the results page displays a link to the country-specific suicide hotline if it was available in that country and external resources for managing suicidal thoughts. The *Thought Diary* tool features a diary entry derived from thought records [32] for the practice of thought restructuring. The *Activities* tool, based on behavioral activation therapy, prompts users to engage in activities to improve their mood. The activities are customizable, and the history page allows users to see which activities provide the most significant boost in subjective mood. The *Videos* tool contains a curated list of YouTube videos such as TED Talks,

guided meditations, and soothing sounds for mindfulness. The *Safety Plan* tool gives users information on coping with suicidal thoughts, allows users to fill out a safety plan, and provides quick access to local urgent care, emergency departments, and national crisis hotlines.

### Data Extraction

Data were derived from mobile analytics data from all unique downloads of the Android version of MoodTools between March 1, 2016, and February 28, 2018. Mobile analytics data were not collected from the iOS app during this study period. Due to the deidentification and data aggregation process, no demographic or personal identifying information was tied to individual user data. All measures were stored in aggregated, anonymized data files on Google Analytics’ storage database.

The Google Analytics software development kit (SDK) was integrated into the Android app in March 2016. We used the SDK to securely capture aggregate usage data and retention information from March 2016 to 2018 (the maximum amount of aggregate data that can be analyzed at a time). The following information was collected:

- Location: The country from which the app was installed on a user’s Android device.
- Number of downloads: The number of individual users who downloaded the app.
- App session: A single period of user interaction within the app. Activity that occurs within 30 minutes of each other is counted as part of the same app session. If there is no activity for 30 minutes, future activity is attributed to a new session.

- Session frequency: The number of app sessions during the data collection.
- Session duration: The number of minutes the app is open or used during an app session.
- Session recency: The amount of elapsed time since a user's last app session.
- Total duration in app: The total amount of time spent across app sessions for an individual user.
- Tools visited: The number of times a user opens the home page for each of the 6 tools within the app.

### Ethics Approval

This study was approved by the Institutional Review Board at Georgia State University as Designation for Not Human Subjects Research (H21568).

## Results

### Downloads

Between March 1, 2016, and February 28, 2018, MoodTools on the Android platform was used by 158,930 people from 198 countries. Data were collected on the percentage of users by

continent, subcontinent, and country (see [Multimedia Appendix 1](#)). The app was downloaded across the Americas (50.46%), Europe (26.46%), Asia (15.48%), Oceania (4.82%), and Africa (2.61%). When categorized by subcontinents, more than half of all users were from Northern America and Northern Europe (46.537% and 13.324%, respectively). Countries with the highest percentage of users included the United States (40.83%), the United Kingdom (10.64%), India (8.47%), Canada (5.60%), and Australia (4.02%). There were downloads in low- and middle-income countries, including India (8.47%), Brazil (1.11%), and South Africa (1.01%). Users whose location could not be determined by Google Analytics were identified as "Not Set" (0.18%).

### Retention

Retention was defined as the percentage of users that return to the app at any point after their initial app session. As shown in [Table 1](#), 51.14% (n=81,277) of MoodTools users returned to the app for a second app session. The retention rate declined with each subsequent app session: session 3 (n=51,382, 32.33%), session 4 (n=35,282, 22.2%), session 5 (n=25,794, 16.23%), session 6 (n=19,787, 12.45%), session 7 (n=15,686, 9.87%), and session 8 (n=12,762, 8.03%).

**Table 1.** User retention across sessions.

	User retention, n (%)
2 sessions	81,277 (51.14)
3 sessions	51,382 (32.33)
4 sessions	35,282 (22.2)
5 sessions	25,794 (16.23)
6 sessions	19,787 (12.45)
7 sessions	15,686 (9.87)
8 sessions	12,762 (8.03)

### Time Spent

Users spent 4 minutes, on average, on each session. About one-third of sessions lasted between 0 and 10 seconds, one-third lasted between 11 seconds and 3 minutes, and the remainder lasted more than 3 minutes ([Table 2](#)). Over one-third of all app

sessions were initiated within the same day of download, and about 1% of sessions occurred following a 3-month to 1-year period of inactivity ([Table 3](#)). After downloading MoodTools, the typical person used the app for 3 sessions for a total of 12 minutes over 90 days.

**Table 2.** Duration of MoodTools sessions.

Session duration	Sessions, %
0-10 s	33.78
11-30 s	9.04
31-60 s	9.07
61-180 s	19.14
180-600 s	17.15
601-1800 s	9.73
≥1801 s	2.08

**Table 3.** The amount of time between MoodTools app sessions.

Days between app sessions	App sessions, n (%)
<1 day	134,987 (37.16)
1 day	41,653 (11.47)
2 days	24,187 (6.66)
3 days	17,256 (4.74%)
4 days	13,365 (3.68%)
5 days	10,972 (3.02)
6 days	9820 (2.70)
7 days	8184 (2.25)
8-14 days	35,334 (9.73)
15-30 days	31,306 (8.62)
31-60 days	20,017 (5.51)
61-120 days	12,174 (3.35)
121-364 days	4041 (1.11)

## Tools Visited

We examined how often users visited each of the 6 tools (Thought Diary, Test, Information, Activities, Videos, and Safety Plan). Visiting a tool was operationalized as opening the home page screen for that tool. The Thought Diary tool and Test tool were tied for the most frequently visited tools, each making up 24.32% (n=393,487) of all home page screens viewed across all app sessions for all users (N=1,618,277 total screen

views; [Table 4](#)). The Information tool (ie, psychoeducation about depression) was the least frequently visited tool (n=1,246,667, 7.7%). Users spent an average of 3 minutes and 5 seconds (185 seconds) on the Thought Diary's entry screen, which allows users to complete a digital thought record. Users spent an average of 49 seconds completing the PHQ-9 screening questionnaire and an average of 53 seconds on the screen that displayed the results of the PHQ-9.

**Table 4.** Home page views by tool.

Tool name	Total screen views (N=1,618,277), n (%)	Average screen views per app session	Average time on home page screen(s)
Thought Diary	393,549 (24.32)	2.24	12.25
Test	393,487 (24.32)	2.00	5.71
Activities	331,961 (20.51)	2.35	10.08
Safety Plan	236,449 (14.61)	2.32	14.20
Videos	138,164 (8.54)	1.40	5.76
Information	124,667 (7.70)	1.23	11.34

## Discussion

### Principal Results

MoodTools was downloaded in 198 countries, suggesting that there is global interest in a free-to-use self-guided smartphone app for depression. It is worth noting that the developers of MoodTools had done no marketing campaigns and that, even though the app is presented exclusively in English, there have been downloads in countries where English is not the primary language. Users from low- and middle-income countries downloaded the app as well. Despite widespread interest, self-guided mental health apps will not make an impact on the global burden of depression if they are not effective. A review of evidence-based apps for anxiety and depression showed that a large majority (74%) were free to download, but only 3% had research to justify claims of effectiveness [33]. Efficacy studies

remain rare in the ever-changing landscape of publicly available mHealth apps. The efficacy and effectiveness of MoodTools (and other mHealth apps for depression) is clearly an important area for future research.

A key challenge for mHealth interventions for depression is to engage and retain users, given the low motivation and behavioral avoidance associated with the condition. Just over half of MoodTools users (n=81,277, 51.14%) returned to the app after the initial download, which is comparable to IntelliCare (about 50%) and PTSD Coach (61.1%) [25,26]. These return rates are similar to the pooled dropout rate (47.8%) reported in a meta-analysis of randomized clinical trials of smartphone apps for depressive symptoms [34]. Return rates of users "in the wild" and dropout rates from efficacy trials for smartphone apps is not, however, an apples-to-apples comparison. Dropout in clinical trials is defined as the incompleteness of end-of-intervention assessments, which are generally collected

outside of the smartphone app of interest, whereas retention in studies of naturalistic user behavior is typically defined as using the app. Users who download and open a free mHealth app for depression on the app marketplace may have different levels of interest and intent compared to participants who agree to use an mHealth app as part of a clinical trial. Given our limited knowledge of the retention rates of self-guided mental health apps in the real world, it is difficult to determine what is considered “standard” retention within this category of mHealth interventions. Future clinical trials could measure engagement-related dropout as well as assessment-related dropout, and real-world mHealth apps could adopt a similar standardized assessment structure to better compare retention and engagement across these methodologies.

It is important to determine how much app use is needed for meaningful improvement in symptoms. Research indicates that there is a relationship between app use and clinically meaningful benefit [35-37], but a minimum “dose” has not been identified. Around one-third of all MoodTools sessions lasted 10 seconds or less. These “touch-and-go” app sessions are likely too short to be meaningful. The high frequency of these app sessions creates noise in the data and highlights the importance of understanding the conditions under which users stay engaged with an app session long enough to have a meaningful interaction with it.

The dose achieved by typical MoodTools users was 3 app sessions (averaging 4 minutes per session), for a total of 12 minutes over 90 days, which can be compared to PTSD Coach users, who initiated 6.3 app sessions (averaging 47 seconds per session) for a total of 5 minutes before discontinuing its use [25]. In psychotherapy, which is typically delivered weekly in 50-minute sessions, about 50% and 75% of patients improve after 8 and 26 sessions, respectively [38,39]. In this context, it is difficult to imagine that typical MoodTools users achieve a dose that would lead to meaningful improvement in depressive symptoms, though this is an empirical question that remains to be tested. There were, however, a small proportion of users who returned to the app for 9 sessions or more, and some who used the app >200 times.

Unlike mental health professionals, everyday users of mental health apps are not trained to use the science of psychopathology to relieve symptoms. It is critical to identify what users of self-guided mental health apps naturally gravitate to, and the results are encouraging. The most frequently visited tools for MoodTools users were Thought Diary (thought record) and Test (mood self-monitoring). These results are consistent with research on naturalistic user behavior of other smartphone apps, despite differences in layout and psychological problems targeted. The most-visited areas of PTSD Coach were Self-Assessment (symptom tracking) and Manage Symptoms (coping skills) [25], and the most downloaded apps from IntelliCare’s suite were Thought Challenger (thought restructuring) and Worry Knot (worry management) [26]. These studies show that the users of publicly available, self-guided mental health apps most often visit cognitive restructuring and symptom tracking tools. These tools may be more popular because they engage users through active participation (eg, recording symptoms) as compared to tools that are more passive

(eg, reading information about the symptoms of depression). Negative cognitions are a core feature of depression, so it is encouraging that users of MoodTools most often visit the Thought Diary tool.

### Limitations

This study of naturalistic user behavior has several notable limitations. First, this study only included users from the Android platform of MoodTools and differences between those who use Android and iOS devices may influence user behavior. iOS users are more likely than Android users to be female, more educated, belong to a higher income group, and have more technical knowledge [40], and there are differences in user reviews between iOS and Android platforms [25]. Second, the deidentified, aggregate nature of data obtained through Google Analytics limits the types of analyses that can be conducted. For example, data on individual user characteristics (eg, gender, age, socioeconomic status) were not collected for this cohort.

### Future Directions

The number of areas for future research on self-guided mHealth apps for reducing depression is seemingly infinite. It is critical to improve our understanding of what happens immediately following the download of the app. Consistent with research on other apps, half of all MoodTools users did not return to the app after their first app session. Research is needed to understand how to maintain users’ engagement from the very first app session and to identify users most at risk for discontinuing use. Additionally, it may be helpful to sort users into low- and high-use comparison groups, as was done in the Wysa study [28], to examine how usage and retention patterns may differ across groups. Virtually nothing is known about how individual user characteristics (eg, gender, age, socioeconomic status) impact engagement or retention. The use of app-based reminders or notifications, as well as incentive structures such as gamification, should be examined to see how these app features impact user engagement. Understanding how a mental health app’s content, approachability, and style affects user engagement is another critical next step. The user version of the Mobile Application Rating Scale (uMARS) [41]—the only self-report scale that is developed for use by the general public [42]—can provide subjective data on user perceptions of a mental health app.

### Conclusions

The scope and impact of depression worldwide is breathtaking. Results show a global interest in a publicly available, free-to-use mHealth app designed to improve depressive symptoms, including in low- and middle-income countries and in countries where English is not the primary language. About half of MoodTools users returned to the app after their initial app session. About one-third of all sessions lasted between 0 and 10 seconds, one-third lasted between 11 seconds and 3 minutes, and the remaining third lasted 3 minutes or longer. The average MoodTools user used the app for 3 sessions for a total of 12 minutes over 90 days. Users tend to spend most time using tools designed for self-monitoring of symptoms and for targeting a core mechanism of depressive psychopathology and negative cognitions. Observational data from this study show that



self-guided mental health apps could be one among the many approaches needed to reduce the global burden of depression; however, research is needed to determine whether app engagement can lead to symptom improvement.

## Conflicts of Interest

LS is the cocreator of MoodTools and one of the owners of Inquiry Health LLC, which publishes MoodTools. PLA has no conflicts of interest to declare.

## Multimedia Appendix 1

MoodTools users by continent, subcontinent, and country.

[[DOCX File , 43 KB - formative\\_v6i10e35538\\_app1.docx](#) ]

## References

1. van Heerden A, Tomlinson M, Swartz L. Point of care in your pocket: a research agenda for the field of m-health. *Bull World Health Org* 2012 May 01;90(5):393-394. [doi: [10.2471/blt.11.099788](#)]
2. GSMA. The Mobile Economy 2020. 2020. URL: [https://www.gsma.com/mobileeconomy/wp-content/uploads/2020/03/GSMA\\_MobileEconomy2020\\_Global.pdf](https://www.gsma.com/mobileeconomy/wp-content/uploads/2020/03/GSMA_MobileEconomy2020_Global.pdf) [accessed 2022-09-14]
3. Ferrari AJ, Charlson FJ, Norman RE, Patten SB, Freedman G, Murray CJL, et al. Burden of depressive disorders by country, sex, age, and year: findings from the global burden of disease study 2010. *PLoS Med* 2013 Nov;10(11):e1001547 [FREE Full text] [doi: [10.1371/journal.pmed.1001547](#)] [Medline: [24223526](#)]
4. Hawton K, Comabella CCI, Haw C, Saunders K. Risk factors for suicide in individuals with depression: a systematic review. *J Affect Disord* 2013 May;147(1-3):17-28. [doi: [10.1016/j.jad.2013.01.004](#)] [Medline: [23411024](#)]
5. Depression and other common mental disorders: global health estimates. World Health Organization. 2017. URL: <https://apps.who.int/iris/handle/10665/254610> [accessed 2022-09-14]
6. Liu JX, Goryakin Y, Maeda A, Bruckner T, Scheffler R. Global health workforce labor market projections for 2030. *Hum Resour Health* 2017 Feb 03;15(1):11 [FREE Full text] [doi: [10.1186/s12960-017-0187-2](#)] [Medline: [28159017](#)]
7. Global health workforce, finances remain low for mental health. World Health Organization. Geneva; 2015. URL: <https://www.who.int/news/item/14-07-2015-global-health-workforce-finances-remain-low-for-mental-health#> [accessed 2022-09-14]
8. Lintvedt OK, Griffiths KM, Eisemann M, Waterloo K. Evaluating the translation process of an internet-based self-help intervention for prevention of depression: a cost-effectiveness analysis. *J Med Internet Res* 2013 Jan 23;15(1):e18 [FREE Full text] [doi: [10.2196/jmir.2422](#)] [Medline: [23343481](#)]
9. Muñoz RF, Bunge EL, Chen K, Schueller SM, Bravin JI, Shaughnessy EA, et al. Massive open online interventions: a novel model for delivering behavioral health services worldwide. *Clin Psychol Sci* 2015 May 13;4(2):194-205. [doi: [10.1177/2167702615583840](#)]
10. 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 14;33(2):217-227. [doi: [10.1017/s0033291702007225](#)] [Medline: [12622301](#)]
11. Proudfoot J, Ryden C, Everitt B, Shapiro DA, Goldberg D, Mann A, et al. Clinical efficacy of computerised cognitive-behavioural therapy for anxiety and depression in primary care: randomised controlled trial. *Br J Psychiatry* 2004 Jul 01;185(1):46-54. [doi: [10.1192/bjp.185.1.46](#)] [Medline: [15231555](#)]
12. Christensen H, Griffiths KM, Jorm AF. Delivering interventions for depression by using the internet: randomised controlled trial. *BMJ* 2004 Jan 31;328(7434):265 [FREE Full text] [doi: [10.1136/bmj.37945.566632.EE](#)] [Medline: [14742346](#)]
13. 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](#)]
14. Waller R, Gilbody S. Barriers to the uptake of computerized cognitive behavioural therapy: a systematic review of the quantitative and qualitative evidence. *Psychol Med* 2009 May;39(5):705-712. [doi: [10.1017/S0033291708004224](#)] [Medline: [18812006](#)]
15. DaPonte D, Talbot F, Titov N, Dear BF, Hadjistavropoulos HD, Hadjistavropoulos T, et al. Facilitating the dissemination of iCBT for the treatment of anxiety and depression: a feasibility study. *Behav change* 2018 Jun 18;35(3):139-151. [doi: [10.1017/bec.2018.14](#)]
16. Gilbody S, Littlewood E, Hewitt C, Brierley G, Tharmanathan P, Araya R, REEACT Team. Computerised cognitive behaviour therapy (cCBT) as treatment for depression in primary care (REEACT trial): large scale pragmatic randomised controlled trial. *BMJ* 2015 Nov 11;351:h5627 [FREE Full text] [doi: [10.1136/bmj.h5627](#)] [Medline: [26559241](#)]
17. Firth J, Torous J, Nicholas J, Carney R, Pratap A, Rosenbaum S, et al. The efficacy of smartphone-based mental health interventions for depressive symptoms: a meta-analysis of randomized controlled trials. *World Psychiatry* 2017 Oct;16(3):287-298 [FREE Full text] [doi: [10.1002/wps.20472](#)] [Medline: [28941113](#)]

18. Poushter J. Smartphone ownership and internet usage continues to climb in emerging economies. Pew Research Center. 2016. URL: <https://www.pewresearch.org/global/2016/02/22/smartphone-ownership-and-internet-usage-continues-to-climb-in-emerging-economies/> [accessed 2022-09-14]
19. Taylor K, Silver L. Smartphone ownership is growing rapidly around the world, but not always equally. Pew Research Center. 2019. URL: <https://www.pewresearch.org/global/2019/02/05/smartphone-ownership-is-growing-rapidly-around-the-world-but-not-always-equally/> [accessed 2022-09-14]
20. Mohr DC, Lyon AR, Lattie EG, Reddy M, Schueller SM. Accelerating digital mental health research from early design and creation to successful implementation and sustainment. *J Med Internet Res* 2017 May 10;19(5):e153 [FREE Full text] [doi: [10.2196/jmir.7725](https://doi.org/10.2196/jmir.7725)] [Medline: [28490417](https://pubmed.ncbi.nlm.nih.gov/28490417/)]
21. 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/)]
22. Carlo AD, Hosseini Ghomi R, Renn BN, Areán PA. By the numbers: ratings and utilization of behavioral health mobile applications. *NPJ Digit Med* 2019 Jun 17;2(1):54 [FREE Full text] [doi: [10.1038/s41746-019-0129-6](https://doi.org/10.1038/s41746-019-0129-6)] [Medline: [31304400](https://pubmed.ncbi.nlm.nih.gov/31304400/)]
23. Anthes E. Mental health: there's an app for that. *Nature* 2016 Apr 07;532(7597):20-23. [doi: [10.1038/532020a](https://doi.org/10.1038/532020a)] [Medline: [27078548](https://pubmed.ncbi.nlm.nih.gov/27078548/)]
24. 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](https://doi.org/10.2196/jmir.9275)] [Medline: [29875089](https://pubmed.ncbi.nlm.nih.gov/29875089/)]
25. 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/)]
26. Lattie EG, Schueller SM, Sargent E, Stiles-Shields C, Tomasino KN, Corden ME, et al. Uptake and usage of IntelliCare: a publicly available suite of mental health and well-being apps. *Internet Interv* 2016 May;4(2):152-158 [FREE Full text] [doi: [10.1016/j.invent.2016.06.003](https://doi.org/10.1016/j.invent.2016.06.003)] [Medline: [27398319](https://pubmed.ncbi.nlm.nih.gov/27398319/)]
27. Kuhn E, van der Meer C, Owen JE, Hoffman JE, Cash R, Carrese P, et al. PTSD Coach around the world. *Mhealth* 2018 May;4:15 [FREE Full text] [doi: [10.21037/mhealth.2018.05.01](https://doi.org/10.21037/mhealth.2018.05.01)] [Medline: [29963560](https://pubmed.ncbi.nlm.nih.gov/29963560/)]
28. Inkster B, Sarda S, Subramanian V. An empathy-driven, conversational artificial intelligence agent (Wysa) for digital mental well-being: real-world data evaluation mixed-methods study. *JMIR Mhealth Uhealth* 2018 Nov 23;6(11):e12106 [FREE Full text] [doi: [10.2196/12106](https://doi.org/10.2196/12106)] [Medline: [30470676](https://pubmed.ncbi.nlm.nih.gov/30470676/)]
29. Howells A, Ivtzan I, Eiroa-Orosa FJ. Putting the 'app' in happiness: a randomized controlled trial of a smartphone-based mindfulness intervention to enhance well-being. *J Happiness Stud* 2014 Oct 29;17(1):163-185. [doi: [10.1007/s10902-014-9589-1](https://doi.org/10.1007/s10902-014-9589-1)]
30. 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/)]
31. Bush NE, Skopp N, Smolenski D, Crumpton R, Fairall J. Behavioral screening measures delivered with a smartphone app: psychometric properties and user preference. *J Nerv Ment Dis* 2013 Nov;201(11):991-995. [doi: [10.1097/NMD.000000000000039](https://doi.org/10.1097/NMD.000000000000039)] [Medline: [24177488](https://pubmed.ncbi.nlm.nih.gov/24177488/)]
32. Beck JS. *Cognitive Behavior Therapy: Basics and Beyond*. 2nd ed. New York: Guilford Press; 2011.
33. Marshall JM, Dunstan DA, Bartik W. The digital psychiatrist: in search of evidence-based apps for anxiety and depression. *Front Psychiatry* 2019 Nov 15;10:831 [FREE Full text] [doi: [10.3389/fpsyt.2019.00831](https://doi.org/10.3389/fpsyt.2019.00831)] [Medline: [31803083](https://pubmed.ncbi.nlm.nih.gov/31803083/)]
34. Torous J, Lipschitz J, Ng M, Firth J. Dropout rates in clinical trials of smartphone apps for depressive symptoms: a systematic review and meta-analysis. *J Affect Disord* 2020 Feb 15;263:413-419. [doi: [10.1016/j.jad.2019.11.167](https://doi.org/10.1016/j.jad.2019.11.167)] [Medline: [31969272](https://pubmed.ncbi.nlm.nih.gov/31969272/)]
35. 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/)]
36. Mattila E, Lappalainen R, Välikkynen P, Sairanen E, Lappalainen P, Karhunen L, et al. Usage and dose response of a mobile acceptance and commitment therapy app: secondary analysis of the intervention arm of a randomized controlled trial. *JMIR Mhealth Uhealth* 2016 Jul 28;4(3):e90 [FREE Full text] [doi: [10.2196/mhealth.5241](https://doi.org/10.2196/mhealth.5241)] [Medline: [27468653](https://pubmed.ncbi.nlm.nih.gov/27468653/)]
37. Zhang R, Nicholas J, Knapp AA, Graham AK, Gray E, Kwasny MJ, et al. Clinically meaningful use of mental health apps and its effects on depression: mixed methods study. *J Med Internet Res* 2019 Dec 20;21(12):e15644 [FREE Full text] [doi: [10.2196/15644](https://doi.org/10.2196/15644)] [Medline: [31859682](https://pubmed.ncbi.nlm.nih.gov/31859682/)]
38. Howard KI, Kopta SM, Krause MS, Orlinsky DE. The dose-effect relationship in psychotherapy. *Am Psychol* 1986;41(2):159-164. [doi: [10.1037/0003-066x.41.2.159](https://doi.org/10.1037/0003-066x.41.2.159)]
39. McNeilly CL, Howard KI. The effects of psychotherapy: a re-evaluation based on dosage. *Psychother Res* 2008 Jan 14;1(1):74-78. [doi: [10.1080/10503309112331334081](https://doi.org/10.1080/10503309112331334081)]
40. Pryss R, Reichert M, Schlee W, Spiliopoulou M, Langguth B, Probst T. Differences between Android and iOS Users of the TrackYourTinnitus mobile crowdsensing mHealth platform. In: 2018 IEEE 31st International Symposium on

- Computer-Based Medical Systems. 2018 Presented at: IEEE 31st International Symposium on Computer-Based Medical Systems; 18-21, 2018; Karlstad, Sweden p. 411-416. [doi: [10.1109/cbms.2018.00078](https://doi.org/10.1109/cbms.2018.00078)]
41. Stoyanov SR, Hides L, Kavanagh DJ, Wilson H. Development and validation of the user version of the Mobile Application Rating Scale (uMARS). JMIR Mhealth Uhealth 2016 Jun 10;4(2):e72 [FREE Full text] [doi: [10.2196/mhealth.5849](https://doi.org/10.2196/mhealth.5849)] [Medline: [27287964](https://pubmed.ncbi.nlm.nih.gov/27287964/)]
  42. Azad-Khaneghah P, Neubauer N, Miguel Cruz A, Liu L. Mobile health app usability and quality rating scales: a systematic review. Disabil Rehabil Assist Technol 2021 Oct 08;16(7):712-721. [doi: [10.1080/17483107.2019.1701103](https://doi.org/10.1080/17483107.2019.1701103)] [Medline: [31910687](https://pubmed.ncbi.nlm.nih.gov/31910687/)]

## Abbreviations

**mHealth:** mobile health

**PHQ-9:** Patient Health Questionnaire

**SDK:** software development kit

**uMARS:** user version of the Mobile Application Rating Scale

*Edited by A Mavragani; submitted 08.12.21; peer-reviewed by B Nelson, A McLean; comments to author 01.03.22; revised version received 25.07.22; accepted 26.08.22; published 25.10.22.*

*Please cite as:*

*Su L, Anderson PL*

*User Behavior of a Publicly Available, Free-to-Use, Self-guided mHealth App for Depression: Observational Study in a Global Sample*  
*JMIR Form Res 2022;6(10):e35538*

*URL: <https://formative.jmir.org/2022/10/e35538>*

*doi: [10.2196/35538](https://doi.org/10.2196/35538)*

*PMID: [36282559](https://pubmed.ncbi.nlm.nih.gov/36282559/)*

©Langting Su, Page Lyn Anderson. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 25.10.2022. 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

# Measuring the Usability of eHealth Solutions for Patients With Parkinson Disease: Observational Study

Jonas Bendig<sup>1</sup>, Dr med; Anja Spanz<sup>1</sup>; Jana Leidig<sup>1</sup>; Anika Frank<sup>1,2</sup>, Dr med; Marcus Stahr<sup>1</sup>, Dr med; Heinz Reichmann<sup>1</sup>, Dr med; Kai F Loewenbrück<sup>1\*</sup>, Dr med; Björn H Falkenburger<sup>1,2\*</sup>, Dr med

<sup>1</sup>Department of Neurology, University Hospital and Faculty of Medicine Carl Gustav Carus, Technische Universität Dresden, Dresden, Germany

<sup>2</sup>German Center for Neurodegenerative Diseases, Dresden, Germany

\*these authors contributed equally

**Corresponding Author:**

Jonas Bendig, Dr med

Department of Neurology

University Hospital and Faculty of Medicine Carl Gustav Carus

Technische Universität Dresden

Fetscherstrasse 74

Dresden, 01307

Germany

Phone: 49 351 458 19675

Email: [Jonas.Bendig@uniklinikum-dresden.de](mailto:Jonas.Bendig@uniklinikum-dresden.de)

## Abstract

**Background:** Parkinson disease (PD) is a neurodegenerative disorder with a variety of motor and nonmotor symptoms. Many of these symptoms can be monitored by eHealth solutions, including smartphone apps, wearable sensors, and camera systems. The usability of such systems is a key factor in long-term use, but not much is known about the predictors of successful use and preferable methods to assess usability in patients with PD.

**Objective:** This study tested methods to assess usability and determined prerequisites for successful use in patients with PD.

**Methods:** We performed comprehensive usability assessments with 18 patients with PD using a mixed methods usability battery containing the System Usability Scale, a rater-based evaluation of device-specific tasks, and qualitative interviews. Each patient performed the usability battery with 2 of 3 randomly assigned devices: a tablet app, wearable sensors, and a camera system. The usability battery was administered at the beginning and at the end of a 4-day testing period. Between usability batteries, the systems were used by the patients during 3 sessions of motor assessments (wearable sensors and camera system) and at the movement disorder ward (tablet app).

**Results:** In this study, the rater-based evaluation of tasks discriminated the best between the 3 eHealth solutions, whereas subjective modalities such as the System Usability Scale were not able to distinguish between the systems. Successful use was associated with different clinical characteristics for each system: eHealth literacy and cognitive function predicted successful use of the tablet app, and better motor function and lower age correlated with the independent use of the camera system. The successful use of the wearable sensors was independent of clinical characteristics. Unfortunately, patients who were not able to use the devices well provided few improvement suggestions in qualitative interviews.

**Conclusions:** eHealth solutions should be developed with a specific set of patients in mind and subsequently tested in this cohort. For a complete picture, usability assessments should include a rater-based evaluation of task performance, and there is a need to develop strategies to circumvent the underrepresentation of poorly performing patients in qualitative usability research.

(*JMIR Form Res* 2022;6(10):e39954) doi:[10.2196/39954](https://doi.org/10.2196/39954)

**KEYWORDS**

eHealth; usability; Parkinson disease; telehealth and telemonitoring; older adults; aging; older population; neurodegenerative disease; digital solution; wearable sensor; mobile health; system usability; eHealth solution

## Introduction

Parkinson disease (PD) is a neurodegenerative disorder characterized by a variety of motor and nonmotor symptoms. Despite the neurodegenerative nature of the disease, dopamine replacement therapy can drastically improve symptoms and quality of life, especially in the early stages of the disease [1]. With longer disease duration, symptoms often begin to fluctuate during the day, making the exact timing and dosage of medication more important [2]. eHealth solutions are becoming increasingly available, offering the potential to remind patients of their medications, assess the extent and timing of motor fluctuations, and ultimately help guide the decision for advanced therapeutic options such as deep brain stimulation and medication pumps [3-5]. In addition, eHealth solutions enable clinicians to assess patients over extended periods of time in their home environment. This method can help improve patient care but also provide more precise and more relevant end points for clinical trials [6].

A wide range of eHealth solutions has been examined in patients with PD, but most studies focus on selected subgroups of patients, such as those in earlier stages of the disease [7]. In the clinical routine, however, patients with PD are distributed across a wide range of age groups with diverse educational backgrounds, distinct motor impairments and—in many patients—important psychiatric and cognitive comorbidities [8]. There is a paucity of studies systematically investigating barriers for the successful implementation of eHealth solutions in the heterogeneous population of patients with PD [9].

In this context, usability research provides a variety of user-based methods that can be categorized into subjective and objective measures and quantitative and qualitative assessments [10]. Quantitative methods primarily include questionnaires and task completions, with questionnaires being the most frequently used method in eHealth research. The questionnaire with the broadest implementation among usability studies is the Systems Usability Scale (SUS) [11], which provides a subjective assessment of usability by the patient. Task completions provide objective information but require a system-specific setup, which can be difficult and time-consuming and potentially limits comparability. Qualitative methods include focus groups, interviews, and think-alouds. In contrast to quantitative methods, they can be more useful in identifying specific usability problems but suffer from a lack of comparability. Moreover, qualitative methods require trained evaluators and laborious data analysis. Think-alouds and qualitative interviews were the most frequently used methods in the usability testing of eHealth solutions [11]. For patients with PD specifically, usability assessments have mainly relied on questionnaires as well as adherence monitoring and mostly reported positive results for sensor systems and smartphone/tablet apps [12-15]. Although mixed methods approaches have become more common recently, there is substantial heterogeneity in the methods, and only a minority of studies focused specifically on the usability. In the broader context of chronic conditions, a recent systematic review concluded that the usability of wearable devices is poorly measured and reported [15]. Furthermore, there is no consensus regarding the methodology to assess usability in older adults,

even though investigations about the sensitivity of different methods have been explicitly recommended [16].

Against this background, we aimed to identify which methods are suitable for comprehensive usability testing in our primarily older cohort of patients with PD and which factors can predict the successful use of devices for telemedicine and home monitoring. For this objective, we designed a mixed methods usability battery based on the most commonly used quantitative and qualitative methods for eHealth solutions and tested the usability of 3 different devices, including (1) a tablet app, (2) wearable sensors, and (3) a camera system.

## Methods

### Study Population

In all, 18 patients were recruited from the ward for movement disorders at the University Hospital Dresden between July 2020 and September 2021. Written informed consent was obtained from all participants before inclusion in the study. Inclusion criteria were the clinically probable diagnosis of idiopathic PD by a specialist for movement disorders according to the current guidelines of the International Movement Disorders Society [17] and sufficient German language skills. Exclusion criteria were the inability to walk and any psychiatric comorbidity that excluded the patients from participating in the study according to the discretion of the investigator.

### Ethics Approval

The study was approved by the institutional review board of Technische Universität Dresden, Germany (BO-EK-212052020).

### Tested Systems

We assessed 3 eHealth solutions that use different paradigms: (1) guided measurements at specific time points by a camera system; (2) continuous, implicit monitoring of mobility by wearable sensors; and (3) a combination of guided and continuous measurements by a tablet app. The systems were described in detail previously [18].

Briefly, the 3D-camera system (Motognosis Amsa; Motognosis GmbH) consisted of a stand-alone PC and a depth camera (Microsoft Kinect; version 2). Patients were instructed to perform motor exercises by prerecorded videos and audio instructions. Kinematic parameters were derived from the exercises to describe patients' mobility and symptoms.

The wearable system (PD Neurotechnology Ltd) consisted of five 9-axis inertial measurement unit sensors, worn on wrists, shanks, and the trunk. The data from the sensors were used to analyze patients' motor status [19]. The PDMonitor mobile app was not used in this study nor was the device used for motor symptom clinical assessment and treatment modification.

The tablet app (TelePark tablet app; Intecsoft group) included a medication alert, questionnaires, fall documentation, activity documentation, and a task reminder.

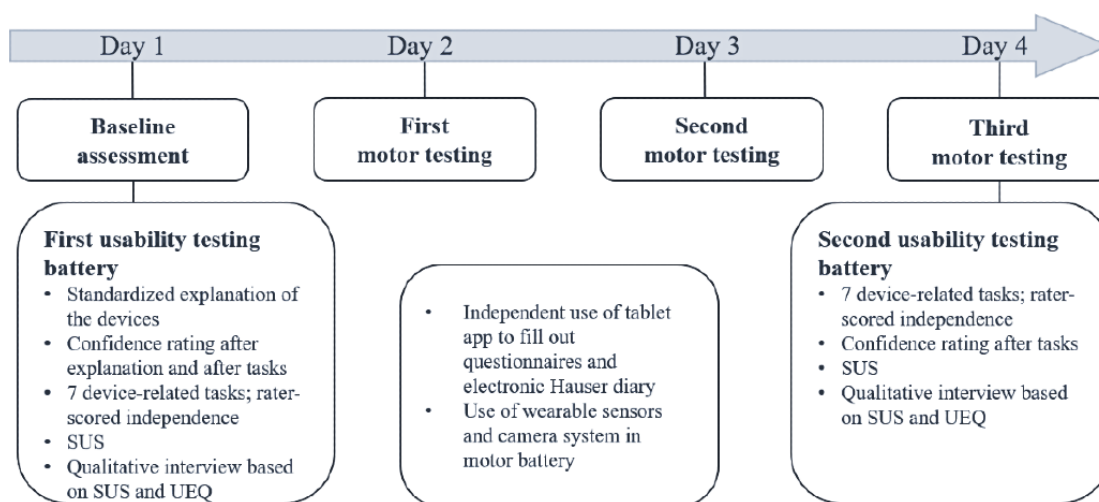
The tablet app represented a system that was still at an early stage of development, whereas the 3D-camera system and the wearable sensors were already fully developed and licensed medical products.

## Study Schedule

Inpatients completed 4 assessments on 4 days within a maximum period of 7 days during their stay at the movement disorders ward (Figure 1). On day 1, patients performed the baseline assessment and the first mixed methods usability testing battery (detailed below). Routine motor testing was carried out on days 2 and 3 (detailed below). Patients were filmed with the camera system and wore the wearable sensors during the motor testing

sessions. Between assessments, patients used the TelePark app to complete questionnaires and an electronic version of the Hauser diary [20]. The wearable sensors and the camera system were only used or worn during the motor tests and usability batteries. Patients were encouraged to put on or remove the sensors independently but received help from the study personnel if requested. On day 4, patients performed a final round of motor testing and the second usability testing battery.

**Figure 1.** Schematic overview of the study schedule. To keep the assessments efficient, only 2 of the 3 devices (tablet app, camera system, and wearable sensors) were tested per patient, resulting in 3 groups of patients that used the same set of devices. UEQ: User Experience Questionnaire; SUS: System Usability Scale.



## Baseline Assessment

Patients were assessed with rater-based scales and self-report questionnaires to evaluate motor and cognitive function as well as eHealth literacy. The questionnaires were filled out digitally by the patients in the TelePark app. If patients were not able to independently complete the questionnaires on the tablet, they were supported by the investigators. The following scales and questionnaires were used in the baseline assessment: the Freezing of Gait Questionnaire (FOG-Q) [21], Hoehn and Yahr scale [22], Unified Parkinson's Disease Rating Scale III (UPDRS III) [23], Beck Depression Inventory-II [24], Montreal Cognitive Assessment (MOCA) [25], and eHealth Literacy Scale (EHEALS) [26]. The total score of the EHEALS ranges between 8 and 40, with higher scores indicating higher self-perceived eHealth literacy.

## Motor Testing

Motor testing consisted of a UPDRS III, a timed up-and-go test [27], a freezing of gait test [28], a Mini-BESTest [29], fast 360° turns, the video-instructed Motognosis Amsa protocol (finger tapping, stand up and sit down, stance with closed feet, comfortable 360° turns, stepping in place, short comfortable speed walk, and short maximum speed walk), and the operator-instructed Motognosis PASS-PD protocol (finger tapping, hand grasping, arm holding, finger-nose test, foot tapping, stand up and sit down, stance with closed feet, comfortable 360° turns, stepping in place, comfortable walk, and maximum speed walk). During the period of the assessment, patients were filmed by the 3D-camera system and wore the wearable sensors.

## Usability Testing Battery

The usability testing battery was performed on the first and the last day of the study. To reduce patient burden, each patient assessed the usability of only 2 of the 3 study components (tablet app and camera, tablet app and wearables, or camera and wearables). The devices were assigned randomly to the patients by a prespecified permuted list. Usability was assessed for each device separately.

First, patients were given a standardized explanation of the device. Patients were then instructed to carry out 7 device-related tasks, which covered all important functions of the systems, as independently as possible. These tasks were setting up the camera and performing different tasks in the Amsa protocol (camera system), putting on the sensors and handling the charging procedure and the data transfer processes (wearable sensors), and using all relevant functions in the app (tablet app). The execution of the tasks was observed by the investigators and rated on a 6-item ordinal scale according to the independence of task execution (ranging from 5="Does not need help; does not consult manual" to 0="Can contribute nothing or almost nothing to the implementation of the task"). The sum of all 7 independence ratings was transformed into a rater-based independence score ranging from 0% (no independent use in any tasks) to 100% (fully independent use in all tasks) with the following formula:



After the task-related device testing, patients filled out the SUS [30] and were asked again how confident they felt now to use the devices alone in a home monitoring setting (confidence

score from 0% to 100%). The SUS is a 10-item Likert scale to assess subjective usability, containing questions such as the perceived complexity of a system, the user's confidence in using a system, or its learnability. The SUS has been widely used, and normative data exist allowing SUS ratings to be positioned relative to other systems [31]. Furthermore, it has been shown that the SUS can provide valid scores even with small sample sizes [32].

Finally, we conducted an interview based on domains of established usability instruments (SUS and user experience questionnaire [33]). This interview took 15 to 30 minutes and consisted of 12 open-ended questions concerning the following domains: attractiveness, independent use, learnability, perspicuity, efficiency, stimulation, and novelty. For each domain, the patients were asked 2 open-ended questions about their opinion on the domain quality and about improvement suggestions in that specific domain. The same procedure was then carried out for the second device.

### Data Analyses and Sample Size

Data are depicted as median with 25th and 75th percentile or as mean with SD depending on data normality as assessed by

a visual inspection of histograms. To assess the differences between the systems, a Kruskal-Wallis test with post hoc Dunn test was used. Due to the small sample size and the exploratory nature of the study, no correction for multiple testing was used. Predictors of successful use were identified by correlation analysis (Spearman  $\rho$ ). Significant correlations ( $P < .05$ ) were visualized in a network graph with the *ForceAtlas2* algorithm [34]. The temporal stability of usability outcomes was assessed by comparing the first and the second measurement with a Wilcoxon signed-rank test. Data visualization and statistical analyses were performed with Python (*Statsmodels*, *Scipy*, *Matplotlib*, and *Seaborn* packages) and Gephi software. The sample size of 12 patients per system was determined using guidelines for conducting qualitative research [35].

## Results

In total, 19 patients were included in the study, and 1 patient dropped out after the first usability battery due to personal reasons (not named). The clinical and demographic data of the remaining 18 patients are summarized in [Table 1](#).

**Table 1.** Clinical and demographic data. Data are presented as mean with SD or median with absolute range.

	Value
Patient, n	18
Age (years), median (range)	69 (37-86)
<b>Sex (N=18), n (%)</b>	
Female	7 (39)
Male	11 (61)
Hoehn and Yahr stage, median (range)	3 (1-4)
Disease duration (years), mean (SD)	11 (7.3)
UPDRS III <sup>a</sup> score, mean (SD)	27 (9.0)
MOCA <sup>b</sup> score, mean (SD)	25 (2.7)
EHEALS <sup>c</sup> score, mean (SD)	23 (8.8)
BDI-II <sup>d</sup> score, mean (SD)	12 (7.4)
FOG-Q <sup>e</sup> score, mean (SD)	11 (5.1)

<sup>a</sup>UPDRS III: Unified Parkinson's Disease Rating Scale III.

<sup>b</sup>MOCA: Montreal Cognitive Assessment.

<sup>c</sup>EHEALS: eHealth Literacy Scale.

<sup>d</sup>BDI-II: Beck Depression Inventory-II.

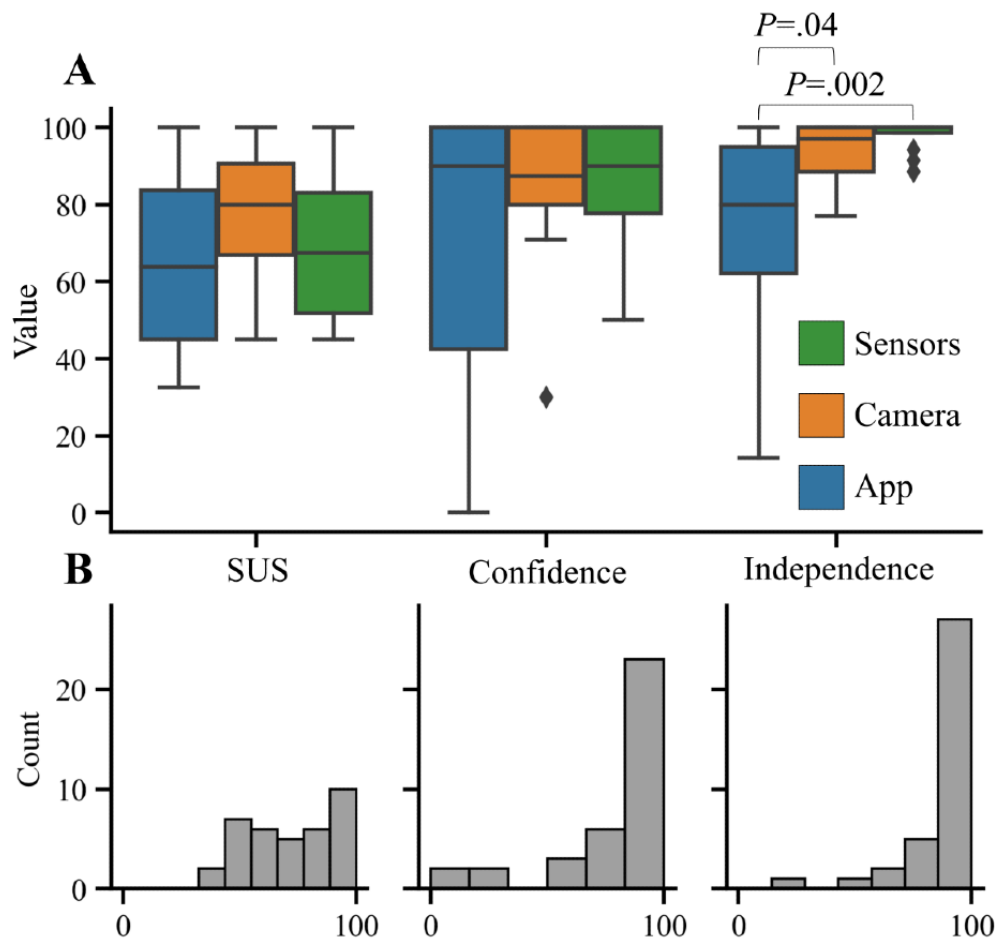
<sup>e</sup>FOG-Q: Freezing of Gait Questionnaire.

### Testing Usability Measures

The SUS is a widely used score for a quick and simple assessment of usability [31]. SUS scores (second usability battery) did not differ significantly between devices ( $P = .34$ , Kruskal-Wallis test; [Figure 2A](#)). In addition, we compared the empirical confidence scores (patient-rated) and the task-based independence scores (investigator-rated) between the 3 systems ([Figure 2A](#)). The confidence scores and the independence scores

showed a pronounced ceiling effect, whereas the SUS scores were more evenly distributed ([Figure 2B](#)). Exclusively, the independence scores differentiated between the app (ie, the device that is still at an early stage of development) and the fully developed and licensed systems ( $P = .006$ , Kruskal-Wallis test; post hoc tests in [Figure 2A](#)). For the subsequent correlation analyses in this study, we therefore selected the objective independence score as the most relevant measure of successful use.

**Figure 2.** Comparison of independence scores, confidence scores, and SUS scores from the second usability battery: (A) Box plots and (B) histograms. *P* values were from Dunn test without correction after significant Kruskal-Wallis test. Box plots depict median (black line), IQR (boxes), range (whiskers), and outliers (diamonds; >75th percentile + 1.5 IQR or <25th percentile – 1.5 IQR). SUS: System Usability Scale.



### Identifying Predictors for Successful Use

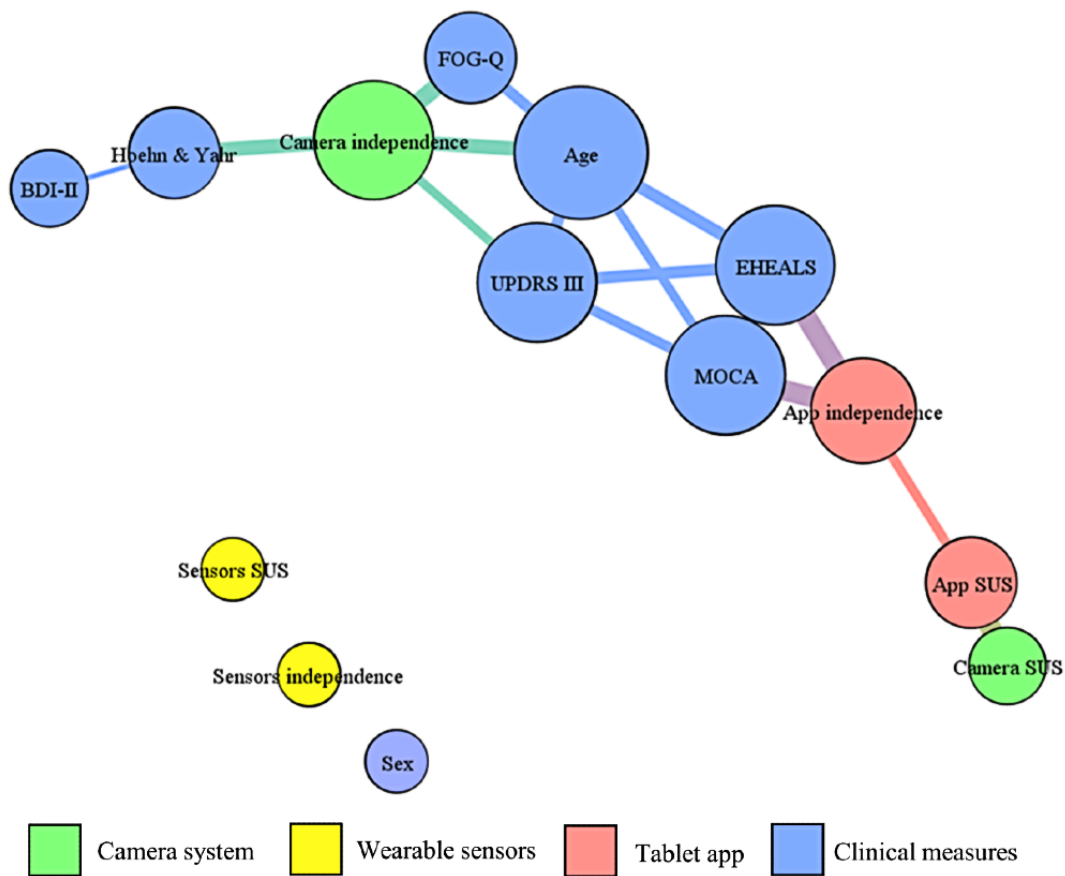
To identify factors that predict whether patients are able to use a device well, we plotted correlation matrices to explore the interdependence between the rater-based independence score, the SUS, and baseline parameters. In addition to the independence scores and SUS scores, the following variables were used in the correlation analysis: age, sex, Hoehn and Yahr stage, UPDRS III, FOG-Q, MOCA, and EHEALS. The network graph of correlations visualizes that the rater-based independence scores for wearable sensors (yellow), camera system (green), and tablet app (red) do not cluster together (Figure 3). This visualization indicates that the prerequisites for successful use differ between the 3 systems. The independent use of the wearable sensors did not correlate significantly with any clinical characteristics (age:  $P=.07$ ; sex:  $P=.38$ ; Hoehn and Yahr stage:  $P=.44$ ; UPDRS III:  $P=.59$ ; FOG-Q:  $P=.94$ ; MOCA:  $P=.40$ ; EHEALS:  $P=.68$ ), but only 3 (25%) out of 12 patients were not able to use the system fully independently. This finding

implies that the sensors were usable for most of the patients regardless of their clinical characteristics. The independent use of the camera system correlated strongly with age and motor scores (FOG-Q, UPDRS III, and Hoehn and Yahr stage), and the independent use of the tablet app showed strong correlations with cognition (MOCA) and eHealth literacy (EHEALS). Table 2 shows the strongest correlations with the rater-based independence score for each system.

In contrast to the rater-based independence scores, we found no significant correlations between the subjective and more variable SUS scores with the clinical measures for the tablet app (age:  $P=.79$ ; sex:  $P=.89$ ; Hoehn and Yahr stage:  $P=.85$ ; UPDRS III:  $P=.92$ ; FOG-Q:  $P=.14$ ; MOCA:  $P=.28$ ; EHEALS:  $P=.07$ ), the wearable sensors (age:  $P=.78$ ; sex:  $P=.15$ ; Hoehn and Yahr stage:  $P=.52$ ; UPDRS III:  $P=.99$ ; FOG-Q:  $P=.12$ ; MOCA:  $P=.96$ ; EHEALS:  $P=.19$ ), or the camera system (age:  $P=.45$ ; sex:  $P=.70$ ; Hoehn and Yahr stage:  $P=.62$ ; UPDRS III:  $P=.16$ ; FOG-Q:  $P=.49$ ; MOCA:  $P=.99$ ; EHEALS:  $P=.26$ ).



**Figure 3.** Network graph of correlations (Spearman  $\rho$ ) between baseline variables, SUS scores, and independence scores from the second usability battery. Only significant correlations ( $P < .05$ , uncorrected values) are included in the network. The relative size of the variables indicates the absolute number of connections. The thickness of the connections indicates the magnitude of the correlation (thicker lines indicate stronger correlations). BDI-II: Beck Depression Inventory; EHEALS: eHealth Literacy Scale; FOG-Q: Freezing of Gait Questionnaire; MOCA: Montreal Cognitive Assessment; UPDRS III: Unified Parkinson's Disease Rating Scale III; SUS: System Usability Scale.



**Table 2.** Correlations of clinical characteristics with independent use. The 3 strongest correlations (Spearman  $\rho$ ) with  $P$  values between clinical characteristics and independence scores for the 3 systems are shown.

Device, clinical characteristic	Spearman $\rho$	$P$ value
<b>Tablet app</b>		
EHEALS <sup>a</sup>	0.90	<.001
MOCA <sup>b</sup>	0.89	<.001
Age	-0.63	.03
<b>Camera system</b>		
FOG-Q <sup>c</sup>	-0.80	.002
Hoehn and Yahr	-0.72	.009
Age	-0.71	.009
<b>Wearable sensors</b>		
— <sup>d</sup>	—	—

<sup>a</sup>EHEALS: eHealth Literacy Scale.

<sup>b</sup>MOCA: Montreal Cognitive Assessment.

<sup>c</sup>FOG-Q: Freezing of Gait Questionnaire.

<sup>d</sup>For the wearable sensors, no significant correlations were found.

## Temporal Change in Usability Outcomes

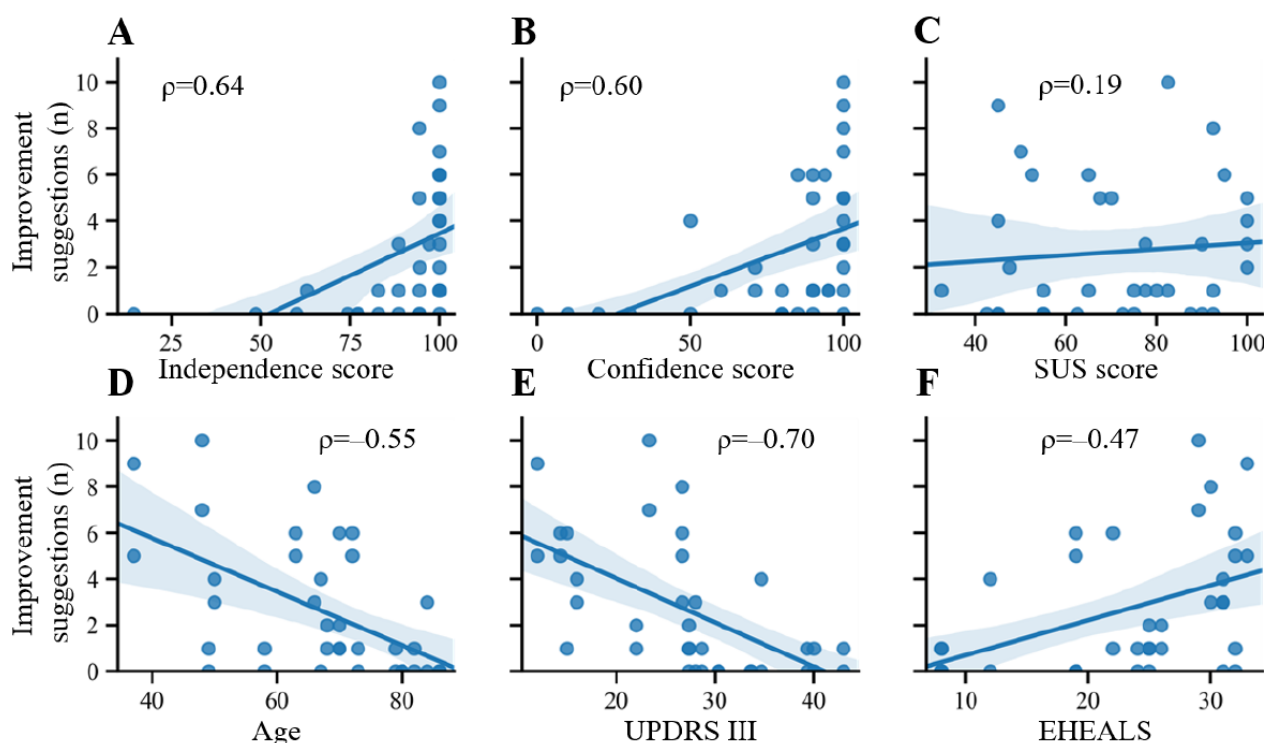
To assess the system-specific learnability and stability of the usability outcomes, we compared usability outcomes between the first and second round of the usability battery on days 1 and 4, respectively. For the tablet app, independence and confidence scores did not differ significantly between the 2 time points (independence: mean 79.5%, SD 25.6% vs 75.5%, SD 25.8%;  $P=.34$ ; confidence: mean 75.4%, SD 24.3% vs 69.6%, SD 38.8%;  $P=.29$ ). The camera system, in contrast, had a significantly higher confidence score in the second usability battery (mean 63.3%, SD 32.5% vs 84.7%, SD 19.8%;  $P=.008$ ); independence scores were high at both time points (mean 89.3%, SD 14.8% vs 93.8%, SD 7.9%;  $P=.12$ ). The wearable sensors showed a significantly higher independence score in the second usability battery (mean 91.4%, SD 9.2% vs 97.9%, SD 4.1%;  $P=.03$ ); confidence ratings did not change (mean 79.5%, SD 25.6% vs 75.5%, SD 25.8%;  $P=.67$ ). SUS scores did not change significantly between the 2 time points for any of the tested systems (tablet app:  $P=.18$ ; camera system:  $P=.20$ ; wearable sensors:  $P=.88$ ). The system-specific changes in usability outcomes indicate a different learnability for each individual system and underscore the importance of longitudinal usability

assessments. Furthermore, they suggest that performance and confidence may differ. The lack of difference in the SUS scores between the 2 time points is consistent with the lack of difference in the SUS scores between the 3 systems (Figure 2), suggesting that the SUS can miss important aspects of usability.

## Influences on Qualitative Interviews

In the qualitative section of the first and second usability batteries, patients were asked about improvement suggestions for the eHealth solutions. To determine predictors of qualitative feedback, we counted the total number of unique improvement suggestions per patient and correlated them with usability outcomes and clinical characteristics. We found moderate-to-strong and highly significant correlations with independence scores, confidence scores, eHealth literacy, motor phenotype, and age (Figure 4). These correlations suggest that patients who were able to use the devices well gave more valid improvement suggestions than patients who did not. Patients giving more feedback were also younger, had lower motor disability, and higher eHealth literacy. SUS scores did not correlate with the number of improvement suggestions, suggesting that the subjective rating of an eHealth solution does not affect the number of improvement suggestions.

**Figure 4.** Linear regression plots of valid improvement suggestions with (A) independence scores, (B) confidence scores, (C) SUS scores, and (D-F) clinical characteristics. The strength of the correlation (Spearman  $\rho$ ) is indicated in the plot. All  $P$  values for the Spearman correlations were  $<.001$  except for the SUS score ( $P=.11$ ). Individual values for each system are plotted. Improvement suggestions are aggregated from both usability measurements for each system. EHEALS: eHealth Literacy Scale; UPDRS III: Unified Parkinson's Disease Rating Scale III; SUS: System Usability Scale.



## Discussion

### Principal Findings

In this study, we performed a comprehensive usability battery on 3 eHealth solutions, using subjective and objective assessments. The objective rater-based evaluation of tasks (independence score) discriminated better between the different

eHealth solutions than the subjective quantitative usability scale (SUS). Moreover, the successful use of each eHealth solution was associated with specific clinical characteristics—notably, cognitive ability and eHealth literacy for the tablet app or motor ability and age for the camera system. Finally, most improvement suggestions were provided by patients who were able to use the eHealth solutions well.

## Comparison Between Usability Measures

There is a paucity of data on the sensitivity of usability testing methods [16], and optimal methods for specific eHealth solutions or cohorts have not been identified [11]. We therefore compared usability as reported by the quantitative and easy-to-use SUS with patient-rated confidence and investigator-rated independence in prespecified settings. Given that the 3 eHealth solutions investigated here (tablet app, camera system, and wearable sensors) differed strongly in complexity and development stage, we expected to find differences in all methods. However, only the rater-based independence scores showed a significant difference between the 3 technologies. The ceiling effect of the independence scores could indicate that the systems were indeed easy to use for many patients. Alternatively, the prespecified tasks were not hard enough. As the tasks were developed to cover all relevant functions of each system, we interpreted this ceiling effect as successful use. The SUS score did not show a ceiling effect, but it did not differentiate between the fully developed systems and the less developed system in our study. Furthermore, the SUS did not reflect the increased confidence and independence between the first and second time point of testing. Collectively, these findings are in line with similar studies, where successful use was not associated with higher SUS scores [36,37]. These findings suggest that this well-established scale could potentially miss important information in the population of patients with PD, and in other populations of older and cognitively impaired people. The recent development of a simplified SUS score for older adults is in line with this interpretation [38].

The improvements in confidence or independence scores for the camera system and the wearable sensors indicate that even in a short period of 4 days, older adults (1) are able to change their perspective toward eHealth solutions and (2) can learn to handle such systems. The lack of improvement for the tablet app shows that learnability is dependent on the eHealth solution, which is in line with previous results from other studies [39]. These results should caution researchers to not rely on a single test to predict successful use. Based on our results, we recommend a short rater-based test, a subjective patient-rating validated in older adults (eg, a questionnaire), and a trial period for each patient and device before applying eHealth solutions in trials or clinical practice.

## Predictors of Successful Use

Predictors of successful use differ strongly between individual eHealth solutions (Figure 3). For the app, the strong associations with cognitive function and eHealth literacy indicate that both constructs need to be considered in the design of such mobile health systems with a largely software-based interface [40,41]. Hence, eHealth solutions should be developed with a specific range of cognitive function and eHealth literacy in mind and then should be tested and marketed for this group of patients. For the camera system, older patients with more severe motor symptoms had more problems, whereas the wearable sensors were usable for most patients independent of clinical characteristics. This finding is not surprising given the mainly physical nature of interacting with the wearable sensors or performing guided tasks in front of the camera. In contrast,

sensors were usable by most patients regardless of their clinical characteristics. Collectively, our findings align well with the MOLD-US framework, where usability prerequisites for the app system fall into the domains of cognition and motivation and prerequisites for the camera system are associated with the physical ability [42].

The subjective aspect reported by the SUS is necessary for a patient to start the use of eHealth solutions to avoid attrition with continual use [10], and indeed, SUS scores varied considerably between participants (Figure 2). We sought to determine predictors of SUS scores. However, we were not able to determine predictors for subjective usability scores as reported by SUS, likely due to the small sample size ( $n=12$  per system) in our study. We only observed for the app an association between the SUS and independence scores in the graph analysis (Figure 3). This analysis, therefore, needs to be addressed in subsequent studies with more participants. Moreover, attrition could not be assessed in this short and highly standardized paradigm.

## Improvement Suggestions

We found a strong positive correlation between successful use and a patient's ability to advise on possible improvements of the tested systems during the qualitative interview (Figure 4A). In other words, suggestions came mainly from individuals that did not have problems using the system. Therefore, established methods such as "think-alouds" or "focus groups" could suffer from an overrepresentation of opinions voiced by well-performing, mildly affected patients. It is not clear whether following these suggestions will improve or worsen usability for those who have trouble using the system successfully. The method of counting the total number of improvement suggestions does not take into account the quality of the suggestions; thus, the presented results should be reinvestigated more thoroughly in future studies. Furthermore, the reported correlation could also be mediated or moderated by the factors age, disease severity, cognitive status, or eHealth literacy (Figure 4D to F). However, inferring causal connections between highly interconnected variables was beyond the scope of this study, and to our knowledge, there are currently no articles that have comprehensively assessed the effects of these variables in the context of qualitative usability research. With an aging population in Western countries and a predicted rise in patients living with neurodegenerative diseases [43,44], a critical assessment of qualitative usability methods in the context of the target group is warranted.

## Limitations

Limitations of our study include the small sample size, only a single recruitment site, and the controlled inpatient setting. Furthermore, the comparison of different methods is based on a subset of the existing tools that does not include techniques such as think-alouds, focus groups, or alternative measures of efficiency (eg, time to complete a task). This limitation reduces the generalizability of our findings and warrants further investigation with different systems, settings, and patient cohorts. Moreover, the high correlations between eHealth literacy, motor symptoms, cognitive impairment, and age limit the causal interpretability of the obtained results.

## Conclusions

The successful use of eHealth solutions in patients with PD is highly dependent on system-specific and patient-specific characteristics. Considering the growing field of digital health and the already existing abundance of different solutions for

patients with PD [4,45,46], researchers and industrial partners need to consider the heterogeneity of patients and design eHealth solution for a specific constellation of age, cognitive and motor function, as well as eHealth literacy, and these criteria can be helpful for physicians in selecting the best solution for each individual patient.

## Acknowledgments

This research was funded by the European regional development fund. Medical devices from the companies PD Neurotechnology Ltd (wearable sensors) and Motognosis GmbH (camera system) were provided free of charge. Motognosis GmbH and PD Neurotechnology Ltd agreed upon cooperation with the Technical University of Dresden for the TelePark study. These agreements include the mutual use of pseudonymized clinical data and the data from each respective system for the development and improvement of algorithms. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

## Authors' Contributions

For the research project, JB, KFL, MS, and BHF contributed to conceptualization; AS, JL, JB, and AF contributed to project administration; and AS, JL, and JB contributed to the investigation. For statistical analysis, JB, AF, KFL, and BHF contributed to the methodology; AS and JB conducted the formal analysis; and BHF and HR contributed to supervision. For manuscript preparation, JB, AF, BHF wrote the original draft and all authors contributed to writing—review and editing. JB and BHF take responsibility for the integrity of the data and the accuracy of the data analysis.

## Conflicts of Interest

None declared.

## References

1. Obeso JA, Stamelou M, Goetz CG, Poewe W, Lang AE, Weintraub D, et al. Past, present, and future of Parkinson's disease: a special essay on the 200th anniversary of the shaking palsy. *Mov Disord* 2017 Sep;32(9):1264-1310 [FREE Full text] [doi: [10.1002/mds.27115](https://doi.org/10.1002/mds.27115)] [Medline: [28887905](https://pubmed.ncbi.nlm.nih.gov/28887905/)]
2. Fox SH, Katzenschlager R, Lim S, Barton B, de Bie RMA, Seppi K, Movement Disorder Society Evidence-Based Medicine Committee. International Parkinson and movement disorder society evidence-based medicine review: update on treatments for the motor symptoms of Parkinson's disease. *Mov Disord* 2018 Aug;33(8):1248-1266. [doi: [10.1002/mds.27372](https://doi.org/10.1002/mds.27372)] [Medline: [29570866](https://pubmed.ncbi.nlm.nih.gov/29570866/)]
3. Sica M, Tedesco S, Crowe C, Kenny L, Moore K, Timmons S, et al. Continuous home monitoring of Parkinson's disease using inertial sensors: a systematic review. *PLoS One* 2021 Apr 04;16(2):e0246528 [FREE Full text] [doi: [10.1371/journal.pone.0246528](https://doi.org/10.1371/journal.pone.0246528)] [Medline: [33539481](https://pubmed.ncbi.nlm.nih.gov/33539481/)]
4. Sibley KG, Girges C, Hoque E, Foltynie T. Video-based analyses of Parkinson's disease severity: a brief review. *J Parkinsons Dis* 2021 Jul 16;11(s1):S83-S93 [FREE Full text] [doi: [10.3233/JPD-202402](https://doi.org/10.3233/JPD-202402)] [Medline: [33682727](https://pubmed.ncbi.nlm.nih.gov/33682727/)]
5. Little MA. Smartphones for remote symptom monitoring of Parkinson's disease. *J Parkinsons Dis* 2021 Jul 16;11(s1):S49-S53 [FREE Full text] [doi: [10.3233/JPD-202453](https://doi.org/10.3233/JPD-202453)] [Medline: [33814462](https://pubmed.ncbi.nlm.nih.gov/33814462/)]
6. Espay AJ, Bonato P, Nahab FB, Maetzler W, Dean JM, Klucken J, Movement Disorders Society Task Force on Technology. Technology in Parkinson's disease: challenges and opportunities. *Mov Disord* 2016 Sep;31(9):1272-1282 [FREE Full text] [doi: [10.1002/mds.26642](https://doi.org/10.1002/mds.26642)] [Medline: [27125836](https://pubmed.ncbi.nlm.nih.gov/27125836/)]
7. Morgan C, Rolinski M, McNaney R, Jones B, Rochester L, Maetzler W, et al. Systematic review looking at the use of technology to measure free-living symptom and activity outcomes in Parkinson's disease in the home or a home-like environment. *J Parkinsons Dis* 2020 Apr 03;10(2):429-454 [FREE Full text] [doi: [10.3233/JPD-191781](https://doi.org/10.3233/JPD-191781)] [Medline: [32250314](https://pubmed.ncbi.nlm.nih.gov/32250314/)]
8. Greenland JC, Williams-Gray CH, Barker RA. The clinical heterogeneity of Parkinson's disease and its therapeutic implications. *Eur J Neurosci* 2019 Feb;49(3):328-338. [doi: [10.1111/ejn.14094](https://doi.org/10.1111/ejn.14094)] [Medline: [30059179](https://pubmed.ncbi.nlm.nih.gov/30059179/)]
9. Cancela J, Pastorino M, Tzallas AT, Tsiouras MG, Rigas G, Arredondo MT, et al. Wearability assessment of a wearable system for Parkinson's disease remote monitoring based on a body area network of sensors. *Sensors (Basel)* 2014 Sep 16;14(9):17235-17255 [FREE Full text] [doi: [10.3390/s140917235](https://doi.org/10.3390/s140917235)] [Medline: [25230307](https://pubmed.ncbi.nlm.nih.gov/25230307/)]
10. Hornbæk K. Current practice in measuring usability: challenges to usability studies and research. *Int J Hum Comput Stud* 2006 Feb;64(2):79-102. [doi: [10.1016/j.ijhcs.2005.06.002](https://doi.org/10.1016/j.ijhcs.2005.06.002)]
11. Maramba I, Chatterjee A, Newman C. Methods of usability testing in the development of eHealth applications: a scoping review. *Int J Med Inform* 2019 Jun;126:95-104. [doi: [10.1016/j.ijmedinf.2019.03.018](https://doi.org/10.1016/j.ijmedinf.2019.03.018)] [Medline: [31029270](https://pubmed.ncbi.nlm.nih.gov/31029270/)]

12. Bouça-Machado R, Pona-Ferreira F, Leitão M, Clemente A, Vila-Viçosa D, Kauppila LA, et al. Feasibility of a mobile-based system for unsupervised monitoring in Parkinson's disease. *Sensors (Basel)* 2021 Jul 21;21(15):4972 [FREE Full text] [doi: [10.3390/s21154972](https://doi.org/10.3390/s21154972)] [Medline: [34372208](https://pubmed.ncbi.nlm.nih.gov/34372208/)]
13. Gatsios D, Antonini A, Gentile G, Marcante A, Pellicano C, Macchiusi L, et al. Feasibility and utility of mHealth for the remote monitoring of Parkinson disease: ancillary study of the PD\_manager randomized controlled trial. *JMIR mHealth uHealth* 2020 Jun 29;8(6):e16414 [FREE Full text] [doi: [10.2196/16414](https://doi.org/10.2196/16414)] [Medline: [32442154](https://pubmed.ncbi.nlm.nih.gov/32442154/)]
14. Mascheroni A, Choe EK, Luo Y, Marazza M, Ferlito C, Caverzasio S, et al. The SleepFit tablet application for home-based clinical data collection in Parkinson disease: user-centric development and usability study. *JMIR mHealth uHealth* 2021 Jun 08;9(6):e16304 [FREE Full text] [doi: [10.2196/16304](https://doi.org/10.2196/16304)] [Medline: [34100767](https://pubmed.ncbi.nlm.nih.gov/34100767/)]
15. Keogh A, Argent R, Anderson A, Caulfield B, Johnston W. Assessing the usability of wearable devices to measure gait and physical activity in chronic conditions: a systematic review. *J Neuroeng Rehabil* 2021 Sep 15;18(1):138 [FREE Full text] [doi: [10.1186/s12984-021-00931-2](https://doi.org/10.1186/s12984-021-00931-2)] [Medline: [34526053](https://pubmed.ncbi.nlm.nih.gov/34526053/)]
16. Silva AG, Caravau H, Martins A, Almeida AMP, Silva T, Ribeiro ?, et al. Procedures of user-centered usability assessment for digital solutions: scoping review of reviews reporting on digital solutions relevant for older adults. *JMIR Hum Factors* 2021 Jan 13;8(1):e22774 [FREE Full text] [doi: [10.2196/22774](https://doi.org/10.2196/22774)] [Medline: [33439128](https://pubmed.ncbi.nlm.nih.gov/33439128/)]
17. Postuma RB, Berg D, Stern M, Poewe W, Olanow CW, Oertel W, et al. MDS clinical diagnostic criteria for Parkinson's disease. *Mov Disord* 2015 Oct;30(12):1591-1601. [doi: [10.1002/mds.26424](https://doi.org/10.1002/mds.26424)] [Medline: [26474316](https://pubmed.ncbi.nlm.nih.gov/26474316/)]
18. Bendig J, Wolf A, Mark T, Frank A, Mathiebe J, Scheibe M, et al. Feasibility of a multimodal telemedical intervention for patients with Parkinson's disease-a pilot study. *J Clin Med* 2022 Feb 18;11(4):1074 [FREE Full text] [doi: [10.3390/jcm11041074](https://doi.org/10.3390/jcm11041074)] [Medline: [35207351](https://pubmed.ncbi.nlm.nih.gov/35207351/)]
19. Kostikis N, Rigas G, Konitsiotis S, Fotiadis D. Motor fluctuations and dyskinesia. *Mov Disord Clin Pract* 2020 Mar 17;7(S2):S6-S44. [doi: [10.1002/mdc3.12923](https://doi.org/10.1002/mdc3.12923)]
20. Hauser RA, Deckers F, Lehert P. Parkinson's disease home diary: further validation and implications for clinical trials. *Mov Disord* 2004 Dec;19(12):1409-1413. [doi: [10.1002/mds.20248](https://doi.org/10.1002/mds.20248)] [Medline: [15390057](https://pubmed.ncbi.nlm.nih.gov/15390057/)]
21. Nilsson MH, Hariz G, Wictorin K, Miller M, Forsgren L, Hagell P. Development and testing of a self administered version of the Freezing of Gait Questionnaire. *BMC Neurol* 2010 Sep 23;10:85 [FREE Full text] [doi: [10.1186/1471-2377-10-85](https://doi.org/10.1186/1471-2377-10-85)] [Medline: [20863392](https://pubmed.ncbi.nlm.nih.gov/20863392/)]
22. Hoehn MM, Yahr MD. Parkinsonism: onset, progression and mortality. *Neurology* 1967 May;17(5):427-442. [doi: [10.1212/wnl.17.5.427](https://doi.org/10.1212/wnl.17.5.427)] [Medline: [6067254](https://pubmed.ncbi.nlm.nih.gov/6067254/)]
23. Movement Disorder Society Task Force on Rating Scales for Parkinson's Disease. The Unified Parkinson's Disease Rating Scale (UPDRS): status and recommendations. *Mov Disord* 2003 Jul;18(7):738-750. [doi: [10.1002/mds.10473](https://doi.org/10.1002/mds.10473)] [Medline: [12815652](https://pubmed.ncbi.nlm.nih.gov/12815652/)]
24. Beck AT, Steer RA, Ball R, Ranieri W. Comparison of Beck Depression Inventories -IA and -II in psychiatric outpatients. *J Pers Assess* 1996 Dec;67(3):588-597. [doi: [10.1207/s15327752jpa6703\\_13](https://doi.org/10.1207/s15327752jpa6703_13)] [Medline: [8991972](https://pubmed.ncbi.nlm.nih.gov/8991972/)]
25. Nasreddine ZS, Phillips NA, Bédirian V, Charbonneau S, Whitehead V, Collin I, et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *J Am Geriatr Soc* 2005 Apr;53(4):695-699. [doi: [10.1111/j.1532-5415.2005.53221.x](https://doi.org/10.1111/j.1532-5415.2005.53221.x)] [Medline: [15817019](https://pubmed.ncbi.nlm.nih.gov/15817019/)]
26. Norman CD, Skinner HA. eHEALS: the eHealth Literacy Scale. *J Med Internet Res* 2006 Nov 14;8(4):e27 [FREE Full text] [doi: [10.2196/jmir.8.4.e27](https://doi.org/10.2196/jmir.8.4.e27)] [Medline: [17213046](https://pubmed.ncbi.nlm.nih.gov/17213046/)]
27. Podsiadlo D, Richardson S. The timed "Up & Go": a test of basic functional mobility for frail elderly persons. *J Am Geriatr Soc* 1991 Feb;39(2):142-148. [doi: [10.1111/j.1532-5415.1991.tb01616.x](https://doi.org/10.1111/j.1532-5415.1991.tb01616.x)] [Medline: [1991946](https://pubmed.ncbi.nlm.nih.gov/1991946/)]
28. Ziegler K, Schroeteler F, Ceballos-Baumann AO, Fietzek UM. A new rating instrument to assess festination and freezing gait in Parkinsonian patients. *Mov Disord* 2010 Jun 15;25(8):1012-1018. [doi: [10.1002/mds.22993](https://doi.org/10.1002/mds.22993)] [Medline: [20310009](https://pubmed.ncbi.nlm.nih.gov/20310009/)]
29. Franchignoni F, Horak F, Godi M, Nardone A, Giordano A. Using psychometric techniques to improve the Balance Evaluation Systems Test: the mini-BESTest. *J Rehabil Med* 2010 Apr;42(4):323-331 [FREE Full text] [doi: [10.2340/16501977-0537](https://doi.org/10.2340/16501977-0537)] [Medline: [20461334](https://pubmed.ncbi.nlm.nih.gov/20461334/)]
30. Brooke J. SUS: a 'quick and dirty' usability scale. In: Jordan PW, Thomas B, McClelland IL, Weerdmeester B, editors. *Usability Evaluation In Industry*. London, UK: Taylor Francis, CRC Press; 1996:6.
31. Lewis JR. The System Usability Scale: past, present, and future. *Int J Hum Comput Interact* 2018 Mar 30;34(7):577-590. [doi: [10.1080/10447318.2018.1455307](https://doi.org/10.1080/10447318.2018.1455307)]
32. Tullis TS, Stetson JN. A comparison of questionnaires for assessing website usability. 2004 Presented at: UPA 2004: 13th Annual Usability Professionals' Association Conference; June 7-11, 2004; Minneapolis, MN p. 1-12 URL: <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.396.3677&rep=rep1&type=pdf>
33. Laugwitz B, Held T, Schrepp M. Construction and evaluation of a user experience questionnaire. 2008 Presented at: USAB 2008: HCI and Usability for Education and Work; November 20-21, 2008; Graz, Austria p. 63-76. [doi: [10.1007/978-3-540-89350-9\\_6](https://doi.org/10.1007/978-3-540-89350-9_6)]
34. Jacomy M, Venturini T, Heymann S, Bastian M. ForceAtlas2, a continuous graph layout algorithm for handy network visualization designed for the Gephi software. *PLoS One* 2014 Jun 10;9(6):e98679 [FREE Full text] [doi: [10.1371/journal.pone.0098679](https://doi.org/10.1371/journal.pone.0098679)] [Medline: [24914678](https://pubmed.ncbi.nlm.nih.gov/24914678/)]

35. Guest G, Bunce A, Johnson L. How many interviews are enough? *Field Methods* 2016 Jul 21;18(1):59-82. [doi: [10.1177/1525822X05279903](https://doi.org/10.1177/1525822X05279903)]
36. Broekhuis M, van Velsen L, Hermens H. Assessing usability of eHealth technology: a comparison of usability benchmarking instruments. *Int J Med Inform* 2019 Aug;128:24-31. [doi: [10.1016/j.ijmedinf.2019.05.001](https://doi.org/10.1016/j.ijmedinf.2019.05.001)] [Medline: [31160008](https://pubmed.ncbi.nlm.nih.gov/31160008/)]
37. Gibson A, Ryan A, Bunting B, McCauley C, Laird L, Ferry F, et al. Assessing usability testing for people living with dementia. 2016 Oct 13 Presented at: REHAB '16: 4th Workshop on ICTs for improving Patients Rehabilitation Research Techniques; October 13-14, 2016; Lisbon, Portugal p. 25-31. [doi: [10.1145/3051488.3051492](https://doi.org/10.1145/3051488.3051492)]
38. Holden RJ. A Simplified System Usability Scale (SUS) for cognitively impaired and older adults. *Proc Int Symp Hum Factors Ergon Heal Care* 2020 Sep 16;9(1):180-182. [doi: [10.1177/2327857920091021](https://doi.org/10.1177/2327857920091021)]
39. Barnard Y, Bradley MD, Hodgson F, Lloyd AD. Learning to use new technologies by older adults: perceived difficulties, experimentation behaviour and usability. *Comput Human Behav* 2013 Jul;29(4):1715-1724. [doi: [10.1016/j.chb.2013.02.006](https://doi.org/10.1016/j.chb.2013.02.006)]
40. Kreps GL. The relevance of health literacy to mHealth. *Stud Health Technol Inform* 2017;240:347-355. [Medline: [28972527](https://pubmed.ncbi.nlm.nih.gov/28972527/)]
41. El Benny M, Kabakian-Khasholian T, El-Jardali F, Bardus M. Application of the eHealth literacy model in digital health interventions: scoping review. *J Med Internet Res* 2021 Jun 03;23(6):e23473 [FREE Full text] [doi: [10.2196/23473](https://doi.org/10.2196/23473)] [Medline: [34081023](https://pubmed.ncbi.nlm.nih.gov/34081023/)]
42. Wildenbos GA, Peute L, Jaspers M. Aging barriers influencing mobile health usability for older adults: a literature based framework (MOLD-US). *Int J Med Inform* 2018 Jun;114:66-75. [doi: [10.1016/j.ijmedinf.2018.03.012](https://doi.org/10.1016/j.ijmedinf.2018.03.012)] [Medline: [29673606](https://pubmed.ncbi.nlm.nih.gov/29673606/)]
43. GBD 2019 Dementia Forecasting Collaborators. Estimation of the global prevalence of dementia in 2019 and forecasted prevalence in 2050: an analysis for the Global Burden of Disease Study 2019. *Lancet Public Health* 2022 Feb 01;7(2):e105-e125 [FREE Full text] [doi: [10.1016/S2468-2667\(21\)00249-8](https://doi.org/10.1016/S2468-2667(21)00249-8)] [Medline: [34998485](https://pubmed.ncbi.nlm.nih.gov/34998485/)]
44. Bach J, Ziegler U, Deuschl G, Dodel R, Doblhammer-Reiter G. Projected numbers of people with movement disorders in the years 2030 and 2050. *Mov Disord* 2011 Oct;26(12):2286-2290. [doi: [10.1002/mds.23878](https://doi.org/10.1002/mds.23878)] [Medline: [22021158](https://pubmed.ncbi.nlm.nih.gov/22021158/)]
45. Del Din S, Kirk C, Yarnall AJ, Rochester L, Hausdorff JM. Body-worn sensors for remote monitoring of Parkinson's disease motor symptoms: vision, state of the art, and challenges ahead. *J Parkinsons Dis* 2021 Jul 16;11(s1):S35-S47 [FREE Full text] [doi: [10.3233/JPD-202471](https://doi.org/10.3233/JPD-202471)] [Medline: [33523020](https://pubmed.ncbi.nlm.nih.gov/33523020/)]
46. Linares-del Rey M, Vela-Desojo L, Cano-de la Cuerda R. Mobile phone applications in Parkinson's disease: a systematic review. *Neurologia (Engl Ed)* 2019 Jan;34(1):38-54 [FREE Full text] [doi: [10.1016/j.nrl.2017.03.006](https://doi.org/10.1016/j.nrl.2017.03.006)] [Medline: [28549757](https://pubmed.ncbi.nlm.nih.gov/28549757/)]

## Abbreviations

**EHEALS:** eHealth Literacy Scale

**FOG-Q:** Freezing of Gait Questionnaire

**MOCA:** Montreal Cognitive Assessment

**PD:** Parkinson Disease

**UPDRS III:** Unified Parkinson's Disease Rating Scale III

**SUS:** System Usability Scale

*Edited by M Focsa; submitted 29.05.22; peer-reviewed by W Maetzler, J Brooke; comments to author 03.07.22; revised version received 28.08.22; accepted 03.09.22; published 25.10.22.*

*Please cite as:*

*Bendig J, Spanz A, Leidig J, Frank A, Stahr M, Reichmann H, Loewenbrück KF, Falkenburger BH*

*Measuring the Usability of eHealth Solutions for Patients With Parkinson Disease: Observational Study*

*JMIR Form Res* 2022;6(10):e39954

URL: <https://formative.jmir.org/2022/10/e39954>

doi: [10.2196/39954](https://doi.org/10.2196/39954)

PMID: [36282558](https://pubmed.ncbi.nlm.nih.gov/36282558/)

©Jonas Bendig, Anja Spanz, Jana Leidig, Anika Frank, Marcus Stahr, Heinz Reichmann, Kai F Loewenbrück, Björn H Falkenburger. Originally published in *JMIR Formative Research* (<https://formative.jmir.org>), 25.10.2022. 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 Food Insecurity Among Individuals Using a Telehealth and Nutrition Platform: Longitudinal Study

Shivani Bakre<sup>1\*</sup>, MHS; Benjamin Shea<sup>1\*</sup>, SM; Kaylee Ortega<sup>1</sup>, MPH; Jared Scharen<sup>1</sup>, MBA; Jason Langheier<sup>1</sup>, MPH, MD; Emily Hu<sup>1</sup>, MHS, PhD

Foodsmart, San Francisco, CA, United States

\* these authors contributed equally

**Corresponding Author:**

Benjamin Shea, SM

Foodsmart

595 Pacific Ave. Fl. 4

San Francisco, CA, 94133

United States

Phone: 1 650 823 9022

Email: [ben.shea@foodsmart.com](mailto:ben.shea@foodsmart.com)

## Abstract

**Background:** Food insecurity is a complex public health problem affecting many individuals in the United States. Digital health interventions that promote behavior change and provide access to affordable and healthy food may help to alleviate food insecurity.

**Objective:** The aim of this study was to characterize food-insecure users of Foodsmart, a telehealth and nutrition platform with meal planning, food ordering, nutrition education, budgeting, and grocery discount features, and to evaluate changes in diet and food insecurity.

**Methods:** We retrospectively analyzed data collected from 4595 adults who used the Foodsmart platform between February and October 2021. Participants self-reported their diet, demographics, biometrics, and food insecurity status in a 56-item questionnaire. Participants were reported to be food insecure if they answered “sometimes” or “often” to the question “How often does the food you buy not last and you don't have money to get more?” from the United States Department of Agriculture’s Household Food Security survey. We examined baseline characteristics of participants by food insecurity status, associations between characteristics and baseline food insecurity, and changes in diet quality and food insecurity status. To evaluate potential causes of reversing food insecurity, the use of 6 Foodsmart features was compared between food-insecure participants who achieved food security versus food-insecure participants who remained food insecure, based on their last response to the food insecurity question.

**Results:** We found that 16% (742/4595) of participants were food insecure at baseline. Participants who were food insecure at baseline were more likely to be obese, to have at least one chronic condition, to have a lower diet quality, to cook less frequently at home, to think healthy food is too expensive, and less likely to order takeout or eat at a restaurant. Among participants who were food insecure at baseline, 61% (451/742) improved their nutrition and 29% (217/742) responded that they were food secure at follow-up, with an increasing percentage achieving food security with longer enrollment time. Using a multivariable logistic regression model, we found that age, diabetes, prediabetes, BMI categories, and diet quality at baseline were statistically significantly associated with the likelihood of being food insecure at baseline. Among those who were food insecure at baseline, there was a higher relative proportion of participants who achieved food security and used the “deals” (28.6% higher), “CookItNow” (36.4% higher), and “telenutrition” (27.5% higher) features compared to those who remained food insecure.

**Conclusions:** This study assesses the characteristics of individuals enrolled on the Foodsmart platform who answered the food insecurity question. We found that a significant number of participants who were food insecure at enrollment achieved food security. This finding shows that telehealth and nutrition platforms may potentially help users improve household food security.

(JMIR Form Res 2022;6(10):e41418) doi:[10.2196/41418](https://doi.org/10.2196/41418)

**KEYWORDS**

food insecurity; tele-nutrition; telehealth; meal planning; SNAP; diet; nutrition; COVID-19

## Introduction

Food insecurity affects many households in the United States, profoundly impacting the health and financial stability of individuals. Recent trends in the last decade have suggested a lower prevalence of food insecurity, dropping from 13% of Americans reported to be food insecure in 2016 to 10.5% in 2019 [1,2]. However, the COVID-19 pandemic has dramatically exacerbated the problem, increasing household food insecurity to 38% in March 2020 [3]. The increase in food insecurity can be attributed to a variety of factors, such as poverty; unemployment; instability and disruptions of the food supply to grocery stores and charitable feeding systems, such as food banks; and lack of eligibility, access, or enrollment in federal programs like the Supplemental Nutrition Assistance Program (SNAP) [4,5]. In a cross-sectional observational study conducted at the end of April 2020, 15.6% of households that were food secure prior to the pandemic experienced low food security during the pandemic, but only 2.3% of households that were food insecure before the pandemic became food secure [6]. A national survey conducted at the end of June 2020 also found that 59% of households in which one member lost a job or income were food insecure [7]. Low food security could also be associated with worsening diet quality and purchasing of unhealthy foods. Adams et al [6] found in a cross-sectional study that about one-third of households that experienced low food security during the pandemic reported increasing the purchasing of high-calorie snack foods, desserts, and sweets. Furthermore, in a cross-sectional analysis of National Health and Nutrition Examination Survey results from 2011 to 2014, food-insecure adults reported a 2.22-unit lower Healthy Eating Index (HEI)-2015 score compared to food-secure adults [8].

The impact of food insecurity on health care expenditures is significant. A model evaluating the impact on health care costs that used 2011 to 2013 National Health Interview Survey/Medical Expenditure Panel Survey data estimated that food-insecure adults spent US \$1834 more on health care annually compared to food-secure adults; this equates to an estimated US \$51.8 billion in excess health care expenditures due to food insecurity in 2016 [1].

A major contributor to exorbitant health care expenditures is the high prevalence of chronic conditions among individuals with food insecurity. A United States Department of Agriculture (USDA) report found that among adults with very low food security, there were 16.4%, 7.2%, 3.6%, and 3.7% higher proportions of hypertension, diabetes, coronary heart disease, and kidney disease, respectively, compared to adults with high food security [9]. Obesity and higher BMI have also been found to be associated with food insecurity [10]. Additionally, a study conducted on 711 patients with diabetes by the American Diabetes Association showed that food-insecure participants were more likely to have poor glycemic control and were more likely to report difficulties affording a diabetes-friendly diet than those who were food secure [11].

There are significant opportunities for interventions to alleviate food insecurity through increasing food access. Prior studies have evaluated the impact of food delivery and online grocery shopping on food insecurity and diet. The Community Servings: Food as Medicine for Diabetes clinical trial, where participants were delivered medically tailored meals in a randomized crossover study, reported that 42% of participants were food insecure during the on-meal-delivery periods versus 62% during the off-meal-delivery periods [12]. The Baltimore Virtual Supermarket Program, a large-scale online grocery ordering system in which participants can pick up their groceries from a local hub, showed that among 93 survey respondents, 93% believed the program made it easier for them to eat healthily and 61% attributed access to healthy foods to the program [13].

Digital technology can serve as an intervention to help alleviate food insecurity. Foodsmart, a telehealth provider platform with a large network of registered dietitians (RDs) across the country that includes a digital nutrition platform, is a potential solution to help address food insecurity, diet quality, and health outcomes. The platform provides personalized recipe recommendations and meal planning and helps users purchase ingredients and compare prices between participating grocers. The platform also allows users to use SNAP benefits to purchase foods online. Previous research on Foodsmart found improvements in clinical metrics, such as weight, lipid levels, and hemoglobin A<sub>1c</sub>, among users with obesity, dyslipidemia, and diabetes, respectively, suggesting clinical benefits of the platform [14-17].

The objective of this study was to characterize demographics, meal planning characteristics, and diet quality among participants who were food insecure compared to participants who were food secure. The study also sought to evaluate longitudinal associations between using Foodsmart and changes in diet quality and food insecurity status among participants during the COVID-19 pandemic.

## Methods

### Study Sample

As of October 2021, 76,506 users of Foodsmart across the United States had answered a question on their food insecurity status. Participants were connected and enrolled into the Foodsmart platform as a service, either through their employer or insurance. We included in our analysis participants who had answered this question at least 2 times since its implementation in February 2021, with their last response being at least 30 days after their first response (n=5798) and for whom complete information was available on demographics, diet, weight, chronic conditions, and meal planning habits (n=4794). Participants who had extreme values for BMI (<15 kg/m<sup>2</sup> or >50 kg/m<sup>2</sup>) were excluded from our sample (n=199). Our final study sample was 4595 participants (Figure 1).



**Figure 1.** Participant selection criteria.

## Foodsmart, a Telehealth and Nutrition Platform

Foodsmart is a platform that includes two components: (1) a digital app and web solution with tools including nutrition assessments, meal planning, and grocery ordering and (2) a telehealth component that lets users virtually meet with an RD. These components are incorporated into 6 features of the Foodsmart platform: “CookItNow,” “recipes,” “meal planner,” “food buying,” “deals,” and “telenutrition.” The CookItNow feature prompts the user to input ingredients they currently have in their kitchen; the Foodsmart platform then provides them tailored recipes based on those ingredients. The recipes feature on the Foodsmart app also allows individuals to filter recipes based on their food preferences or by category, such as budget-friendly recipes, recipes with leafy greens, or recipes fit for a SNAP budget. The meal-plan feature allows users to create a personalized meal plan based on the user’s dietary assessment, built from thousands of recipes. The food-buying feature then allows the user to have their meal plan automatically transferred to a grocer for online grocery pickup or delivery. The Food and Nutrition Service SNAP Online Purchasing Pilot [18] has also enabled individuals to pay for groceries with SNAP benefits, encouraging individuals to spend SNAP dollars on nutritious food. Foodsmart also offers medically tailored meals on its platform that users can order or that health plans can subsidize. The deals feature helps individuals save approximately 34% on each grocery order via (1) providing digital coupons from the grocery store they shop at and (2) finding the store with the cheapest price for their groceries. Finally, the telenutrition feature allows participants to meet with RDs who can help them by providing recommendations and support based on their eating and health goals and by providing technology assistance. Foodsmart also includes a food insecurity screening question to help identify users who are food insecure; this helps organizations determine which users need the most help and measure these users’ food insecurity status over time.

## Assessment of Diet and Other Characteristics

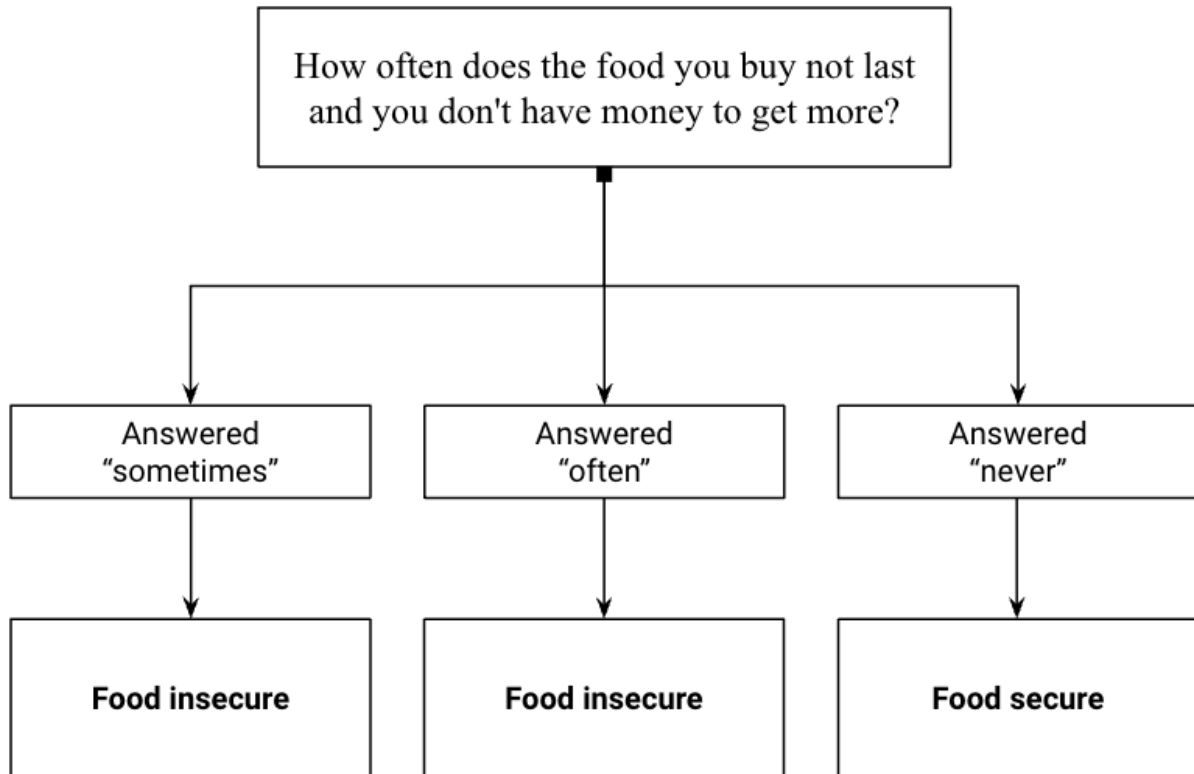
Participants report their age, gender, usual dietary intake, meal planning habits, weight, chronic conditions, and physical activity on Foodsmart’s platform by filling out a 56-item modified food frequency questionnaire called Nutriquiz (adapted from the National Cancer Institute Diet History Questionnaire I) upon registration [19]. In addition to dietary questions, the Nutriquiz

also ascertains demographic information, height, weight, chronic conditions (including diabetes, prediabetes, high blood pressure, dyslipidemia, or the absence of these conditions), frequency of physical activity (light, moderate, or vigorous), and meal planning habits (frequency of home-cooked meals, frequency of purchasing groceries from a grocery store, and how groceries are bought). We collapsed and renamed the categories of responses for each question to “rarely,” “weekly,” and “daily.” The “rarely” category contains the survey responses “never,” “1 time/month,” and “2-3 times/month.” The weekly category contains the survey responses “1 time/week,” “2 times/week,” “3-4 times/week,” and “5-6 times/week.” The daily category contains survey responses for “1 time/day,” “2-3 times/day,” “4-5 times/day,” and “6+ times/day.”

Based on a participant’s responses to the Nutriquiz questions assessing diet, a score called the Nutriscore was calculated to assess overall diet quality; this score is based on the Alternative HEI-2010 and the Commonwealth Scientific and Industrial Research Organization Healthy Diet Score [20,21]. The Nutriscore has 7 components (called Nutriscore Essentials): fruits, vegetables, protein ratio, fat ratio, carbohydrate ratio, hydration, and sodium. Each component is scored on a scale of 0 to 10, and the scores are then added together for an overall Nutriscore (ranging from 0 to 70). Higher scores indicate higher diet quality. Change in Nutriscore was calculated by taking the difference between a participant’s last Nutriscore and first Nutriscore, given that they were taken at least 30 days apart. An increase in Nutriscore demonstrates that the participant has improved their diet quality.

## Ascertainment of Food Insecurity Status

In the Nutriquiz, participants answered a question about their level of food insecurity. This question is adapted from one of 2 questions in a shortened food security screener validated by Hager et al [22] that is valid when compared to the USDA 18-item Household Food Security Survey (HFSS) [23]. The question asked was “How often does the food you buy not last and you don’t have money to get more?” The answer choices were “sometimes,” “often,” and “never.” If participants answered “sometimes” or “often” they would be considered food insecure; if they answered “never” to the question, they would be considered food secure, as shown in [Figure 2](#).

**Figure 2.** Decision tree for assessing food insecurity status.

### Statistical Analysis

We used descriptive statistics to evaluate participant demographic and clinical characteristics, diet quality, physical activity, and meal planning habits. Categorical variables are presented as percentages of the study population and continuous variables are presented as the mean (SD). We used the chi-square test to determine if there was a statistically significant difference in categorical variables with food insecurity status. We used a 2-sample, 2-tailed *t* test to determine if there was a statistically significant difference in continuous variables with food insecurity status.

We used a multivariable logistic regression model to estimate the odds ratios (ORs) and 95% CIs of baseline food insecurity status for several variables, including gender, age, diabetes, prediabetes, high blood pressure, dyslipidemia, good health (ie, no conditions), baseline BMI category, and baseline Nutriscore. For those who were categorized as food insecure, we calculated the percentage of participants whose status changed to food secure in total and by cumulative length of enrollment in Foodsmart (with cutoffs at  $\geq 2$ ,  $\geq 4$ , and  $\geq 6$  months). To evaluate possible causes for reduction in food insecurity among food-insecure participants at baseline, we compared the difference in the percentage of participants who used 6 Foodsmart features (CookItNow, deals, telenutrition, meal planner, recipes, and food buying) among those who achieved

food security versus those who remained food insecure. The absolute difference was calculated by taking the proportion of participants who were food insecure at baseline and became food secure and subtracting the proportion of participants who were food insecure at baseline and remained food insecure. The relative difference was calculated by dividing the absolute difference by the proportion of participants who were food insecure at baseline and remained food insecure.

*P* values of .05 or less were considered to be statistically significant. Stata (version 16; StataCorp) was used for all statistical analyses.

### Ethical Considerations

The study was declared exempt from institutional review board oversight by the Pearl Institutional Review Board given the retrospective design of the study and the less than minimal risk to participants (Protocol #20-ZIPO-101).

## Results

### Participant Characteristics

In order to better understand baseline demographic and clinical characteristics, as well as dietary habits and behaviors, we conducted descriptive analyses stratified by baseline food insecurity (Table 1).

**Table 1.** Baseline characteristics by baseline food insecurity status. The chi-square test was used to test differences in categorical variables.

Characteristics	Total (N=4595), n (%)	Food insecure (N=742), n (%)	Food secure (N=3853), n (%)	P value
Female	3558 (77.4)	609 (82.1)	2949 (76.5)	.001
<b>Age</b>				<.001
<40 years old	1499 (32.6)	285 (38.4)	1214 (31.5)	
40-59 years old	2417 (52.6)	380 (51.2)	2037 (52.9)	
≥60 years old	679 (14.8)	77 (10.4)	602 (15.6)	
<b>BMI category</b>				<.001
Normal (<25 kg/m <sup>2</sup> )	1459 (31.8)	149 (20.1)	1310 (34)	
Overweight (25-29.9 kg/m <sup>2</sup> )	1393 (30.3)	213 (28.7)	1180 (30.6)	
Obese (≥30 kg/m <sup>2</sup> )	1743 (37.9)	380 (51.2)	1363 (35.4)	
<b>Chronic conditions</b>				
Diabetes	250 (5.4)	61 (8.2)	189 (4.9)	<.001
Prediabetes	239 (5.2)	63 (8.5)	176 (4.6)	<.001
High blood pressure	765 (16.6)	159 (21.4)	606 (15.7)	<.001
Dyslipidemia	888 (19.3)	156 (21)	732 (19)	.2
Healthy (defined as having no conditions)	3048 (66.3)	446 (60.1)	2602 (67.5)	<.001
<b>Physical activity</b>				
<b>Light</b>				<.001
Almost never	602 (13.1)	130 (17.5)	472 (12.3)	
Weekly	2379 (51.8)	373 (50.3)	2006 (52.1)	
Daily	1614 (35.1)	239 (32.2)	1375 (35.7)	
<b>Moderate</b>				<.001
Almost never	877 (19.1)	192 (25.9)	685 (17.8)	
Weekly	2820 (61.4)	434 (58.5)	2386 (61.9)	
Daily	898 (19.5)	116 (15.6)	782 (20.3)	
<b>Vigorous</b>				.01
Almost never	1537 (33.4)	285 (38.4)	1252 (32.5)	
Weekly	2621 (57)	395 (53.2)	2226 (57.8)	
Daily	437 (9.5)	62 (8.4)	375 (9.7)	
<b>Alcohol use</b>				<.001
Almost never	2619 (57)	483 (65.1)	2136 (55.4)	
Weekly	1663 (36.2)	206 (27.8)	1457 (37.8)	
Daily	313 (6.8)	53 (7.1)	260 (6.7)	
<b>Responses to questions on meal planning habits</b>				
<b>In the last 6 months, how often were meals home-cooked by you or a household member?</b>				<.001
Almost never	98 (2.1)	28 (3.8)	70 (1.8)	
Weekly	2216 (48.2)	405 (54.6)	1811 (47)	
Daily	2281 (49.6)	309 (41.6)	1972 (51.2)	
<b>In the last 6 months, how often did you or your household purchase groceries from a grocery store?</b>				.001
Almost never	635 (13.8)	134 (18.1)	501 (13)	
Weekly	3960 (86.2)	608 (81.9)	3352 (87)	
<b>How do you typically buy groceries?</b>				.1

Characteristics	Total (N=4595), n (%)	Food insecure (N=742), n (%)	Food secure (N=3853), n (%)	P value
At a store	3655 (79.5)	568 (76.5)	3087 (80.1)	
Online	690 (15)	123 (16.6)	567 (14.7)	
Both	250 (5.4)	51 (6.9)	199 (5.2)	
With an Electronic Benefits Transfer card	161 (3.5)	65 (8.7)	96 (2.5)	<.001
<b>How often do you feel that healthy food is too expensive to buy?</b>				<.001
Never	1883 (41)	48 (6.5)	1835 (47.6)	
Sometimes	1978 (43)	321 (43.3)	1657 (43)	
Often	734 (16)	373 (50.3)	361 (9.4)	
<b>How often do you order take out/visit a restaurant/fast food establishment?</b>				.001
Less than once a week	1917 (41.7)	352 (47.4)	1565 (40.6)	
At least once a week	2678 (58.3)	390 (52.6)	2288 (59.4)	

We found that 16.2% (742/4595) of participants were categorized as food insecure at baseline. In our sample of 4595 participants, participants who were food insecure at baseline were more likely to be female, be aged 40 to 59 years, be obese, have diabetes, have prediabetes, and have high blood pressure compared to those who were food secure. Additionally, food-insecure participants were less likely to exercise or to drink alcohol. In regard to meal planning habits, participants who were food insecure were less likely to make home-cooked meals on a daily basis, less likely to purchase groceries on a weekly basis, more likely to think that healthy food is too expensive, and less likely to order takeout food or eat at a restaurant or fast-food establishment compared to those who were food secure at baseline. Lastly, a higher percentage of food-insecure participants reported having an Electronic Benefits Transfer card compared to food-secure participants at baseline.

### Changes in Diet Quality by Food Security Status

To better understand how baseline diet quality and change in diet quality were associated with baseline food insecurity status, we show the first and last Nutriscores in [Table 2](#). First and last Nutriscores were significantly higher for those who were food secure compared to food insecure ( $P<.001$ ). However, there was no significant difference in the change in Nutriscore between those who were food insecure and food secure (1.9 points vs 1.4 points).

We show a comparison of the average change in Nutriscore and the subcomponents of the Nutriscore in [Table 3](#). We found that on average, participants who were food insecure at baseline had a greater improvement in overall Nutriscore, vegetable intake, fruit intake, carbohydrate ratio, fat ratio, and protein ratio.

**Table 2.** Baseline and changes in diet quality by food insecurity status. Numbers in parentheses in row headings indicate score ranges; higher numbers indicate higher quality.

Nutriscore	Total (N=4595)	Food insecure (N=742)	Food secure (N=3853)	P value
Baseline Nutriscore (0-70), mean (SD)	33.9 (8.6)	31.9 (8.3)	34.3 (8.6)	<.001
<b>Baseline Nutriscore component scores (0-10), mean (SD)</b>				
Vegetables	3.4 (1.4)	3.3 (1.5)	3.5 (1.4)	.002
Fruits	3.1 (2.2)	2.9 (2.2)	3.2 (2.2)	.003
Carbohydrate ratio <sup>a</sup>	7.7 (1.7)	7.3 (1.7)	7.7 (1.7)	<.001
Fat ratio <sup>b</sup>	3.0 (2.8)	2.6 (2.6)	3.1 (2.9)	<.001
Protein ratio <sup>c</sup>	7.0 (3.2)	6.8 (3.2)	7.1 (3.2)	.03
Sodium	7.1 (4.0)	6.6 (4.3)	7.2 (3.9)	<.001
Hydration	6.5 (2.3)	6.3 (2.4)	6.6 (2.2)	.01
Final Nutriscore (0-70), mean (SD)	35.4 (8.5)	33.8 (8.6)	35.7 (8.5)	<.001
Change in Nutriscore, mean (SD)	1.5 (6.9)	1.9 (7.2)	1.4 (6.9)	.1
Participants who improved nutrition at all, n (%)	2690 (58.5)	451 (60.8)	2239 (58.1)	.2
Participants who improved nutrition by 5% or more, n (%)	2173 (47.3)	374 (50.4)	1799 (46.7)	.1

<sup>a</sup>Carbohydrate ratio: total carbohydrate intake divided by total fiber intake.

<sup>b</sup>Fat ratio: polyunsaturated fatty acid intake divided by the sum of intake of saturated and trans fats.

<sup>c</sup>Protein ratio: white meat and plant protein intake divided by red meat and processed meat intake.

**Table 3.** Mean group percentage change in Nutriscore and Nutriscore component scores by food insecurity status at baseline.

Scores	Change in score in food-secure group, %	Change in score in food-insecure group, %
Nutriscore	4	6
<b>Nutriscore component scores</b>		
Vegetables	3.2	3.7
Fruits	5.1	7.9
Carbohydrate ratio <sup>a</sup>	1.4	1.6
Fat ratio <sup>b</sup>	7.2	11.8
Protein ratio <sup>c</sup>	1	2.2
Sodium	0.6	-1.7
Hydration	2.7	0.4

<sup>a</sup>Carbohydrate ratio: total carbohydrate intake divided by total fiber intake.

<sup>b</sup>Fat ratio: polyunsaturated fatty acid intake divided by the sum of intake of saturated and trans fats.

<sup>c</sup>Protein ratio: white meat and plant protein intake divided by red meat and processed meat intake.

### Changes in Food Insecurity Status

To evaluate how food insecurity status changed over time, we determined the proportion of participants who were categorized as food insecure at baseline and were also food secure at the time of their last response among the total population and by cumulative length of enrollment ( $\geq 2$ ,  $\geq 4$ , and  $\geq 6$  months). Of the 742 food-insecure participants at baseline, 29.2% (217/742) were food secure at the time of their last response. When subsetting participants by cumulative length of enrollment ( $\geq 2$  months,  $\geq 4$  months, and  $\geq 6$  months), 30.9% (182/590), 33% (70/212), and 42.4% (25/59) of participants, respectively,

changed from food insecure to secure. We also evaluated the proportion of participants who were categorized as food secure at baseline and were subsequently food insecure in their last response. Of the 3853 food-secure participants at baseline, 6.7% (259) were food insecure at the time of their last response.

### Characteristics Associated With Food Insecurity Status at Baseline

In order to evaluate which participant characteristics were associated with food insecurity at baseline, we used a multivariable logistic regression model adjusted for demographics, chronic conditions, and baseline diet quality

(Table 4). We found that females were 44% more likely to be food insecure compared to males (OR 1.44, 95% CI 1.17-1.77;  $P=.001$ ). Compared to those who were younger than 40 years, those aged 40 to 59 years were 34% less likely to be food insecure (OR 0.66, 95% CI 0.55-0.79;  $P<.001$ ), and those aged at least 60 years were 56% less likely to be food insecure (OR 0.44, 95% CI 0.33-0.59;  $P<.001$ ). Participants with diabetes were 65% more likely to be food insecure compared to those without diabetes (OR 1.65, 95% CI 1.17-2.32;  $P=.004$ ).

Participants with prediabetes were 61% more likely to be food insecure compared to those without prediabetes (OR 1.61, 95% CI 1.14-2.27;  $P=.01$ ). Additionally, those who were overweight or obese were, respectively, 60% (OR 1.60, 95% CI 1.28-2.02;  $P<.001$ ) and 118% (OR 2.18, 95% CI 1.75-2.71;  $P<.001$ ) more likely to be food insecure than those who had a normal BMI. Finally, for each 5-point increase in a participant's baseline Nutriscore, they were 10% less likely to be food insecure (OR 0.90, 95% CI 0.86-0.95;  $P<.001$ ).

**Table 4.** Multivariable logistic regression assessing factors associated with baseline food insecurity status.

Characteristics	Odds ratios (95% CI)	P value
Gender (female)	1.44 (1.17-1.77)	.001
<b>Age</b>		
<40 years	1 (reference)	
40-59 years	0.66 (0.55-0.79)	<.001
≥60 years	0.44 (0.33-0.59)	<.001
<b>Conditions</b>		
Diabetes	1.65 (1.17-2.32)	.004
Prediabetes	1.61 (1.14-2.27)	.01
High blood pressure	1.25 (0.94-1.67)	.1
Dyslipidemia	0.99 (0.75-1.32)	.9
Healthy (no conditions)	1.01 (0.72-1.41)	.9
<b>BMI category</b>		
Normal	1 (reference)	
Overweight	1.60 (1.28-2.02)	<.001
Obese	2.18 (1.75-2.71)	<.001
Baseline Nutriscore (per 5 points)	0.90 (0.86-0.95)	<.001

### Foodsmart Platform Usage Among Participants Who Were Food Insecure at Baseline

To evaluate possible causes of overall reduction of food insecurity among food-insecure participants, we compared the percentage of users who used specific Foodsmart features among food-insecure participants who achieved food security and

food-insecure participants who remained food insecure (Table 5). There was a 36.4% higher relative proportion of participants who used the CookItNow feature, a 28.6% higher relative proportion of participants who used the deals feature, and a 27.5% higher relative proportion of participants who used the telenutrition feature among participants who achieved food security versus participants who remained food insecure.

**Table 5.** Proportion of participants who were food insecure at baseline who used the Foodsmart feature by last food security status.

Foodsmart features	Users who achieved food security who used the feature (N=217), n (%)	Users who remained food insecure who used the feature (N=525), n (%)	Absolute difference, percentage points	Relative difference, %
CookItNow	48 (22.1)	85 (16.2)	5.9	36.4
Deals	43 (19.8)	81 (15.4)	4.4	28.6
Telenutrition	11 (5.1)	21 (4)	1.1	27.5
Meal planner	64 (29.5)	155 (29.5)	0	0
Recipes	192 (88.5)	477 (90.9)	-2.4	-2.6
Food buying	145 (66.8)	372 (70.9)	-4	-5.8

## Discussion

### Principal Findings

In this study of 4595 participants, 16% (742) were identified as being food insecure at baseline. Participants who were food insecure at baseline were more likely to be female, older, have a preexisting condition (overweight or obesity, diabetes, or prediabetes), have a lower quality diet, and believe healthy food is too expensive. Additionally, those who were food insecure were less likely to exercise, consume alcohol, prepare home-cooked meals daily, order takeout or eat at a restaurant at least once a week, or buy groceries compared to those who were food secure at baseline. Among participants who were food insecure at baseline, 29% (217/742) identified as being food secure at the end of follow-up.

Diet quality, assessed using the Nutriscore, improved in both the food-insecure and food-secure groups by 1.9 and 1.4 points on average, respectively; this suggests that the platform can benefit food-insecure and food-secure participants equally in improving nutrition. Furthermore, both groups improved in most of the components that make up the Nutriscore, with greater improvements in the food-insecure group most likely due to lower scores on average at baseline compared to scores in the food-secure group.

In our multivariable logistic regression evaluating the association between various characteristics and food insecurity at baseline, we found that those who were female, were younger than 40 years old, had diabetes or prediabetes, or were overweight or obese were more likely to have food insecurity. For every 5-unit increase in baseline diet quality as assessed by Nutriscore, the likelihood of a participant being food insecure was 10% lower. This aligns with a study conducted by Leung et al [24], which found that food insecurity was associated with lower diet quality. Furthermore, these results are consistent with another study assessing trends in food insecurity in the United States, which found that individuals aged 65 years or older were less likely to be food insecure than individuals who were 18 to 34 years old, and females were 23% more likely to be food insecure than males [25].

In evaluating which Foodsmart features participants used, we found that a higher proportion of participants who were food insecure at baseline and achieved food security used the CookItNow, deals, and telenutrition features than participants who were food insecure at baseline and remained food insecure. The CookItNow feature promotes cooking at home more often with ingredients participants already have at home, showing that there could be a potential association between cooking at home and food security. The use of the deals feature to find the best prices for grocery ingredients can also help participants spend less on their food purchasing, which could also potentially lead to reducing food insecurity. The use of telenutrition visits with an RD to better understand how to cook and meal plan on a budget while eating healthily could also potentially support reducing food insecurity.

These results align with those of prior studies, which demonstrated a higher prevalence of obesity and fewer

home-cooked meals among those who were food insecure [26,27]. Pan et al [26] found that among adults in 12 states in the United States, the prevalence of obesity was significantly higher among food-insecure adults than among food-secure adults (35.1% vs 25.2%;  $P<.001$ ). A hypothesis for explaining this association is that food-insecure individuals are more likely to consume inexpensive, energy-dense foods. Furthermore, a cycle of having an abundance of food at the beginning of the month followed by food scarcity at the end of the month, due to the monthly distribution of SNAP benefits, may contribute to weight gain [28]. In a study of 1171 SNAP-eligible adults in Texas, Ranjit et al [27] found that food-secure participants had more days where they ate a home-cooked meal (difference of  $-0.26$  days,  $P=.03$ ) and cooked more days per week (difference of  $-0.30$  days,  $P=.01$ ) than food-insecure participants. A lower frequency of cooking and eating at home among food-insecure participants compared to food-secure participants may be due to less access to cooking supplies and resources and less time. This highlights the importance of increasing nutrition and cooking knowledge and increasing resources to help promote cooking within a limited budget and time constraints.

In line with our findings, a prior study showed that medically tailored meal delivery for participants with diabetes improved food insecurity and dietary quality [12]. The Food as Medicine for Diabetes clinical trial used a 24-week randomized cross-over design to study 42 adults with type 2 diabetes who reported food insecurity based on a 2-item screener questionnaire [22]. Dietary quality improved (the mean HEI-2010 score was 71.3 during the on-meal-delivery period, compared to 39.9 during the off-meal-delivery period;  $P<.001$ ). Furthermore, participants reported a 20% lower prevalence of food insecurity during on-meal-delivery periods versus off-meal-delivery periods ( $P=.047$ ). While our study was a 1-sample pre-post analysis, our results were in line with this randomized cross-over clinical trial. However, our study did not specifically use medically tailored, specific delivery. With the cost of food per month being on average US \$281 per individual in the United States [29], it is also uncertain whether these medically tailored meal plans are cost-effective for a larger population, given that the cost of the meals and delivery was approximately US \$350 per individual per month in this study [12].

Addressing food and nutrition insecurity requires solving problems for both food access and affordability. However, affordability and access need to be applied not just to food in general, but specifically to nutritious food, as ultraprocessed foods with poor nutritional quality exacerbate the prevalence of chronic conditions [30]. Affordability has been enhanced by a record SNAP increase in 2021 [31], but many families still struggle with affording food outside of the SNAP allowance. Additionally, many of these funds are often allocated toward nonnutritious foods [32]. Many individuals with food insecurity struggle with disabilities or lack of transportation, limiting their ability to get to a physical grocery store [33]. Many major online grocers have begun accepting SNAP benefits online; however, for many of these users, the online grocery delivery fees are still prohibitive. Additionally, others are unaware if they qualify for SNAP and struggle to navigate the SNAP enrollment process. A prior study in California showed that newly enrolled

participants in SNAP were less likely to report food insecurity (83.1% vs 67.5%,  $P < .001$ ), showing that SNAP enrollment can reduce food insecurity [34]. Many health plans have also attempted to deliver meals to solve food insecurity, but unfortunately, this has been difficult to operationalize and scale in a cost-effective manner; for example, the Blue Cross Blue Shield Association terminated its FoodQ pilot after 1 year [35]. In our study, 29% of food-insecure users at baseline reported becoming food secure, and diet quality improved over time; thus, the Foodsmart platform is addressing these food insecurity challenges at scale. Foodsmart RDs have also since begun enrolling members in SNAP, and we plan to evaluate the impact of SNAP enrollment on long-term food insecurity status in a future study.

### Strengths and Limitations

This study had several important strengths. To our knowledge, no longitudinal study has been conducted at this scale to assess the impact of a telehealth platform or nutrition and meal planning app on food insecurity and diet quality. A systematic review of interventions addressing food insecurity in a health care setting found that very few studies that evaluated health outcomes had high enrollment or a significant follow-up period [36]. Our study differs from these studies in that it evaluated postintervention outcomes in dietary intake and food insecurity status and had high enrollment and a long-term follow-up period (mean length 3.3, SD 1.5, months). Given the broad range of follow-up times, we were able to measure changes in food insecurity over differing lengths of time. The question assessing food insecurity was added in February 2021, and almost 5000 users responded to it twice within 8 months, showing the power of digital technology to collect large amounts of data in real time in response to public health crises. To assess food insecurity, we used a question adapted from the USDA's 18-item HFSS that has been previously validated. Last, we were able to draw associations between food insecurity and meal planning

and eating habits, which are important factors in food security and nutrition.

There are some important limitations to note for this study. This study only used 1 of 18 questions on the 18-item HFSS to evaluate food insecurity [23]. Due to the length of time needed to complete the 18-item HFSS, the first question was chosen as an efficient method to assess food insecurity. Hager et al [22] found that 92.5% of respondents from food-insecure households agreed with the statement "Within the past 12 months we worried whether our food would run out before we got money to buy more." While our study only used the first question, this question was still a strong indicator of food insecurity, and the second question, as used by Hager et al [22], has been recently added. Since we did not have exact data on when participants left the program, we used the last response to the food insecurity question as an approximate end date. Furthermore, it is challenging to draw firm conclusions on how duration of usage was associated with change in food insecurity because we are using real-world data without the consistent setting of a controlled study; participants were free to start and stop usage of the app whenever they wanted to. Due to the observational design of this study, we cannot draw any causal conclusions on whether usage of the platform leads to transitioning to food security.

### Conclusions

This study evaluated changes in self-reported food insecurity status and diet quality among participants with food insecurity who used a telehealth and nutrition platform with personalized recipe recommendations, meal planning, food ordering, grocery discounts, and price comparisons. While associations can be drawn between the use of Foodsmart features and achieving food security and better nutrition, future research, including a randomized controlled trial, will be needed to assess the causal effect of the Foodsmart platform on dietary changes and reduction in food insecurity.

---

### Acknowledgments

This study was funded by Foodsmart.

---

### Authors' Contributions

SB acquired data, analyzed data, interpreted results, and drafted the manuscript. BS acquired data, analyzed data, interpreted the results, and drafted the manuscript. KO drafted the manuscript. JL and JS interpreted results. EH designed the study, interpreted results, and drafted the manuscript. All authors reviewed and approved the final version of the manuscript and take responsibility for the manuscript.

---

### Conflicts of Interest

SB, BS, JS, JL, are employees of Foodsmart. KO and EH were previously employees of Foodsmart.

---

### References

1. Berkowitz SA, Basu S, Gundersen C, Seligman HK. State-level and county-level estimates of health care costs associated with food insecurity. *Prev Chronic Dis* 2019 Jul 11;16:E90 [FREE Full text] [doi: [10.5888/pcd16.180549](https://doi.org/10.5888/pcd16.180549)] [Medline: [31298210](https://pubmed.ncbi.nlm.nih.gov/31298210/)]
2. Coleman-Jensen A, Rabbitt M, Gregory C, Singh A. Household Food Security in the United States in 2019, ERR-275. Economic Research Service, U.S. Department of Agriculture. URL: <https://www.ers.usda.gov/webdocs/publications/99282/err-275.pdf?v=6570.3> [accessed 2022-10-01]



3. Fitzpatrick K, Harris C, Drawve G. Assessing U.S. Food Insecurity in the United States during COVID-19 Pandemic. University of Arkansas. URL: <https://fulbright.uark.edu/departments/sociology/research-centers/community-family-institute/resources/community-and-family-institute/revised-assessing-food-insecurity-brief.pdf> [accessed 2022-10-01]
4. Pryor S, Dietz W. The COVID-19, obesity, and food insecurity syndemic. *Curr Obes Rep* 2022 Sep;11(3):70-79 [FREE Full text] [doi: [10.1007/s13679-021-00462-w](https://doi.org/10.1007/s13679-021-00462-w)] [Medline: [35138590](https://pubmed.ncbi.nlm.nih.gov/35138590/)]
5. Lakhani N. 'A perfect storm': US facing hunger crisis as demand for food banks soars. *The Guardian*. URL: <https://www.theguardian.com/environment/2020/apr/02/us-food-banks-coronavirus-demand-unemployment> [accessed 2022-10-01]
6. Adams EL, Caccavale LJ, Smith D, Bean MK. Food insecurity, the home food environment, and parent feeding practices in the era of COVID-19. *Obesity (Silver Spring)* 2020 Nov;28(11):2056-2063 [FREE Full text] [doi: [10.1002/oby.22996](https://doi.org/10.1002/oby.22996)] [Medline: [32762129](https://pubmed.ncbi.nlm.nih.gov/32762129/)]
7. Wolfson JA, Leung CW. Food insecurity during COVID-19: an acute crisis with long-term health implications. *Am J Public Health* 2020 Dec;110(12):1763-1765. [doi: [10.2105/AJPH.2020.305953](https://doi.org/10.2105/AJPH.2020.305953)] [Medline: [32970451](https://pubmed.ncbi.nlm.nih.gov/32970451/)]
8. Leung CW, Tester JM. The association between food insecurity and diet quality varies by race/ethnicity: an analysis of National Health and Nutrition Examination Survey 2011-2014 results. *J Acad Nutr Diet* 2019 Oct;119(10):1676-1686 [FREE Full text] [doi: [10.1016/j.jand.2018.10.011](https://doi.org/10.1016/j.jand.2018.10.011)] [Medline: [30579633](https://pubmed.ncbi.nlm.nih.gov/30579633/)]
9. Gregory C, Coleman-Jensen A. Food Insecurity, Chronic Disease, and Health Among Working-Age Adults, ERR-235. Economic Research Service, U.S. Department of Agriculture. URL: <https://www.ers.usda.gov/webdocs/publications/84467/err-235.pdf> [accessed 2022-10-01]
10. Larson NI, Story MT. Food insecurity and weight status among U.S. children and families: a review of the literature. *Am J Prev Med* 2011 Feb;40(2):166-173. [doi: [10.1016/j.amepre.2010.10.028](https://doi.org/10.1016/j.amepre.2010.10.028)] [Medline: [21238865](https://pubmed.ncbi.nlm.nih.gov/21238865/)]
11. Seligman H, Jacobs E, López A, Tschann J, Fernandez A. Food insecurity and glycemic control among low-income patients with type 2 diabetes. *Diabetes Care* 2012 Feb;35(2):233-238 [FREE Full text] [doi: [10.2337/dc11-1627](https://doi.org/10.2337/dc11-1627)] [Medline: [22210570](https://pubmed.ncbi.nlm.nih.gov/22210570/)]
12. Berkowitz S, Delahanty L, Terranova J, Steiner B, Ruazol MP, Singh R, et al. Medically tailored meal delivery for diabetes patients with food insecurity: a randomized cross-over trial. *J Gen Intern Med* 2019 Mar;34(3):396-404 [FREE Full text] [doi: [10.1007/s11606-018-4716-z](https://doi.org/10.1007/s11606-018-4716-z)] [Medline: [30421335](https://pubmed.ncbi.nlm.nih.gov/30421335/)]
13. Lagisetty P, Flamm L, Rak S, Landgraf J, Heisler M, Forman J. A multi-stakeholder evaluation of the Baltimore City virtual supermarket program. *BMC Public Health* 2017 Oct 23;17(1):837 [FREE Full text] [doi: [10.1186/s12889-017-4864-9](https://doi.org/10.1186/s12889-017-4864-9)] [Medline: [29061141](https://pubmed.ncbi.nlm.nih.gov/29061141/)]
14. Hu E, Pasupuleti M, Nguyen V, Langheier J, Shurney D. Sustaining weight loss among adults with obesity using a digital meal planning and food purchasing platform for 12, 24, and 36 months: a longitudinal study. *Nutr J* 2021 Jan 21;20(1):8 [FREE Full text] [doi: [10.1186/s12937-021-00666-9](https://doi.org/10.1186/s12937-021-00666-9)] [Medline: [33478516](https://pubmed.ncbi.nlm.nih.gov/33478516/)]
15. Hu EA, Nguyen V, Langheier J, Shurney D. Weight reduction through a digital nutrition and food purchasing platform among users with obesity: longitudinal study. *J Med Internet Res* 2020 Sep 02;22(9):e19634 [FREE Full text] [doi: [10.2196/19634](https://doi.org/10.2196/19634)] [Medline: [32792332](https://pubmed.ncbi.nlm.nih.gov/32792332/)]
16. Hu EA, Scharen J, Nguyen V, Langheier J. Evaluating the impact of a digital nutrition platform on cholesterol levels in users with dyslipidemia: longitudinal study. *JMIR Cardio* 2021 Jun 10;5(1):e28392 [FREE Full text] [doi: [10.2196/28392](https://doi.org/10.2196/28392)] [Medline: [34110291](https://pubmed.ncbi.nlm.nih.gov/34110291/)]
17. Shea B, Bakre S, Carano K, Scharen J, Langheier J, Hu EA. Changes in glycemic control among individuals with diabetes who used a personalized digital nutrition platform: longitudinal study. *JMIR Diabetes* 2021 Oct 18;6(4):e32298 [FREE Full text] [doi: [10.2196/32298](https://doi.org/10.2196/32298)] [Medline: [34661545](https://pubmed.ncbi.nlm.nih.gov/34661545/)]
18. Stores Accepting SNAP Online. U.S. Department of Agriculture. URL: <https://www.fns.usda.gov/snap/online-purchasing-pilot> [accessed 2022-10-07]
19. Subar A, Thompson F, Kipnis V, Midthune D, Hurwitz P, McNutt S, et al. Comparative validation of the Block, Willett, and National Cancer Institute food frequency questionnaires : the Eating at America's Table Study. *Am J Epidemiol* 2001 Dec 15;154(12):1089-1099. [doi: [10.1093/aje/154.12.1089](https://doi.org/10.1093/aje/154.12.1089)] [Medline: [11744511](https://pubmed.ncbi.nlm.nih.gov/11744511/)]
20. Chiuvè SE, Fung TT, Rimm EB, Hu FB, McCullough ML, Wang M, et al. Alternative dietary indices both strongly predict risk of chronic disease. *J Nutr* 2012 Jun;142(6):1009-1018 [FREE Full text] [doi: [10.3945/jn.111.157222](https://doi.org/10.3945/jn.111.157222)] [Medline: [22513989](https://pubmed.ncbi.nlm.nih.gov/22513989/)]
21. Hendrie GA, Baird D, Golley RK, Noakes M. The CSIRO Healthy Diet Score: an online survey to estimate compliance with the Australian Dietary Guidelines. *Nutrients* 2017 Jan 09;9(1):47 [FREE Full text] [doi: [10.3390/nu9010047](https://doi.org/10.3390/nu9010047)] [Medline: [28075355](https://pubmed.ncbi.nlm.nih.gov/28075355/)]
22. Hager E, Quigg A, Black M, Coleman S, Heeren T, Rose-Jacobs R, et al. Development and validity of a 2-item screen to identify families at risk for food insecurity. *Pediatrics* 2010 Jul;126(1):e26-e32. [doi: [10.1542/peds.2009-3146](https://doi.org/10.1542/peds.2009-3146)] [Medline: [20595453](https://pubmed.ncbi.nlm.nih.gov/20595453/)]
23. Nord M, Andrews M, Carlson S. Household food security in the United States. Economic Research Service, U.S. Department of Agriculture. 2007. URL: <https://www.ers.usda.gov/media/8271/hh2012.pdf> [accessed 2022-10-01]
24. Leung C, Epel E, Ritchie L, Crawford P, Laraia B. Food insecurity is inversely associated with diet quality of lower-income adults. *J Acad Nutr Diet* 2014 Dec;114(12):1943-53.e2. [doi: [10.1016/j.jand.2014.06.353](https://doi.org/10.1016/j.jand.2014.06.353)] [Medline: [25091796](https://pubmed.ncbi.nlm.nih.gov/25091796/)]

25. Walker RJ, Garacci E, Dawson AZ, Williams JS, Ozieh M, Egede LE. Trends in food insecurity in the United States from 2011-2017: disparities by age, sex, race/ethnicity, and income. *Popul Health Manag* 2021 Aug 01;24(4):496-501 [FREE Full text] [doi: [10.1089/pop.2020.0123](https://doi.org/10.1089/pop.2020.0123)] [Medline: [32941115](https://pubmed.ncbi.nlm.nih.gov/32941115/)]
26. Pan L, Sherry B, Njai R, Blanck HM. Food insecurity is associated with obesity among US adults in 12 states. *J Acad Nutr Diet* 2012 Sep;112(9):1403-1409 [FREE Full text] [doi: [10.1016/j.jand.2012.06.011](https://doi.org/10.1016/j.jand.2012.06.011)] [Medline: [22939441](https://pubmed.ncbi.nlm.nih.gov/22939441/)]
27. Ranjit N, Macias S, Hoelscher D. Factors related to poor diet quality in food insecure populations. *Transl Behav Med* 2020 Dec 31;10(6):1297-1305 [FREE Full text] [doi: [10.1093/tbm/ibaa028](https://doi.org/10.1093/tbm/ibaa028)] [Medline: [33421089](https://pubmed.ncbi.nlm.nih.gov/33421089/)]
28. Laraia BA. Food insecurity and chronic disease. *Adv Nutr* 2013 Mar 01;4(2):203-212 [FREE Full text] [doi: [10.3945/an.112.003277](https://doi.org/10.3945/an.112.003277)] [Medline: [23493536](https://pubmed.ncbi.nlm.nih.gov/23493536/)]
29. Map the Meal Gap 2020. Feeding America. URL: <https://www.feedingamerica.org/sites/default/files/2020-06/Map%20the%20Meal%20Gap%202020%20Combined%20Modules.pdf> [accessed 2022-10-07]
30. Lane MM, Davis JA, Beattie S, Gómez-Donoso C, Loughman A, O'Neil A, et al. Ultra-processed food and chronic noncommunicable diseases: A systematic review and meta-analysis of 43 observational studies. *Obes Rev* 2021 Mar;22(3):e13146. [doi: [10.1111/obr.13146](https://doi.org/10.1111/obr.13146)] [Medline: [33167080](https://pubmed.ncbi.nlm.nih.gov/33167080/)]
31. SNAP Benefits - COVID-19 Pandemic and Beyond. Food and Nutrition Service, U.S. Department of Agriculture. URL: <https://www.fns.usda.gov/snap/benefit-changes-2021> [accessed 2022-10-01]
32. Leung CW, Ding EL, Catalano PJ, Villamor E, Rimm EB, Willett WC. Dietary intake and dietary quality of low-income adults in the Supplemental Nutrition Assistance Program. *Am J Clin Nutr* 2012 Nov;96(5):977-988 [FREE Full text] [doi: [10.3945/ajcn.112.040014](https://doi.org/10.3945/ajcn.112.040014)] [Medline: [23034960](https://pubmed.ncbi.nlm.nih.gov/23034960/)]
33. Access to affordable and nutritious food: Measuring and understanding food deserts and their consequences. Economic Research Service, U.S. Department of Agriculture. 2009. URL: [https://www.ers.usda.gov/webdocs/publications/42711/12698\\_ap036fm\\_1.pdf?v=41055](https://www.ers.usda.gov/webdocs/publications/42711/12698_ap036fm_1.pdf?v=41055) [accessed 2022-10-01]
34. Wang M, Levi R, Seligman H. New SNAP eligibility in California associated with improved food security and health. *Prev Chronic Dis* 2021 Apr 01;18:E28 [FREE Full text] [doi: [10.5888/pcd18.200587](https://doi.org/10.5888/pcd18.200587)] [Medline: [33793394](https://pubmed.ncbi.nlm.nih.gov/33793394/)]
35. Jackson C. Blue Cross Blue Shield Subsidiary to End Food Delivery. Specialty Food Association. URL: <https://www.specialtyfood.com/news/article/blue-cross-blue-shield-subsidiary-end-food-delivery-program/> [accessed 2022-10-01]
36. De Marchis EH, Torres JM, Benesch T, Fichtenberg C, Allen IE, Whitaker EM, et al. Interventions addressing food insecurity in health care settings: a systematic review. *Ann Fam Med* 2019 Sep;17(5):436-447 [FREE Full text] [doi: [10.1370/afm.2412](https://doi.org/10.1370/afm.2412)] [Medline: [31501207](https://pubmed.ncbi.nlm.nih.gov/31501207/)]

## Abbreviations

- HEI:** Healthy Eating Index  
**HFSS:** Household Food Security Survey  
**OR:** odds ratio  
**RD:** registered dietitian  
**SNAP:** Supplemental Nutrition Assistance Program  
**USDA:** United States Department of Agriculture

*Edited by A Mavragani; submitted 25.07.22; peer-reviewed by O Rahaman, E Brown; comments to author 17.08.22; revised version received 09.09.22; accepted 22.09.22; published 25.10.22.*

*Please cite as:*

Bakre S, Shea B, Ortega K, Scharen J, Langheier J, Hu E  
*Changes in Food Insecurity Among Individuals Using a Telehealth and Nutrition Platform: Longitudinal Study*  
*JMIR Form Res* 2022;6(10):e41418  
URL: <https://formative.jmir.org/2022/10/e41418>  
doi: [10.2196/41418](https://doi.org/10.2196/41418)  
PMID: [36282563](https://pubmed.ncbi.nlm.nih.gov/36282563/)

©Shivani Bakre, Benjamin Shea, Kaylee Ortega, Jared Scharen, Jason Langheier, Emily Hu. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 25.10.2022. 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

# Colocating Teleophthalmology Within Primary Care Settings to Improve Access to Diabetic Retinopathy Screening: Retrospective Descriptive Evaluation

Tiffany Wandy<sup>1</sup>, MPH, MBA; Shuja Rayaz<sup>1</sup>, MD; Jenna Ashly Levenson Brager<sup>1</sup>, RN, PhD; Michael Kiritsy<sup>1</sup>, MBA, MD; Elizabeth Offermann<sup>1</sup>, BA, MSHS; Daniel Durand<sup>1</sup>, MD

LifeBridge Health, Baltimore, MD, United States

**Corresponding Author:**

Shuja Rayaz, MD  
LifeBridge Health  
2401 West Belvedere Avenue  
Baltimore, MD, 21215  
United States  
Phone: 1 410 601 4388  
Email: [shuja.rayaz@gmail.com](mailto:shuja.rayaz@gmail.com)

## Abstract

**Background:** Annual retinal exams for patients with diabetes are critical as diabetic retinopathy is the number one cause of preventable blindness in working-age adults in the United States. Currently, most patients with diabetes in the United States receive a referral from their primary care provider to see an ophthalmologist for their annual dilated eye exam, which can be an added inconvenience and expense. As such, there is a need for alternative screening strategies within an outpatient network. The use of a telemedicine platform in a primary care setting serves as a novel strategy to increase diabetic retinopathy screening rates. In order to provide better access to diabetic retinopathy screening for our patients, cameras were placed in 3 primary care practices in October 2017 as part of an 8-month pilot program. Specialized cameras from Intelligent Retinal Imaging Systems (IRIS) were used to acquire images that could be interpreted remotely by ophthalmologists within the LifeBridge Health network for the diagnosis of diabetic retinopathy and the detection of other types of pathology (eg, macular edema).

**Objective:** The aim of this retrospective descriptive study was to examine whether a telemedicine platform can be used as a cost-effective way to increase diabetic retinopathy screening rates in the primary care setting.

**Methods:** Aggregate screening volume and diagnostic data were collected for each of the 3 practice locations for the 8-month pilot period (October 30, 2017, through June 30, 2018). Additionally, payor reimbursement data and equipment cost data were used to determine the payback period for each of the 3 practice locations.

**Results:** The pilot program proved the business case that implementation of the IRIS camera in 3 practice locations could result in enough patients being screened to pay for the cost of the camera within a maximum of 2 years. The 3 practices showed increased diabetic retinopathy screening rates of 1%, 6%, and 24%, respectively, and were all able to screen enough patients to be on track to pay off the cost of the camera within 2 years of implementation. Aggregate data from the pilot period showed that of the 1213 patients who were screened, approximately 17.1% (n=207) were diagnosed with diabetic retinopathy and an additional 17.7% (n=215) were suspected of having some other form of pathology. Of note, 10.1% (n=123) were also identified as being “IRIS saves,” defined as having pathology identified that was severe enough to be considered an imminent threat to their vision.

**Conclusions:** This retrospective descriptive study suggests that a telemedicine platform can be used to improve diabetic retinopathy screening rates in the primary care setting within a large health care system in a cost-effective way that allows for the cost of the equipment to be recouped through billing within a maximum of 2 years.

(*JMIR Form Res* 2022;6(10):e17838) doi:[10.2196/17838](https://doi.org/10.2196/17838)

**KEYWORDS**

diabetic retinopathy; screening; telemedicine; teleretinopathy; population health; value-based care; quality measure performance; accountable care organization; ACO performance

## Introduction

Diabetic retinopathy is the leading cause of blindness among working-age adults in the United States [1]. Early detection and treatment can prevent or delay blindness due to diabetic retinopathy in 90% of patients with diabetes, but 50% or more of patients do not obtain an eye exam or are diagnosed too late for effective treatment [1]. As a result, early screening is crucial as retinopathy is often asymptomatic but can progress over time to vision loss [2].

The American Diabetes Association recommends an annual eye exam for all adult patients with diabetes [3]. Annual eye exams are rather burdensome in that they involve a referral to an ophthalmologist for a standard, in-person dilated retinal screening each year. A preliminary analysis within our health system suggested that primary care providers are usually diligent in providing patients with referrals to ophthalmologists, but many patients do not follow their primary care provider's advice and get screened. This lack of follow-through was due to a number of challenges and barriers, including the patient's lack of understanding of risk, the burdens involved in additional doctor visits for patients who were already under the care of multiple physicians, the costs associated with specialty care, obstacles to accessing care, and the confusing and fragmented nature of the US health care environment.

In recent years, practitioners have looked to alternative methods for eye screening to overcome patient- and system-level barriers. Telemedicine for diabetic retinopathy screening can be advantageous compared to a traditional dilated eye exam because it does not involve dilation of the eye, can usually be done in a short amount of time (total time <15 minutes), does not require a specialist copay cost on behalf of the patient, does not require a specialty provider to acquire the images, and does not require additional travel for patients when the exam is performed in a primary care setting during a regularly scheduled appointment [4,5]. Using telemedicine gives the opportunity for health systems to increase screening rates and provide the proper standard of care for their patients with diabetes.

Intelligent Retinal Imaging Systems (IRIS) is a company that provides a secure platform for capturing and grading retinal images that can be interpreted remotely by an ophthalmologist. The IRIS platform was identified based on the value of the platform, the quality of the retinal images, and the company's commitment to implementation success (in-person and virtual

assistance). In this study, specialized IRIS cameras were used to acquire retinal images at 3 outpatient practices, which were then automatically uploaded to an online platform that served as a data warehouse for these images [6]. The IRIS online platform is currently approved by the US Food and Drug Administration for the diagnosis of diabetic retinopathy and diabetic macular edema [7]. LifeBridge Health primarily used the platform for the diagnosis of diabetic retinopathy, though ophthalmologists interpreting the images at LifeBridge Health's Krieger Eye Institute were also given the option to identify images for macular edema, glaucoma, macular degeneration, vein occlusion, and hypertensive retinopathy. The objective of this study was to examine the rate of diabetic retinopathy screening with the novel implementation of a telemedicine platform in the primary care clinical setting and show that implementation of the cameras could be cost-effective, even with the small profit margins experienced in primary care.

## Methods

### Study Design

The 8-month pilot program began in late October 2017 with the placement of 2 tabletop TopCon IRIS cameras and one handheld, mobile Volk IRIS camera at 3 primary care practices. Both types of cameras obtain similar quality images, but the handheld unit requires additional training for optimal image acquisition. The tabletop camera costs almost twice as much as the handheld camera but is less subject to damage, for example, from dropping the unit, as can happen with the handheld unit.

The primary care practices included in the pilot were identified using quantitative and qualitative measures. The 3 pilot locations were large, multiprovider practices with a high number of patients with diabetes. Additionally, the practices were enthusiastic about participating in a new technology pilot program, and each identified a physician champion to assist with the implementation. As an indication of the overall patient population, [Table 1](#) describes the insurance payor mix for the 3 pilot primary care practices. Practice #1 is located in suburban Baltimore in Carroll County and includes patients from a mix of payors, with a heavy emphasis (61%) on commercial payor patients. Practice #2 is located in an affluent suburb of Baltimore in Howard County and mainly includes commercial payor and Medicare patients, with no Medicaid. Practice #3 is located in the city of Baltimore and includes the lowest percentage of commercially insured patients.

**Table 1.** Overall insurance payor mix per practice.

Practices	Commercial or private, %	Medicare, %	Medicaid, %	Uninsured, %	Other payor, %
Practice #1	61	15	17	7	0
Practice #2	61	36	0	0	3
Practice #3	43	38	14	5	0

At these practices, patients who had a current diagnosis of diabetes documented in their electronic medical record and who had not had a diabetic retinopathy screening exam documented within the last 12 months were identified during their primary care visits. These patients were then given the opportunity to

have a nondilated diabetic retinopathy exam as part of the pilot study. Voluntary participation in the pilot program was offered to all patients who met the inclusion criteria. There were no specific exclusion criteria for this study because the focus of the program was to improve diabetic retinopathy screening rates

(specifically whether a patient had a screening or not) among all patients, regardless of payor, in outpatient practices. The diabetic retinopathy screening exam was accepted by all payors in the pilot region, and claims were submitted accordingly for fundus photography with interpretation and report (Current Procedural Terminology [CPT] 92250).

Working with a dedicated LifeBridge Health team, the practices created and implemented workflows, documented processes, and established best practices for camera use in conjunction with primary care physicians and other health care team members at each specific practice. Medical assistants were trained and held responsible for performing the exams and obtaining the images. For each patient screened, 2 fundus images of the retina without the need for dilation were captured via the IRIS cameras within the primary care practices. This is in contrast to the workflow in practices not participating in the pilot where patients were merely referred to an ophthalmologist for a traditional dilated retinal exam, to be performed at a separate visit on another date.

Following the screening, the images were automatically transmitted electronically via the secure online IRIS platform. These images were then read by ophthalmologists at LifeBridge Health's Krieger Eye Institute via the IRIS platform. Ophthalmologists could manipulate the images using enhancement tools and then conduct a manual diagnostic grading of each image. LifeBridge Health's Krieger Eye Institute ophthalmologists diagnosed patients as having diabetic retinopathy based on the image review and filled out a templated report to send back to the primary care provider. The ophthalmologists also had the option to designate images as being suspicious for other forms of pathology (including macular edema, glaucoma, macular degeneration, vein occlusion, and hypertensive retinopathy) or flag patients who were identified as having pathology that was severe enough to put them at risk for imminent vision loss for immediate follow-up. If an image was diagnosed as diabetic retinopathy or found to be suspicious for some other form of pathology, the patient was called to schedule an in-person visit with an ophthalmologist for further diagnostics and treatment as needed.

The pre- and postscreening rates for each of the participating practices in the pilot program were obtained by manual chart review. This chart review involved randomly selecting for each provider at least 30 patients with diabetes with office visits in a given month who were eligible for retinopathy screening. The screening rate from the chart review samples was then used to extrapolate the screening rates for the entire practice for all patients with diabetes. The data set for the period from October 2017 to December 2019 used data collected from the IRIS online platform. The IRIS platform provides a dashboard with aggregated data on the number of patients screened, the number of patients diagnosed with diabetic retinopathy or some other form of pathology, and the number of patients who were identified as being at risk for imminent vision loss.

### Data Analysis

The main return-on-investment analysis was performed to determine whether it would be possible for each of the primary care practices to generate enough revenue from screening

patients using the IRIS cameras to pay back the cost of the camera within a set period of time. The number of screenings needed to pay off the cost of the camera within 2 years was determined by taking the overall cost of the camera (including shipping and warranty costs) and dividing by the average reimbursement for CPT 92250 for each practice, which varied according to payor mix. The screening goal for each practice was determined by prorating the total number of screenings needed to pay off the camera within 2 years to reflect the 8-month length of the pilot study.

The key clinical analysis focused on whether the diabetic retinopathy screening rates improved with the utilization of IRIS cameras within the 3 pilot practices. This analysis compared the screening rates at the 3 pilot practices before and after the implementation of the IRIS cameras. This analysis compared screening rates from January 2018 to September 2018 to the screening rates at those practices from January 2017 to December 2017, when the practices did not have the cameras or use any type of digital screening (although the pilot was conducted from the end of October 2017 to June 2018, sufficient time was given as a "ramp-up period" for the use of the cameras to be integrated into the practice workflow). Of note, since diagnosis via captured images, compared to in-person examination, depends on the quality of the image, not all images were suitable for diagnosis. The IRIS online platform deemed 88% of all acquired images to be gradable for interpretation by LifeBridge Health's Krieger Eye Institute ophthalmologists. The analysis presented in this paper includes only those images deemed gradable.

### Ethics Approval

The LifeBridge Health institutional review board (IRB) determined that the study is exempt from ethics approval as it did not meet the definition of human subjects research and did not need a formal review.

### Results

**Table 2** shows the total number of screenings conducted during the pilot at each of the 3 practices and breaks down the total screening volumes by the goal (the number of screenings needed to be on track to pay off the cost of the camera within 2 years) and the number of screenings conducted in excess of the goal. All 3 practices screened enough patients to reach their goal, regardless of the type of camera they were using or the payor mix of their patient population.

During the 8-month pilot, practice #1 screened a total of 98 patients. This practice had a handheld camera and needed to screen 30 patients during the pilot to be on track to pay off the cost of the camera within 2 years. Practice #2 screened a total of 164 patients. This practice had a tabletop camera and needed to screen 114 during the pilot to be on track to pay off the cost of the camera within 2 years. Practice #3 screened a total of 196 patients. This practice also had a tabletop camera and needed to screen 138 patients to pay off the cost of the camera within 2 years. The number of screenings needing to be performed to pay off the cost of the camera within 2 years depended on the type of camera the practice had (the tabletop camera costs nearly

twice as much as the handheld version) and the average reimbursement based upon payor mix. For context, the average reimbursement across all 3 practices and across all payors was US \$56.45, with commercial payors tending to reimburse at the highest rates, followed by Medicare, with Medicaid reimbursing at the lowest rate.

Table 3 depicts the screening rates before and after the initiation of the pilot program at each of the 3 practices, as well as the corresponding percentage change. The 3 practices showed increased diabetic retinopathy screening rates of 1%, 6%, and 24%, respectively.

In total, 1213 patients were screened for diabetic retinopathy. Of these 1213 patients, approximately 17.1% (n=207) were diagnosed with diabetic retinopathy and 17.7% (n=215) were

suspected of having some other form of pathology. In addition, 10.1% (n=123) were suspected of having pathology that was deemed serious enough to put them at risk for imminent vision loss. For all patients requiring follow-up (those with any form of pathology), direct referrals were made to an in-network ophthalmologist at the Krieger Eye Institute for further evaluation and treatment. Table 3 also shows the impact of the telemedicine initiative and the improved screening rates achieved at the 3 primary care practices. Practice #3 showed a particularly large improvement in screening rate (24%), most likely due to a best-practice workflow that was implemented at the practice location, specifically one in which medical assistants were able to seamlessly identify and direct patients with diabetes to obtain a retinopathy exam via a standing order.

**Table 2.** Volume of screenings by practice (October 30, 2017, to June 30, 2018).

	Goal (screenings needed to pay off camera in 2 years), n	Screenings in excess of goal, n	Total screenings conducted during pilot, N
Practice #1	30	68	98
Practice #2	114	50	164
Practice #3	138	58	196

**Table 3.** Screening rates before and after implementation of Intelligent Retinal Imaging Systems cameras in the 3 pilot practices.

Practices	Screening rate before pilot <sup>a</sup> , %	Screening rate during pilot <sup>b</sup> , %	Percentage change
Practice #1	39	40	+1
Practice #2	32	38	+6
Practice #3	40	64	+24

<sup>a</sup>Data calculated from January 2017 to December 2017.

<sup>b</sup>Data calculated from January 2018 to September 2018.

## Discussion

### Principal Findings

Our study shows that the implementation of a telemedicine platform for diabetic retinopathy screening in primary care practices can increase screening rates among adult patients with diabetes and be feasible within a primary care environment with low profit margins. Furthermore, integrating camera utilization within the provider workflow allowed for better adoption of its use and therefore increased opportunities for screening patients. Of the 1213 patients who underwent screening, 17.1% (207 patients) were diagnosed as having diabetic retinopathy and 10.1% (123 patients) were found to be at risk for imminent vision loss. The 123 patients at risk for imminent vision loss were subsequently contacted for an urgent follow-up visit with an ophthalmologist. Thus, there was a direct and meaningful impact on patient care due to the implementation of the IRIS cameras and use of the online platform within the LifeBridge Health network.

Diabetic retinopathy is the most common and serious ocular complication of diabetes and is one of the leading causes of vision loss in many developed countries [2,8,9]. Screening for diabetic retinopathy traditionally involves an in-person dilated eye exam with an ophthalmologist; however, even with support

from governmental and nongovernmental agencies, medical societies, and various global organizations, a large portion of patients with diabetes do not receive the recommended annual eye exam [10,11]. Telemedicine provides an alternative method for diabetic retinopathy screening and has been shown to be effective for the diagnosis of new diabetic retinopathy as well as other ophthalmologic diseases [12,13]. A large-scale retrospective study involving over 15,000 patients conducted in the United States showed that the use of the IRIS cameras and platform yielded high sensitivity and a low false-negative rate for the diagnosis of diabetic retinopathy [14].

From a population health standpoint, current literature has examined several patient-level barriers to retinopathy screening, including patients having competing priorities, anxiety about the screening, disengagement with diabetes care, misinformation about the screening, and forgetting to attend the screening [15]. Using telemedicine can help curb some of these issues and increase access to care. Patients can have their screening performed during a regular office visit with their primary care physician, eliminating the need for an additional appointment with an ophthalmologist. Retinal images can be interpreted remotely by an ophthalmologist; if necessary, patients can be contacted by a specialist for further evaluation. This workflow directly involves the specialist in the patient's care after their primary care visit, rather than placing the responsibility on the

patient to follow up with the specialist. A multicenter randomized controlled trial conducted over 5 years in the United Kingdom also showed that patients who received digital diabetic retinopathy screening at a primary care office were more likely to receive diabetic retinopathy screening compared to those who underwent screening at an additional office visit with an eye specialist [16]. This finding implies that the use of telemedicine can increase access and adherence to screening, which is particularly important for those patients in lower socioeconomic communities who tend to have a higher prevalence of diabetes [17]. Additionally, this form of retinopathy disproportionately affects the vision of lower-income households, for whom the subsequent disability can be economically devastating. In this study, practice #3, which has the highest percentage of patients with government-subsidized (Medicare and Medicaid) insurance or no insurance and the highest percentage of patients of lower socioeconomic status, showed the highest increase in screening rates after the implementation of the IRIS cameras among the pilot practices (Table 3).

### Strengths

This study is retrospective and descriptive in nature as it examines the novel deployment of a telemedicine system within an outpatient primary care clinical practice network in a major metropolitan area. The original intent of the pilot program was to improve access to diabetic retinopathy screening among our patient population. Overall, we have been able to screen over 1200 patients, many of whom may have not been screened otherwise. It is our view that the implementation of the IRIS cameras within the 3 practices provided a convenient alternative for our patients within the community. Furthermore, 123 out of 1213 (10%) screened patients were found to be at risk for imminent vision loss, and we were able to swiftly identify those patients and refer them to our ophthalmologists in a timely manner for urgent follow-up. Due to the successful implementation in the pilot practices, we have continued to expand the program to additional practices within our primary care network. Since the conclusion of the pilot study, we have conducted a full electronic medical record integration project to ensure an even smoother clinical process for conducting diabetic retinopathy screening and sending clinical grading back to the primary care providers. This initial study provides

the foundation for future work, including a possible randomized controlled trial to better elucidate the impact of a telemedicine system in the outpatient primary care setting.

### Limitations

This study was not able to report statistical comparisons or inferences due to a small sample size. However, we were able to identify a clinical benefit with the implementation and use of the IRIS cameras in the outpatient practices as evidenced by the aggregate data mentioned in the Results section. System-level limitations exist when using a telemedicine platform for diabetic retinopathy screening and were shown to affect this pilot study. For example, the use of the cameras requires additional time and effort by the medical assistants in the primary care practices, which can be challenging to operationalize in busy outpatient practices. Though this study was small and had its limitations, the overall results highlight a noteworthy improvement in screening rates with the potential to impact a larger population should this process be rolled out to additional primary care practice locations.

### Conclusions

This study suggests that the use of a telemedicine platform in the primary care setting can be an effective alternative to dilated eye exams performed by ophthalmologists for adult patients with diabetes. Implementation of a telemedicine platform allowed our primary care providers to offer comprehensive care to their patients with diabetes, reducing the need for multiple appointments and specialist co-pays. Our most successful implementation of the telemedicine platform was in the practice with the highest percentage of patients of lower socioeconomic status and government-subsidized insurance or no insurance (practice #3). Telemedicine can provide increased access to care for this type of patient population in particular. Timely screening for retinopathy also had a direct benefit on the clinical care of our patients. Since the implementation of the telemedicine platform, 123 patients were found to be at risk for imminent vision loss and were immediately referred to an ophthalmologist for time-sensitive treatment. Overall, this approach of using medical devices and specialized software to increase screening rates will allow us greater opportunities to provide sight-saving treatments and help to prevent blindness in our patients.

### Conflicts of Interest

None declared.

### References

1. Keep an eye on your vision health. Centers for Disease Control and Prevention. URL: <https://www.cdc.gov/visionhealth/resources/features/keep-eye-on-vision-health.html> [accessed 2019-12-08]
2. Fong D, Aiello L, Gardner T, King G, Blankenship G, Cavallerano J, American Diabetes Association. Retinopathy in diabetes. *Diabetes Care* 2004 Jan;27 Suppl 1(Supplement 1):S84-S87. [doi: [10.2337/diacare.27.2007.s84](https://doi.org/10.2337/diacare.27.2007.s84)] [Medline: [14693935](https://pubmed.ncbi.nlm.nih.gov/14693935/)]
3. American Diabetes Association. Standards of medical care in diabetes-2015 abridged for primary care providers. *Clin Diabetes* 2015 Apr;33(2):97-111 [FREE Full text] [doi: [10.2337/diaclin.33.2.97](https://doi.org/10.2337/diaclin.33.2.97)] [Medline: [25897193](https://pubmed.ncbi.nlm.nih.gov/25897193/)]
4. Kim J, Driver DD. Teleophthalmology for first nations clients at risk of diabetic retinopathy: a mixed methods evaluation. *JMIR Med Inform* 2015 Feb 23;3(1):e10 [FREE Full text] [doi: [10.2196/medinform.3872](https://doi.org/10.2196/medinform.3872)] [Medline: [25705814](https://pubmed.ncbi.nlm.nih.gov/25705814/)]

5. Abramoff MD, Suttorp-Schulten MS. Web-based screening for diabetic retinopathy in a primary care population: the EyeCheck project. *Telemed J E Health* 2005 Dec;11(6):668-674. [doi: [10.1089/tmj.2005.11.668](https://doi.org/10.1089/tmj.2005.11.668)] [Medline: [16430386](https://pubmed.ncbi.nlm.nih.gov/16430386/)]
6. Intelligent Retinal Imaging Systems. URL: <https://retinalscreenings.com/solution> [accessed 2019-12-15]
7. US Department of Health and Human Services. 510(k) premarket notification. US Food and Drug Administration. 2015. URL: <https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpmn/pmnm.cfm?ID=K141922> [accessed 2021-09-01]
8. Congdon NG, Friedman DS, Lietman T. Important causes of visual impairment in the world today. *JAMA* 2003 Oct 15;290(15):2057-2060. [doi: [10.1001/jama.290.15.2057](https://doi.org/10.1001/jama.290.15.2057)] [Medline: [14559961](https://pubmed.ncbi.nlm.nih.gov/14559961/)]
9. Jeganathan VSE, Wang JJ, Wong TY. Ocular associations of diabetes other than diabetic retinopathy. *Diabetes Care* 2008 Sep 01;31(9):1905-1912 [FREE Full text] [doi: [10.2337/dc08-0342](https://doi.org/10.2337/dc08-0342)] [Medline: [18753669](https://pubmed.ncbi.nlm.nih.gov/18753669/)]
10. Mukamel BD, Bresnick HG, Wang Q, Dickey FC. Barriers to compliance with screening guidelines for diabetic retinopathy. *Ophthalmic Epidemiol* 1999 Mar 08;6(1):61-72. [doi: [10.1076/oep.6.1.61.1563](https://doi.org/10.1076/oep.6.1.61.1563)] [Medline: [10384685](https://pubmed.ncbi.nlm.nih.gov/10384685/)]
11. Gower EW, Silverman E, Cassard SD, Williams SK, Baldonado K, Friedman DS. Barriers to attending an eye examination after vision screening referral within a vulnerable population. *J Health Care Poor Underserved* 2013 Aug;24(3):1042-1052. [doi: [10.1353/hpu.2013.0134](https://doi.org/10.1353/hpu.2013.0134)] [Medline: [23974379](https://pubmed.ncbi.nlm.nih.gov/23974379/)]
12. Bhargava M, Cheung CY, Sabanayagam C, Kawasaki R, Harper CA, Lamoureux EL, et al. Accuracy of diabetic retinopathy screening by trained non-physician graders using non-mydratric fundus camera. *Singapore Med J* 2012 Nov;53(11):715-719 [FREE Full text] [Medline: [23192497](https://pubmed.ncbi.nlm.nih.gov/23192497/)]
13. Philip S, Fleming AD, Goatman KA, Fonseca S, McNamee P, Scotland GS, et al. The efficacy of automated "disease/no disease" grading for diabetic retinopathy in a systematic screening programme. *Br J Ophthalmol* 2007 Nov 15;91(11):1512-1517 [FREE Full text] [doi: [10.1136/bjo.2007.119453](https://doi.org/10.1136/bjo.2007.119453)] [Medline: [17504851](https://pubmed.ncbi.nlm.nih.gov/17504851/)]
14. Walton OB, Garoon RB, Weng CY, Gross J, Young AK, Camero KA, et al. Evaluation of automated teleretinal screening program for diabetic retinopathy. *JAMA Ophthalmol* 2016 Feb 01;134(2):204-209. [doi: [10.1001/jamaophthalmol.2015.5083](https://doi.org/10.1001/jamaophthalmol.2015.5083)] [Medline: [26720694](https://pubmed.ncbi.nlm.nih.gov/26720694/)]
15. Kashim R, Newton P, Ojo O. Diabetic retinopathy screening: a systematic review on patients' non-attendance. *Int J Environ Res Public Health* 2018 Jan 19;15(1):157 [FREE Full text] [doi: [10.3390/ijerph15010157](https://doi.org/10.3390/ijerph15010157)] [Medline: [29351207](https://pubmed.ncbi.nlm.nih.gov/29351207/)]
16. Mansberger SL, Sheppler C, Barker G, Gardiner SK, Demirel S, Wooten K, et al. Long-term comparative effectiveness of telemedicine in providing diabetic retinopathy screening examinations: a randomized clinical trial. *JAMA Ophthalmol* 2015 May 01;133(5):518-525 [FREE Full text] [doi: [10.1001/jamaophthalmol.2015.1](https://doi.org/10.1001/jamaophthalmol.2015.1)] [Medline: [25741666](https://pubmed.ncbi.nlm.nih.gov/25741666/)]
17. Xu Z, Yu D, Yin X, Zheng F, Li H. Socioeconomic status is associated with global diabetes prevalence. *Oncotarget* 2017 Jul 04;8(27):44434-44439 [FREE Full text] [doi: [10.18632/oncotarget.17902](https://doi.org/10.18632/oncotarget.17902)] [Medline: [28574844](https://pubmed.ncbi.nlm.nih.gov/28574844/)]

## Abbreviations

**CPT:** Current Procedural Terminology

**IRIS:** Intelligent Retinal Imaging Systems

*Edited by G Eysenbach; submitted 15.01.20; peer-reviewed by G Lim, G Capocasale; comments to author 23.09.20; revised version received 19.02.21; accepted 03.05.22; published 26.10.22.*

*Please cite as:*

Wandy T, Rayaz S, Brager JAL, Kiritsy M, Offermann E, Durand D

*Colocating Teleophthalmology Within Primary Care Settings to Improve Access to Diabetic Retinopathy Screening: Retrospective Descriptive Evaluation*

*JMIR Form Res* 2022;6(10):e17838

URL: <https://formative.jmir.org/2022/10/e17838>

doi: [10.2196/17838](https://doi.org/10.2196/17838)

PMID: [36287608](https://pubmed.ncbi.nlm.nih.gov/36287608/)

©Tiffany Wandy, Shuja Rayaz, Jenna Ashly Levenson Brager, Michael Kiritsy, Elizabeth Offermann, Daniel Durand. Originally published in *JMIR Formative Research* (<https://formative.jmir.org>), 26.10.2022. 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 Smartphone-Based Self-administered Tool (R+ Dietitian) for Nutritional Risk Screening and Dietary Assessment in Hospitalized Patients With Cancer: Evaluation and Diagnostic Accuracy Study

Zhiwen Long<sup>1\*</sup>, MPH; Shan Huang<sup>2,3\*</sup>, MD; Jie Zhang<sup>2,3</sup>, MM; Deng Zhang<sup>2</sup>, MM; Jun Yin<sup>1</sup>, MM; Chengyuan He<sup>1</sup>, BSc; Qinqiu Zhang<sup>1,4</sup>, MSA; Huilin Xu<sup>1</sup>, AAS; Huimin He<sup>1</sup>, BSc; Ho Ching Sun<sup>1</sup>, BSc; Ke Xie<sup>2,3</sup>, MD

<sup>1</sup>Recovery Plus Clinic, Chengdu, China

<sup>2</sup>Department of Oncology, Sichuan Academy of Medical Sciences and Sichuan Provincial People's Hospital, Chengdu, China

<sup>3</sup>School of Medicine, University of Electronic Science and Technology of China, Chengdu, China

<sup>4</sup>College of Food Science, Sichuan Agricultural University, Ya'an, China

\*these authors contributed equally

**Corresponding Author:**

Ke Xie, MD

Department of Oncology, Sichuan Academy of Medical Sciences and Sichuan Provincial People's Hospital

Number 32, West 2nd Section, First Ring Road

Qingyang District

Chengdu, 610072

China

Phone: 86 18981838382

Email: [xieke@med.uestc.edu.cn](mailto:xieke@med.uestc.edu.cn)

## Abstract

**Background:** Malnutrition is a common and severe problem in patients with cancer that directly increases the incidence of complications and significantly deteriorates quality of life. Nutritional risk screening and dietary assessment are critical because they are the basis for providing personalized nutritional support. No digital smartphone-based self-administered tool for nutritional risk screening and dietary assessment among hospitalized patients with cancer has been developed and evaluated.

**Objective:** This study aims to develop a digital smartphone-based self-administered mini program for nutritional risk screening and dietary assessment for hospitalized patients with cancer and to evaluate the validity of the mini program.

**Methods:** We have developed the R+ Dietitian mini program, which consists of 3 parts: (1) collection of basic information of patients, (2) nutritional risk screening, and (3) dietary energy and protein assessment. The face-to-face paper-based Nutritional Risk Screening (NRS-2002), the Patient-Generated Subjective Global Assessment Short Form (PG-SGA-SF), and 3 days of 24-hour dietary recall (3d-24HRs) questionnaires were administered according to standard procedure by 2 trained dietitians as the reference methods. Sensitivity, specificity, positive predictive value, negative predictive value,  $\kappa$  value, and correlation coefficients (CCs) of nutritional risk screened in R+ Dietitian against the reference methods, as well as the difference and CCs of estimated dietary energy and protein intakes between R+ Dietitian and 3d-24HRs were calculated to evaluate the validity of R+ Dietitian.

**Results:** A total of 244 hospitalized patients with cancer were recruited to evaluate the validity of R+ Dietitian. The NRS-2002 and PG-SGA-SF tools in R+ Dietitian showed high accuracy, sensitivity, and specificity (77.5%, 81.0%, and 76.7% and 69.3%, 84.5%, and 64.5%, respectively), and fair agreement ( $\kappa=0.42$  and  $0.37$ , respectively; CC  $0.62$  and  $0.56$ , respectively) with the NRS-2002 and PG-SGA-SF tools administered by dietitians. The estimated intakes of dietary energy and protein were significantly higher ( $P<.001$  for both) in R+ Dietitian (mean difference of energy intake:  $144.2$  kcal, SD  $454.8$ ; median difference of protein intake:  $10.7$  g, IQR  $9.5$ - $39.8$ ), and showed fair agreement (CC  $0.59$  and  $0.47$ , respectively), compared with 3d-24HRs performed by dietitians.

**Conclusions:** The identified nutritional risk and assessment of dietary intakes of energy and protein in R+ Dietitian displayed a fair agreement with the screening and assessment conducted by dietitians. R+ Dietitian has the potential to be a tool for nutritional risk screening and dietary intake assessment among hospitalized patients with cancer.

**Trial Registration:** Chinese Clinical Trial Registry ChiCTR1900026324; <https://www.chictr.org.cn/showprojen.aspx?proj=41528>

(*JMIR Form Res* 2022;6(10):e40316) doi:[10.2196/40316](https://doi.org/10.2196/40316)

## KEYWORDS

digital tool; nutritional risk screening; dietary assessment; validity; cancer patients

## Introduction

Cancer has become a primary public health problem in China over the past several decades [1]. Malnutrition is a common and severe problem in patients with cancer that directly increases the incidence of complications and significantly deteriorates their quality of life [2,3]. Hospitalized malnourished patients with cancer comprise 30%-80% of all patients with cancer [4]. Nausea, vomiting, and diarrhea induced by chemotherapy and radiotherapy can aggravate the nutritional status of patients with cancer [5,6], but nutritional support can improve their clinical outcomes [7-9]. Nutritional risk screening and dietary assessment are critical because they are the basis for providing personalized nutritional support [10].

Guidelines from the Chinese Society for Parenteral and Enteral Nutrition (CSPEN) [11], the American Society for Parenteral and Enteral Nutrition (ASPEN) [12], and the European Society for Clinical Nutrition and Metabolism (ESPEN) [13] recommended that hospitalized patients with cancer should be screened for nutritional risk at admission to enable timely recognition and treatment of nutritional derangements. For those who are classified as having nutritional risk, further objective and quantitative assessment of nutritional intake should be undertaken, especially their energy and protein intakes, which are 2 of the most important nutrients for patients with cancer [14-16].

However, it is difficult to screen all hospitalized patients for nutritional risk in hospitals in China, not to mention assessment of nutritional intake, which is more complex. On the one hand, there is a shortage of nutritional specialists in hospitals in China. Even some tertiary hospitals, which usually employ most medical specialists and provide full component of medical services, do not have departments of nutrition. On the other hand, clinicians and nurses are often under pressure to perform a variety of tasks. Nutritional risk screening and dietary assessment can significantly increase the burden of clinicians and nurses and so they do not perform these tasks routinely. In the coming years, there will still be a shortage of nutritional specialists in hospitals in China, meaning clinicians and nurses will continue to work in a busy clinical environment, but importantly patients' need for nutritional care should not be neglected. Thus, there is an extremely urgent need for more effective, less time-consuming, and less people-demanding tools.

Modern advancements in digital technologies provide a feasible solution for this problem. For example, a computer-based electronic version of the Malnutrition Universal Screening Tool (MUST), which is used for outpatient self-screening, displayed high validity for nutritional risk self-screening. A dietary assessment app (MyFood) showed good ability to estimate

dietary intake [17] and another electronic system improved the documentation of nutritional intake, treatment, and nutritional care plans in hospitalized patients [18]. Several other digital tools based on computers or smartphones were also proved to enhance the efficacy in identifying patients at nutritional risk and assessing dietary intake [19-24]. However, to the best of our knowledge, no study has been conducted on developing a digital smartphone-based self-administered tool for nutritional risk screening and dietary assessment for hospitalized patients with cancer.

We developed a mini program, R+ Dietitian, for smartphones as a support system for nutritional risk screening and dietary assessment for hospitalized patients with cancer. It is a self-administered tool and requires patients to input data about their disease, weight, dietary intake, among other variables. Further nutritional assessment and individualized nutritional care plan would be customized based on the data entered. Therefore, this mini program should be validated to ensure clinicians can provide appropriate nutritional care for patients.

The aims of this study were (1) to develop a digital smartphone-based self-administered mini program for nutritional risk screening and dietary assessment of hospitalized patients with cancer, (2) to evaluate the validity of the mini program for patients' self-screening nutritional risk compared with dietitians' professional screening, and (3) to evaluate the validity of the mini program for estimating dietary energy and protein intakes compared with dietitians' professional estimation using 3 days of 24-hour dietary recalls (3d-24HRs).

## Methods

### Development of the R+ Dietitian Mini Program

#### Overview

R+ Dietitian was developed by dietitians, developers, and interaction designers at Recovery Plus Inc. (R+). Clinicians and nurses at the Oncology Department of Sichuan People's Hospital were involved in the literary design process.

The initial paper draft, including content and algorithm, was designed by dietitians at R+. Clinicians and nurses at the Oncology Department of Sichuan People's Hospital reviewed the draft and then gave feedback to dietitians at R+. The content and language of the mini program were then modified before starting the technical development. Next, the beta version of R+ Dietitian was developed and then tested by 4 dietitians, 2 nurses, and 10 patients with cancer. Their feedback was used for further modification prior to commencing this study.

Given the popularity of WeChat, which is China's most popular messaging app with a monthly user base of more than 1 billion people [25], we developed R+ Dietitian as a subapp within the

WeChat ecosystem. This was intended to save users' time, as this avoids the need to additionally download a new app on their cell phones. Instead, users just need to open their WeChat profile to run R+ Dietitian quickly and smoothly.

All contents displayed in R+ Dietitian were in Chinese. R+ Dietitian consists of the following 3 parts.

**Figure 1.** Collection of basic information of patients.

Field	Value
Sex	Woman
Age	28
Height	158 cm
Weight	55 kg

[Next](#)

## Part 2: Nutritional Risk Screening

### Design

This part was designed based on the Nutritional Risk Screening (NRS-2002) tool and the Patient-Generated Subjective Global Assessment Short Form (PG-SGA-SF) tool. The NRS-2002 tool is recommended by the CSPEN to screen for nutritional risk in the hospital settings [11], whereas the PG-SGA-SF tool is usually used to screen for nutritional risk in patients with cancer [26-28]. Therefore, the estimated NRS-2002 and PG-SGA-SF scores would be given to participants automatically and separately, while their nutritional risk was separately assessed based on these 2 scores.

### Design and Nutritional Risk Score Estimation of NRS-2002 in R+ Dietitian

The NRS-2002 tool includes 2 components, the initial screening and the final screening. The initial screening has 4 questions. If the answer is "Yes" to any question on the initial screening, the final screening is performed [29]. In China, the initial screening is skipped and only the final screening is performed according to the standard practice [30]. Therefore, the NRS-2002 in R+ Dietitian was aligned with the final screening of the NRS-2002, which consists of 3 sections: impaired nutritional status (scores 0-3), severity of the disease (scores 0-3), and age (scores 0-1), with a total score of 7. Age was already recorded by patients in part 1.

## Part 1: Collection of Basic Information of Patients

In the first 2 interfaces of R+ Dietitian (Figure 1), patients recorded their basic information, including name, inpatient admission number, bed number, age (in years), gender, height (in centimeters), and weight (in kilograms).

Severity of the disease was evaluated by the question "Please choose the disease you have (multiple choice)" with a disease list covering all types of cancer, common chronic disease, and other diseases that are displayed in the NRS-2002.

Impaired nutritional status involved 3 indicators: degree of weight loss, BMI, and degree of food intake reduction. BMI of patients was calculated as patients' weight in kilogram (kg) divided by the square of height in meters (m). Degree of weight loss was evaluated by the question (Figure 2) "Has your weight changed over the past three months?" with 3 options, "slight weight gain or substantially unchanged (weight loss less than 5%)," "slight weight loss (weight loss of 5%-15%)," and "severe weight loss (weight loss exceeding 15%)." The exact values of 5% and 15% weight loss would be calculated automatically through the algorithm set in the back end of R+ Dietitian based on the weight recorded in part 1. The calculated weight loss will be displayed within parentheses next to each option so that patients could clearly choose the right option. For example, if a patient recorded the weight as 58 kg, then options such as "slight weight gain or substantially unchanged (weight loss less than 5 kg)," "slight weight loss (weight loss of 5-17 kg)," and "severe weight loss (weight loss exceeding 17 kg)." would be presented. If patients chose 1 of the latter 2 options, an additional question was presented, namely, "Weight loss in" with 3 options, "3 months," "2 months," and "1 month." Degree of food intake reduction was evaluated by the question "Has there been a

reduction in your food intake recently?” with 4 options, “substantially unchanged,” “slight reduction (reduction of about 25%-50%),” “moderate reduction (reduction of about 50%-75%),” “severe reduction (reduction of about 75%-100%).”

Based on age, severity of the disease, and impaired nutritional status reported by patients, the estimated NRS-2002 score was automatically calculated through the algorithm set in the back end of R+ Dietitian and presented in the results interface. The algorithm was consistent with the scoring rules of the NRS-2002 [29].

**Figure 2.** Evaluation for degree of weight loss.

The screenshot shows a mobile application interface for 'R+ Dietitian'. The main question is 'Has your weight changed over the past three months?'. There are three selectable options, each with an emoji icon: a yellow smiley face for 'slightly weight gain or substantially unchanged', a yellow sad face for 'slightly weight loss (2.5-8 kg)', and a red sad face for 'severe weight loss (over 8 kg)'. The second option is highlighted in green. Below the options, there is a question 'The weight loss was in \_\_\_\_.' followed by three radio button options: '3 months', '2 months', and '1 month'. At the bottom of the screen, there are two buttons: a green 'Back' button and a grey 'Next' button.

### Design and Nutritional Risk Score Estimation of the PG-SGA-SF Tool in R+ Dietitian

The PG-SGA-SF tool comprises the first 4 options of the Patient-Generated Subjective Global Assessment (PG-SGA), including weight history, food intake, symptoms, and activities and function, and is designed to be used by patients independently [31]. Studies have shown that inclusion of the first 4 options did not add further value as a screening tool compared with inclusion of the first 3 options [32]. Therefore, the PG-SGS-SF tool in R+ Dietitian only evaluates weight history, food intake, and symptoms.

Weight history was already evaluated by the question “Has your weight changed over the past three months?,” which was also used to evaluate the weight loss in the NRS-2002 tool. Food intake in the PG-SGS-SF tool was evaluated by the question “Has there been a reduction in your food intake recently?,” which was also used to evaluate the degree of food intake reduction in the NRS-2002 tool. Food intake was additionally evaluated by the question “What does your dietary look like recently?,” with the following 4 options: “fasting,” “liquid diet,” “soft food,” and “normal food.” Besides, a picture showing what the food looks like was presented next to the latter 3 options to help patients make choice. Symptoms were evaluated by the question “Have you had a bad appetite lately?” with 3 options, “never,” “occasionally,” and “frequently.”

Based on weight history, food intake, and symptoms reported by patients, the estimated PG-SGA-SF score was automatically calculated through the algorithm set in the back end of R+ Dietitian and presented in the results interface. The algorithm was consistent with the scoring rules of the PG-SGS-SF tool.

### Part 3: Dietary Energy and Protein Assessment

The 24-hour dietary recall (24HR) is a traditional method for assessing dietary nutrient intake and is widely used in nutrition research [33-35]. However, it requires patients to recall any food, beverages, and water that they consumed, as well as extensive training of the interviewer. These make it burdensome and time-consuming, limiting its utility in clinical practice. For describing the mean usual intake of target nutrients, a “short dietary screener” is feasible. Blalock and colleagues [36] developed a short instrument for assessing dietary intakes of calcium and vitamin D, which only includes 22 foods and beverages that are rich in these 2 nutrients. Similar short tools for assessing dietary intake of cholesterol [37] and saturated fat are also found in the literature [38]. In this mini program, we only wanted to estimate the dietary energy and protein intakes in patients, so, as for a quick assessment, patients’ recalling all the food they consumed is not necessary because some foods contribute less to the energy and protein intakes. Therefore, we shortened the 24HR method to make it suitable for self-reporting and easier to complete, thereby allowing the daily energy and protein intakes of patients to be estimated quickly.

The China Health and Nutrition Surveys (CHNS) revealed that the main sources of dietary energy and protein among Chinese residents were cereals and animal foods [39-41]. Rice is the primary cereal type in China and milk is an important source of protein. Therefore, we designed 3 questions to assess dietary energy and protein intakes, “How much rice have you recently consumed on a daily basis?,” “How much meat have you recently consumed on a daily basis?,” and “How much milk have you recently consumed on a daily basis?.” An adjustable ruler (unit: gram or milliliter) presented under the 3 questions allowed the patients to report the amount of rice/meat/milk they consumed.

Before this study began, we conducted a pilot nutrition survey using 3 days of 24HR data (3d-24HRs) among hospitalized patients. Data on foods and nutrients derived from the survey were used to create the algorithm for calculating energy and protein intakes in R+ Dietitian. Based on the rice/meat/milk consumption reported, the estimated dietary energy and protein intakes were automatically calculated through the algorithm set in the back end of R+ Dietitian. Besides, the patients' requirements for energy and protein were automatically estimated by the algorithm, which was based on the guideline from the CSPEN [11]. The difference between the estimation of dietary intakes of energy and protein and patients' requirements was calculated automatically and was presented on the results screen.

## Validity Evaluation of R+ Dietitian

### *Study Design, Setting, and Participants*

This was a prospective diagnostic accuracy study conducted at the Oncology Department of Sichuan People's Hospital, Chengdu, China, from March 2021 to April 2021. Eligible patients were adults aged 18-80 years with pathologically confirmed tumors who were able to communicate normally. We excluded patients with mental or psychological disorders, those with incomplete data, and those who were unwilling or unable to provide written informed consent. All patients registering to the Oncology Department in March 2021 and April 2021 were evaluated for study eligibility by a researcher, and consequently, this was a convenience sample of oncology patients.

### *Ethics Approval*

This study was performed according to the Declaration of Helsinki and was approved by the Medical Ethics Committee of Sichuan Provincial People's Hospital (2019/243). Written informed consent was obtained from all participants.

### *Test Methods*

#### **Index Test**

Two dietitians (HX and QZ) were trained to use the R+ Dietitian program before the study started so that they can assist participants in using R+ Dietitian when needed. A clinician in the Oncology Department identified patients that met the inclusion and exclusion criteria and then informed 1 of the 2 dietitians about their eligibility.

On the day of hospital admission, eligible participants were asked by a dietitian (HX or QZ) to use R+ Dietitian for nutritional risk screening and dietary assessment of energy and protein. Participants or their family members used smartphones to open the WeChat app and then scanned a QR code to run the R+ Dietitian mini program. Registering in R+ Dietitian was not mandatory, so participants can use the mini program directly. The 3 parts of R+ Dietitian were completed by participants one by one. Estimations of the NRS-2002 score, the PG-SGA-SF score, and dietary energy and protein intakes were immediately presented on the results interface of R+ Dietitian once participants completed the program.

#### **Reference Test**

We used face-to-face interviews as the reference method. On the day of hospital admission, after patients self-evaluated R+ Dietitian, paper-based NRS-2002, PG-SGA-SF, and 3d-24HRs tools were administered according to the standard procedure by 2 trained dietitians (HX and QZ). 3d-24HRs collected data on participants' food intake in the 3 days before the day of hospital admission. Intake of all foods consumed by participants in the 3 days and corresponding cooking methods, including stir-fry, braising, stew, etc., were recorded. If the food was a composite dish, its composition and the corresponding proportion would be further asked and recorded. To ensure the precision of intake recalled by participants, dietitians showed pictures of standard cutlery to the participants to help them assess the intake when 3d-24HRs were performed.

### *Data Collection*

The baseline demographic and medical characteristics of the participants, including age, sex, means of paying medical costs, occupation, marital status, residence, education level, chronic diseases, cancer type, and family history of cancer, were obtained from the electronic medical records by a researcher (JZ).

Estimations of the NRS-2002 score, the PG-SGA-SF score, and dietary energy and protein intakes in R+ Dietitian were retrieved from the back end of R+ Dietitian by another researcher (CH). The NRS-2002 and PG-SGA-SF scores of the reference method were calculated according to the scoring rules of the 2 questionnaires and then entered into EpiData (EpiData Association) by the 2 dietitians (HX and QZ).

Dietary data obtained from 3d-24HRs were managed by the 2 dietitians. First, the individual food and the corresponding specific amount consumed by the participants each day were entered into MS Excel (Microsoft Corporation). Next, the raw weight of every individual food was calculated based on the raw-to-cook ratio of the food. Then, if the dish was cooked with oil, the approximate amount of the oil consumed was estimated based on the intake of the dish of the participants. Finally, the energy and protein intakes from each individual food were calculated based on China Food Composition Tables. The total energy and protein intakes each day were then calculated. The final estimated daily dietary energy and protein intakes was the mean of 3-day intake and was entered into EpiData.

### Data Assessment and Statistical Analysis

The NRS-2002 scores of 3 or above identified patients at nutritional risk [29], whereas the PG-SGA-SF scores of 4 or above identified patients at nutritional risk [27,42]. The difference in dietary energy and protein intakes estimated through R+ Dietitian from those measured by dietitians was calculated.

All statistical analyses were performed using SAS 9.4 software (SAS Institute Inc.). First, continuous variables were analyzed for normality using the Kolmogorov–Smirnov test. Data were then described as mean (SD) or median (IQR). The Student *t* test or the rank sum test was applied accordingly. Categorical variables were described as frequencies and percentages, and then the  $\chi^2$  test was applied. Accuracy, sensitivity, specificity, positive predictive value, negative predictive value (NPV), and  $\kappa$  values of the NRS-2002 and PG-SGA-SF tools in R+ Dietitian were calculated using the McNemar test for correlated proportions, respectively. The consistency of R+ Dietitian for dietary assessment with 3d-24HRs was tested using the Pearson or Spearman correlation test. All tests were 2-sided, and a significant level of 5% (*P* .05) was applied.

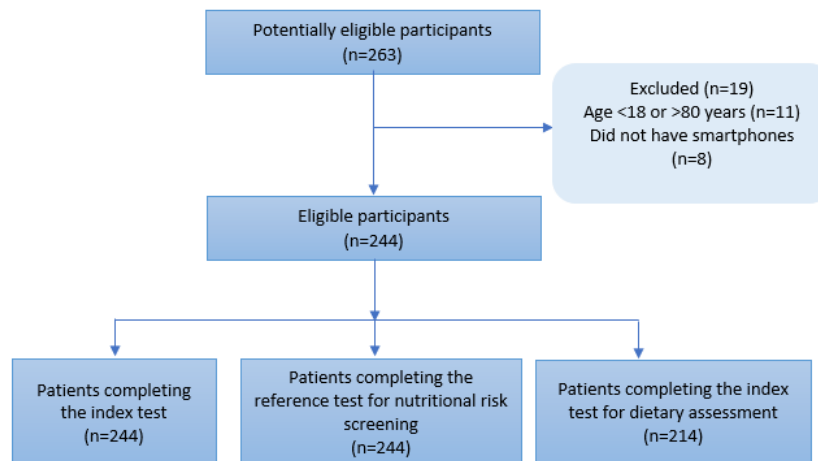
The validity of R+ Dietitian was evaluated using the cutoffs recommended by van Bokhorst–de van der Schueren et al [43], which were based on sensitivity and specificity. “Good” indicated both sensitivity and specificity exceeding 80%; “fair” indicated sensitivity or specificity exceeding 50% but lower than 80%; and “poor” indicated sensitivity or specificity lower than 50%. For other validation indices, the cutoffs of the correlation coefficients (CCs) followed Guilford’s [44] description as follows: “good,” CC  $\geq 0.75$ ; “fair,” CC  $\geq 0.4$  and  $< 0.75$ ; and “poor,” CC  $< 0.4$ . The cutoffs for good, fair, and poor  $\kappa$  values [45] were  $\geq 0.6$ ,  $\geq 0.4$  and  $< 0.6$ , and  $< 0.4$ , respectively.

## Results

### Participants

Figure 3 illustrates the flow of participants through the study. From March 2021 to April 2021, 263 patients were assessed for eligibility and were invited to participate in this study. Patients who were excluded or who withdrew from this study and the corresponding reasons were noted. Overall, 244 patients were recruited to evaluate the validity of R+ Dietitian for nutritional risk screening, and 214 patients were included to evaluate the validity of R+ Dietitian for dietary intake assessment.

**Figure 3.** Flow of participants.



### Characteristics of Participants

The baseline characteristics of the participants included in this study are presented in Table 1. The median (IQR) age of participants was 59 (51–68) years and the mean (SD) weight

and BMI were 58.6 (8.9) kg and 22.2 (2.9) kg/m<sup>2</sup>, respectively. Most participants were male (156/244, 63.9%). The most common diagnosis was gastrointestinal tumor (107/244, 43.9%) and the least common diagnosis was head and neck cancer (17/244, 6.9%).

**Table 1.** Baseline characteristics of participants (N=244).

Variables	Values
Age (years), median (IQR)	59 (51-68)
Weight (kg), mean (SD)	58.6 (8.9)
BMI (kg/m <sup>2</sup> ), mean (SD)	22.2 (2.9)
<b>Gender, n (%)</b>	
Male	156 (63.9)
Female	88 (36.1)
<b>Means of paying medical costs, n (%)</b>	
Social security	230 (94.3)
Self-paid	14 (5.7)
<b>Occupation, n (%)</b>	
Famer/worker	27 (11.1)
Retirement	45 (18.4)
Other	172 (70.5)
<b>Marital status, n (%)</b>	
Married	241 (98.8)
Unmarried	3 (1.2)
<b>Residence, n (%)</b>	
Rural	130 (53.3)
City	114 (46.7)
<b>Education, n (%)</b>	
Primary	28 (11.5)
Secondary	197 (80.7)
Senior	19 (7.8)
<b>Hypertension, n (%)</b>	
Yes	40 (16.4)
No	204 (83.6)
<b>Diabetes mellitus, n (%)</b>	
Yes	22 (9)
No	222 (91)
<b>Cancer family history, n (%)</b>	
Yes	16 (6.6)
No	228 (93.4)
<b>Cancer type, n (%)</b>	
Head and neck tumor	17 (6.9)
Gastrointestinal tumor	107 (43.9)
Respiratory tumor	76 (31.1)
Other	44 (18)

### Agreement Between the NRS-2002 Tool in R+ Dietitian and the NRS-2002 Tool Administered by Dietitians

Table 2 presents the screening results of the NRS-2002 tool in R+ Dietitian in comparison to dietitians' screening. According

to the NRS-2002 tool in R+ Dietitian, 33.2% (81/244) of patients were at nutritional risk, which was significantly higher than in the screening by dietitians (42/244, 17.2%;  $P < .001$ ; McNemar test). The NRS-2002 tool in R+ Dietitian was in fair agreement with the NRS-2002 tool administered by dietitians (Table 3;

both sensitivity and specificity >50%,  $\kappa > 0.42$ , and  $CC > 0.4$ ). In addition, self-screening by patients using the NRS-2002 tool in R+ Dietitian had a high NPV (95.1%), which indicates that self-screening by patients using the NRS-2002 in R+ Dietitian can strongly predict those that are not at nutritional risk.

**Table 2.** Cross tabulation of nutritional risk according to the NRS-2002<sup>a</sup> tool in R+ Dietitian and the NRS-2002 tool administered by dietitians.

R+ Dietitian (patients' self-screening)	Dietitians' screening		Total (N=244)
	Positive (n=42)	Negative (n=202)	
Positive, n (%)	34 (13.9)	47 (19.3)	81 (33.2)
Negative, n (%)	8 (3.3)	155 (63.5)	163 (66.8)
Total (n=244)	42 (17.2)	202 (82.8)	244 (100)

<sup>a</sup>Nutritional Risk Screening.

**Table 3.** Sensitivity, specificity, positive predictive value, negative predictive value,  $\kappa$  value, and correlation coefficient of patients' self-screening using the NRS-2002 tool in R+ Dietitian.

Index	NRS-2002 <sup>a</sup> in R+ Dietitian (patients' self-screening), % (95% CI)	P value
Accuracy	77.5 (71.7-82.5)	N/A <sup>b</sup>
Sensitivity	81.0 (65.4-90.9)	N/A
Specificity	76.7 (70.2-82.2)	N/A
Positive predictive value	42.0 (31.3-53.5)	N/A
Negative predictive value	95.1 (90.2-97.7)	N/A
$\kappa$ value	0.42 (0.30-0.54)	N/A
Correlation coefficient	0.62 (0.54-0.70)	<.001

<sup>a</sup>Nutritional Risk Screening.

<sup>b</sup>Not applicable.

### Agreement Between the PG-SGA-SF Tools in R+ Dietitian and the PG-SGA-SF Tool Administered by Dietitians

Table 4 shows the screening results of the PG-SGS-SF tool in R+ Dietitian in comparison to dietitians' screening. According to the PG-SGS-SF tool in R+ Dietitian, 47.1% (115/244) of

patients were at nutritional risk, which was significantly higher than in the screening by dietitians (58/244, 23.8%;  $P < .001$ ) according to the McNemar test. Further, the PG-SGS-SF tool in R+ Dietitian was fairly in agreement with the NRS-2002 tool administered by dietitians (Table 5; both sensitivity and specificity >50% and  $CC > 0.4$ ). Besides, as with the NRS-2002 tool in R+ Dietitian, patients' self-screening using the PG-SGS-SF tool in R+ Dietitian had a high NPV (95.1%).

**Table 4.** Cross tabulation of nutritional risk according to the PG-SGA-SF<sup>a</sup> tool in R+ Dietitian and the PG-SGS-SF tool administered by dietitians.

R+ Dietitian (patients' self-screening)	Dietitians' screening, n (%)		Total (N=244)
	Positive (n=58)	Negative (n=186)	
Positive, n (%)	49 (20.1)	66 (27.0)	115 (47.1)
Negative, n (%)	9 (3.7)	120 (49.2)	129 (52.9)
Total (n=244)	58 (23.8)	186 (76.2)	244 (100)

<sup>a</sup>Patient-Generated Subjective Global Assessment Short Form.



**Table 5.** Sensitivity, specificity, positive predictive value, negative predictive value,  $\kappa$  value, and correlation coefficient of patients' self-screening using the PG-SGA-SF<sup>a</sup> tool in R+ Dietitian.

Index	PG-SGA-SF in R+ Dietitian (patients' self-screening), % (95% CI)	P value
Accuracy	69.3 (63.1-75.0)	N/A <sup>b</sup>
Sensitivity	84.5 (72.1-92.2)	N/A
Specificity	64.5 (57.1-71.3)	N/A
Positive predictive value	42.6 (33.5-52.2)	N/A
Negative predictive value	93.0 (86.8-96.6)	N/A
$\kappa$ value	0.37 (0.26-0.47)	N/A
Correlation coefficient	0.56 (0.47-0.64)	<.001

<sup>a</sup>Patient-Generated Subjective Global Assessment Short Form.

<sup>b</sup>Not applicable.

### Agreement Between Dietary Energy and Protein Intake Assessment in R+ Dietitian and Assessment Performed by Dietitians

The dietary energy and protein intakes estimated by R+ Dietitian and dietitians are presented in Table 6. Both the estimated energy and protein intakes were significantly higher ( $P<.001$

for both) in R+ Dietitian, compared with the 3d-24HRs. Estimations of energy and protein intakes in R+ Dietitian were 10.7% and 29.0% higher than those in 3d-24HRs, respectively (Table 7). Nevertheless, R+ Dietitian was still in moderate agreement with 3d-24HRs for both energy and protein intake assessment (Table 7; both CC >0.4).

**Table 6.** Estimation of energy and protein intakes in R+ Dietitian compared with the 3d-24HRs administered by dietitians.

Dietary intake	R+ Dietitian	3d-24HRs <sup>a</sup>	P value
Energy (kcal), mean (SD)	1578.3 (468.4)	1434.1 (528.8)	<.001
Protein (g), median (IQR)	79.0 (62.7-95.3)	61.7 (43.0-82.8)	<.001

<sup>a</sup>Three days of 24-hour dietary recall.

**Table 7.** Correlation coefficient and absolute and relative differences of dietary energy and protein intake estimation in R+ Dietitian against 3d-24HRs<sup>a</sup>.

Index	Energy	P value	Protein	P value
Correlation coefficient (95% CI)	0.59 (0.49-0.67)	<.001	0.47 (0.36-0.57)	<.001
Absolute difference (kcal or g), mean (SD)	144.2 (454.8)	<.001	14.7 (29.2)	<.001
Relative difference (%), median (IQR)	10.7 (9.5-39.8)	<.001	29 (3.1-68.1)	<.001

<sup>a</sup>Three days of 24-hour dietary recall.

## Discussion

### Principal Findings

The R+ Dietitian mini program is developed for use among hospitalized patients with cancer. This prospective study demonstrates the validity of R+ Dietitian for nutritional risk screening and dietary energy and protein intake assessment in oncology department. Overall, the NRS-2002, PG-SGA-SF, and dietary energy and protein intake assessments in R+ Dietitian were in fair agreement with those performed by trained dietitians, although the number of patients identified at nutritional risk was more and the estimated dietary energy and protein intakes were higher in R+ Dietitian, compared with the reference methods. Both the NRS-2002 and the PG-SGA-SF tools in R+ Dietitian showed an excellent ability to predict those who were not at nutritional risk.

To the best of our knowledge, no similar mini program has been developed and this is the first prospective study to examine the validity of a digital smartphone-based self-administered tool for nutritional risk screening and dietary assessment in the oncology department.

### The Validity of R+ Dietitian's Nutritional Risk Screening

For nutritional risk screening in hospitals, the NRS-2002 is recommended by the CSPEN [11] and the PG-SGA-SF is widely used among patients with cancer [26-28]. Therefore, we developed nutritional risk screening in R+ Dietitian based on both the NRS-2002 and the PG-SGA-SF. Our study shows that nutritional risk evaluated by the NRS-2002 and PG-SGA-SF tools in R+ Dietitian had fair validity.

No similar smartphone-based app for nutritional risk self-screening has been developed, although a computer-based self-screening app is available [21,46]. This app is the electronic

version of the MUST and was developed for hospital outpatients. It showed high agreement with the MUST administered by health care professionals ( $\kappa=0.74-1.00$ ). These values are higher than the NRS-2002 and PG-SGA-SF tools in R+ Dietitian. The different degrees of agreement between the electronic version of the MUST and R+ Dietitian can be explained as follows: First, in the study of the electronic version of the MUST, the weight and height were measured using a weighing scale and a portable height measure, respectively, in the outpatients' waiting room by patients. Consequently, the weight and height reported by patients are more likely to be objective than those in R+ Dietitian. Both dietitians in our study and health care professionals in the study of the electronic version of the MUST used weight and height scales to measure the weight and height, respectively, of participants. However, there was a higher deviation in the weight and height reported by patients and professionals in our study, compared with those in the study of the electronic version of the MUST. BMI calculated based on weight and height is a factor for nutritional risk scoring in both R+ Dietitian and the electronic version of the MUST. Weight was also used to calculate the degree of weight loss in the past 3-6 months, which is also a criterion for evaluating nutritional risk. Hence, the different methods applied to measure weight and height by participants may be one of the reasons for the difference in agreement between these 2 studies. However, bias induced from recall is common in surveys [47-51] and can hardly be avoided. Second, the NRS-2002 tool in R+ Dietitian additionally evaluates the degree of food intake reduction using the question, "Has there been a reduction in your food intake recently?" with 3 options, "slight reduction (reduction of about 25%-50%)," "moderate reduction (reduction of about 50%-75%)," and "severe reduction (reduction of about 75%-100%)." For this question, some patients may compare their recent food intake with food intake several weeks ago rather than their normal requirement of food intake, whereas dietitians compared patients' recent food intake with their normal requirement. Third, the PG-SGS-SF tool in R+ Dietitian did not evaluate the symptoms affecting patients' nutritional intake, but dietitians did. The electronic version of the MUST was developed for hospital outpatients and requires patients to use the weight scale and the portable height measure tool, which is less feasible for inpatients, as they are busy with completing their admission procedures while being admitted to the hospital. In addition, the electronic version of the MUST was computer based, which makes it less flexible compared with R+ Dietitian, which is a smartphone-based app for self-screening.

Nutritional risk screening in R+ Dietitian showed a fair agreement with the dietitians' screening. As the first smartphone-based tool for nutritional risk self-screening, it displayed potential ability to be used in clinical practice for nutritional risk screening.

### The Validity of R+ Dietitian's Dietary Energy and Protein Intake Assessment

The energy intake was overestimated by 144 kcal/day in R+ Dietitian compared with the estimation from 3d-24HRs performed by dietitians. Our finding is different from other self-administered dietary assessment apps. Compared with 24HRs or the Food Frequency Questionnaire (FFQ), most of

these apps underestimated energy intake from -8 to -466 kcal/day [24,52-60], and 1 app overestimated energy intake by 55 kcal/day [17]. The absolute mean difference of estimated energy intake between R+ Dietitian and 3d-24HRs was higher in our study compared with the absolute difference in 5 studies (from 8.1 to 101 kcal/day) [24,55-57,60] but was lower compared with the absolute difference reported in 6 studies (from 145.1 to 466.0 kcal/day) [22,52-54,58,59]. The following might have contributed to the difference between R+ Dietitian and other apps. First, all the other validity studies required participants to use the experimental app to record their food intake for several days [17,23,24,53,54,56,57,59,60] or even few months [52,58]. Based on the data inputted, the daily energy intake and the mean energy intake over the study period would be calculated automatically by the app. By contrast, in our study, patients self-reported their mean food intake but the mean energy intake was calculated automatically by R+ Dietitian. The different methods of collecting data on food intake between R+ Dietitian and other tools may result in differences in the energy intake estimated. Second, other apps recorded all the foods consumed by participants, whereas only 3 food groups were reported by patients in R+ Dietitian. R+ Dietitian was developed to be a digital tool that can quickly assess energy and protein intakes of patients, thereby allowing clinicians, who generally have limited time to analyze these, to have an approximate overview of the intake level of the 2 nutrients (ie, energy and protein). Further and more complex and detailed dietary assessment is needed if patients are at nutritional risk. Third, patients in most validity studies for self-administered dietary assessment apps were young. For example, their sample population was recruited in high school [52] or in a university setting [24,53,55-57,59]. The median age of participants in our study was 59 years, which is much higher than that in these studies. Young and old people may have different tendencies in self-reporting studies. For the 2 studies that recruited old people or hospitalized patients [17,60], the absolute mean difference in estimated energy intake between the experimental app and the reference method was lower (55 kcal/day and 101 kcal/day, respectively), but in another study [58] this was higher (408.8 kcal/day) than that in our study (144 kcal/day).

The estimation of protein intake in R+ Dietitian was higher than that from 3d-24HRs by 14.7 g/day. Mescoloto and colleagues [61] and Bucher Della Torre and coworkers [60] also reported overestimation of protein intake in 2 digital dietary recording apps by 2.1 and 2.0 g/day, respectively, compared with measured energy expenditure and 2 unannounced 24-h phone dietary recalls on overlapping days. In contrast to our study and these 2 studies [60,61], 7 studies found an underestimation of protein intake in phone-based dietary recording apps from -2.6 to -28 g/day [17,24,52,53,55,56,58], compared with 24HRs or FFQs or paper-based food records. The absolute difference between the experimental app and the reference method in our study was higher than that in 4 studies (from -2.6 to -10.5 g/day) [17,24,55,56], but was lower than that in 3 other studies (from -21.5 to 28.0 g/day) [52,53,58]. For the estimation of protein intake, just like the estimation of energy intake, other studies asked participants to use the dietary recording apps to record all the foods they consumed during the study period in real time, whereas we asked participants in our study to recall their intake

of 3 recent food groups for a quick estimation during their hospital admission. In general, real-time recording may be more precise than recall in dietary assessment [62]. But still, the difference in the estimation of protein intake between R+ Dietitian and 3d-24HRs was lower than that between 3 dietary recording apps and 24HRs or paper-based food records [52,53,55].

Nowadays, similar to traditional dietary intake assessment tools, including 24HRs, FFQs, and weighed or nonweighed food records, all smartphone-based digital dietary intake assessing apps require users to report on all the foods or food groups they consume [17,22-24,52-60,63]. However, this approach may not be suitable for all hospitalized patients. On the one hand, these apps require users to record their food intake for several days, or even few months, which cannot help clinicians to understand their patients' nutritional intake within the short period available for initial assessment. The only 1 study that recruited hospitalized patients [17] asked them to use the experimental app to record food intake for 2 days; however, MyFood was developed for inpatients with nutritional risk rather than all hospitalized patients, which may limit its usage in other patient groups. In this case, the nutritional intake of other nonnutritional-risk patients can be missed by the clinicians. On the other hand, recording all foods or food groups is also time-consuming and burdensome for patients, so some may forget to record their food intake at times. R+ Dietitian was used on the day of hospital admission and only includes 3 food groups, reducing patients' burden and helping clinicians to have a quick and approximate overview of the patients' energy and protein intakes. If the patients are diagnosed to have nutritional risk, further comprehensive dietary assessments are needed.

### Strengths and Limitations

To our knowledge, R+ Dietitian is the first digital smartphone-based self-administered tool for both nutritional risk screening and dietary assessment for hospitalized patients with cancer. This is a strong strength of this study. Another strength is that R+ Dietitian was developed based on the NRS-2002 and PG-SGA-SF tools. The NRS-2002 is the only tool validated by a retrospective analysis of 128 randomized controlled trials [29] and is recommended by several professional communities [11,64]. The PG-SGS-SF is specific for patients with cancer and has been proved to be an effective tool to screen nutritional risk for patients with cancer [28,31,42,65,66]. Experienced dietitians, nurses, and patients were involved in the development of R+ Dietitian, which is also an important strength. R+ Dietitian was developed as a subapp

within the WeChat ecosystem, which is China's most popular messaging app with a monthly user base of more than 1 billion people [25]. This makes it commonly available and easily accessible. All participants in our study had installed WeChat on their smartphones and thus the need to download a new app was not required, thereby saving their time.

A limitation of this study is that we only recruited hospitalized patients with cancer in 1 hospital, which may limit the generalizability of our findings. However, R+ Dietitian was developed based on the NRS-2002 tool, which has been used among various types of patients. Hence, we propose that R+ Dietitian can also be used for other types of patients, but this needs further research. Besides, usability is one of the important factors determining the tool's actual usefulness in practical settings [67]. The usability of R+ Dietitian was not evaluated in this study. The 2 dietitians responsible for asking the patients to use R+ Dietitian reported that none of the patients reported difficulties while using the tool.

### R+ Dietitian's Potential for Nutritional Risk Screening and Dietary Intake Assessment Among Hospitalized Patients With Cancer

Based on the findings in this study, we propose that R+ Dietitian has a large potential to be a tool for nutritional risk screening and dietary intake assessment among hospitalized patients with cancer. R+ Dietitian may provide support for nurses and clinicians to perform nutritional risk screening and dietary assessment among hospitalized patients with cancer, enhancing the rate and efficiency of nutritional risk screening and dietary assessment among this patient group. In addition, R+ Dietitian is a WeChat-based tool, which makes it commonly available and may potentially increase its use among hospitalized patients. Further validity of this study for other types of patients may be helpful to expand the underlying use of R+ Dietitian in hospital settings.

### Conclusions

We have developed a digital smartphone-based self-administered instrument for nutritional risk screening and dietary assessment among hospitalized patients with cancer. The instrument enables the evaluation of estimated dietary intake of energy and protein against individual's requirements. The identified nutritional risk and assessment of dietary energy and protein intakes in R+ Dietitian displayed a fair agreement with the screening and assessment conducted by dietitians. R+ Dietitian has the potential to be a tool for nutritional risk screening and dietary intake assessment among hospitalized patients with cancer.

### Acknowledgments

We acknowledge all study participants for supporting this study. The authors thank Enago for the English language review. This study was funded by Recovery Plus Inc., Chengdu, China.

### Authors' Contributions

ZL contributed to the study design, data collection and analysis, and wrote the manuscript. SH contributed to the study design and patient recruitment. JZ contributed to patient recruitment and data collection. JY contributed to data collection and read and edited the manuscript. DZ contributed to data collection. HX and QZ contributed to calculation of dietary nutrients. HCS and

HH contributed to data collection. CH and KX contributed to the study design, data analysis, and manuscript preparation and review.

### Conflicts of Interest

This study was funded by Recovery Plus Inc., Chengdu, China, and the R+ Dietitian mini program used in this study was provided by this company. ZL, JY, HX, QZ, HCS, and HH are employees of Recovery Plus Inc.

### References

1. Cao W, Chen H, Yu Y, Li N, Chen W. Changing profiles of cancer burden worldwide and in China: a secondary analysis of the global cancer statistics 2020. *Chin Med J (Engl)* 2021 Mar 17;134(7):783-791 [FREE Full text] [doi: [10.1097/CM9.0000000000001474](https://doi.org/10.1097/CM9.0000000000001474)] [Medline: [33734139](https://pubmed.ncbi.nlm.nih.gov/33734139/)]
2. Clinton SK, Giovannucci EL, Hursting SD. The World Cancer Research Fund/American Institute for Cancer Research Third Expert Report on Diet, Nutrition, Physical Activity, and Cancer: Impact and Future Directions. *J Nutr* 2020 Apr 01;150(4):663-671 [FREE Full text] [doi: [10.1093/jn/nxz268](https://doi.org/10.1093/jn/nxz268)] [Medline: [31758189](https://pubmed.ncbi.nlm.nih.gov/31758189/)]
3. Rhondali W, Chisholm GB, Daneshmand M, Allo J, Kang D, Filbet M, et al. Association between body image dissatisfaction and weight loss among patients with advanced cancer and their caregivers: a preliminary report. *J Pain Symptom Manage* 2013 Jun;45(6):1039-1049 [FREE Full text] [doi: [10.1016/j.jpainsymman.2012.06.013](https://doi.org/10.1016/j.jpainsymman.2012.06.013)] [Medline: [23102755](https://pubmed.ncbi.nlm.nih.gov/23102755/)]
4. Zhang X, Pang L, Sharma SV, Li R, Nyitray AG, Edwards BJ. Malnutrition and overall survival in older patients with cancer. *Clin Nutr* 2021 Mar;40(3):966-977. [doi: [10.1016/j.clnu.2020.06.026](https://doi.org/10.1016/j.clnu.2020.06.026)] [Medline: [32665101](https://pubmed.ncbi.nlm.nih.gov/32665101/)]
5. Seo SH, Kim S, Kang Y, Ryoo B, Ryu M, Jeong JH, et al. Association of nutritional status-related indices and chemotherapy-induced adverse events in gastric cancer patients. *BMC Cancer* 2016 Nov 18;16(1):900 [FREE Full text] [doi: [10.1186/s12885-016-2934-5](https://doi.org/10.1186/s12885-016-2934-5)] [Medline: [27863481](https://pubmed.ncbi.nlm.nih.gov/27863481/)]
6. Sánchez-Lara K, Ugalde-Morales E, Motola-Kuba D, Green D. Gastrointestinal symptoms and weight loss in cancer patients receiving chemotherapy. *Br J Nutr* 2013 Mar 14;109(5):894-897. [doi: [10.1017/S0007114512002073](https://doi.org/10.1017/S0007114512002073)] [Medline: [22691288](https://pubmed.ncbi.nlm.nih.gov/22691288/)]
7. Tan S, Meng Q, Jiang Y, Zhuang Q, Xi Q, Xu J, et al. Impact of oral nutritional supplements in post-discharge patients at nutritional risk following colorectal cancer surgery: A randomised clinical trial. *Clin Nutr* 2021 Jan;40(1):47-53. [doi: [10.1016/j.clnu.2020.05.038](https://doi.org/10.1016/j.clnu.2020.05.038)] [Medline: [32563599](https://pubmed.ncbi.nlm.nih.gov/32563599/)]
8. Ukovic B, Porter J. Nutrition interventions to improve the appetite of adults undergoing cancer treatment: a systematic review. *Support Care Cancer* 2020 Oct;28(10):4575-4583. [doi: [10.1007/s00520-020-05475-0](https://doi.org/10.1007/s00520-020-05475-0)] [Medline: [32451701](https://pubmed.ncbi.nlm.nih.gov/32451701/)]
9. Capra S, Ferguson M, Ried K. Cancer: impact of nutrition intervention outcome--nutrition issues for patients. *Nutrition* 2001 Sep;17(9):769-772. [doi: [10.1016/s0899-9007\(01\)00632-3](https://doi.org/10.1016/s0899-9007(01)00632-3)] [Medline: [11527676](https://pubmed.ncbi.nlm.nih.gov/11527676/)]
10. Reber E, Schönerberger KA, Vasiloglou MF, Stanga Z. Nutritional Risk Screening in Cancer Patients: The First Step Toward Better Clinical Outcome. *Front Nutr* 2021;8:603936 [FREE Full text] [doi: [10.3389/fnut.2021.603936](https://doi.org/10.3389/fnut.2021.603936)] [Medline: [33898493](https://pubmed.ncbi.nlm.nih.gov/33898493/)]
11. Chinese Society for Parenteral/Enteral Nutrition. [Guidelines on nutritional support in patients with tumor]. *Zhonghua Wai Ke Za Zhi* 2017 Nov 01;55(11):801-829. [doi: [10.3760/cma.j.issn.0529-5815.2017.11.001](https://doi.org/10.3760/cma.j.issn.0529-5815.2017.11.001)] [Medline: [29136728](https://pubmed.ncbi.nlm.nih.gov/29136728/)]
12. Mueller C, Compher C, Ellen DM, American Society for Parenteral/Enteral Nutrition (A.S.P.E.N.) Board of Directors. A.S.P.E.N. clinical guidelines: Nutrition screening, assessment, and intervention in adults. *JPEN J Parenter Enteral Nutr* 2011 Jan;35(1):16-24. [doi: [10.1177/0148607110389335](https://doi.org/10.1177/0148607110389335)] [Medline: [21224430](https://pubmed.ncbi.nlm.nih.gov/21224430/)]
13. Arends J, Bachmann P, Baracos V, Barthelemy N, Bertz H, Bozzetti F, et al. ESPEN guidelines on nutrition in cancer patients. *Clin Nutr* 2017 Feb;36(1):11-48. [doi: [10.1016/j.clnu.2016.07.015](https://doi.org/10.1016/j.clnu.2016.07.015)] [Medline: [27637832](https://pubmed.ncbi.nlm.nih.gov/27637832/)]
14. Arends J. Energy and protein intake may have an impact on survival in patients with advanced cancer. *Clin Nutr* 2022 Jan;41(1):266-267. [doi: [10.1016/j.clnu.2021.11.003](https://doi.org/10.1016/j.clnu.2021.11.003)] [Medline: [34865908](https://pubmed.ncbi.nlm.nih.gov/34865908/)]
15. Pimentel GD, Pichard C, Laviano A, Fernandes RC. High protein diet improves the overall survival in older adults with advanced gastrointestinal cancer. *Clin Nutr* 2021 Mar;40(3):1376-1380. [doi: [10.1016/j.clnu.2020.08.028](https://doi.org/10.1016/j.clnu.2020.08.028)] [Medline: [32919817](https://pubmed.ncbi.nlm.nih.gov/32919817/)]
16. Stobäus N, Müller MJ, Küpferling S, Schulzke J, Norman K. Low Recent Protein Intake Predicts Cancer-Related Fatigue and Increased Mortality in Patients with Advanced Tumor Disease Undergoing Chemotherapy. *Nutr Cancer* 2015;67(5):818-824. [doi: [10.1080/01635581.2015.1040520](https://doi.org/10.1080/01635581.2015.1040520)] [Medline: [25996582](https://pubmed.ncbi.nlm.nih.gov/25996582/)]
17. Paulsen MM, Hagen MLL, Frøyen MH, Foss-Pedersen RJ, Bergsager D, Tangvik RJ, et al. A Dietary Assessment App for Hospitalized Patients at Nutritional Risk: Development and Evaluation of the MyFood App. *JMIR Mhealth Uhealth* 2018 Sep 07;6(9):e175 [FREE Full text] [doi: [10.2196/mhealth.9953](https://doi.org/10.2196/mhealth.9953)] [Medline: [30194059](https://pubmed.ncbi.nlm.nih.gov/30194059/)]
18. Paulsen MM, Paur I, Gjестland J, Henriksen C, Varsi C, Tangvik RJ, et al. Effects of using the MyFood decision support system on hospitalized patients' nutritional status and treatment: A randomized controlled trial. *Clin Nutr* 2020 Dec;39(12):3607-3617 [FREE Full text] [doi: [10.1016/j.clnu.2020.03.012](https://doi.org/10.1016/j.clnu.2020.03.012)] [Medline: [32241711](https://pubmed.ncbi.nlm.nih.gov/32241711/)]
19. Krznarić, Bender DV, Laviano A, Cuerda C, Landi F, Monteiro R, et al. A simple remote nutritional screening tool and practical guidance for nutritional care in primary practice during the COVID-19 pandemic. *Clin Nutr* 2020 Jul;39(7):1983-1987 [FREE Full text] [doi: [10.1016/j.clnu.2020.05.006](https://doi.org/10.1016/j.clnu.2020.05.006)] [Medline: [32425292](https://pubmed.ncbi.nlm.nih.gov/32425292/)]

20. Karagiozoglou-Lampoudi T, Daskalou E, Lampoudis D, Apostolou A, Agakidis C. Computer-based malnutrition risk calculation may enhance the ability to identify pediatric patients at malnutrition-related risk for unfavorable outcome. *JPEN J Parenter Enteral Nutr* 2015 May;39(4):418-425. [doi: [10.1177/0148607114529161](https://doi.org/10.1177/0148607114529161)] [Medline: [24718785](https://pubmed.ncbi.nlm.nih.gov/24718785/)]
21. McGurk P, Jackson JM, Elia M. Rapid and reliable self-screening for nutritional risk in hospital outpatients using an electronic system. *Nutrition* 2013 Apr;29(4):693-696. [doi: [10.1016/j.nut.2012.12.020](https://doi.org/10.1016/j.nut.2012.12.020)] [Medline: [23466054](https://pubmed.ncbi.nlm.nih.gov/23466054/)]
22. Lozano-Lozano M, Galiano-Castillo N, Martín-Martín L, Pace-Bedetti N, Fernández-Lao C, Arroyo-Morales M, et al. Monitoring Energy Balance in Breast Cancer Survivors Using a Mobile App: Reliability Study. *JMIR Mhealth Uhealth* 2018 Mar 27;6(3):e67 [FREE Full text] [doi: [10.2196/mhealth.9669](https://doi.org/10.2196/mhealth.9669)] [Medline: [29588273](https://pubmed.ncbi.nlm.nih.gov/29588273/)]
23. Svensson, Larsson C. A Mobile Phone App for Dietary Intake Assessment in Adolescents: An Evaluation Study. *JMIR Mhealth Uhealth* 2015 Nov 03;3(4):e93 [FREE Full text] [doi: [10.2196/mhealth.4804](https://doi.org/10.2196/mhealth.4804)] [Medline: [26534783](https://pubmed.ncbi.nlm.nih.gov/26534783/)]
24. Rangan AM, O'Connor S, Giannelli V, Yap ML, Tang LM, Roy R, et al. Electronic Dietary Intake Assessment (e-DIA): Comparison of a Mobile Phone Digital Entry App for Dietary Data Collection With 24-Hour Dietary Recalls. *JMIR Mhealth Uhealth* 2015 Oct 27;3(4):e98 [FREE Full text] [doi: [10.2196/mhealth.4613](https://doi.org/10.2196/mhealth.4613)] [Medline: [26508282](https://pubmed.ncbi.nlm.nih.gov/26508282/)]
25. Science on WeChat. *Nat Methods* 2020 Sep;17(9):863. [doi: [10.1038/s41592-020-0954-1](https://doi.org/10.1038/s41592-020-0954-1)] [Medline: [32873983](https://pubmed.ncbi.nlm.nih.gov/32873983/)]
26. Zhang Q, Li X, Zhang X, Ding J, Liu T, Qian L, Investigation on Nutrition Statusits Clinical Outcome of Common Cancers (INSCOC) Group. PG-SGA SF in nutrition assessment and survival prediction for elderly patients with cancer. *BMC Geriatr* 2021 Dec 10;21(1):687 [FREE Full text] [doi: [10.1186/s12877-021-02662-4](https://doi.org/10.1186/s12877-021-02662-4)] [Medline: [34893024](https://pubmed.ncbi.nlm.nih.gov/34893024/)]
27. Dewansingh P, Euwes M, Krijnen WP, Strijbos JH, van der Schans CP, Jager-Wittenaar H. Patient-Generated Subjective Global Assessment Short Form better predicts length of stay than Short Nutritional Assessment Questionnaire. *Nutrition* 2021;91-92:111366 [FREE Full text] [doi: [10.1016/j.nut.2021.111366](https://doi.org/10.1016/j.nut.2021.111366)] [Medline: [34274655](https://pubmed.ncbi.nlm.nih.gov/34274655/)]
28. Carriço M, Guerreiro CS, Parreira A. The validity of the Patient-Generated Subjective Global Assessment Short-form© in cancer patients undergoing chemotherapy. *Clin Nutr ESPEN* 2021 Jun;43:296-301. [doi: [10.1016/j.clnesp.2021.03.037](https://doi.org/10.1016/j.clnesp.2021.03.037)] [Medline: [34024530](https://pubmed.ncbi.nlm.nih.gov/34024530/)]
29. Kondrup J, Rasmussen HH, Hamberg O, Stanga Z, Ad Hoc ESPEN Working Group. Nutritional risk screening (NRS 2002): a new method based on an analysis of controlled clinical trials. *Clin Nutr* 2003 Jun;22(3):321-336. [doi: [10.1016/s0261-5614\(02\)00214-5](https://doi.org/10.1016/s0261-5614(02)00214-5)] [Medline: [12765673](https://pubmed.ncbi.nlm.nih.gov/12765673/)]
30. National Health Commission of the People's Republic of China. Hygiene Standard: Nutritional risk screening. National Health Commission of the People's Republic of China. 2013. URL: <http://www.nhc.gov.cn/wjw/yingyang/201308/9c70beb3461e4676aa4afaed8cdaab5.shtml> [accessed 2022-10-04]
31. Gomes-Neto AW, van Vliet IMY, Osté MCJ, de Jong MFC, Bakker SJL, Jager-Wittenaar H, et al. Malnutrition Universal Screening Tool and Patient-Generated Subjective Global Assessment Short Form and their predictive validity in hospitalized patients. *Clin Nutr ESPEN* 2021 Oct;45:252-261 [FREE Full text] [doi: [10.1016/j.clnesp.2021.08.015](https://doi.org/10.1016/j.clnesp.2021.08.015)] [Medline: [34620325](https://pubmed.ncbi.nlm.nih.gov/34620325/)]
32. Abbott J, Teleni L, McKavanagh D, Watson J, McCarthy AL, Isenring E. Patient-Generated Subjective Global Assessment Short Form (PG-SGA SF) is a valid screening tool in chemotherapy outpatients. *Support Care Cancer* 2016 Sep;24(9):3883-3887. [doi: [10.1007/s00520-016-3196-0](https://doi.org/10.1007/s00520-016-3196-0)] [Medline: [27095352](https://pubmed.ncbi.nlm.nih.gov/27095352/)]
33. Shan Z, Guo Y, Hu FB, Liu L, Qi Q. Association of Low-Carbohydrate and Low-Fat Diets With Mortality Among US Adults. *JAMA Intern Med* 2020 Apr 01;180(4):513-523 [FREE Full text] [doi: [10.1001/jamainternmed.2019.6980](https://doi.org/10.1001/jamainternmed.2019.6980)] [Medline: [31961383](https://pubmed.ncbi.nlm.nih.gov/31961383/)]
34. Huang K, Yu D, Guo Q, Yang Y, Wei X, Zhao L, et al. Validation of the MSM and NCI Method for Estimating the Usual Intake of Nutrients and Food According to Four Seasons of Seven Consecutive Daily 24 Hour Dietary Recalls in Chinese Adults. *Nutrients* 2022 Jan 19;14(3):445 [FREE Full text] [doi: [10.3390/nu14030445](https://doi.org/10.3390/nu14030445)] [Medline: [35276804](https://pubmed.ncbi.nlm.nih.gov/35276804/)]
35. Htet MK, Fahmida U, Do TT, Dibley MJ, Ferguson E. The Use of Tablet-Based Multiple-Pass 24-Hour Dietary Recall Application (MP24Diet) to Collect Dietary Intake of Children under Two Years Old in the Prospective Cohort Study in Indonesia. *Nutrients* 2019 Nov 27;11(12):2889 [FREE Full text] [doi: [10.3390/nu11122889](https://doi.org/10.3390/nu11122889)] [Medline: [31783608](https://pubmed.ncbi.nlm.nih.gov/31783608/)]
36. Blalock SJ, Norton LL, Patel RA, Cabral K, Thomas CL. Development and assessment of a short instrument for assessing dietary intakes of calcium and vitamin D. *J Am Pharm Assoc (2003) 2003*;43(6):685-693. [doi: [10.1331/154434503322642615](https://doi.org/10.1331/154434503322642615)] [Medline: [14717265](https://pubmed.ncbi.nlm.nih.gov/14717265/)]
37. Stuhldreher WL, Orchard TJ, Donahue RP, Kuller LH, Gloninger MF, Drash AL. Cholesterol screening in childhood: sixteen-year Beaver County Lipid Study experience. *J Pediatr* 1991 Oct;119(4):551-556. [doi: [10.1016/s0022-3476\(05\)82403-5](https://doi.org/10.1016/s0022-3476(05)82403-5)] [Medline: [1919885](https://pubmed.ncbi.nlm.nih.gov/1919885/)]
38. Kris-Etherton P, Eissenstat B, Jaax S, Srinath U, Scott L, Rader J, et al. Validation for MEDFACTS, a dietary assessment instrument for evaluating adherence to total and saturated fat recommendations of the National Cholesterol Education Program Step 1 and Step 2 diets. *J Am Diet Assoc* 2001 Jan;101(1):81-86. [doi: [10.1016/S0002-8223\(01\)00020-7](https://doi.org/10.1016/S0002-8223(01)00020-7)] [Medline: [11209589](https://pubmed.ncbi.nlm.nih.gov/11209589/)]
39. Ju L, Yu D, Fang H, Guo Q, Xu X, Li S, et al. [Trends and food sources composition of energy, protein and fat in Chinese residents, 1992-2012]. *Wei Sheng Yan Jiu* 2018 Sep;47(5):689-704. [Medline: [30593290](https://pubmed.ncbi.nlm.nih.gov/30593290/)]
40. Jiang H, Zhang J, Du W, Su C, Zhang B, Wang H. Energy intake and energy contributions of macronutrients and major food sources among Chinese adults: CHNS 2015 and CNTCS 2015. *Eur J Clin Nutr* 2021 Feb;75(2):314-324. [doi: [10.1038/s41430-020-0698-0](https://doi.org/10.1038/s41430-020-0698-0)] [Medline: [32724092](https://pubmed.ncbi.nlm.nih.gov/32724092/)]

41. Zhou C, Liu C, Zhang Z, Liu M, Zhang Y, Li H, et al. Variety and quantity of dietary protein intake from different sources and risk of new-onset diabetes: a Nationwide Cohort Study in China. *BMC Med* 2022 Jan 13;20(1):6 [FREE Full text] [doi: [10.1186/s12916-021-02199-8](https://doi.org/10.1186/s12916-021-02199-8)] [Medline: [35022027](https://pubmed.ncbi.nlm.nih.gov/35022027/)]
42. Jager-Wittenaar H, de Bats HF, Welink-Lamberts BJ, Gort-van Dijk D, van der Laan BFAM, Ottery FD, et al. Self-Completion of the Patient-Generated Subjective Global Assessment Short Form Is Feasible and Is Associated With Increased Awareness on Malnutrition Risk in Patients With Head and Neck Cancer. *Nutr Clin Pract* 2020 Apr;35(2):353-362 [FREE Full text] [doi: [10.1002/ncp.10313](https://doi.org/10.1002/ncp.10313)] [Medline: [31134665](https://pubmed.ncbi.nlm.nih.gov/31134665/)]
43. van Bokhorst-de van der Schueren MAE, Guaitoli PR, Jansma EP, de Vet HCW. A systematic review of malnutrition screening tools for the nursing home setting. *J Am Med Dir Assoc* 2014 Mar;15(3):171-184. [doi: [10.1016/j.jamda.2013.10.006](https://doi.org/10.1016/j.jamda.2013.10.006)] [Medline: [24290910](https://pubmed.ncbi.nlm.nih.gov/24290910/)]
44. Guilford JP. *Fundamental Statistics in psychology and Education*. Vol. 565. New York, NY: McGraw-Hill Book Company; 2018:1-565.
45. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977 Mar;33(1):159-174. [Medline: [843571](https://pubmed.ncbi.nlm.nih.gov/843571/)]
46. Cawood AL, Walters ER, Sharp SKE, Elia M, Stratton RJ. 'Self-screening' for malnutrition with an electronic version of the Malnutrition Universal Screening Tool ('MUST') in hospital outpatients: concurrent validity, preference and ease of use. *Br J Nutr* 2018 Sep;120(5):528-536. [doi: [10.1017/S000711451800185X](https://doi.org/10.1017/S000711451800185X)] [Medline: [30058522](https://pubmed.ncbi.nlm.nih.gov/30058522/)]
47. Ryu H, Choi Y, Kim E, Park J, Lee S, Yoon J, et al. Misclassification and characterization of exposure to humidifier disinfectants using a questionnaire. *BMC Public Health* 2021 Jul 27;21(1):1458 [FREE Full text] [doi: [10.1186/s12889-021-11459-4](https://doi.org/10.1186/s12889-021-11459-4)] [Medline: [34315451](https://pubmed.ncbi.nlm.nih.gov/34315451/)]
48. Gotlin MJ, Kingery MT, Baron SL, McCafferty J, Jazrawi LM, Meislin RJ. Recall Bias in Retrospective Assessment of Preoperative Patient-Reported American Shoulder and Elbow Surgeons Scores in Arthroscopic Rotator Cuff Repair Surgery. *Am J Sports Med* 2020 May;48(6):1471-1475. [doi: [10.1177/0363546520913491](https://doi.org/10.1177/0363546520913491)] [Medline: [32255675](https://pubmed.ncbi.nlm.nih.gov/32255675/)]
49. Lizama N, Heyworth J, Thomson A, Slevin T, Fritschi L. Self-reported shift work, recall bias, and belief about disease causation in a case-control study of breast cancer. *Cancer Epidemiol* 2017 Oct;50(Pt A):9-15. [doi: [10.1016/j.canep.2017.07.007](https://doi.org/10.1016/j.canep.2017.07.007)] [Medline: [28756308](https://pubmed.ncbi.nlm.nih.gov/28756308/)]
50. Sinnott S, Bengtson R, Brown A, Glass AP, Johnson MA, Lee JS. The validity of Nutrition Screening Initiative DETERMINE Checklist responses in older Georgians. *J Nutr Elder* 2010 Oct;29(4):393-409. [doi: [10.1080/01639366.2010.521031](https://doi.org/10.1080/01639366.2010.521031)] [Medline: [21104511](https://pubmed.ncbi.nlm.nih.gov/21104511/)]
51. Keller HH, Goy R, Kane S. Validity and reliability of SCREEN II (Seniors in the community: risk evaluation for eating and nutrition, Version II). *Eur J Clin Nutr* 2005 Oct;59(10):1149-1157. [doi: [10.1038/sj.ejcn.1602225](https://doi.org/10.1038/sj.ejcn.1602225)] [Medline: [16015256](https://pubmed.ncbi.nlm.nih.gov/16015256/)]
52. Lee J, Song S, Ahn JS, Kim Y, Lee JE. Use of a Mobile Application for Self-Monitoring Dietary Intake: Feasibility Test and an Intervention Study. *Nutrients* 2017 Jul 13;9(7):748 [FREE Full text] [doi: [10.3390/nu9070748](https://doi.org/10.3390/nu9070748)] [Medline: [28703758](https://pubmed.ncbi.nlm.nih.gov/28703758/)]
53. Chen J, Berkman W, Bardouh M, Ng CYK, Allman-Farinelli M. The use of a food logging app in the naturalistic setting fails to provide accurate measurements of nutrients and poses usability challenges. *Nutrition* 2019 Jan;57:208-216. [doi: [10.1016/j.nut.2018.05.003](https://doi.org/10.1016/j.nut.2018.05.003)] [Medline: [30184514](https://pubmed.ncbi.nlm.nih.gov/30184514/)]
54. Wellard-Cole L, Chen J, Davies A, Wong A, Huynh S, Rangan A, et al. Relative Validity of the Eat and Track (EaT) Smartphone App for Collection of Dietary Intake Data in 18-to-30-Year Olds. *Nutrients* 2019 Mar 14;11(3):621 [FREE Full text] [doi: [10.3390/nu11030621](https://doi.org/10.3390/nu11030621)] [Medline: [30875772](https://pubmed.ncbi.nlm.nih.gov/30875772/)]
55. Teixeira V, Voci SM, Mendes-Netto RS, da Silva DG. The relative validity of a food record using the smartphone application MyFitnessPal. *Nutr Diet* 2018 Apr;75(2):219-225. [doi: [10.1111/1747-0080.12401](https://doi.org/10.1111/1747-0080.12401)] [Medline: [29280547](https://pubmed.ncbi.nlm.nih.gov/29280547/)]
56. Carter MC, Burley VJ, Nykjaer C, Cade JE. 'My Meal Mate' (MMM): validation of the diet measures captured on a smartphone application to facilitate weight loss. *Br J Nutr* 2013 Feb 14;109(3):539-546. [doi: [10.1017/S0007114512001353](https://doi.org/10.1017/S0007114512001353)] [Medline: [22717334](https://pubmed.ncbi.nlm.nih.gov/22717334/)]
57. Ambrosini GL, Hurworth M, Giglia R, Trapp G, Strauss P. Feasibility of a commercial smartphone application for dietary assessment in epidemiological research and comparison with 24-h dietary recalls. *Nutr J* 2018 Jan 09;17(1):5 [FREE Full text] [doi: [10.1186/s12937-018-0315-4](https://doi.org/10.1186/s12937-018-0315-4)] [Medline: [29316930](https://pubmed.ncbi.nlm.nih.gov/29316930/)]
58. Recio-Rodriguez JI, Rodriguez-Martin C, Gonzalez-Sanchez J, Rodriguez-Sanchez E, Martin-Borras C, Martínez-Vizcaino V, EVIDENT Investigators. EVIDENT Smartphone App, a New Method for the Dietary Record: Comparison With a Food Frequency Questionnaire. *JMIR Mhealth Uhealth* 2019 Feb 08;7(2):e11463 [FREE Full text] [doi: [10.2196/11463](https://doi.org/10.2196/11463)] [Medline: [30735141](https://pubmed.ncbi.nlm.nih.gov/30735141/)]
59. Pendergast FJ, Ridgers ND, Worsley A, McNaughton SA. Evaluation of a smartphone food diary application using objectively measured energy expenditure. *Int J Behav Nutr Phys Act* 2017 Mar 14;14(1):30 [FREE Full text] [doi: [10.1186/s12966-017-0488-9](https://doi.org/10.1186/s12966-017-0488-9)] [Medline: [28288657](https://pubmed.ncbi.nlm.nih.gov/28288657/)]
60. Bucher Della Torre S, Carrard I, Farina E, Danuser B, Kruseman M. Development and Evaluation of e-CA, an Electronic Mobile-Based Food Record. *Nutrients* 2017 Jan 18;9(1):76. [doi: [10.3390/nu9010076](https://doi.org/10.3390/nu9010076)] [Medline: [28106767](https://pubmed.ncbi.nlm.nih.gov/28106767/)]
61. Mescoloto SB, Caivano S, Domene S. Avaliação de um aplicativo móvel para a estimativa de ingestão alimentar [Evaluation of a mobile application for estimation of food intake]. *Revista de Nutrição* 2017 Feb;30:91-98 [FREE Full text] [doi: [10.1590/1678-98652017000100009](https://doi.org/10.1590/1678-98652017000100009)]

62. Shim J, Oh K, Kim HC. Dietary assessment methods in epidemiologic studies. *Epidemiol Health* 2014;36:e2014009 [FREE Full text] [doi: [10.4178/epih/e2014009](https://doi.org/10.4178/epih/e2014009)] [Medline: [25078382](https://pubmed.ncbi.nlm.nih.gov/25078382/)]
63. Rangan AM, Tieleman L, Louie JCY, Tang LM, Hebden L, Roy R, et al. Electronic Dietary Intake Assessment (e-DIA): relative validity of a mobile phone application to measure intake of food groups. *Br J Nutr* 2016 Jun;115(12):2219-2226. [doi: [10.1017/S0007114516001525](https://doi.org/10.1017/S0007114516001525)] [Medline: [27121045](https://pubmed.ncbi.nlm.nih.gov/27121045/)]
64. Kondrup J, Allison SP, Elia M, Vellas B, Plauth M, Educational Clinical Practice Committee, European Society of Parenteral and Enteral Nutrition (ESPEN). ESPEN guidelines for nutrition screening 2002. *Clin Nutr* 2003 Aug;22(4):415-421. [doi: [10.1016/s0261-5614\(03\)00098-0](https://doi.org/10.1016/s0261-5614(03)00098-0)] [Medline: [12880610](https://pubmed.ncbi.nlm.nih.gov/12880610/)]
65. Ma C, Cruz C, Eslamparast T, Taylor L, Farhat K, Shommu N, et al. The patient generated subjective global assessment short form is a useful screening tool to detect risk for malnutrition in patients with cirrhosis. *Clin Nutr ESPEN* 2022 Aug;50:330-333. [doi: [10.1016/j.clnesp.2022.05.010](https://doi.org/10.1016/j.clnesp.2022.05.010)] [Medline: [35871945](https://pubmed.ncbi.nlm.nih.gov/35871945/)]
66. De Groot LM, Lee G, Ackerie A, van der Meij BS. Malnutrition Screening and Assessment in the Cancer Care Ambulatory Setting: Mortality Predictability and Validity of the Patient-Generated Subjective Global Assessment Short form (PG-SGA SF) and the GLIM Criteria. *Nutrients* 2020 Jul 30;12(8):2287. [doi: [10.3390/nu12082287](https://doi.org/10.3390/nu12082287)] [Medline: [32751724](https://pubmed.ncbi.nlm.nih.gov/32751724/)]
67. Trtovac D, Lee J. The Use of Technology in Identifying Hospital Malnutrition: Scoping Review. *JMIR Med Inform* 2018 Jan 19;6(1):e4 [FREE Full text] [doi: [10.2196/medinform.7601](https://doi.org/10.2196/medinform.7601)] [Medline: [29351894](https://pubmed.ncbi.nlm.nih.gov/29351894/)]

## Abbreviations

**3d-24HRs:** 3 days of 24-hour dietary recall

**24HR:** 24-hour dietary recall

**ASPEN:** American Society for Parenteral and Enteral Nutrition

**CC:** correlation coefficient

**CHNS:** China Health and Nutrition Surveys

**CSPEN:** Chinese Society for Parenteral and Enteral Nutrition

**ESPEN:** European Society for Clinical Nutrition and Metabolism

**FFQ:** Food Frequency Questionnaire

**MUST:** Malnutrition Universal Screening Tool

**NPV:** negative predictive value

**NRS-2002:** Nutritional Risk Screening

**PG-SGA-SF:** Patient-Generated Subjective Global Assessment Short Form

*Edited by A Mavragani; submitted 15.06.22; peer-reviewed by M Paulsen; comments to author 21.07.22; revised version received 22.09.22; accepted 23.09.22; published 26.10.22.*

### *Please cite as:*

Long Z, Huang S, Zhang J, Zhang D, Yin J, He C, Zhang Q, Xu H, He H, Sun HC, Xie K

*A Digital Smartphone-Based Self-administered Tool (R+ Dietitian) for Nutritional Risk Screening and Dietary Assessment in Hospitalized Patients With Cancer: Evaluation and Diagnostic Accuracy Study*

*JMIR Form Res* 2022;6(10):e40316

URL: <https://formative.jmir.org/2022/10/e40316>

doi: [10.2196/40316](https://doi.org/10.2196/40316)

PMID: [36287601](https://pubmed.ncbi.nlm.nih.gov/36287601/)

©Zhiwen Long, Shan Huang, Jie Zhang, Deng Zhang, Jun Yin, Chengyuan He, Qinqiu Zhang, Huilin Xu, Huimin He, Ho Ching Sun, Ke Xie. Originally published in *JMIR Formative Research* (<https://formative.jmir.org/>), 26.10.2022. 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

# Relationships Between Blood Pressure Reduction, Weight Loss, and Engagement in a Digital App–Based Hypertension Care Program: Observational Study

OraLee H Branch<sup>1</sup>, PhD; Mohit Rikhy<sup>1</sup>, MSc; Lisa A Auster-Gussman<sup>1</sup>, PhD; Kimberly G Lockwood<sup>1</sup>, PhD; Sarah A Graham<sup>1</sup>, PhD

Lark Technologies, Inc, Mountain View, CA, United States

**Corresponding Author:**

Sarah A Graham, PhD  
Lark Technologies, Inc  
2570 El Camino Real  
Mountain View, CA, 94040  
United States  
Phone: 1 650 300 1755  
Email: [sarah.graham@lark.com](mailto:sarah.graham@lark.com)

## Abstract

**Background:** Home blood pressure (BP) monitoring is recommended for people with hypertension; however, meta-analyses have demonstrated that BP improvements are related to additional coaching support in combination with self-monitoring, with little or no effect of self-monitoring alone. High-contact coaching requires substantial resources and may be difficult to deliver via human coaching models.

**Objective:** This observational study assessed changes in BP and body weight following participation in a fully digital program called Lark Hypertension Care with coaching powered by artificial intelligence (AI).

**Methods:** Participants (N=864) had a baseline systolic BP (SBP)  $\geq 120$  mm Hg, provided their baseline body weight, and had reached at least their third month in the program. The primary outcome was the change in SBP at 3 and 6 months, with secondary outcomes of change in body weight and associations of changes in SBP and body weight with participant demographics, characteristics, and program engagement.

**Results:** By month 3, there was a significant drop of  $-5.4$  mm Hg (95% CI  $-6.5$  to  $-4.3$ ;  $P < .001$ ) in mean SBP from baseline. BP did not change significantly (ie, the SBP drop maintained) from 3 to 6 months for participants who provided readings at both time points ( $P = .49$ ). Half of the participants achieved a clinically meaningful drop of  $\geq 5$  mm Hg by month 3 (178/349, 51.0%) and month 6 (98/199, 49.2%). The magnitude of the drop depended on starting SBP. Participants classified as hypertension stage 2 had the largest mean drop in SBP of  $-12.4$  mm Hg (SE 1.2 mm Hg) by month 3 and  $-13.0$  mm Hg (SE 1.6 mm Hg) by month 6; participants classified as hypertension stage 1 lowered by  $-5.2$  mm Hg (SE 0.8 mm Hg) by month 3 and  $-7.3$  mm Hg (SE 1.3 mm Hg) by month 6; participants classified as elevated lowered by  $-1.1$  mm Hg (SE 0.7 mm Hg) by month 3 but did not drop by month 6. Starting SBP ( $\beta = .11$ ;  $P < .001$ ), percent weight change ( $\beta = -.36$ ;  $P = .02$ ), and initial BMI ( $\beta = -.56$ ;  $P < .001$ ) were significantly associated with the likelihood of lowering SBP  $\geq 5$  mm Hg by month 3. Percent weight change acted as a mediator of the relationship between program engagement and drop in SBP. The bootstrapped unstandardized indirect effect was  $-0.0024$  (95% CI  $-0.0052$  to 0;  $P = .002$ ).

**Conclusions:** A hypertension care program with coaching powered by AI was associated with a clinically meaningful reduction in SBP following 3 and 6 months of program participation. Percent weight change was significantly associated with the likelihood of achieving a  $\geq 5$  mm Hg drop in SBP. An AI-powered solution may offer a scalable approach to helping individuals with hypertension achieve clinically meaningful reductions in their BP and associated risk of cardiovascular disease and other serious adverse outcomes via healthy lifestyle changes such as weight loss.

(JMIR Form Res 2022;6(10):e38215) doi:[10.2196/38215](https://doi.org/10.2196/38215)

**KEYWORDS**

high blood pressure; obesity; weight loss; conversational artificial intelligence; lifestyle coaching



## Introduction

The American Heart Association (AHA) defines high blood pressure (BP), called hypertension, as systolic BP (SBP)  $\geq 130$  mm Hg or diastolic BP (DBP)  $\geq 80$  mm Hg that remains elevated over time [1]. Nearly half of US adults have hypertension or are taking medication for hypertension, and only 1 in 4 have their BP under control [2]. Effective strategies to improve self-management of BP are critical since hypertension is a leading modifiable risk factor for cardiovascular disease [3], ranks as the leading cause of mortality in the United States [4], and is associated with a higher risk for other serious and costly conditions such as stroke, kidney disease, dementia, and eye damage [5]. In a large meta-analysis, lowering SBP by at least 5 mm Hg reduced the risk of major cardiovascular events by 10% even at normal ( $<120$  mm Hg) and high-normal (120-129 mm Hg) values of SBP [6].

The AHA recommends home monitoring of BP as a part of self-management for all people with hypertension because it provides a better estimate of BP under “normal” conditions and may help improve BP control [7]. However, meta-analyses have provided strong evidence that BP improvements are related to cointerventions involving individually tailored coaching support in combination with self-monitoring, with little or no effect of self-monitoring alone [7,8]. Coaching can provide personalized support for increasing health-promoting lifestyle behaviors that are known to reduce BP, such as reaching and maintaining a healthy weight, eating a healthy diet, limiting alcohol consumption, avoiding smoking, adhering to prescribed medications, and being physically active [9]. Compared to a healthy weight, prior research has attributed an estimated 32% excess risk of hypertension to being overweight and 47% to being obese [10]. There is a well-characterized linear relationship between SBP and BMI, with a higher prevalence of hypertension in individuals in higher BMI classes [11]. Weight loss is a particularly important focus for individuals with hypertension who are overweight or obese because it is associated with improvements in BP control [12,13].

The US Preventive Task Force recommends moderate-to-high-intensity behavioral coaching for adults with cardiovascular risk factors including high BP and being overweight or obese [14]. However, behavioral coaching is most effective when it includes personalized content and feedback as well as frequent and timely interactions [15,16]. This type of coaching is highly time and resource intensive and may be difficult to achieve via human coaching models. Fully digital programs powered by artificial intelligence (AI) represent one solution for hypertension care that combines self-monitoring with highly personalized, automated coaching. An AI-powered coaching platform enables the delivery of continuous,

synchronous coaching and feedback and offers a scalable, high-touch, and long-term solution to help people make lifestyle changes and sustain healthy behaviors. However, there is little evidence of the effectiveness of AI-powered solutions for facilitating reductions in BP and body weight for individuals with hypertension.

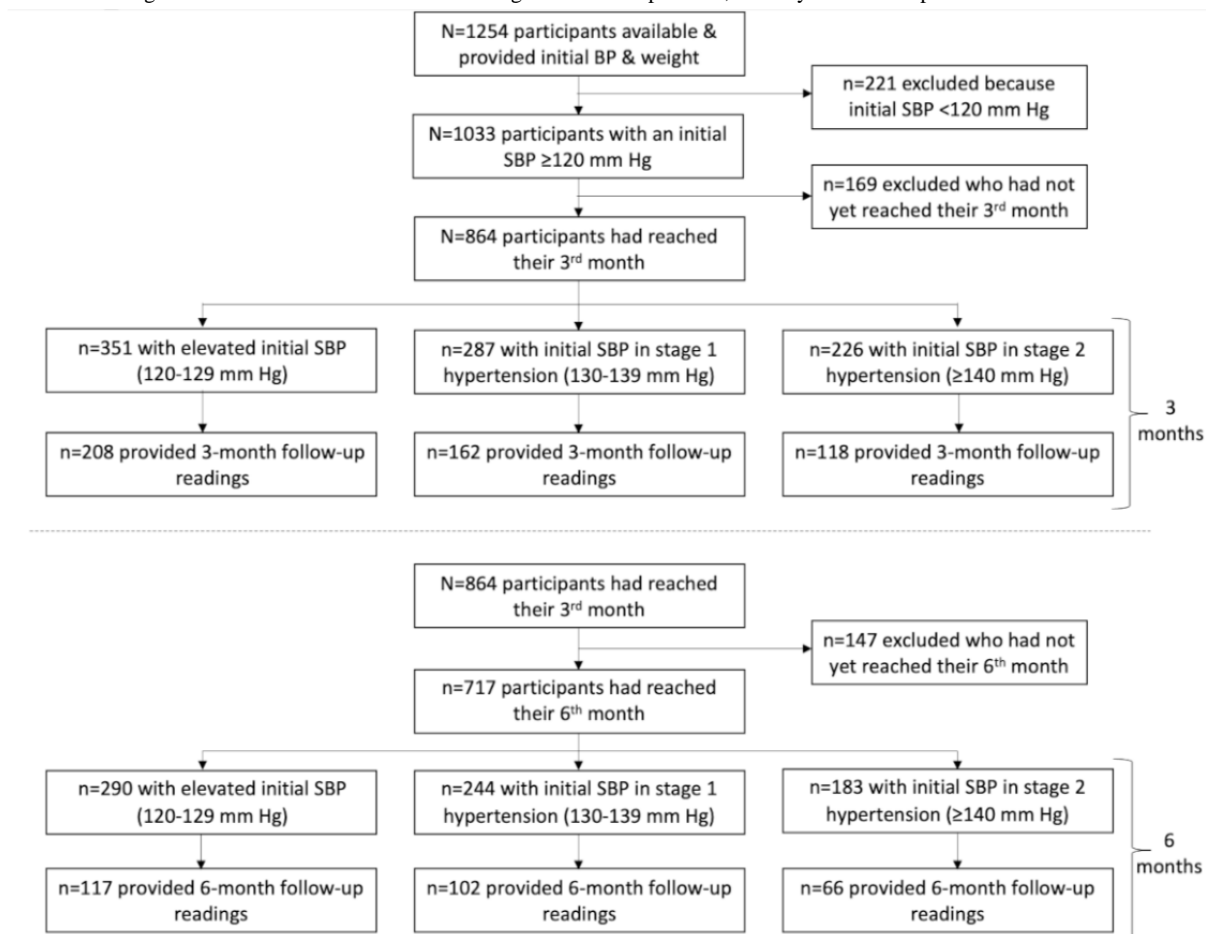
This observational study assessed changes in BP and body weight following participation in a fully digital hypertension care program powered by AI. This program used self-monitoring of BP coupled with conversational AI delivered on a participant's smartphone to coach participants in lowering their BP, losing weight, and making other healthy lifestyle changes. The primary study objective was to evaluate the change in BP over time (baseline, 3 months, and 6 months), with secondary objectives of assessing changes in weight and associations of BP and weight change with AI-powered coaching and activities. The primary hypothesis was that participants with elevated or greater SBP (ie,  $\geq 120$  mm Hg) at baseline would lower their SBP on average by at least 5 mm Hg, which is commonly considered a clinically meaningful threshold for reduced cardiovascular disease risk [6,17]. We expected a greater reduction for participants with a higher starting SBP. The secondary hypothesis was that a greater reduction in BP and weight loss would be associated with greater participation in AI coaching and activities within the mobile app.

## Methods

### Participants and Recruitment

This was a study of participants enrolled on a rolling basis (beginning January 1, 2019, and concluding on November 4, 2021) in a commercially available program called Lark Hypertension Care offered via existing partnerships between the company and health insurance providers, employers, and other organizations. The program recruits eligible participants via direct referrals from health plans or digital advertising through email campaigns and social media platforms such as Facebook. All participants received a link via SMS text messages that prompted them to download the mobile app, agree to the app's privacy policy, and give permission to use their deidentified data for research. Included program participants were  $\geq 18$  years of age at enrollment, English speaking, owned an Android or iPhone smartphone, and had a respective health plan that identified them as having hypertension or being at risk for hypertension. Excluded participants did not provide initial BP or body weight readings; had an initial SBP reading  $<120$  mm Hg, indicating that their BP was controlled; or had not yet reached at least their third month in the program (see Figure 1). A subset of participants had also reached their sixth month in the program for analysis of BP and weight change at 6 months.

**Figure 1.** Flow of participants through the study. Participants categorized based on SBP as elevated, hypertension stage 1, or hypertension stage 2. Participants with readings at 3 and 6 months could be BP or weight. BP: blood pressure; SPB: systolic blood pressure.



## Ethical Considerations

The study received exemption status from Advarra Institutional Review Board (protocol #Pro00047181) for retrospective analysis of previously collected and deidentified data.

## Description of the Hypertension Care Program

The hypertension care program consisted of educational lessons and fully automated, personalized coaching on healthy lifestyle behaviors powered by conversational AI. Conversational AI technologies facilitate humanlike interactions between a robot (computer) and a human via a text-based interface. See [Multimedia Appendix 1](#) [1,14,18-22] for a detailed description. After enrollment, participants completed a brief orientation on how to obtain accurate BP measurements, set medication reminders, and select an optional weight loss goal. Participants then progressed through weekly educational lessons spread over 26 weeks.

Participants in the program could opt to receive connected devices (digital BP cuff or weight scale) to measure BP and weight and could enter BP data in a variety of ways in the app, including wirelessly or manually. If using a connected BP monitor, participants could take a measurement through a guided coaching exchange and sync the measurement immediately. Those who already had a home BP cuff could use their existing device and manually enter BP readings in the app. Regardless of the measurement method, participants received detailed

instructions on taking at-home BP measurements, as outlined by the AHA [18]. The program has built-in safety mechanisms: in the case of extremely high readings (SBP >180 mm Hg or DBP >110 mm Hg) or low readings (<90 mm Hg or <60 mm Hg) and symptoms like dizziness, the AI coach prompts participants to seek assistance or call their medical provider and assists them in taking these actions.

## Primary Outcome of Change in SBP

The primary outcome was the change in SBP from the start of the program to 3 and 6 months, respectively. Starting BP was a participant's average measurement within the first week of the program, and 3- and 6-month BPs were a participant's average measurement during the third and sixth month in the program, respectively. A few participants had readings that occurred soon after the sixth month, and we included these data points as well to maximize the sample size available for analysis. This was necessary for a real-world study where participants were not aware that they were supposed to provide readings at a particular time point. We considered a clinically meaningful improvement to be a drop of ≥5 mm Hg at any time point. We also conducted subgroup analyses on participants classified as having elevated SBP at baseline (SBP 120-129 mm Hg), stage 1 hypertension (SBP 130-139 mm Hg), and stage 2 hypertension (SBP ≥140 mm Hg). We assessed corresponding changes in DBP based on starting SBP classification.

## Secondary Outcomes of Change in Body Weight and Associations With Program Engagement

The secondary outcome was the percent weight change at month 3 and month 6, respectively. We calculated the percent weight change at 3 months as follows: (first weight – 3-month nadir weight)/first weight. We calculated the percent weight change at 6 months as follows: (first weight – 6-month nadir weight)/first weight. We removed any abnormal weigh-ins indicating a weight loss or gain rate of >7 lbs/week unless confirmed by the user to be a correct measurement. We assessed associations between the change in SBP and body weight at month 3 with participant demographics, characteristics, and engagement metrics using 2 separate regression models. For the regression with change in SBP as the dependent variable, independent variables included participant demographics (age, sex), characteristics (initial BMI, starting SBP, percent weight change), and program engagement metrics (number of sessions with the AI coach, number of BP measurements). For the regression with percent weight change as the dependent variable, independent variables included participant demographics (age, sex), characteristics (initial BMI, starting SBP), and program engagement metrics (number of sessions with the AI coach, number of weight measurements). We examined these associations at month 3 instead of month 6 due to the larger sample size available at month 3 (for statistical power) and since most of the change in BP occurred by month 3 and was maintained through month 6 for participants who provided both measurements.

### Statistical Analyses

We conducted all statistical analyses in RStudio 4.0.5. We compared participant demographic and characteristic data

between subgroups categorized by baseline SBP classification. We used paired *t* tests to evaluate the change in BP and body weight between each pair of time points (baseline, 3 months, 6 months) to maximize the sample size available for each comparison (more participants had made it to 3 months in the program [N=864] compared to 6 months [n=717]). We conducted two separate regression analyses: (1) a multiple logistic regression to assess the effects of participant demographics, characteristics, and program engagement on participants' likelihood of achieving a clinically meaningful drop in SBP of  $\geq 5$  mm Hg by month 3 in the program; and (2) a linear regression to assess the effects of participant demographics, characteristics, and program engagement on percent weight change by month 3 in the program. We conducted these analyses separately to consider each outcome independently and because not all participants had both BP and weight data available at 3 months. Engagement variables in both regressions had no issues of multicollinearity; all variance inflation factors in both models were <2. The results of the regression analyses suggested that percent weight change was a mediator of the relationship between program engagement (number of sessions with the AI coach) and SBP drop, so we conducted an exploratory full mediation analysis [23] to confirm this observation. The a priori  $\alpha$  was  $\leq .05$  for all statistical tests.

## Results

### Participant Demographics and Characteristics

There were significant differences across participants based on the SBP category for initial BMI (Table 1). Initial BMI increased, on average, by 1 unit for each increase in the SBP classification category.

**Table 1.** Baseline demographics and characteristics for all participants and grouped by starting systolic blood pressure.

	BP <sup>a</sup> category				<i>F</i> test/ chi-square ( <i>df</i> ) <sup>b</sup>	<i>P</i> value
	All, mean (SE) <sup>c</sup>	Elevated, mean (SE) <sup>c</sup>	Stage 1 hypertension, mean (SE) <sup>c</sup>	Stage 2 hypertension, mean (SE) <sup>c</sup>		
Age (years)	51.5 (0.34)	52.3 (0.54)	51.2 (0.55)	50.7 (0.70)	1.9 (2,847)	.14
Initial BMI (kg/m <sup>2</sup> )	34.0 (0.25)	33.2 (0.37)	34.2 (0.39)	35.2 (0.56)	5.4 (2,861)	.005
No. of BP readings included in baseline BP	3.6 (0.11)	3.5 (0.16)	3.9 (0.21)	3.3 (0.19)	2.6 (2,861)	.07
Baseline systolic BP (mm Hg)	134.5 (0.38)	125.0 (0.15)	134.3 (0.17)	149.5 (0.60)	1506.0 (2,861)	<.001
Baseline diastolic BP (mm Hg)	85.3 (0.30)	81.4 (0.38)	85.4 (0.43)	91.4 (0.63)	110.5 (2,861)	<.001
Female sex <sup>c</sup>	500/861 (58.1)	216/351 (61.5)	162/286 (56.6)	122/224 (54.5)	3.2 (2)	.21
White race <sup>c</sup>	370/499 (74.1)	157/207 (75.8)	128/168 (76.2)	85/124 (68.5)	2.7 (2)	.26
Taking BP meds <sup>c</sup>	530/578 (91.7)	228/245 (93.1)	156/177 (88.1)	146/156 (93.6)	4.3 (2)	.12

<sup>a</sup>BP: blood pressure.

<sup>b</sup>Chi-square is applicable only to female sex, White race, and taking BP meds. For the other demographics, *F* test is applicable.

<sup>c</sup>Mean (SE) here is not applicable to categories of female sex, White race, and taking BP meds; for these categories, data are expressed as n/N (%).

## Changes in BP

Participants provided a mean of 3.6 (SE 0.1) BP readings for the calculation of average starting BP, a mean of 15.0 (SE 1.2) readings for the calculation of average BP at month 3, and a mean of 17.0 (SE 1.6) BP readings for the calculation of average BP at month 6.

There was a significant overall drop of  $-5.4$  mm Hg in mean SBP following 3 months ( $t=9.5_{348}$ ;  $P<.001$ ; 95% CI  $-6.5$  to  $-4.3$ ), and no change in SBP from 3 to 6 months for those who provided readings at both time points ( $t=0.7_{139}$ ;  $P=.49$ ; [Table 2](#)).

2). Participants with a starting SBP classified as hypertension stage 2 had the greatest change in SBP at both time points, with a drop of  $-12.4$  mm Hg (SE 1.2 mm Hg) by month 3 and a drop of  $-13.0$  mm Hg (SE 1.6 mm Hg) by month 6.

Approximately half of the overall sample achieved a clinically meaningful SBP drop of  $\geq 5$  mm Hg by month 3 (178/349, 51%) and by month 6 (98/199, 49.2%). The drop in SBP resulted in 47.6% (166/349) of participants lowering their SBP by at least 1 classification category (eg, hypertension stage 2 to hypertension stage 1; hypertension stage 1 to elevated) by month 3.

**Table 2.** Change in blood pressure from baseline to 3 and 6 months.

	Change in BP <sup>a</sup> at 3 months, $\Delta$ mean (95% CI) <sup>b</sup>	<i>t</i> test ( <i>df</i> )	<i>P</i> value	With $\geq 5$ mm Hg SBP <sup>c</sup> drop at month 3, n/N (%)	Change in SBP at 6+ months, $\Delta$ mean (95% CI) <sup>b</sup>	<i>t</i> test ( <i>df</i> )	<i>P</i> value	With $\geq 5$ mm Hg SBP drop at month 6, n/N (%)
<b>SBP<sup>b</sup> (mm Hg)</b>								
Overall	$-5.4$ ( $-6.5$ to $-4.3$ )	9.5 (348)	<.001	178/349 (51.0)	$-5.3$ ( $-6.9$ to $-3.6$ )	6.3 (198)	<.001	98/199 (49.2)
120-129	$-1.1$ ( $-2.5$ to $0.4$ )	1.5 (148)	.14	51/149 (34.2)	0.6 ( $-1.7$ to $2.8$ )	$-0.5$ (85)	.62	25/86 (29.1)
130-139	$-5.2$ ( $-6.8$ to $-3.7$ )	6.6 (106)	<.001	53/107 (49.5)	$-7.3$ ( $-9.8$ to $-4.8$ )	5.8 (64)	<.001	36/65 (55.4)
$\geq 140$	$-12.4$ ( $-14.9$ to $-10.0$ )	10.2 (92)	<.001	74/93 (79.6)	$-13.0$ ( $-16.2$ to $-9.8$ )	8.2 (47)	<.001	37/48 (77.1)
<b>DBP<sup>d</sup> (mm Hg)<sup>e</sup></b>								
Overall	$-1.3$ ( $-2.1$ to $-0.5$ )	3.1 (348)	.002	N/A <sup>e</sup>	$-1.2$ ( $-2.3$ to $-0.2$ )	2.3 (198)	.02	N/A <sup>f</sup>
Elevated	0.6 ( $-0.6$ to $1.7$ )	$-1.0$ (148)	.34	N/A	1.0 ( $-0.6$ to $2.5$ )	$-1.2$ (85)	.22	N/A
Stage 1	$-0.9$ ( $-2.4$ to $0.5$ )	1.3 (106)	.21	N/A	$-2.4$ ( $-4.2$ to $-0.7$ )	2.8 (64)	.007	N/A
Stage 2	$-4.6$ ( $-6.2$ to $-3.0$ )	5.8 (92)	<.001	N/A	$-3.5$ ( $-5.7$ to $-1.3$ )	3.2 (47)	.002	N/A

<sup>a</sup>BP: blood pressure.

<sup>b</sup>Negative  $\Delta$  values indicate a drop in BP and positive values an increase.

<sup>c</sup>SBP: systolic blood pressure.

<sup>d</sup>DBP: diastolic blood pressure.

<sup>e</sup>DBP categories based on initial SBP classification: elevated, hypertension stage 1, or hypertension stage 2.

<sup>f</sup>N/A: not applicable.

## Associations With BP Drop and Weight Change

Results of the multiple logistic regression for BP revealed associations of participant demographics, characteristics, and program engagement metrics with the likelihood of achieving a clinically meaningful drop of  $\geq 5$  mm Hg in SBP by month 3. The overall regression was statistically significant (log-likelihood  $-152.3$ ; McFadden's pseudo  $R^2=0.18$ ;  $P<.001$ ). Starting SBP, initial BMI, and percent weight change at month 3 were significantly associated with the likelihood of achieving a drop of  $\geq 5$  mm Hg in SBP ([Table 3](#)).

Of the participants who provided weigh-ins in the third month, 90.1% (374/415) remained weight stable or lost weight over the first 3 months of the program ([Figure 2](#)).

Results of the multiple linear regression for weight change revealed associations of participant demographics, characteristics, and program engagement metrics with the magnitude of percent weight change by month 3. The overall regression was statistically significant ( $F_{7,397}=5.97$ ;  $R^2=0.10$   $P<.001$ ). Initial BMI, the number of sessions with the AI coach, and the number of weigh-ins recorded in the first 3 months were significantly associated with percent weight change ([Table 4](#)).

The 2 regression models together demonstrated that percent weight change at month 3 was significantly associated with the likelihood of achieving a  $\geq 5$  mm Hg drop in SBP, and program engagement variables were significantly associated with the magnitude of percent weight change but not SBP drop. Thus, it was important to consider whether percent weight change acted as a statistical mediator between program engagement and SBP drop. The results of the mediation analysis indeed demonstrated that the effect of program engagement on the drop in SBP was fully mediated by percent weight change at month 3.

As Figure 3 illustrates, the regression coefficient between percent weight change at month 3 and the drop in SBP was significant, even though the regression coefficient between the number of sessions with the AI coach and the drop in SBP was not. Although the total effect was therefore not significant, this is not considered a requirement for statistical mediation [23]. We tested the significance of the unstandardized indirect effect of the number of sessions with the AI coach on the change in SBP that occurred via the mediator percent weight change at month 3 using 1000 bootstrapped samples. The bootstrapped unstandardized indirect or average causal mediation effect was  $-0.0024$  (95% CI  $-0.0052$  to  $0$ ;  $P=.002$ ).

**Table 3.** Regression results of the likelihood of lowering  $\geq 5$  mm Hg in SBP by 3 months ( $n=268$ )<sup>a</sup>.

Variable	Standardized coefficient ( $\beta$ )	SE	Z value	P value
Constant	.12	0.14	0.87	.38
Age	-.03	0.15	-0.24	.81
Male sex	-.19	0.14	-1.31	.19
Initial BMI ( $\text{kg}/\text{m}^2$ )	-.56	0.15	-3.57	<.001
Percent weight change by 3 months <sup>b</sup>	-.36	0.15	2.43	.02
Starting SBP <sup>c</sup>	1.11	0.18	6.13	<.001
No. of sessions with AI <sup>d</sup> coach in first 3 months	-.04	0.16	-0.27	.79
No. of BP measurements recorded in first 3 months	.05	0.16	0.28	.78

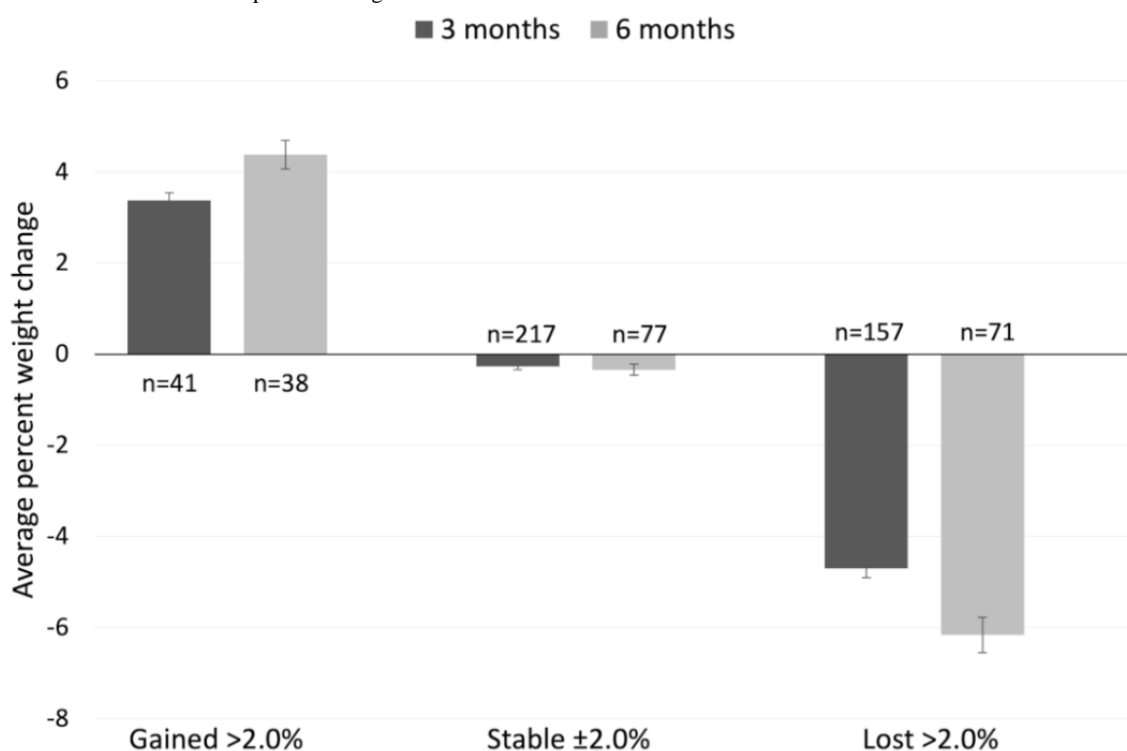
<sup>a</sup>Included participants had to have both blood pressure and weight data available at 3 months.

<sup>b</sup>A negative sign for weight change indicates greater weight loss.

<sup>c</sup>SBP: systolic blood pressure.

<sup>d</sup>AI: artificial intelligence.

**Figure 2.** The number of participants and mean (SE) for each category of weight change: gained, stable, and lost for the 415 participants who provided weigh-ins at 3 months and the 186 who provided weigh-ins at 6 months.



**Table 4.** Regression results for associations of participant demographics, characteristics, and program engagement with the dependent variable percent weight change. For the dependent variable, weight change, a negative value indicates weight loss (n=400)<sup>a</sup>.

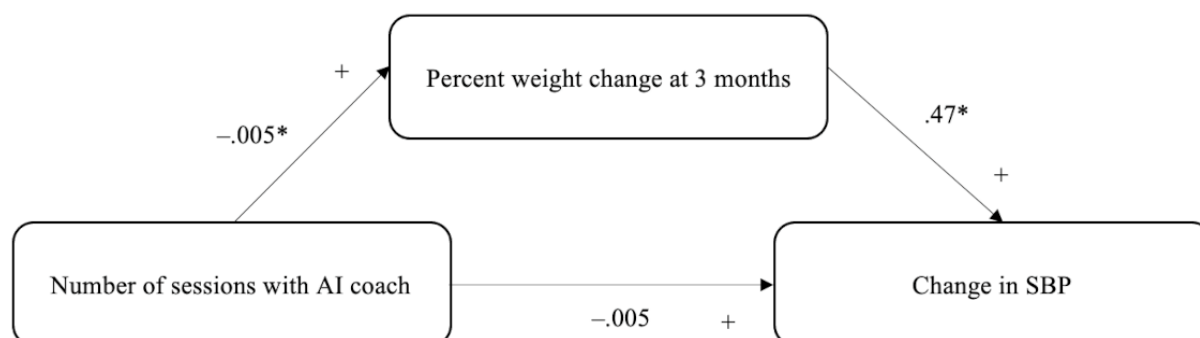
Variable	Standardized coefficient ( $\beta$ )	SE	t value	P value
Constant	-1.57	0.15	10.26	<.001
Age	.05	0.17	-0.30	.77
Male sex	.09	0.15	-0.59	.56
Initial BMI (kg/m <sup>2</sup> )	-.48	0.16	3.02	.003
Starting SBP <sup>b</sup>	.17	0.15	-1.09	.27
No. of sessions with AI <sup>c</sup> coach in first 3 months	-.62	0.16	3.81	<.001
No. of weight measurements recorded in first 3 months	-.44	0.16	2.78	.006

<sup>a</sup>Larger sample size due to removed requirement for blood pressure data at 3 months.

<sup>b</sup>SBP: systolic blood pressure.

<sup>c</sup>AI: artificial intelligence.

**Figure 3.** Percent weight change at 3 months as a mediator of the relationship between program engagement and change in SBP. A negative coefficient indicates that a greater number of sessions with the AI coach was significantly related to greater weight loss (represented by a negative sign). The positive coefficient for the significant relationship between percent weight change and change in SPB is because drops (improvements) in both are represented with a negative sign. AI: artificial intelligence; SBP: systolic blood pressure. \*Unstandardized coefficients and significance level  $\leq .001$ .



## Discussion

### Principal Findings

The primary aim of this study was to assess changes in SBP following participation in an AI-powered hypertension care program. We further assessed changes in weight and associations of BP and weight change with participant characteristics and program engagement. In support of the primary hypothesis, participation in the AI-powered hypertension care program was associated with clinically meaningful reductions in SBP over 3 to 6 months, with larger drops observed in the subgroups of participants classified as stage 1 or 2 hypertension at baseline. More than half of all participants achieved a clinically meaningful drop in SBP of  $\geq 5$  mm Hg. The percentage of participants that achieved a clinically meaningful reduction was higher for those classified as hypertensive at baseline, with 79.6% (74/93) of participants with a starting SBP  $\geq 140$  mm Hg achieving a drop of  $\geq 5$  mm Hg by month 3 in the program. Percent weight change at month 3 was significantly associated with program engagement and the likelihood of achieving a clinically meaningful drop in SBP, and percent weight change mediated the relationship between program engagement and change in SBP.

The overall average reduction in SBP of  $-5.4$  mm Hg in this study corresponded to roughly half of the participants lowering their initial SBP classification by at least 1 classification category by month 3. Key aspects of the program included reminders to monitor BP, medication adherence support, personalized in-the-moment feedback about progress, hypertension-specific nutrition coaching, stress-reduction coaching, and educational material about hypertension. Prior research has shown that this type of personalized and multifaceted intervention is critical for success in digital hypertension care programs [24]. Compared to individuals with normotensive BP, individuals with treated but uncontrolled hypertension are at a higher risk of cardiovascular, cerebrovascular, and all-cause mortality [25]. Lowering BP is of well-established importance for those with hypertension stage 2; however, lowering BP for individuals with elevated BP or hypertension stage 1 is also clinically meaningful because absolute reductions in the risk of stroke, major cardiovascular events, and cardiovascular and all-cause mortalities have been shown to be progressively lower with a lower attained value of SBP [26].

The observed change in systolic BP in this study is comparable to the drops reported in published human coach-led behavioral

lifestyle interventions. In the meta-analysis by Tucker et al [8], only 1 of 5 interventions that investigated self-monitoring alone showed a significant drop in SBP, with a pooled average of  $-1.0$  mm Hg (95% CI  $-3.3$  to  $1.2$ ). In contrast, self-monitoring with intensive coaching support via counseling or telecounseling showed statistically significant drops in SBP compared to a control group, with a pooled average reduction of  $-6.1$  mm Hg (95% CI  $-9.0$  to  $-3.2$ ). The present study provides new evidence that members of a fully digital program powered by AI coaching experienced improvements in their BP while enrolled in the program.

Participants in this study were obese on average, with the overall initial BMI of  $34.0$  kg/m<sup>2</sup> falling into class I obesity [27]. There was a 1-unit increase in initial BMI for each increasing classification of starting SBP, with participants classified as elevated having an average BMI of  $33.2$  kg/m<sup>2</sup>, hypertension stage 1 an average BMI of  $34.2$  kg/m<sup>2</sup>, and hypertension stage 2 an average BMI of  $35.2$  kg/m<sup>2</sup> (class II obesity). Given that weight loss is associated with improvements in BP control [12,13], weight loss was a particularly important target for participants in this study. Percent weight change at month 3 was significantly associated with the likelihood of achieving a clinically meaningful drop in SBP of  $\geq 5$  mm Hg. Participants with a higher initial BMI were less likely to achieve this clinically meaningful drop. However, in the regression with percent weight change as the dependent variable, participants with a higher initial BMI lost a greater amount of weight. Taken together, it appears that there were some participants with a higher initial BMI that did not achieve a clinically meaningful reduction in BP. However, for the larger number of users analyzed in the weight change regression, having a higher initial BMI was associated with a greater percent weight loss.

There are multiple reasons for the observed relationship between percent weight change and achieving a clinically meaningful drop in SBP in this study. There are well-established physiological benefits of weight loss for hypertension, such as improvement in insulin sensitivity and a decrease in sympathetic nervous system activity and inflammation [12,13,28]. However, given that percent weight change was significantly associated with program engagement and statistically mediated the relationship between program engagement and SBP drop in this study, it may also be that percent weight change was an indicator of those participants who were more closely adhering to the recommendations of the AI coach and adopting the healthy lifestyle changes and behaviors (eg, diet, exercise) that are also known to lower BP [9,29]. Indeed, prior research has linked weight loss to behaviors such as frequent tracking of exercise and weight [30].

### Study Strengths and Limitations

This was a single-arm study, preventing any determination of cause and effect. However, participants were real-world users of a commercially available digital health program designed for hypertension management; thus, this study provides evidence for the effectiveness of an AI-powered behavioral coaching

program for lowering BP in the target population of interest. Participants were not required to provide socioeconomic information (eg, income, education), which limited insight into potential socioeconomic disparities. Although retention was lower than that in clinical efficacy studies, this is expected for digital health [31], and retention was substantially higher than what is commonly reported for similar programs in the literature [32]. Prior investigations have demonstrated that self-monitoring along with self-titration of medications in collaboration with a treating physician yields robust drops in BP and good retention [33]. Engaging physicians in the member experience could be one way to improve member retention. The exploratory mediation analysis had some inherent limitations: without the ability to infer causal relationships or directionality from the results of this study, this was not “true” mediation. Given the timing of the measurements, we cannot state that engagement caused weight loss, which then caused BP reductions. However, the alternative model (engagement  $\rightarrow$  BP reduction  $\rightarrow$  weight loss) was not significant, which supports the directionality of the relationship between engagement, percent weight loss, and BP change proposed in this study.

### Future Directions

This study was an important first step in demonstrating changes in BP and body weight that occurred during a fully digital hypertension management program. Although we had information on other important BP management strategies at baseline (eg, medication status), we did not have the ability to track changes made to participants' care management that might have occurred during the duration of the study. In future investigations, we intend to examine interactions between program participation and additional aspects of care. Evaluating medication adherence is a new feature within the app, and future investigations will examine whether participating in a fully digital hypertension management program results in improved adherence to prescribed medications. Finally, we did not separately consider coaching sessions per topic area (eg, diet) in the regression analyses due to collinearity issues. However, certain types of coaching may be more important than others. We plan to explore the different factors related to AI coaching in future investigations.

### Conclusions

Members enrolled in a fully digital hypertension care program with coaching powered by AI who provided BP readings during their third and sixth months of program participation achieved clinically meaningful reductions in SBP. The magnitude of the drop depended on starting SBP, with participants classified as hypertension stage 2 experiencing the greatest drop. Most participants remained weight stable or lost weight by month 3 in the program, and the percent weight change at month 3 was significantly associated with program engagement and the likelihood of achieving a drop of  $\geq 5$  mm Hg in SBP. Taken together, these results provide formative evidence that members enrolled in an AI-powered hypertension care program who remain engaged experience clinically meaningful reductions in their systolic BP and body weight.

## Acknowledgments

We thank the team at Lark for their assistance in managing the data collection.

## Authors' Contributions

OHB designed the study and edited the manuscript; MR performed data analyses and edited the manuscript; LAG wrote and edited the manuscript; KGL wrote and edited the manuscript; and SAG designed the study, performed data analyses, and wrote and edited the manuscript.

## Conflicts of Interest

OHB, MR, LAG, KGL, and SAG are employed by Lark.

## Multimedia Appendix 1

Description of the Lark Hypertension Care Program.

[[DOCX File, 2245 KB - formative\\_v6i10e38215\\_app1.docx](#)]

## References

1. Whelton P, Carey R, Aronow W, Casey D, Collins K, Dennison Himmelfarb C, et al. 2017 ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PCNA Guideline for the Prevention, Detection, Evaluation, and Management of High Blood Pressure in Adults: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. *J Am Coll Cardiol* 2018 May 15;71(19):e127-e248 [FREE Full text] [doi: [10.1016/j.jacc.2017.11.006](https://doi.org/10.1016/j.jacc.2017.11.006)] [Medline: [29146535](https://pubmed.ncbi.nlm.nih.gov/29146535/)]
2. Facts about hypertension. Centers for Disease Control and Prevention. URL: <https://www.cdc.gov/bloodpressure/facts.htm> [accessed 2021-09-05]
3. Yusuf S, Joseph P, Rangarajan S, Islam S, Mente A, Hystad P, et al. Modifiable risk factors, cardiovascular disease, and mortality in 155 722 individuals from 21 high-income, middle-income, and low-income countries (PURE): a prospective cohort study. *The Lancet* 2020 Mar;395(10226):795-808. [doi: [10.1016/s0140-6736\(19\)32008-2](https://doi.org/10.1016/s0140-6736(19)32008-2)]
4. Ahmad FB, Anderson RN. The leading causes of death in the US for 2020. *JAMA* 2021 May 11;325(18):1829-1830 [FREE Full text] [doi: [10.1001/jama.2021.5469](https://doi.org/10.1001/jama.2021.5469)] [Medline: [33787821](https://pubmed.ncbi.nlm.nih.gov/33787821/)]
5. High blood pressure. National Heart, Lung, and Blood Institute. URL: <https://tinyurl.com/yrs8j4zm> [accessed 2021-09-05]
6. Blood Pressure Lowering Treatment Trialists' Collaboration. Pharmacological blood pressure lowering for primary and secondary prevention of cardiovascular disease across different levels of blood pressure: an individual participant-level data meta-analysis. *Lancet* 2021 May 01;397(10285):1625-1636 [FREE Full text] [doi: [10.1016/S0140-6736\(21\)00590-0](https://doi.org/10.1016/S0140-6736(21)00590-0)] [Medline: [33933205](https://pubmed.ncbi.nlm.nih.gov/33933205/)]
7. Shimbo D, Artinian NT, Basile JN, Krakoff LR, Margolis KL, Rakotz MK, American Heart Associationthe American Medical Association. Self-measured blood pressure monitoring at home: a joint policy statement from the American Heart Association and American Medical Association. *Circulation* 2020 Jul 28;142(4):e42-e63. [doi: [10.1161/CIR.0000000000000803](https://doi.org/10.1161/CIR.0000000000000803)] [Medline: [32567342](https://pubmed.ncbi.nlm.nih.gov/32567342/)]
8. Tucker KL, Sheppard JP, Stevens R, Bosworth HB, Bove A, Bray EP, et al. Self-monitoring of blood pressure in hypertension: A systematic review and individual patient data meta-analysis. *PLoS Med* 2017 Sep;14(9):e1002389 [FREE Full text] [doi: [10.1371/journal.pmed.1002389](https://doi.org/10.1371/journal.pmed.1002389)] [Medline: [28926573](https://pubmed.ncbi.nlm.nih.gov/28926573/)]
9. Carey RM, Whelton PK, 2017 ACC/AHA Hypertension Guideline Writing Committee. Prevention, detection, evaluation, and management of high blood pressure in adults: synopsis of the 2017 American College of Cardiology/American Heart Association Hypertension Guideline. *Ann Intern Med* 2018 Mar 06;168(5):351-358 [FREE Full text] [doi: [10.7326/M17-3203](https://doi.org/10.7326/M17-3203)] [Medline: [29357392](https://pubmed.ncbi.nlm.nih.gov/29357392/)]
10. Poorolajal J, Hooshmand E, Bahrami M, Ameri P. How much excess weight loss can reduce the risk of hypertension? *J Public Health (Oxf)* 2017 Sep 01;39(3):e95-e102. [doi: [10.1093/pubmed/fdw077](https://doi.org/10.1093/pubmed/fdw077)] [Medline: [27521927](https://pubmed.ncbi.nlm.nih.gov/27521927/)]
11. Landi F, Calvani R, Picca A, Tosato M, Martone AM, Ortolani E, et al. Body mass index is strongly associated with hypertension: results from the longevity check-up 7+ study. *Nutrients* 2018 Dec 13;10(12):1976-1988 [FREE Full text] [doi: [10.3390/nu10121976](https://doi.org/10.3390/nu10121976)] [Medline: [30551656](https://pubmed.ncbi.nlm.nih.gov/30551656/)]
12. Hall ME, Cohen JB, Ard JD, Egan BM, Hall JE, Lavie CJ, American Heart Association Council on Hypertension, Thrombosis Vascular Biology, Stroke Council, Council on Lifestyle Cardiometabolic Health. Weight-loss strategies for prevention and treatment of hypertension: a scientific statement from the American Heart Association. *Hypertension* 2021 Nov;78(5):e38-e50. [doi: [10.1161/HYP.0000000000000202](https://doi.org/10.1161/HYP.0000000000000202)] [Medline: [34538096](https://pubmed.ncbi.nlm.nih.gov/34538096/)]
13. Mertens I, Van Gaal LF. Overweight, obesity, and blood pressure: the effects of modest weight reduction. *Obes Res* 2000 May;8(3):270-278 [FREE Full text] [doi: [10.1038/oby.2000.32](https://doi.org/10.1038/oby.2000.32)] [Medline: [10832771](https://pubmed.ncbi.nlm.nih.gov/10832771/)]
14. O'Connor EA, Evans CV, Rushkin MC, Redmond N, Lin JS. Behavioral counseling to promote a healthy diet and physical activity for cardiovascular disease prevention in adults with cardiovascular risk factors: updated evidence report and



- systematic review for the US Preventive Services Task Force. *JAMA* 2020 Nov 24;324(20):2076-2094. [doi: [10.1001/jama.2020.17108](https://doi.org/10.1001/jama.2020.17108)] [Medline: [33231669](https://pubmed.ncbi.nlm.nih.gov/33231669/)]
15. Brandt CJ, Søgaaard GI, Clemensen J, Søndergaard J, Nielsen JB. Determinants of successful eHealth coaching for consumer lifestyle changes: qualitative interview study among health care professionals. *J Med Internet Res* 2018 Jul 05;20(7):e237 [FREE Full text] [doi: [10.2196/jmir.9791](https://doi.org/10.2196/jmir.9791)] [Medline: [29980496](https://pubmed.ncbi.nlm.nih.gov/29980496/)]
  16. Lv N, Azar KM, Rosas LG, Wulfovich S, Xiao L, Ma J. Behavioral lifestyle interventions for moderate and severe obesity: A systematic review. *Prev Med* 2017 Jul;100:180-193 [FREE Full text] [doi: [10.1016/j.ypmed.2017.04.022](https://doi.org/10.1016/j.ypmed.2017.04.022)] [Medline: [28450123](https://pubmed.ncbi.nlm.nih.gov/28450123/)]
  17. Thomopoulos C, Parati G, Zanchetti A. Effects of blood pressure lowering on outcome incidence in hypertension: 7. Effects of more vs. less intensive blood pressure lowering and different achieved blood pressure levels - updated overview and meta-analyses of randomized trials. *J Hypertens* 2016 Apr;34(4):613-622. [doi: [10.1097/HJH.0000000000000881](https://doi.org/10.1097/HJH.0000000000000881)] [Medline: [26848994](https://pubmed.ncbi.nlm.nih.gov/26848994/)]
  18. Monitoring your blood pressure at home. American Heart Association. 2017. URL: <https://www.heart.org/en/health-topics/high-blood-pressure/understanding-blood-pressure-readings/monitoring-your-blood-pressure-at-home> [accessed 2021-11-21]
  19. Jacob A, Moullec G, Lavoie KL, Laurin C, Cowan T, Tishaw C, et al. Impact of cognitive-behavioral interventions on weight loss and psychological outcomes: A meta-analysis. *Heal Psychol* 2018;37(7):417-432. [doi: [10.1037/hea0000576](https://doi.org/10.1037/hea0000576)] [Medline: [29698017](https://pubmed.ncbi.nlm.nih.gov/29698017/)]
  20. Celano CM, Gianangelo TA, Millstein RA, Chung WJ, Wexler DJ, Park ER, et al. A positive psychology–motivational interviewing intervention for patients with type 2 diabetes: Proof-of-concept trial. *Int J Psychiatry Med* 2019;54(2):97-114. [doi: [10.1177/0091217418791448](https://doi.org/10.1177/0091217418791448)] [Medline: [30114958](https://pubmed.ncbi.nlm.nih.gov/30114958/)]
  21. Barnes RD, Barnes V. A systematic review of motivational interviewing for weight loss among adults in primary care. *Obesity reviews* 2015 Mar 05;16(4):304-318. [doi: [10.1111/obr.12264](https://doi.org/10.1111/obr.12264)] [Medline: [25752449](https://pubmed.ncbi.nlm.nih.gov/25752449/)]
  22. Löwe B, Kroenke K, Gräfe K. Detecting and monitoring depression with a two-item questionnaire (PHQ-2). *J Psychosom Res* 2005 Feb;58(2):163-171. [doi: [10.1016/j.jpsychores.2004.09.006](https://doi.org/10.1016/j.jpsychores.2004.09.006)] [Medline: [5820844](https://pubmed.ncbi.nlm.nih.gov/5820844/)]
  23. Fairchild AJ, McDaniel HL. Best (but oft-forgotten) practices: mediation analysis. *Am J Clin Nutr* 2017 Jun;105(6):1259-1271 [FREE Full text] [doi: [10.3945/ajcn.117.152546](https://doi.org/10.3945/ajcn.117.152546)] [Medline: [28446497](https://pubmed.ncbi.nlm.nih.gov/28446497/)]
  24. Etminani K, Tao Engström A, Göransson C, Sant'Anna A, Nowaczyk S. How behavior change strategies are used to design digital interventions to improve medication adherence and blood pressure among patients with hypertension: systematic review. *J Med Internet Res* 2020 Apr 09;22(4):e17201 [FREE Full text] [doi: [10.2196/17201](https://doi.org/10.2196/17201)] [Medline: [32271148](https://pubmed.ncbi.nlm.nih.gov/32271148/)]
  25. Zhou D, Xi B, Zhao M, Wang L, Veeranki SP. Uncontrolled hypertension increases risk of all-cause and cardiovascular disease mortality in US adults: the NHANES III Linked Mortality Study. *Sci Rep* 2018 Jun 20;8(1):9418 [FREE Full text] [doi: [10.1038/s41598-018-27377-2](https://doi.org/10.1038/s41598-018-27377-2)] [Medline: [29925884](https://pubmed.ncbi.nlm.nih.gov/29925884/)]
  26. Bundy JD, Li C, Stuchlik P, Bu X, Kelly TN, Mills KT, et al. Systolic blood pressure reduction and risk of cardiovascular disease and mortality: a systematic review and network meta-analysis. *JAMA Cardiol* 2017 Jul 01;2(7):775-781 [FREE Full text] [doi: [10.1001/jamacardio.2017.1421](https://doi.org/10.1001/jamacardio.2017.1421)] [Medline: [28564682](https://pubmed.ncbi.nlm.nih.gov/28564682/)]
  27. Defining Adult Overweight & Obesity. Centers for Disease Control and Prevention. URL: <https://www.cdc.gov/obesity/adult/defining.html> [accessed 2021-11-19]
  28. Aronow WS. Association of obesity with hypertension. *Ann Transl Med* 2017 Sep;5(17):350 [FREE Full text] [doi: [10.21037/atm.2017.06.69](https://doi.org/10.21037/atm.2017.06.69)] [Medline: [28936444](https://pubmed.ncbi.nlm.nih.gov/28936444/)]
  29. Ozemek C, Tiwari S, Sabbahi A, Carbone S, Lavie CJ. Impact of therapeutic lifestyle changes in resistant hypertension. *Prog Cardiovasc Dis* 2020;63(1):4-9 [FREE Full text] [doi: [10.1016/j.pcad.2019.11.012](https://doi.org/10.1016/j.pcad.2019.11.012)] [Medline: [31756356](https://pubmed.ncbi.nlm.nih.gov/31756356/)]
  30. Pourzanjani A, Quisel T, Foschini L. Adherent use of digital health trackers is associated with weight loss. *PLoS One* 2016;11(4):e0152504 [FREE Full text] [doi: [10.1371/journal.pone.0152504](https://doi.org/10.1371/journal.pone.0152504)] [Medline: [27049859](https://pubmed.ncbi.nlm.nih.gov/27049859/)]
  31. Amagai S, Pila S, Kaat AJ, Nowinski CJ, Gershon RC. Challenges in participant engagement and retention using mobile health apps: literature review. *J Med Internet Res* 2022 Apr 26;24(4):e35120 [FREE Full text] [doi: [10.2196/35120](https://doi.org/10.2196/35120)] [Medline: [35471414](https://pubmed.ncbi.nlm.nih.gov/35471414/)]
  32. Pratap A, Neto EC, Snyder P, Stepnowsky C, Elhadad N, Grant D, et al. Indicators of retention in remote digital health studies: a cross-study evaluation of 100,000 participants. *NPJ Digit Med* 2020;3:21 [FREE Full text] [doi: [10.1038/s41746-020-0224-8](https://doi.org/10.1038/s41746-020-0224-8)] [Medline: [32128451](https://pubmed.ncbi.nlm.nih.gov/32128451/)]
  33. McManus RJ, Mant J, Haque MS, Bray EP, Bryan S, Greenfield SM, et al. Effect of self-monitoring and medication self-titration on systolic blood pressure in hypertensive patients at high risk of cardiovascular disease: the TASMIN-SR randomized clinical trial. *JAMA* 2014 Aug 27;312(8):799-808. [doi: [10.1001/jama.2014.10057](https://doi.org/10.1001/jama.2014.10057)] [Medline: [25157723](https://pubmed.ncbi.nlm.nih.gov/25157723/)]

## Abbreviations

- AHA:** American Heart Association  
**AI:** artificial intelligence  
**BP:** blood pressure  
**DBP:** diastolic blood pressure

**SBP:** systolic blood pressure

*Edited by A Mavragani; submitted 23.03.22; peer-reviewed by K Margolis, PH Lin; comments to author 30.04.22; revised version received 09.09.22; accepted 26.09.22; published 27.10.22.*

*Please cite as:*

*Branch OH, Rikhy M, Auster-Gussman LA, Lockwood KG, Graham SA*

*Relationships Between Blood Pressure Reduction, Weight Loss, and Engagement in a Digital App-Based Hypertension Care Program: Observational Study*

*JMIR Form Res 2022;6(10):e38215*

*URL: <https://formative.jmir.org/2022/10/e38215>*

*doi: [10.2196/38215](https://doi.org/10.2196/38215)*

*PMID: [36301618](https://pubmed.ncbi.nlm.nih.gov/36301618/)*

©OraLee H Branch, Mohit Rikhy, Lisa A Auster-Gussman, Kimberly G Lockwood, Sarah A Graham. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 27.10.2022. 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

# Screening for Generalized Anxiety Disorder From Acoustic and Linguistic Features of Impromptu Speech: Prediction Model Evaluation Study

Bazen Gashaw Teferra<sup>1</sup>, BSc, MSc; Sophie Borwein<sup>2</sup>, BA, MPP, PhD; Danielle D DeSouza<sup>3,4</sup>, MSc, PhD; Jonathan Rose<sup>1,5</sup>, BSc, MSc, PhD

<sup>1</sup>The Edward S Rogers Sr Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON, Canada

<sup>2</sup>School of Public Policy, Simon Fraser University, Vancouver, BC, Canada

<sup>3</sup>Winterlight Labs, Toronto, ON, Canada

<sup>4</sup>Department of Neurology and Neurological Sciences, Stanford University, Palo Alto, CA, United States

<sup>5</sup>The Centre for Addiction and Mental Health, Toronto, ON, Canada

**Corresponding Author:**

Bazen Gashaw Teferra, BSc, MSc

The Edward S Rogers Sr Department of Electrical and Computer Engineering

University of Toronto

10 King's College Road

Toronto, ON, M5S 3G4

Canada

Phone: 1 4169786992

Email: [bazen.teferra@mail.utoronto.ca](mailto:bazen.teferra@mail.utoronto.ca)

## Abstract

**Background:** Frequent interaction with mental health professionals is required to screen, diagnose, and track mental health disorders. However, high costs and insufficient access can make frequent interactions difficult. The ability to assess a mental health disorder passively and at frequent intervals could be a useful complement to the conventional treatment. It may be possible to passively assess clinical symptoms with high frequency by characterizing speech alterations collected using personal smartphones or other wearable devices. The association between speech features and mental health disorders can be leveraged as an objective screening tool.

**Objective:** This study aimed to evaluate the performance of a model that predicts the presence of generalized anxiety disorder (GAD) from acoustic and linguistic features of impromptu speech on a larger and more generalizable scale than prior studies did.

**Methods:** A total of 2000 participants were recruited, and they participated in a single web-based session. They completed the Generalized Anxiety Disorder-7 item scale assessment and provided an impromptu speech sample in response to a modified version of the Trier Social Stress Test. We used the linguistic and acoustic features that were found to be associated with anxiety disorders in previous studies along with demographic information to predict whether participants fell above or below the screening threshold for GAD based on the Generalized Anxiety Disorder-7 item scale threshold of 10. Separate models for each sex were also evaluated. We reported the mean area under the receiver operating characteristic (AUROC) from a repeated 5-fold cross-validation to evaluate the performance of the models.

**Results:** A logistic regression model using only acoustic and linguistic speech features achieved a significantly greater prediction accuracy than a random model did (mean AUROC 0.57, SD 0.03;  $P < .001$ ). When separately assessing samples from female participants, we observed a mean AUROC of 0.55 (SD 0.05;  $P = .01$ ). The model constructed from the samples from male participants achieved a mean AUROC of 0.57 (SD 0.07;  $P = .002$ ). The mean AUROC increased to 0.62 (SD 0.03;  $P < .001$ ) on the all-sample data set when demographic information (age, sex, and income) was included, indicating the importance of demographics when screening for anxiety disorders. The performance also increased for the female sample to a mean of 0.62 (SD 0.04;  $P < .001$ ) when using demographic information (age and income). An increase in performance was not observed when demographic information was added to the model constructed from the male samples.

**Conclusions:** A logistic regression model using acoustic and linguistic speech features, which have been suggested to be associated with anxiety disorders in prior studies, can achieve above-random accuracy for predicting GAD. Importantly, the addition of basic demographic variables further improves model performance, suggesting a role for speech and demographic information to be used as automated, objective screeners of GAD.

(*JMIR Form Res* 2022;6(10):e39998) doi:[10.2196/39998](https://doi.org/10.2196/39998)

## KEYWORDS

mental health; generalized anxiety disorder; impromptu speech; acoustic features; linguistic features; anxiety prediction; mobile phone

## Introduction

### Background

Anxiety disorders are characterized by an excessive and uncontrollable fear of what is to come and are associated with preparation for possible future adverse events [1]. Although anxiety is an important emotion that helps us prepare for future events, it limits the performance of day-to-day tasks when it becomes uncontrollable. Anxiety disorders are one of the most common mental health issues with an incidence of approximately 10% in the Canadian population [2]. Unfortunately, many Canadians affected by anxiety are unable to access psychological and psychiatric resources [3] due in part to the cost [4] and the general lack of availability [5]. Some of this deficit may be addressed using methods that automate certain aspects of the measurement and diagnosis of anxiety disorders.

In this study, we focused on generalized anxiety disorder (GAD) [6] and sought to automatically detect GAD from speech. We monitored speech because it is possible to passively and frequently sample ambient speech, ensuring that the privacy and confidentiality of the participants are appropriately handled. The capability to detect anxiety from ambient speech could be part of a system to automatically screen for anxiety, monitor treatment, and detect relapse.

We anticipated the following scenario for a system that included the capability to automatically predict anxiety from speech. Such a system would sample a sequence of the participant's speech throughout the day and produce multiple predictions. Depending on the accuracy of individual predictions, the system could use multiple predictions to increase the overall accuracy of the final screening result. Note that this approach uses passively collected speech, which gives rise to its own challenges, including the need for a process of speaker identification to select words spoken by the participant.

Another motivation for pursuing the automatic detection of anxiety from speech is to avoid the subjectivity normally present in the screening and diagnosis of GAD. The current gold standard diagnosis for GAD is influenced by both the subject information supplied by a patient to a clinician and the subjective judgment of that information by the clinician. This subjectivity can lead to inaccurate diagnostic outcomes in patients [7]. There is a potential benefit of an objective marker of anxiety. In this study, we explored how well such a biomarker could be obtained from a person's speech. Prior studies suggest that anxiety influences the acoustic features of speech [8], as

these features are difficult for a person to control [9]. Moreover, anxiety may also manifest in the choice of words, which we refer to as linguistic features [10].

In an earlier study [11], we identified acoustic and linguistic features of speech that were correlated with anxiety as measured by the Generalized Anxiety Disorder 7-item (GAD-7) scale. Building on our earlier study and using the same participants and data, in this study, we aimed to build and measure the performance of a model that predicts whether participants are above or below the screening threshold for GAD on a much larger scale than what prior studies did. Previous studies validating the GAD-7 scale have shown that using a cut point of  $\geq 10$  optimizes sensitivity and specificity for identifying individuals with a diagnosis of GAD [12]. The model makes use of the previously identified features of speech together with the demographics of the participants.

This paper is organized as follows: the next subsection summarizes related work on anxiety prediction and proposes a hypothesis. The Methods section describes the speech sample collection methods, set of features used, and construction and evaluation of predictive models. The Results section presents the demographics of the participants and performance of the prediction model, while the Discussion section discusses the results and their implications for future research on anxiety detection.

### Related Work

Several previous studies have measured the association between speech features and various forms of anxiety, and other efforts sought to automatically detect anxiety from acoustic and linguistic features of speech. The studies explored in this section examine a broader class of anxiety disorders, including internalizing disorders, social phobia or social anxiety disorders (SADs), panic disorder, agoraphobia, and GAD.

McGinnis et al [13] identified several acoustic characteristics of speech that can be used to detect anxiety disorders in children. They studied the speech of 71 participants between the ages of 3 and 8 years and were able to detect internalizing disorders (a collective term for anxiety and depression). The authors extracted and selected several acoustic features from the speech produced in a 3-minute task based on the Trier Social Stress Test (TSST) for children [14]. These features included the zero-crossing rate, Mel frequency cepstral coefficients [15], zero-crossing rate of the  $z$  score of the power spectral density, dominant frequency, mean frequency, perceptual spectral centroid, spectral flatness, and the skew and kurtosis of the power spectral density. Several models were built to predict

whether children had an internalizing disorder (43/71, 61%) or were healthy. Both logistic regression (LR) and a support vector machine (SVM) analyses [16] achieved a classification accuracy of 80%.

Weeks et al [17] found a relationship between anxiety and voice alterations. Their study showed a link between vocal pitch (characterized by fundamental frequency [F0]) and SAD. They collected impromptu speech samples from 46 undergraduate students, 25 (54%) with a diagnosis of SAD and 21 (46%) healthy controls. The participants also completed the Beck Anxiety Scale as a measure of self-reported anxiety severity [18]. Their results indicated that the mean F0 was positively correlated ( $r=0.72$ ;  $P=.002$ ) with anxiety severity in all male participants. However, the correlation in female participants was weaker ( $r=0.02$ ;  $P=.92$ ), indicating possible sex differences in the relationship between anxiety severity and vocal pitch. In a related continuing study [19], the authors attempted to classify men with SAD using the mean F0. Using a mean F0 value of 122.78 Hz, they achieved a sensitivity of 89% (8/9 male patients with SAD correctly classified) and a specificity of 100% (4/4 male healthy controls correctly classified).

Salekin et al [20] explored methods for detecting social anxiety and depression from an audio clip of a person's speech. Their data set included a 3-minute speech sample from each of the 105 participants describing what they liked and disliked about college or their hometown. The participants were asked to report their peak levels of anxiety during the speech. The authors presented and used a novel feature modeling technique called NN2Vec that can identify the relationship between a participant's speech and affective states. Using the features from NN2Vec and a bidirectional Long Short-Term Memory Multiple Instance Learning network, they were able to detect speakers with high social anxiety with an  $F_1$ -score of 90.1% and speakers with depression symptoms with an  $F_1$ -score of 85.4%.

Baird et al [21] explored the effect of anxiety on speech by attempting to predict anxiety using sustained vowels. Their data set comprised 239 speakers (69 male participants) aged 18 to 68 years who performed various vocal exercises, which included sustained vowel sounds. They used the Beck Anxiety Inventory (BAI) [22] questionnaire as a label for each participant. The BAI is also one of the scales used to screen for GAD. They used 4 classes of sustained (*a*) vowels from each participant: a sad phonation, a smiling phonation, a comfortable phonation, and a powerful or loud phonation. From the sustained vowels, they extracted acoustic features such as the SD of F0, intensity, and harmonic-to-noise ratio. Using these features and a BAI label, they trained a support vector regressor with a linear kernel and used Spearman correlation between the predicted and the actual label to evaluate the performance of their model. They split their data into training and test sets and achieved a Spearman correlation of 0.243 on the test data set. They reported a better performance of a Spearman correlation of 0.59 when they only considered the group with high BAI scores, indicating that the symptoms of anxiety are more observable in individuals with high anxiety.

Rook et al [23] hypothesized that the worrying behavior in GAD comes from the verbal linguistic process. They attempted to

predict GAD using only linguistic patterns. A total of 142 undergraduate students (56 male and 86 female participants) were recruited and asked to recall and write down an anxious experience during their university life. Each participant filled out the GAD-7 scale score and the behavioral inhibition/behavioral approach scale (BIS/BAS) [24]. The Linguistic Inquiry and Word Count (LIWC) [25] method was used to extract features from the texts written by the participants. Another set of features was also used by combining the LIWC features with BIS/BAS scores. Several machine learning models were explored, including SVM with linear kernel, LR, naïve Bayes, and random forest. Their results showed that all the models performed significantly better than a random model. In addition, better performance was obtained from all the models except the SVM when the LIWC and BIS/BAS features were used together as inputs compared with using only the LIWC features.

Di Matteo et al [26] examined the relationship between passively collected audio data and anxiety and depression. Their study continued for 2 weeks, where 84 participants installed an app on their smartphone that collected the average volume of sounds (the average of 15-second audio collected every 5 minutes) and the presence or absence of speech in the environment. They then extracted 4 environmental audio-based features: daily similarity, sleep disturbance (on all nights and weeknights only), and speech-presence ratio. Their results showed that none of the extracted features were significantly correlated with anxiety. However, these features were significantly correlated with depression: daily similarities ( $r=-0.37$ ;  $P<.001$ ), sleep disturbance on weeknights ( $r=0.23$ ;  $P=.03$ ), and speech presence ( $r=-0.37$ ;  $P<.001$ ).

Di Matteo et al [27] also explored the relationship between linguistic features of speech and anxiety. They used passively collected intermittent samples of audio data from participants' smartphones, which they converted to text. The authors used the LIWC approach [25] to classify words into 67 categories. They calculated correlations using 4 self-report measures: SAD, GAD, depression, and functional impairment. They observed a significant correlation between words related to perceptual process (*See* in the LIWC) with SAD ( $r=0.31$ ;  $P=.003$ ) and words related to rewards with GAD ( $r=-0.29$ ;  $P=.007$ ).

In their third study, using the data collected from the 84 participants, Di Matteo et al [28] attempted to predict GAD, SAD, and depression from the smartphone-collected data. The features used in this study included daily similarity, speech presence, weeknight sleep disturbance, death-related words, number of locations visited, number of exits from home, screen use, and time in darkness. Although the models built on these features achieved an above-random prediction accuracy for SAD and depression, they did not observe above-random prediction accuracy for GAD.

Overall, prior studies suggest that it is possible to detect anxiety disorders from speech. However, the largest sample size among these previous studies was a total of 239, with an average of 115 participants, which limits the generalizability of the results. In addition, the number of participants might not be the only factor affecting generalizability. Apart from the studies by Di

Matteo et al [28] and Baird et al [21], the prior studies were mostly limited to very specific demographics: McGinnis et al [13] focused on children; Weeks et al [17], Salekin et al [20], and Rook et al [23] focused on undergraduate students at a university or college.

We hypothesized that by recruiting a substantially larger cohort (N=2000) with broader demographic characteristics than that in prior studies, it is possible to achieve above-random prediction accuracy in screening for GAD using acoustic and linguistic features that have been previously suggested.

## Methods

### Data Collection

#### Recruitment and Demographics

We must note that the participants recruited and the data used in this study are the same as those in our earlier study [11], which focused solely on the correlations between acoustic and linguistic features of speech and the GAD-7. This study used those features and additional demographics to construct a predictive model. Participants were recruited from a nonclinical population using Prolific [29], a web-based human participant recruitment platform. The inclusion criteria for this study were an age range of 18 to 65 years, fluency in English, English as a first language, and at least 10 previous studies completed on Prolific, with 95% of these previous Prolific tasks completed satisfactorily (as labeled by the study author). The Prolific platform also provided several relevant demographics of the participants, including their age and income. The data set was also balanced for sex (50% female and 50% male).

Participants who completed the study were paid £2 (approximately CAD \$3.41; US \$2.74) for approximately 15 minutes of work. They completed the entire study remotely, using their PCs.

#### Study Procedure

Participants were recruited on Prolific for a 10- to 15-minute task implemented through a custom website. Our earlier paper on the correlates of anxiety [11] described the data collection procedure in detail. Parts of the data collection procedure that are relevant for the purpose of this study are described in the following sections.

On the Prolific platform, individuals who met the inclusion criteria were presented with the opportunity to participate in this study. Those who wished to participate clicked on the study link, which brought them to a consent form that described the procedure and goals of the study and provided information on data privacy. If a participant provided consent, a hyperlink brought them to an external web app that implemented the tasks described in further sections.

Participants were asked to fill out the standard GAD-7 questionnaire [12] described in more detail in the Anxiety Measures section. They were then asked to perform a speech task, which was recorded using their computer microphone. The speech task followed a modified version of the widely used TSST [30], which aimed to evoke a moderate amount of stress

from each participant. Prior studies [31,32] have shown higher activation (cardiovascular, skin conductance, and plasma levels of norepinephrine and testosterone) in participants with relatively higher anxiety after exposure to moderate stress induced by the TSST.

In the modified version of TSST, participants were told to imagine that they were job applicants invited for an interview with a hiring manager. They were told to imagine that it was a job that they really wanted—their so-called “dream” job. They were given a few minutes to prepare—to choose their “dream” job—and to think about how they would convince an interviewer that they were the right person for that position. Participants were also told that the recorded video would be viewed by researchers studying their behavior and language. Participants were then asked to speak for 5 minutes, making the case for themselves to be hired for that dream job.

Note that in the original TSST [30], participants would normally deliver their speech in front of a live panel of judges. If a participant finished their delivery in <5 minutes, the judges in the original TSST design would encourage the participant to speak for the full 5 minutes. For example, a statement of encouragement in the original TSST was, “What are your personal strengths?” In the modified TSST, we implemented a similar method to encourage participants to speak for the full 5 minutes; when our system detected silence (the absence of speech for >6 seconds), it would display several different prompts inviting participants to keep speaking on different topics related to the task. Finally, the modified TSST only included the first part of the original TSST, not the second task, which involved the performance of mental arithmetic.

#### Anxiety Measures

We aimed to predict, based on features of the speech, if a participant is above or below the screening threshold for GAD based on the GAD-7 scale. The GAD-7 [12] scale is a 7-item questionnaire that asks participants how often they were bothered by anxiety-related problems during the previous 2 weeks. Although the 2-week time period suggests that the GAD-7 measures a temporary condition, this is in contrast to the fact that a GAD diagnosis requires a 6-month duration of symptoms [33,34]. However, the GAD-7 has been validated as a diagnostic tool for GAD using a cutoff threshold of 10, with a sensitivity of 89% and specificity of 82% [12]. Thus, we chose to use the GAD-7 threshold of 10 to obtain a binary label of GAD as the indicator of anxiety.

Each of the 7 questions on the GAD-7 has 4 options for the participant to select, indicating how often they have been bothered by the 7 problems in the scale. These options and their numerical ratings are as follows: 0=not at all, 1=several days, 2=more than half the days, and 3=nearly every day. The final GAD-7 score is a summation of the values for each question, giving a severity measure for GAD ranging from 0 (no anxiety symptoms) to 21 (severe anxiety symptoms).

#### Separation of Data for Analysis

Certain demographic attributes were directly indicative of anxiety. For example, sex is known to influence the prevalence of anxiety [35]. In addition, both age [36] and income [37]

influence anxiety. Owing to the strong effect of sex and our interest in analyzing the effect of anxiety on both the sexes separately, we created a separate data set for female and male participants, in addition to the combined data set.

### Inputs to the Classification Model

The inputs to our models were acoustic and linguistic features that were determined in a previous study [11] to have a statistically significant correlation with the GAD-7. These features were found to be correlated with the GAD-7 after controlling for demographic variables such as age, sex, and

personal income. These features are presented in [Table 1](#) as all-sample, female sample, and male sample data sets. The definitions of these features are presented in [Multimedia Appendix 1](#).

In addition to the acoustic and linguistic features, we explored the use of demographic information, such as age, sex, and personal income, as input features to the model. We decided to use these demographics as features in the model because they were available to us from the Prolific recruitment platform [29]. Should the model be used as a diagnostic screener in the future, it should be possible to obtain these demographics.

**Table 1.** Correlation of statistically significant acoustic and linguistic features with Generalized Anxiety Disorder-7 item (GAD-7) scale—results taken from earlier study (N=1744).

Feature	Data set					
	All-sample		Female sample (n=862)		Male sample (n=882)	
	<i>r</i>	<i>P</i> value	<i>r</i>	<i>P</i> value	<i>r</i>	<i>P</i> value
AllPunc	0.13	<.001	0.14	<.001	0.13	<.001
WC	-0.12	<.001	-0.13	<.001	-0.12	<.001
Speaking duration	-0.12	<.001	-0.11	<.001	-0.13	<.001
Period	0.12	<.001	0.16	<.001	0.08	.02
assent	0.10	<.001	0.10	.004	0.11	.001
negemo	0.10	<.001	0.11	<.001	0.08	.01
relativ	-0.09	<.001	-0.09	.006	-0.10	.002
motion	-0.08	<.001	-0.10	.003	-0.07	.048
Shimmer	0.08	<.001	0.10	.004	0.07	.04
swear	0.08	<.001	— <sup>a</sup>	—	0.10	.004
anger	0.08	<.001	0.11	.002	—	—
mfcc_std_2	-0.08	.002	—	—	-0.09	.005
mfcc_std_3	-0.07	.002	-0.10	.002	—	—
focusfuture	-0.07	.003	—	—	-0.08	.02
mfcc_mean_2	-0.07	.004	-0.09	.01	—	—
adverb	-0.07	.004	-0.11	<.001	—	—
time	-0.07	.004	—	—	-0.10	.004
function	-0.07	.005	—	—	—	—
negate	0.07	.006	—	—	0.08	.01
prep	-0.06	.007	—	—	-0.08	.02
WPS	-0.06	.007	-0.07	.03	—	—
anx	0.06	.008	—	—	0.08	.01
f0_std	0.06	.01	—	—	0.07	.04
hear	0.06	.01	—	—	0.10	.003
mfcc_std_5	-0.06	.01	—	—	-0.09	.01
death	0.06	.01	0.07	.04	—	—
ipron	-0.06	.01	—	—	-0.07	.04
see	-0.06	.01	-0.09	.006	—	—
affect	0.06	.02	—	—	0.07	.04
i	0.05	.02	—	—	—	—
family	0.05	.02	—	—	0.08	.02
mfcc_std_4	-0.05	.03	—	—	-0.07	.04
sad	0.05	.03	0.08	.01	—	—
ppron	0.05	.03	—	—	0.09	.01
space	-0.05	.04	—	—	—	—
article	-0.05	.04	—	—	-0.08	.01
leisure	0.05	.04	—	—	0.10	.002
friend	0.05	.047	—	—	—	—
lpcc_std_6	—	—	-0.09	.008	—	—



Feature	Data set					
	All-sample		Female sample (n=862)		Male sample (n=882)	
	<i>r</i>	<i>P</i> value	<i>r</i>	<i>P</i> value	<i>r</i>	<i>P</i> value
lpcc_std_4	—	—	-0.09	.008	—	—
intensity_mean	—	—	-0.09	.01	—	—
mfcc_mean_1	—	—	-0.09	.01	—	—
Dic	—	—	-0.08	.02	—	—
power	—	—	0.07	.03	-0.09	.01
lpcc_std_10	—	—	-0.07	.03	—	—
intensity_std	—	—	-0.07	.03	—	—
lpcc_std_12	—	—	-0.07	.04	—	—
mfcc_mean_8	—	—	0.07	.04	—	—
percept	—	—	-0.07	.046	—	—
lpcc_mean_4	—	—	0.07	.049	—	—
Apostro	—	—	—	—	0.09	.005
Sixltr	—	—	—	—	-0.09	.01
mfcc_mean_5	—	—	—	—	-0.08	.02
mfcc_std_11	—	—	—	—	-0.07	.045
f1_mean	—	—	—	—	0.07	.047

<sup>a</sup>Not available because the correlation was not significant.

## Construction and Evaluation of Classification Models

In this study, we aimed to evaluate the performance of a binary classifier that predicts if a person's speech sample is in the "anxious" or "nonanxious" class based on the features of speech. The binary classification label is determined by processing the GAD-7 scale (which ranges from 0 to 21 in value) into 2 classes, anxious (GAD-7 $\geq$ 10) and nonanxious (GAD-7 $<$ 10), where 10 is a well-established screening threshold [12] for GAD.

An LR model was trained on the training data to make predictions between the 2 classes. The construction and evaluation steps were as follows. First, the input features were normalized, so that each feature would have a mean of 0 and an SD of 1. Next, the data were undersampled to equalize the representations from both the anxious and nonanxious classes. This avoided the problem of class imbalance, which, if occurred, caused low predictive accuracy for the minority class (which was the anxious class in our case). Therefore, samples were randomly selected and removed from the majority class until the majority class had an equal number of samples to the minority class.

The model construction and training step used 3 data sets: a training data set, which was used to train the model; a validation data set, which was used to select the best hyperparameters during training; and a test data set, which was used to evaluate the performance of the trained model using area under the receiver operating characteristic (AUROC) metrics. These data sets were created within each sampling of the cross-validation (CV) scheme described next.

The CV scheme used a nested resampling with 2-level nested CVs—one CV nested within another [38]. In the outer loop, the data were split into 20% test data and 80% training and validation data. In the inner loop, 80% of the training and validation data were further split into 20% validation data and 80% training data. The inner loop was repeated 5 times, each with a different sampling to obtain a different 20/80 split. For each such split, the best hyperparameters were selected to maximize the accuracy of the validation data after training on the training data of the inner loop. After selecting the best hyperparameters from the inner CV loop, training was once again performed on the entire 80% of the outer loop training plus validation data, and the mean AUROC results were reported on the test data of the outer loop. The outer loop was iterated 5 times, each time selecting a different 20% for test data, until all the samples were left out and tested. This whole process was repeated 7 times, each with different random undersampling seeds, where in each of the 7 iterations, 5 AUROC were reported from the outer CV loop, giving a total of 35 values. The mean and SD of the AUROC values were used as the final metrics in this study to measure performance.

## Feature Selection

During the construction of the model, a subset of features was selected from the features listed in Table 1. The goal of feature selection was to avoid using duplicate information (where the same information was present in different features) and maximize the prediction performance of our model.

To avoid the use of duplicate information, we first calculated the intercorrelations between all the features presented in Table

1. We then used only one of each pair of the highly correlated ( $r > 0.8$ ) features.

In the model construction, it might not always be true that using all the available features maximizes the prediction accuracy; doing so may actually reduce accuracy owing to overfitting [39]. Thus, to maximize the prediction performance of our model, we selected a subset of features using the following method: we began with the single feature that had the highest correlation with the GAD-7 and then measured the prediction performance of a trained model (on a validation data) using only that feature. Subsequent features were then added one-by-one in order of correlation until all the significant features (presented in Table 1) were used (ie, until adding 1 more feature no longer improved the prediction performance).

### Statistical Analysis

To evaluate the performance of the prediction models, the mean AUROC of the 35 models was compared with the mean AUROC of a model that made a random prediction (ie, mean AUROC close to 0.5) using a modified 1-tailed  $t$  test developed by Bouckaert and Frank [40]. The modified  $t$  test considers the fact that the individual AUROC values are not independent from each other, whereas in the original  $t$  test, the samples are expected to be independent. In our case, because the AUROC generated from a model shared some training data (owing to multiple undersampling and the 5-fold CV) with another, the AUROC values were not independent of each other. In our results, we considered a statistically significant difference at a  $P$  value significance level of .05.

### Ethics Approval

This study was approved by the University of Toronto Research Ethics Board (protocol #37584).

## Results

### Recruitment and Data Inclusion

A total of 2000 participants provided acceptable submissions from November 23, 2020, to May 28, 2021, and thus received payments. We reviewed the input data and audio for quality and included 1744 participants in the analysis. A detailed description of recruitment and data quality filtering was provided in our previous study [11].

### Data Overview and Demographics of Participants

Of the 1744 participants, 540 (30.96%) were above the GAD-7 screening threshold of 10 and 1204 (69.04%) were below the GAD-7 screening threshold of 10. Hereon, we will refer to those participants with a GAD-7 score  $\geq 10$  as the *anxious class* and those with a GAD-7 score  $< 10$  as the *nonanxious class*.

Table 2 shows participant demographics obtained from the Prolific recruitment platform. Column 1 of Table 2 provide the names of demographic attributes and each category, while columns 2 and 3 give the number (and percentage) of participants with that attribute in the anxious and nonanxious groups, respectively. The last column gives the  $P$  values for the chi-square test of the null hypothesis that the difference in categories is independent, to determine if there is a significant difference between the anxious and nonanxious groups, for each categorical factor.

**Table 2.** Demographics split by anxious and nonanxious label and chi-square test (N=1744).

Demographic factors	Anxious, n (%) (n=540)	Nonanxious, n (%) (n=1204)	Chi-square test, <i>P</i> value
<b>Sex</b>			<.001
Male	229 (26)	653 (74)	
Female	311 (36.1)	551 (63.9)	
<b>Self-reported ongoing mental health illness or condition</b>			<.001
Yes	297 (48.8)	311 (51.2)	
No	243 (21.4)	893 (78.6)	
<b>Personal income (£<sup>a</sup>)</b>			<.001
<10,000	181 (39.2)	281 (60.8)	
10,000-19,999	112 (35)	208 (65)	
20,000-29,999	92 (26.2)	259 (73.8)	
30,000-39,999	60 (24.6)	184 (75.4)	
40,000-49,999	36 (24.8)	109 (75.2)	
50,000-59,999	20 (21.3)	74 (78.7)	
≥60,000	39 (30.5)	89 (69.5)	
<b>Age (years)</b>			<.001
18-19	27 (38)	44 (62)	
20-29	239 (38.7)	379 (61.3)	
30-39	162 (32.7)	334 (67.3)	
40-49	67 (23.4)	219 (76.6)	
50-59	39 (22.8)	132 (77.2)	
≥60	6 (5.9)	96 (94.1)	

<sup>a</sup>1 £=US \$1.37.

### Classification Model Performance

In this section, the mean AUROC of a binary classification model that classified between anxious and nonanxious classes is presented. The following subsections summarize our main empirical results for different types of inputs to the classification models.

### Acoustic and Linguistic Features as Input

The mean AUROC for the model constructed using a subset of the acoustic and linguistic features selected using the feature selection method described in the Methods section is reported. As described in the Methods section, the features that contain very similar information were not used based on the intercorrelation between the features. Table 3 presents the features with high intercorrelation ( $r > 0.8$ ) between the features presented in Table 1. We only considered using one of each pair of the highly intercorrelated features.

**Table 3.** Features with high intercorrelation (similar features) with each other.

Sample and feature	<i>r</i>	<i>P</i> value
<b>All samples</b>		
AllPunc, Period	0.93	<.001
i, ppron	0.81	<.001
<b>Female samples</b>		
AllPunc, Period	0.93	<.001
Intensity_mean, intensity_std	0.93	<.001
<b>Male samples</b>		
AllPunc, Period	0.93	<.001

As described in the *Methods* section, the acoustic and linguistic features starting with the feature with the highest correlation were included in the model, incrementally, if they showed improvement in the performance of the model. Table 4 shows the subset of features used for the 3 data sets, and Figure 1 shows the mean AUROC as a function of the number of selected features. Using the feature selection method discussed in the *Methods* section, the number of features required to produce

the maximum mean AUROC was 11 for the all-sample, 7 for the female sample, and 11 for the male sample data set, as shown in Table 4. The best model is the one that included all the features listed in Table 4 (according to the data set).

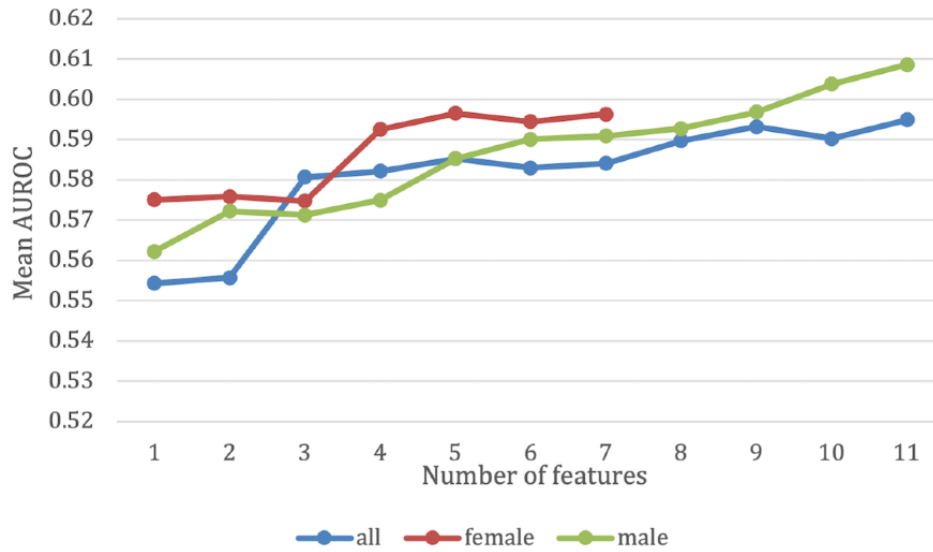
Table 5 shows the mean AUROC across the 35 data splits, as described in the *Methods* section. It also provides results of the 1-tailed *t* test comparison of the best model with that of a random model.

**Table 4.** Subset of acoustic and linguist features used in the 3 models after feature selection.

Feature	Data set					
	All-sample		Female sample		Male sample	
	<i>r</i>	<i>P</i> value	<i>r</i>	<i>P</i> value	<i>r</i>	<i>P</i> value
AllPunc	0.13	<.001	— <sup>a</sup>	—	—	—
Assent	0.1	<.001	—	—	0.11	.001
Relativ	-0.09	<.001	—	—	-0.1	.002
Motion	-0.08	<.001	-0.1	.003	—	—
mfcc_std_2	-0.08	.002	—	—	—	—
mfcc_std_3	-0.07	.002	—	—	—	—
Focusfuture	-0.07	.003	—	—	-0.08	.02
mfcc_std_5	-0.06	.01	—	—	—	—
Death	0.06	.01	0.07	.04	—	—
See	-0.06	.01	-0.09	.006	—	—
mfcc_std_4	-0.05	.045	—	—	—	—
Period	—	—	0.16	<.001	—	—
Dic	—	—	-0.08	.02	—	—
Power	—	—	0.07	.03	—	—
lpcc_std_10	—	—	-0.07	.03	—	—
speaking_duration	—	—	—	—	-0.13	<.001
Leisure	—	—	—	—	0.1	.002
Time	—	—	—	—	-0.1	.004
Ppron	—	—	—	—	0.09	.01
Negemo	—	—	—	—	0.08	.01
Article	—	—	—	—	-0.08	.01
mfcc_mean_5	—	—	—	—	-0.08	.01
f1_mean	—	—	—	—	0.07	.047

<sup>a</sup>Not available because the correlation was not significant.

**Figure 1.** Mean area under the receiver operating characteristic (AUROC) as a function of the number of selected features.



**Table 5.** Mean area under the receiver operating characteristic (AUROC) of a model trained using the subset of features (N=1744).

Data set	AUROC, mean (SD)	t test (df)	P value
All-sample	0.59 (0.02)	4.93 (34)	<.001
Female sample (n=862)	0.60 (0.04)	4.25 (34)	<.001
Male sample (n=882)	0.61 (0.06)	4.21 (34)	<.001

**Using Participants’ Demographics**

This section presents the performance of the model when augmented with age, sex, and income demographic information. Table 6 shows the mean AUROC of the LR model that used both demographic information and acoustic and linguistic features. It also included a modified t test comparison with a random model.

Table 7 shows the results of the t test between the model with only acoustic or linguistic features and the model that also used demographic information. Table 8 separates each of the demographic features and shows the mean AUROC of these models when using a single demographic at a time, together with the acoustic and linguistic features.

**Table 6.** Mean area under the receiver operating characteristic (AUROC) of a model trained using demographic information (age, sex, and income) in addition to the acoustic and linguistic features and comparison with a random model (N=1744).

Data set	AUROC, mean (SD)	t test (df)	P value
All-sample	0.64 (0.03)	6.21 (34)	<.001
Female sample (n=862)	0.66 (0.04)	5.89 (34)	<.001
Male sample (n=882)	0.62 (0.07)	4.36 (34)	<.001

**Table 7.** Comparison of model trained using only acoustic or linguistic features with model that also uses demographic information (N=1744).

Data set	AUROC <sup>a</sup> , mean (SD)		t test (df)	P value
	Acoustic and linguistic features	Demographics, acoustic and linguistic features		
All-sample	0.59 (0.02)	0.64 (0.03)	4.01 (34)	<.001
Female sample (n=862)	0.60 (0.04)	0.66 (0.04)	4.21 (34)	<.001
Male sample (n=882)	0.61 (0.06)	0.62 (0.07)	0.76 (34)	.45

<sup>a</sup>AUROC: area under the receiver operating characteristic.

**Table 8.** Mean area under the receiver operating characteristic (AUROC) of the model when adding a single demographic characteristic to the acoustic and linguistic features (N=1744).

Data set	AUROC of model using acoustic and linguistic features and including <i>only one of the demographics</i> , mean (SD)		
	Age	Income	Sex
All-sample	0.64 (0.03)	0.6 (0.02)	0.59 (0.02)
Female sample (n=862)	0.66 (0.04)	0.6 (0.05)	N/A <sup>a</sup>
Male sample (n=882)	0.62 (0.07)	0.61 (0.06)	N/A

<sup>a</sup>N/A: not applicable.

## Discussion

The main objective of this study was to investigate the prediction performance of a model that screens for GAD from acoustic and linguistic features of impromptu speech. To do so, we have explored an LR model, and in the following subsections, we discuss the findings presented in the *Results* section, as well as the limitations of this study.

### Principal Findings

#### Recruitment and Data Inclusion

As the study continued from November 23, 2020, to May 28, 2021, the recruitment took place during the global COVID-19 pandemic. We speculated that this might have resulted in an above-normal number of participants who work remotely using their personal computers, hence making web-based recruitments relatively quicker.

#### Demographics of Participants

The percentage of anxious and nonanxious participants shows that the anxious group made up to 30.96% (540/1744) of the total, which is much higher than the general population rate of 10% [2]. Previous studies [11,26-28,41] that also used participants recruited from Prolific showed a higher number of anxious participants in the recruitment pool. Table 2 sheds some light on this difference, showing that a similarly high fraction of participants self-reported on their Prolific profile that they have an ongoing mental health condition.

We also aimed to obtain broader demographics than those in the prior study. Most prior studies focused on a certain type of demographics, such as the study by McGinnis et al [13], which focused on children, and the studies by Week et al [17], Salekin et al [20], and Rook et al [23], which focused on undergraduate students. Both these types significantly limited the age range of the participants. The data presented in Table 2 show that the age range of our participants had a broader distribution. The same is true for personal income, which showed a range of economic status in our participant pool.

#### Acoustic and Linguistic Features as Input

The LR used statistically significant acoustic and linguistic features selected by the feature selection method discussed in the *Methods* section and presented in Table 4. Although the correlation between the features used and the GAD-7 was very small (the highest being 0.13), the model built using these features was able to perform significantly better (with  $P < .05$ ) than a random model. The mean AUROC results presented in

Table 5 suggest that there is some signal to be detected from the combined effect of the acoustic and linguistic features of speech.

Although it is possible to use the GAD-7 scale to screen for GAD (it has a sensitivity of 89% and a specificity of 82% [12]), it cannot serve the purpose of our study, which is a continuous and passive monitoring of a participant. By contrast, an automated screener that listens passively to speech has the potential to frequently monitor speech samples from participants. Furthermore, the probability of correct prediction can be improved by using multiple measurements under the assumption that each measurement from different speech is relatively independent.

This enhanced accuracy could be achieved by considering the model's native accuracy as follows: let the accuracy of a correct prediction from a single measurement be  $a$ , and we take  $N$  successive measurements, based on  $N$  successive speech samples, using our model. As a decision procedure, we would decide that most of the measurement is correct—whichever result, anxious or nonanxious, happens in more than  $N/2$  of the measurements. We were interested in the probability that this decision procedure will produce a correct result. The probability that  $n$  or more of the  $N$  measurements would have a correct prediction can be calculated using the cumulative binomial distribution function (Inline graphic 1). Given the decision procedure of taking the correct result to be the majority result of the  $N$  trials, we set the value of  $n$  to be  $N/2$ , which computes the probability of more than  $N/2$  correct answers. As long as the single prediction  $a$  is  $>0.5$ , the computed probability  $A$  will be  $>a$ . It should be noted that this result does rely on the assumption that the measurements are independent when, in reality, they are not because the measurements were taken from the same person. However, the set of words coming from the person was different, and more spaced-out measurements might have reduced the dependency between the samples. To summarize, it is possible to increase the accuracy of correct predictions by taking multiple measurements and taking the class (anxious or nonanxious) that has been predicted most of the time as the final predicted value.



#### Participants' Demographics as Input Features

In a scenario in which an anxiety screening or prediction model can be deployed, an individual's demographic information can easily be collected. Thus, it is reasonable to explore the predictive capability of this additional information. Table 6

shows that a model built using demographic information as input in addition to the acoustic and linguistic features was still able to perform significantly better than a random model. [Table 7](#) compares this model with a model that used only acoustic and linguistic features as input. The results show that the demographic information significantly improved the mean AUROC of the models built on the all-sample and female sample data sets but failed on the model built on the male sample data set.

The impact of each demographic variable was explored separately. [Table 8](#) shows the mean AUROC of the model when only one of the demographics was used together with the acoustic and linguistic features. The addition of age affected the prediction performance of the model, whereas the addition of either sex or income did not show a significant improvement. In addition, the fact that the addition of age affected the prediction performance for the model built on female participants suggests that their anxiety depends on age compared with anxiety in male participants.

Comparing our results with a prior study that aimed to predict GAD, there were studies that achieved the above-random prediction for GAD [[21,23](#)], and there were studies that did not [[28](#)]. Our models performed significantly better than the random model, and we speculated that this might be attributed to the larger and demographically broad sample size that enabled our model to learn a large amount of information in predicting GAD. We also note that a prior study that did not succeed in predicting GAD [[28](#)] did, in fact, succeed in predicting SAD. They believed that the symptoms of SAD might be more manifested in the participants' behavior and, therefore speech, compared with GAD. Other studies focusing on SAD have also been successful in above-random prediction [[17,20](#)].

### Limitations

A limitation of this study arises from the data collection method used with respect to the scenarios of use that were described in the Introduction section. We suggested that the prediction of anxiety from speech could be applied to passively collected speech data gathered while the patient is going through their daily activities. This could help in automated anxiety screening, treatment monitoring, and relapse detection. However, the data used in this study were actively collected when the participants spoke in front of a camera, and it may be substantially different from such passively collected speech. Future studies could investigate the models that we suggest using passively collected speech.

Another limitation was the use of a web-based participant recruitment method. Individuals willing to work on a web-based participant recruitment platform may be limited to a particular type of demographics in a certain society. For example, we noted that in our recruitment pool, there was a higher percentage of anxious participants compared with the general population. In our study, we sought generalizability, and even though our participants were more diverse in terms of demographics compared with prior studies, it could be more generalizable if we recruited participants from sources other than the web-based recruitment.

Another limitation was the artificial setup used to replicate the original TSST. In the original TSST, participants described their dream job in front of a live panel of judges. Owing to the restrictions that the COVID-19 pandemic had caused, we were not able to recruit participants for an in-person study; instead, we had participants describe their dream job in front of a camera at their own location (with different recording devices). Despite its limitations, this approach also had important benefits because it enabled us to recruit a large number of participants, which would otherwise have been extremely difficult for an in-person study.

### Conclusions

In this study, we developed a model to predict the presence or absence of GAD based on the speech features. These speech features were chosen because prior studies have suggested that they are associated with other types of anxiety disorders including GAD. Our results have shown that it is possible to achieve the above-random prediction accuracy for GAD from the acoustic and linguistic features of speech while using a larger and more generalizable sample size. Prediction accuracy can also be further improved by adding basic demographic information. Even though we have investigated adding 3 different types of demographic variables (age, sex, and income), the most influential variable that showed improvement in prediction accuracy was age.

Furthermore, we have discussed that the results from multiple measurements have the possibility to improve prediction accuracy. Therefore, we recommend that future studies explore the collection of multiple speech samples sampled throughout the day or week and investigate the extent to which the prediction accuracy can be improved. This will allow for the acoustic and linguistic features of speech, together with basic demographic information, to be used in a system to trigger early intervention, monitor treatment responses, or detect relapses.

---

### Acknowledgments

This research was funded by a University of Toronto XSeed grant, Natural Sciences and Engineering Research Council of Canada Discovery grant RGPIN-2019-04395, and Social Sciences and Humanities Research Council Partnership Engage grant 892-2019-0011.

The authors are also grateful to Professor Ludovic Rheault for his energy and assistance with the launch of this project and the data collection and advice.

---

## Conflicts of Interest

None declared.

## Multimedia Appendix 1

Description of the acoustic and linguistic features.

[[PDF File \(Adobe PDF File\). 65 KB - formative\\_v6i10e39998\\_app1.pdf](#)]

## References

1. Barlow DH. Anxiety and Its Disorders: The Nature and Treatment of Anxiety and Panic. New York, NY, USA: Guilford Press; 1988.
2. Mental Health - Anxiety Disorders. Public Health Canada. 2009 Jul. URL: <https://www.canada.ca/en/health-canada/services/healthy-living/your-health/diseases/mental-health-anxiety-disorders.html> [accessed 2022-01-24]
3. Roberge P, Fournier L, Duhoux A, Nguyen CT, Smolders M. Mental health service use and treatment adequacy for anxiety disorders in Canada. *Soc Psychiatry Psychiatr Epidemiol* 2011 Apr;46(4):321-330. [doi: [10.1007/s00127-010-0186-2](https://doi.org/10.1007/s00127-010-0186-2)] [Medline: [20217041](https://pubmed.ncbi.nlm.nih.gov/20217041/)]
4. Koerner N, Dugas MJ, Savard P, Gaudet A, Turcotte J, Marchand A. The economic burden of anxiety disorders in Canada. *Can Psychol* 2004 Aug;45(3):191-201. [doi: [10.1037/h0088236](https://doi.org/10.1037/h0088236)]
5. Elliott KP, Hunsley J. Evaluating the measurement of mental health service accessibility, acceptability, and availability in the Canadian Community Health Survey. *Am J Orthopsychiatry* 2015 May;85(3):238-242. [doi: [10.1037/ort0000062](https://doi.org/10.1037/ort0000062)] [Medline: [25894307](https://pubmed.ncbi.nlm.nih.gov/25894307/)]
6. Hidalgo RB, Sheehan DV. Generalized anxiety disorder. *Handb Clin Neurol* 2012;106:343-362. [doi: [10.1016/B978-0-444-52002-9.00019-X](https://doi.org/10.1016/B978-0-444-52002-9.00019-X)] [Medline: [22608630](https://pubmed.ncbi.nlm.nih.gov/22608630/)]
7. Brown TA, Di Nardo PA, Lehman CL, Campbell LA. Reliability of DSM-IV anxiety and mood disorders: implications for the classification of emotional disorders. *J Abnorm Psychol* 2001 Feb;110(1):49-58. [doi: [10.1037//0021-843x.110.1.49](https://doi.org/10.1037//0021-843x.110.1.49)] [Medline: [11261399](https://pubmed.ncbi.nlm.nih.gov/11261399/)]
8. Thompson AR. Pharmacological agents with effects on voice. *Am J Otolaryngol* 1995;16(1):12-18. [doi: [10.1016/0196-0709\(95\)90003-9](https://doi.org/10.1016/0196-0709(95)90003-9)] [Medline: [7717466](https://pubmed.ncbi.nlm.nih.gov/7717466/)]
9. Hoehn-Saric R, McLeod DR. The peripheral sympathetic nervous system. Its role in normal and pathologic anxiety. *Psychiatr Clin North Am* 1988 Jun;11(2):375-386. [Medline: [3047706](https://pubmed.ncbi.nlm.nih.gov/3047706/)]
10. Pennebaker JW, Mehl MR, Niederhoffer KG. Psychological aspects of natural language. Use: our words, our selves. *Annu Rev Psychol* 2003;54:547-577. [doi: [10.1146/annurev.psych.54.101601.145041](https://doi.org/10.1146/annurev.psych.54.101601.145041)] [Medline: [12185209](https://pubmed.ncbi.nlm.nih.gov/12185209/)]
11. Teferra BG, Borwein S, DeSouza DD, Simpson W, Rheault L, Rose J. Acoustic and linguistic features of impromptu speech and their association with anxiety: validation study. *JMIR Ment Health* 2022 Jul 08;9(7):e36828 [FREE Full text] [doi: [10.2196/36828](https://doi.org/10.2196/36828)] [Medline: [35802401](https://pubmed.ncbi.nlm.nih.gov/35802401/)]
12. Spitzer RL, Kroenke K, Williams JB, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch Intern Med* 2006 May 22;166(10):1092-1097. [doi: [10.1001/archinte.166.10.1092](https://doi.org/10.1001/archinte.166.10.1092)] [Medline: [16717171](https://pubmed.ncbi.nlm.nih.gov/16717171/)]
13. McGinnis EW, Anderau SP, Hruschak J, Gurchiek RD, Lopez-Duran NL, Fitzgerald K, et al. Giving voice to vulnerable children: machine learning analysis of speech detects anxiety and depression in early childhood. *IEEE J Biomed Health Inform* 2019 Nov;23(6):2294-2301 [FREE Full text] [doi: [10.1109/JBHI.2019.2913590](https://doi.org/10.1109/JBHI.2019.2913590)] [Medline: [31034426](https://pubmed.ncbi.nlm.nih.gov/31034426/)]
14. Buske-Kirschbaum A, Jobst S, Wustmans A, Kirschbaum C, Rauh W, Hellhammer D. Attenuated free cortisol response to psychosocial stress in children with atopic dermatitis. *Psychosom Med* 1997;59(4):419-426. [doi: [10.1097/00006842-199707000-00012](https://doi.org/10.1097/00006842-199707000-00012)] [Medline: [9251162](https://pubmed.ncbi.nlm.nih.gov/9251162/)]
15. Ali S, Tanweer S, Khalid SS, Rao N. Mel frequency cepstral coefficient: a review. In: Proceedings of the 2nd International Conference on ICT for Digital, Smart, and Sustainable Development. 2020 Presented at: ICIDSSD '20; February 27-28, 2020; New Delhi, India. [doi: [10.4108/eai.27-2-2020.2303173](https://doi.org/10.4108/eai.27-2-2020.2303173)]
16. Cortes C, Vapnik V. Support-vector networks. *Mach Learn* 1995 Sep;20(3):273-297. [doi: [10.1007/BF00994018](https://doi.org/10.1007/BF00994018)]
17. Weeks JW, Lee CY, Reilly AR, Howell AN, France C, Kowalsky JM, et al. "The Sound of Fear": assessing vocal fundamental frequency as a physiological indicator of social anxiety disorder. *J Anxiety Disord* 2012 Dec;26(8):811-822. [doi: [10.1016/j.janxdis.2012.07.005](https://doi.org/10.1016/j.janxdis.2012.07.005)] [Medline: [23070030](https://pubmed.ncbi.nlm.nih.gov/23070030/)]
18. Julian LJ. Measures of anxiety: State-Trait Anxiety Inventory (STAI), Beck Anxiety Inventory (BAI), and Hospital Anxiety and Depression Scale-Anxiety (HADS-A). *Arthritis Care Res (Hoboken)* 2011 Nov;63 Suppl 11:S467-S472 [FREE Full text] [doi: [10.1002/acr.20561](https://doi.org/10.1002/acr.20561)] [Medline: [22588767](https://pubmed.ncbi.nlm.nih.gov/22588767/)]
19. Weeks JW, Srivastav A, Howell AN, Menatti AR. "Speaking more than words": classifying men with social anxiety disorder via vocal acoustic analyses of diagnostic interviews. *J Psychopathol Behav Assess* 2015 May 30;38(1):30-41. [doi: [10.1007/s10862-015-9495-9](https://doi.org/10.1007/s10862-015-9495-9)]
20. Salekin A, Eberle JW, Glenn JJ, Teachman BA, Stankovic JA. A weakly supervised learning framework for detecting social anxiety and depression. *Proc ACM Interact Mob Wearable Ubiquitous Technol* 2018 Jun;2(2):81 [FREE Full text] [doi: [10.1145/3214284](https://doi.org/10.1145/3214284)] [Medline: [31187083](https://pubmed.ncbi.nlm.nih.gov/31187083/)]



21. Baird A, Cummins N, Schnieder S, Krajewski J, Schuller BW. An evaluation of the effect of anxiety on speech — computational prediction of anxiety from sustained vowels. In: Proceedings of 2020 Interspeech. 2020 Presented at: InterSpeech '20; October 25-29, 2020; Shanghai, China p. 4951-4955. [doi: [10.21437/interspeech.2020-1801](https://doi.org/10.21437/interspeech.2020-1801)]
22. Beck AT, Epstein N, Brown G, Steer R. Beck Anxiety Inventory. Washington, DC, USA: American Psychological Association; 1988.
23. Rook L, Mazza MC, Lefter I, Brazier F. Toward linguistic recognition of generalized anxiety disorder. *Front Digit Health* 2022 Apr 15;4:779039 [FREE Full text] [doi: [10.3389/fdgth.2022.779039](https://doi.org/10.3389/fdgth.2022.779039)] [Medline: [35493530](https://pubmed.ncbi.nlm.nih.gov/35493530/)]
24. Carver CS, White TL. Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: the BIS/BAS scales. *J Pers Soc Psychol* 1994 Aug;67(2):319-333. [doi: [10.1037/0022-3514.67.2.319](https://doi.org/10.1037/0022-3514.67.2.319)]
25. Pennebaker JW, Boyd RL, Jordan K, Blackburn K. The development and psychometric properties of LIWC2015. The University of Texas at Austin. 2015. URL: [https://repositories.lib.utexas.edu/bitstream/handle/2152/31333/LIWC2015\\_LanguageManual.pdf?sequence](https://repositories.lib.utexas.edu/bitstream/handle/2152/31333/LIWC2015_LanguageManual.pdf?sequence) [accessed 2022-01-24]
26. Di Matteo D, Fotinos K, Lokuge S, Yu J, Sternat T, Katzman MA, et al. The relationship between smartphone-recorded environmental audio and symptomatology of anxiety and depression: exploratory study. *JMIR Form Res* 2020 Aug 13;4(8):e18751 [FREE Full text] [doi: [10.2196/18751](https://doi.org/10.2196/18751)] [Medline: [32788153](https://pubmed.ncbi.nlm.nih.gov/32788153/)]
27. Di Matteo D, Wang W, Fotinos K, Lokuge S, Yu J, Sternat T, et al. Smartphone-detected ambient speech and self-reported measures of anxiety and depression: exploratory observational study. *JMIR Form Res* 2021 Jan 29;5(1):e22723 [FREE Full text] [doi: [10.2196/22723](https://doi.org/10.2196/22723)] [Medline: [33512325](https://pubmed.ncbi.nlm.nih.gov/33512325/)]
28. Di Matteo D, Fotinos K, Lokuge S, Mason G, Sternat T, Katzman MA, et al. Automated screening for social anxiety, generalized anxiety, and depression from objective smartphone-collected data: cross-sectional study. *J Med Internet Res* 2021 Aug 13;23(8):e28918 [FREE Full text] [doi: [10.2196/28918](https://doi.org/10.2196/28918)] [Medline: [34397386](https://pubmed.ncbi.nlm.nih.gov/34397386/)]
29. Palan S, Schitter C. Prolific.ac—a subject pool for online experiments. *J Behav Exp Finance* 2018 Mar;17:22-27. [doi: [10.1016/j.jbef.2017.12.004](https://doi.org/10.1016/j.jbef.2017.12.004)]
30. Kirschbaum C, Pirke KM, Hellhammer DH. The 'Trier Social Stress Test'—a tool for investigating psychobiological stress responses in a laboratory setting. *Neuropsychobiology* 1993;28(1-2):76-81. [doi: [10.1159/000119004](https://doi.org/10.1159/000119004)] [Medline: [8255414](https://pubmed.ncbi.nlm.nih.gov/8255414/)]
31. Gerra G, Zaimovic A, Zambelli U, Timpano M, Reali N, Bernasconi S, et al. Neuroendocrine responses to psychological stress in adolescents with anxiety disorder. *Neuropsychobiology* 2000;42(2):82-92. [doi: [10.1159/000026677](https://doi.org/10.1159/000026677)] [Medline: [10940763](https://pubmed.ncbi.nlm.nih.gov/10940763/)]
32. Jezova D, Makatsori A, Duncko R, Moncek F, Jakubek M. High trait anxiety in healthy subjects is associated with low neuroendocrine activity during psychosocial stress. *Prog Neuropsychopharmacol Biol Psychiatry* 2004 Dec;28(8):1331-1336. [doi: [10.1016/j.pnpbp.2004.08.005](https://doi.org/10.1016/j.pnpbp.2004.08.005)] [Medline: [15588760](https://pubmed.ncbi.nlm.nih.gov/15588760/)]
33. Endler NS, Kocovski NL. State and trait anxiety revisited. *J Anxiety Disord* 2001;15(3):231-245. [doi: [10.1016/s0887-6185\(01\)00060-3](https://doi.org/10.1016/s0887-6185(01)00060-3)] [Medline: [11442141](https://pubmed.ncbi.nlm.nih.gov/11442141/)]
34. First MB, Spitzer RL, Gibbon M, Williams JB. Structured Clinical Interview for DSM-IV Axis I Disorders (SCID-I), Clinician Version. Washington, DC, USA: American Psychiatric Press; 1997.
35. Cameron OG, Hill EM. Women and anxiety. *Psychiatr Clin North Am* 1989 Mar;12(1):175-186. [doi: [10.1016/s0193-953x\(18\)30459-3](https://doi.org/10.1016/s0193-953x(18)30459-3)]
36. Krasucki C, Howard R, Mann A. The relationship between anxiety disorders and age. *Int J Geriatr Psychiatry* 1998 Feb;13(2):79-99. [doi: [10.1002/\(sici\)1099-1166\(199802\)13:2<79::aid-gps739>3.0.co;2-g](https://doi.org/10.1002/(sici)1099-1166(199802)13:2<79::aid-gps739>3.0.co;2-g)] [Medline: [9526178](https://pubmed.ncbi.nlm.nih.gov/9526178/)]
37. Dijkstra-Kersten SM, Biesheuvel-Leliefeld KE, van der Wouden JC, Penninx BW, van Marwijk HW. Associations of financial strain and income with depressive and anxiety disorders. *J Epidemiol Community Health* 2015 Jul;69(7):660-665. [doi: [10.1136/jech-2014-205088](https://doi.org/10.1136/jech-2014-205088)] [Medline: [25636322](https://pubmed.ncbi.nlm.nih.gov/25636322/)]
38. Bischl B, Mersmann O, Trautmann H, Weihs C. Resampling methods for meta-model validation with recommendations for evolutionary computation. *Evol Comput* 2012;20(2):249-275. [doi: [10.1162/EVCO\\_a\\_00069](https://doi.org/10.1162/EVCO_a_00069)] [Medline: [22339368](https://pubmed.ncbi.nlm.nih.gov/22339368/)]
39. Belkin M, Hsu D, Ma S, Mandal S. Reconciling modern machine-learning practice and the classical bias-variance trade-off. *Proc Natl Acad Sci U S A* 2019 Aug 06;116(32):15849-15854 [FREE Full text] [doi: [10.1073/pnas.1903070116](https://doi.org/10.1073/pnas.1903070116)] [Medline: [31341078](https://pubmed.ncbi.nlm.nih.gov/31341078/)]
40. Bouckaert RR, Frank E. Evaluating the replicability of significance tests for comparing learning algorithms. In: Proceedings of the 8th Pacific-Asia Conference on Advances in Knowledge Discovery and Data Mining. 2004 Presented at: PAKDD '04; May 26-28, 2004; Sydney, Australia p. 3-12. [doi: [10.1007/978-3-540-24775-3\\_3](https://doi.org/10.1007/978-3-540-24775-3_3)]
41. Di Matteo D. Inference of anxiety and depression from smartphone-collected data. University of Toronto. 2021. URL: [https://tspace.library.utoronto.ca/bitstream/1807/108894/1/Di\\_Matteo\\_Daniel\\_202111\\_PhD\\_thesis.pdf](https://tspace.library.utoronto.ca/bitstream/1807/108894/1/Di_Matteo_Daniel_202111_PhD_thesis.pdf) [accessed 2022-01-24]

## Abbreviations

- AUROC:** area under the receiver operating characteristic
- BAI:** Beck Anxiety Inventory
- BIS/BAS:** behavioral inhibition/behavioral approach scale
- CV:** cross-validation

**F0:** fundamental frequency  
**GAD:** generalized anxiety disorder  
**GAD-7:** Generalized Anxiety Disorder 7-item  
**LIWC:** Linguistic Inquiry and Word Count  
**LR:** logistic regression  
**SAD:** social anxiety disorder  
**SVM:** support vector machine  
**TSST:** Trier Social Stress Test

*Edited by A Mavragani; submitted 01.06.22; peer-reviewed by B Teachman, C Corcoran; comments to author 10.08.22; revised version received 29.09.22; accepted 30.09.22; published 28.10.22.*

*Please cite as:*

*Teferra BG, Borwein S, DeSouza DD, Rose J*

*Screening for Generalized Anxiety Disorder From Acoustic and Linguistic Features of Impromptu Speech: Prediction Model Evaluation Study*

*JMIR Form Res 2022;6(10):e39998*

*URL: <https://formative.jmir.org/2022/10/e39998>*

*doi: [10.2196/39998](https://doi.org/10.2196/39998)*

*PMID: [36306165](https://pubmed.ncbi.nlm.nih.gov/36306165/)*

©Bazen Gashaw Teferra, Sophie Borwein, Danielle D DeSouza, Jonathan Rose. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 28.10.2022. 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

# Informing mHealth and Web-Based Eating Disorder Interventions: Combining Lived Experience Perspectives With Design Thinking Approaches

Hannah K Jarman<sup>1,2</sup>; Siân A McLean<sup>3</sup>; Rachel Rodgers<sup>4,5</sup>; Matthew Fuller-Tyszkiewicz<sup>1,2</sup>; Susan Paxton<sup>3</sup>; Beth O'Gorman<sup>6</sup>; Emily Harris<sup>6</sup>; Adrian Shatte<sup>7</sup>; Katie Bishop<sup>1</sup>; Tahlia Baumann<sup>8</sup>; Danielle Mahoney<sup>8</sup>; Melissa-Claire Daugelat<sup>9,10</sup>; Zali Yager<sup>11</sup>

<sup>1</sup>School of Psychology, Deakin University, Geelong, Australia

<sup>2</sup>Centre for Social and Early Emotional Development, School of Psychology, Deakin University, Melbourne, Australia

<sup>3</sup>School of Psychology and Public Health, La Trobe University, Melbourne, Australia

<sup>4</sup>Applied Psychology Program for Eating and Appearance Research, Department of Applied Psychology, Northeastern University, Boston, MA, United States

<sup>5</sup>Department of Psychiatric Emergency & Acute Care, Lapeyronie Hospital, Centre Hospitalier Regional Universitaire Montpellier, Montpellier, France

<sup>6</sup>Melbourne School of Psychological Sciences, University of Melbourne, Melbourne, Australia

<sup>7</sup>Department of Planning, Performance & Analytics, James Cook University, Townsville, Australia

<sup>8</sup>School of Psychology, University of Queensland, Brisbane, Australia

<sup>9</sup>Department of Psychosomatic Medicine and Psychotherapy, Medical University Hospital Tübingen, Tübingen, Germany

<sup>10</sup>Centre of Excellence for Eating Disorders, University of Tübingen, Tübingen, Germany

<sup>11</sup>Institute for Health and Sport, Victoria University, Melbourne, Australia

**Corresponding Author:**

Hannah K Jarman

School of Psychology

Deakin University

1 Gheringhap St

Geelong, 3220

Australia

Phone: 61 3 9251 7777

Email: [h.jarman@deakin.edu.au](mailto:h.jarman@deakin.edu.au)

## Abstract

**Background:** App-based interventions designed to prevent and treat eating disorders have considerable potential to overcome known barriers to treatment seeking. Existing apps have shown efficacy in terms of symptom reduction; however, uptake and retention issues are common. To ensure that apps meet the needs and preferences of those for whom they were designed, it is critical to understand the lived experience of potential users and involve them in the process of design, development, and delivery. However, few app-based interventions are pretested on and co-designed with end users before randomized controlled trials.

**Objective:** To address the issue, this study used a highly novel design thinking approach to provide the context and a lived experience perspective of the end user, thus allowing for a deeper level of understanding.

**Methods:** In total, 7 young women (mean age 25.83, SD 5.34, range 21-33 years) who self-identified as having a history of body image issues or eating disorders were recruited. Participants were interviewed about their lived experience of body image and eating disorders and reported their needs and preferences for app-based eating disorder interventions. Traditional (thematic analysis) and novel (empathy mapping; visually depicting and empathizing with the user's personal experience) analyses were performed, providing a lived experience perspective of eating disorders and identifying the needs and preferences of this population in relation to app-based interventions for eating disorders. Key challenges and opportunities for app-based eating disorder interventions were also identified.

**Results:** Findings highlighted the importance of understanding and identifying problematic eating disorder symptoms for the user, helpful practices for recovery that identify personal values and goals, the role of social support in facilitating hope, and aspects of usability to promote continued engagement and recovery.

**Conclusions:** Practical guidance and recommendations are described for those developing app-based eating disorder interventions. These findings have the potential to inform practices to enhance participant uptake and retention in the context of app-based interventions for this population.

(*JMIR Form Res* 2022;6(10):e38387) doi:[10.2196/38387](https://doi.org/10.2196/38387)

## KEYWORDS

eating disorders; app-based intervention; lived experience; design thinking; interviews; young women; co-design; mobile health; mHealth

## Introduction

### Background

Eating disorders are serious, complex, and potentially fatal conditions. Up to 18.6% of women and 6.5% of men experience a diagnosed eating disorder in their lifetime, a figure that has increased over the past 18 years [1], particularly throughout the COVID-19 pandemic [2]. Subclinical or prodromal and clinically diagnosed eating disorders are associated with high levels of medical and psychological comorbidities and poor quality of life and have one of the highest mortality rates among all mental disorders [3-5]. Alongside considerable health costs for individuals and their families, eating disorders also substantially impede productivity, with the cost of disease burden for eating disorders each year estimated at Aus \$69.7 billion (US \$43.6 billion) in 2012 [6] and approximately US \$64.7 billion in 2021 [7].

Early response and treatment are recommended to ensure the best outcomes for individuals [8]; however, there are system-wide and individual-level issues that prevent detection, recognition, diagnosis, and access to treatment for eating disorders [9,10]. Alongside the stigma associated with eating disorders, symptomatic individuals may not realize the severity of their symptoms or may be afraid of change, hence obstructing help seeking [9,11]. Depending on geographic location, health systems vary widely in relation to pathways, costs, and access to diagnosis and treatment [11,12]. Combined, these factors mean that as few as 25% of individuals with eating disorder seek help [13], one of the lowest treatment-seeking rates across mental disorders [14].

Self-help psychoeducational and treatment resources, delivered via the web through digital health platforms (ie, accessed via an app or computer), represent an opportunity for cost-effective access to intervention in a convenient and private manner. These approaches have demonstrated promise in the field of psychiatry (eg, depression and anxiety) [15] and specifically within eating disorders [16,17]. Studies of web-based self-help interventions for eating disorders report moderate acceptability and accessibility [18], with some evidence suggesting that they may be at least as effective as in-person interventions [19,20]. Despite the potential and realized advantages of digital health interventions, limitations exist. Individuals may be ambivalent to change [9] or express concerns regarding data privacy and quality of content within eating disorder apps [16]. For these reasons, mental health and eating disorder apps typically experience low engagement and high attrition rates [15,19].

Although these concerns are present in digital health more broadly, this study will focus specifically on mobile health. Given that app development as a whole is resource-intensive, researchers are encouraged to use user engagement practices to ensure that these investments translate to outcomes [21], particularly given the high levels of attrition in many digital health interventions. Focusing and prioritizing efforts to determine user needs and meeting them in app development will ensure that this investment of time and money will result in increased uptake and use. Although studies have started to explore eating disorder users' preferences [16], reviews have concluded that studies involving end users in the design of such apps is lacking [19] and the few examples that exist do not adhere to best practices (eg, not using succinct and clear communication to end users [22]).

Incorporating voices of lived experience through participatory research and co-design is recognized as being critical in the intervention design process [23]. Design thinking is a novel approach for creative problem-solving, which involves empathizing with future end users, defining the problem, ideating and testing early concepts, and iterating until the product meets accessibility and usability needs [23-25]. This study focused on empathizing with potential end users of eating disorder apps and identifying possible challenges and opportunities within this population.

Empathizing with potential future users is the first phase within all design thinking frameworks. This is typically done by interviewing individuals with lived experience of the condition or topic that is being addressed. The purpose is to better understand the nature of their experiences, implications of the problem, and their attitudes toward possible solutions. Following the interview, the researcher independently develops an empathy map for each user to visually depict the user's personal experience, including what they see, hear, say, do, think, and feel. Then, these are used to elicit deeper values, needs, and motivational drivers that may inspire more innovative solutions [23,26]. Although similar in structure to more traditional qualitative analysis, empathy maps are advantageous in that they provide empathy regarding what the individual sees, feels, and so on, thus allowing the researcher to contextualize participant data relative to the individual's lived experience, rather than coming from the perspective of the researcher. In addition, the empathy mapping approach can be used to identify both challenges faced by the population of interest on a particular issue (ie, pain points) and possible opportunities, including how to overcome these obstacles (ie, gain points). Therefore, empathy maps allow for better understanding of the

context, problems, and needs of the population, while also providing opportunities for addressing the topic or issue.

Design thinking approaches, such as empathy interviews and empathy maps, may be especially helpful when considering individuals with significant body image and eating concerns, including eating disorders, who often have a fear of change or losing control and have ambivalent attitudes toward change [9]. Empathy maps may contribute to knowledge gain regarding users' wants and needs and inform the marketing and dissemination of these apps to address barriers and increase uptake, retention, engagement, and efficacy. Design thinking approaches have been used to inform technology-enabled eating disorder services [27-29] and in the design of a body image and eating disorder prevention program for adolescent boys [30]. However, no study has yet conducted an in-depth examination of the lived experience of individuals with eating disorders to inform app-based intervention development.

## Objectives

Given the opportunities that design thinking may provide [31], this exploratory study aimed to use a novel empathy mapping process to provide insights into end users' perspectives of app-based interventions for eating disorders, including identifying key challenges and opportunities of app-based interventions among this population.

## Methods

### Recruitment

The study forms part of a large project (Web-Based Interventions to Reduce Eating Disorders [WIRED]), which involves the development and evaluation of an eating disorder app. Participants were recruited via posts on the WIRED project's social media accounts, highlighting the value of sharing lived experience, focus on involving end user experiences in app development, and nature and aims of the project. In some instances, these posts were also reshared by partner organizations (eg, the Butterfly Foundation). Young adults were asked to express their interest in participating in the study if they were aged 18 to 35 years, lived in Australia, and had experienced body image issues or eating disorders. Recruitment focused on young adults because (1) eating disorders are prevalent in this age group and (2) this group is likely to be the highest users of an app. Therefore, *lived experience* was self-identified and not reliant on past or current formal diagnoses. Individuals who have not received an eating disorder diagnosis but nonetheless experience significant symptoms are at high risk of not receiving appropriate treatment and are therefore a key group that the to-be-developed app aims to engage. In an adjacent study, young adults were invited to participate in a quantitative survey, in which they were asked to provide demographic details and respond to questions focused on their needs and preferences for an eating disorder app (results presented elsewhere; refer to study by Daugelat et al, unpublished data, May 2020). However, participants who were interviewed for this paper did not necessarily complete that questionnaire.

### Participants

Young Australian women (N=7) who identified as having lived experience of an eating disorder or significant body image concerns provided informed consent to participate in this study. Given the exploratory nature of this study, the sample was not intended to be representative; therefore, a small sample size was deemed acceptable [32]. Furthermore, this sample size is aligned with the number of participants typically engaged in empathy interviews for user experience studies [33]. Of the 7 women, 1 (14%) woman chose not to provide demographic details; therefore, these are reported for the remaining participants (6/7, 86%). This 86% (6/7) of the participants was aged between 21 and 33 (mean 25.83, SD 5.34) years. Of the 6 participants, 5 (83%) lived in Melbourne, Victoria, and 1 (17%) lived in regional New South Wales. All except 1 participant had received a formal eating disorder diagnosis from a health professional (5/6, 83%) and had engaged in treatment, either previously (5/6, 83%) or currently (2/6, 33%).

This study adopted a combined methodology, which incorporated a design thinking approach with traditional qualitative interview research. The semistructured interview schedule is provided in [Multimedia Appendix 1](#). The interviewer (female, Masters of Health Psychology student, aged 22 years, and with recent experience of a subclinical eating disorder) initially asked participants to share their experience of disordered eating and significant body image concerns and how it affected them personally. Following this, questions focused on the web-based environment and the development and use of digital resources such as apps. The aim of the interviews was two-fold: to hear the story and body image or eating disorder journey of the participant and to ask participants what they would have needed or wanted from an app-based intervention around that time.

### Ethics Approval

The Victoria University Human Research Ethics Committee approved all the procedures for conducting this research project (HRE 20-010). Participants provided active consent before completing a web-based survey to collect their demographic and contact details.

### Procedure

Interviews were conducted through the web via Zoom and were transcribed using web-based transcription software. Participants were provided a Aus \$50 (US \$30) gift card to thank them for their time and for sharing their lived experience.

### Data Analysis Strategy

Empathy mapping is the first stage in the design thinking process [23]. Using the interview transcripts, an empathy map was created for each participant to empathize with the individual by viewing the world through their eyes and visually articulating what is known about them and their experience through four key questions: what did they see, what did they say and do, what did they hear, and what did they think and feel. This process provides context with which one can reflect on the deeper needs of individuals, as expressed by the individuals, rather than coming from the perspective of the researcher. Given that the timeline for app development on the project was short,

empathy maps were initially drafted by the interviewer and then extended to provide more detail by the lead author (female, aged 29 years, and with no lived experience of an eating disorder). The last author (female, aged 39 years, and with lived experience of a subclinical eating disorder), who is experienced in both design thinking and eating disorder research, ensured congruence between these stages.

Next, to ensure a lived experience perspective, an empathic lens was used to interpret the interview data, and inductive thematic analysis was performed [34]. This process meant that the lead author used the empathy maps to inform the interpretation of the data, to ensure that the lived experience of the participants was front of mind during the analyses, rather than their own perspective. These analyses were focused on the data that were relevant to intervention and recovery, specifically app-based interventions. An iterative, reflexive process was used, meaning that ideas were formed and developed over time. The lead author identified the initial themes [34]. Then, these were shared with the authors and multiple members of the research team (who were in the early stages of content development for the app). Following discussion and further development, themes were finalized by the lead author.

Alongside each theme, a single pain and gain point was identified to encapsulate key considerations regarding the challenges faced (ie, pain points) and how apps may overcome these issues (ie, gain points). Although pain and gain points traditionally form part of the empathy mapping process, given that the aims of the paper were focused on app-based interventions, interpretation of the pain and gain points was focused specifically on app-based interventions.

## Results

### Empathy Maps

An example of an empathy map from a participant is shown in Figure 1. As described previously, the aim of this process was for the lead author to ensure that the lived experience perspective was considered during the thematic analysis. Furthermore, given that the purpose of the study was focused on app-based interventions rather than describing the lived experience (which is presented in detail elsewhere [35,36]), an overview of the common elements across the empathy maps is provided in Multimedia Appendix 2.

**Figure 1.** An example empathy map, which identifies the lived experience of 1 participant.

<p><b>What did they hear?</b></p> <ul style="list-style-type: none"> <li>• Counsellor told her she looks like she’s in a concentration camp – happy people noticing but unhappy with comparison</li> <li>• Nobody telling her they’re proud of her achievements or that she’s good enough</li> <li>• Self-critical voice: “I need to be better”</li> <li>• Others: “try harder”</li> <li>• Coaches telling her what to eat/not eat</li> <li>• Friends worried for her</li> </ul>	<p><b>What did they say and do?</b></p> <ul style="list-style-type: none"> <li>• “I need to look like that otherwise no one will think I’m as good as them”</li> <li>• “Nobody wants to care for me”</li> <li>• Hiding it from others</li> <li>• Exercises – netball, running, training</li> <li>• Competition within self to do better than before</li> <li>• No eating disorder clinic so in and out of hospital</li> </ul>
<p><b>What did they see?</b></p> <ul style="list-style-type: none"> <li>• Influencers on social media</li> <li>• Model at her school dieting and passing out</li> <li>• Hashtags online, eg, #healthyeating, #fitness</li> <li>• Skinny girls in the netball team</li> <li>• Social media comparison – comments saying “I wish I could look like you” confirmed what she was doing was right</li> <li>• Friends, family, and treatment team giving up on her</li> </ul>	<p><b>What did they think and feel?</b></p> <ul style="list-style-type: none"> <li>• Isolated and lonely</li> <li>• Doesn’t identify with ‘sick’</li> <li>• Thinks she’s just being healthy</li> <li>• Thought her body was the reason she wasn’t good</li> <li>• She wanted support from others with lived experience so she wouldn’t feel so alone and to give hope and encouragement</li> </ul>

### Thematic Analysis and Pain and Gain Points

#### Overview

Following the empathy map process and attending to this lived experience perspective, 4 themes were identified relating to insights into the needs and preferences for app-based eating

disorder interventions among this population. Following each theme, a pain and gain point is described to summarize the personal challenges faced by users and opportunities to overcome these and meet their needs through app-based interventions. The findings are also presented in Table 1. Pseudonyms are used to maintain participants’ anonymity.

**Table 1.** Summary of the pain and gain points, themes, and practical recommendations.

Consideration	Themes	Pain points	Gain points	Recommendations for future apps
Marketing	Recognizing symptoms and attitudes toward accessing help	May not feel as if they are <i>sick enough</i> to warrant attention or treatment	Focus on early intervention to educate and encourage users that they may benefit from an app	<ul style="list-style-type: none"> <li>• Increase awareness about eating disorder symptoms and benefits of early intervention</li> <li>• Self-assessments can provide feedback and indicate potential benefits of early intervention</li> <li>• Careful marketing of an app to encourage, and not discourage, potential users</li> </ul>
Approach	Effective content that reflects key processes in recovery	Fear of losing the control that symptoms provide	Identify personal values and goals to support motivation and recovery	<ul style="list-style-type: none"> <li>• Incorporate reflective questioning</li> <li>• Encourage development of more holistic self-worth</li> <li>• Identify values and future goals</li> </ul>
Inclusions	Social connection to facilitate hope and support	Feelings of isolation and loneliness are common	Create a safe environment, which facilitates a sense of connection and belonging	<ul style="list-style-type: none"> <li>• Use opportunities to help people feel a sense of connection and belonging</li> <li>• Provide data on concurrent users and progress that can help to convey a sense of being in it together</li> <li>• Incorporate stories and journeys from others that can give people hope</li> </ul>
Engagement and retention	Usability and language are key	Fear of failure coupled with self-critical evaluations and low worth	Provide praise and encourage self-compassion to promote continued engagement and recovery	<ul style="list-style-type: none"> <li>• Use compassionate language and framing</li> <li>• Provide a careful balance between establishing credibility and presenting content in a friendly, informal manner</li> <li>• Make the app visually appealing and ensure that the app is easy to navigate</li> <li>• Facilitate flexible access</li> </ul>

### **Theme 1: Recognizing Symptoms and Attitudes Toward Accessing Help**

Findings revealed that none of the participants who had been diagnosed with an eating disorder (0/7, 0%) identified with their diagnosis at the time they were experiencing symptoms. Participants did not think they were “sick enough” to warrant or deserve treatment and, on reflection, had incomplete knowledge of eating disorders and their associated symptoms. For example, media portrayals of eating disorders seemed to form the basis of their understanding of such conditions, where symptoms were perceived as more severe than their own. Furthermore, some participants appeared to normalize these behaviors, often downplaying their symptoms and severity:

*I think I was 16 or 15 when I started engaging in behaviors and stuff, and I didn't really think anything of it either. I hadn't really read much about it or knew much about those things unless it was really severe, and I was like, whoa, that's definitely not me. Like you see on media and stuff, where it's all skeletons and stuff. And then I was just engaging in behaviors for a few years, but not all the time to extremes or anything. [Sophia]*

*So then I kind of battled with it, being like, no, I can't have an eating disorder, I'm just healthy. Maybe I've got orthorexia, that unhealthy obsession with healthy eating. [Sarah; after being diagnosed with anorexia nervosa]*

Other participants also commented that they did not want to take steps to address their eating problems. This stemmed either from fear of making changes and the daunting nature of beginning a recovery journey or from concerns about losing the connection to self and identity that their eating disorder had provided—making them feel special, important, and powerful, often for the first time in their lives:

*Well, I don't think it was that I didn't want to admit that I had a problem, it was that I didn't want to fix it. The anxiety around eating and gaining weight and doing all that sort of stuff you have to do and not being able to exercise and eating around people and things like that was too scary. So, it's easier to just stay in the disorder than to do all of that. [Sophia]*

*And feeling special, I could do something different and lots of people comment on the weight loss, self-control and that reinforces it. I knew that I wasn't normal, but it felt powerful or special for having such good self-control or able to lose weight. All of this type of stuff. The benefits of not being normal were so regularly reinforced that I didn't care. [Lisa]*

This lack of desire to address disordered thoughts and behaviors was a barrier to seeking help or support, and, concerningly, a couple of participants (2/7, 29%) admitted that they did not think they would have downloaded or used an eating disorder app around the time of their diagnosis. Instead, participants suggested that this type of app should be targeted during the early stages of the life course of body image or eating concerns, before the existing symptoms worsened:

*At the start, if I'd been told earlier, I would've downloaded an app like that to try and help. But in the space between when I was already like that, between getting bad and getting treatment, I wouldn't have, because I wouldn't have accepted within myself that there was something wrong. But in treatment, if I had a psychologist or my dietician told me to download an app like that, I would have. [Beatrice]*

*I think the app would be so beneficial really early on. And not when they're engaging in disordered eating, or really that early start. And identifying what disordered eating looks like...I definitely think the psycho ed side of things would have been so helpful, to realize when it was going down a very dangerous track. I think that would be particularly beneficial. [Lisa]*

### **Pain and Gain 1: From Not Wanting Help to Acknowledgment of a Problem**

It appears that this denial of a problem or ambivalence to change may act as a significant barrier to app-based eating disorder interventions (pain 1). This suggests the importance of early intervention and positioning of the app for young people early in the emergence of disordered eating attitudes and behaviors, before symptoms become very severe or the eating disorder becomes part of the individuals' identity. The gain point that emerges from this theme is that young adults may need to be enlightened and educated about the extent of their symptoms to demonstrate that they may be in need of early intervention. For example, questionnaires may be used in marketing, screening, and tailoring of app content to highlight and reinforce the relevance and potential of such a resource to them (gain 1).

### **Theme 2: Effective Content That Reflects Key Processes in Recovery**

Participants discussed a range of approaches that were helpful in their recovery journey and recommended that these could be incorporated into an early intervention app. Of the 7 participants, 6 (86%) reported that using reflective questioning to identify unhelpful thoughts or behaviors was critical for their recovery. This technique provided participants with deeper understanding of the connections between their thoughts and behaviors and allowed them to identify that their eating behavior was disordered. This practice also helped participants to see how these attitudes and behaviors may have been affecting their lives and how they may start to address it:

*Mind mapping...it definitely helped, and that was a lot of what my therapy was based on and trying to find the exits from those loops, the continuous loops...Because I didn't know that was the process going on in my head, until we actually wrote it out. [Beatrice]*

*Understanding your thought processes about how you get to that point about feeling bad about your body...having information about how a certain image or situation can kind of...lead you to having issues with your body and issues with eating. [Alex; when asked what they would like to see in an app]*

Through reflective questioning in recovery, participants realized that their eating disorder stemmed from issues beyond just food or exercise. Eating and exercising behaviors were often being used to exert some control or manage negative emotions. For example, a participant described how bullying had led her to start to look for control in other aspects of her life:

*Basically it began with getting some bullying in primary school and high school. High school bullying is a lot different because girls can be vicious. But I think a friendship breakdown at the start of year 10 sort of all contributed to sort of me wanting some sort of control in my life. [Rachel]*

In addition to identifying problematic thoughts and behaviors, participants also emphasized the need to replace these with more positive and helpful thoughts and behaviors. Providing ideas for how to think differently about the self and body was suggested as particularly valuable. It was clear that many participants (5/7, 71%) judged themselves based solely on their appearance, typically resulting in negative self-evaluations. In contrast, during and after recovery, participants had started to take a more holistic view of themselves and their bodies, which allowed them to feel more comfortable. Having a broader understanding and experience of their bodies led them to a range of helpful perspectives including aspects of positive body image, body neutrality, and enhancing self-worth:

*I know that it can help people who are struggling like I was, with seeing people being comfortable in their bodies and living healthily at any weight...I thought health had to be within the BMI and what you ate and how you exercised. And seeing a lot more information, and people advocating the health at every size, that has also brought a lot of comfort in me, especially now that I'm also studying health science. [Beatrice; talking about body diversity]*

*One thing that actually really helped me I think, [was] being actually neutral about your body and not being so black and white about it. I basically...And I always endeavor every day to just sort of see that this is what my body is and it's fine...Like being able to sort of give resources to help people realize that bodies are not just what they look like. It's a vessel for yourself...Your body is a lot more than what the mainstream media promotes. [Rachel]*

These quotes emphasize the value of focusing on alternative views of the self and one's appearance that does not focus solely on esthetics. By doing so, participants were able to identify and value aspects of themselves that were distinct from their appearance, and this provided them with opportunities to enjoy or be proud of their body, a feeling that was new to them. Other elements that participants found to be helpful in their journey of recovery were identifying other critical aspects of their identity beyond the eating disorder, incorporating a values-based approach, and focusing on the future and potential goals:

*And once I started getting back to uni and having a job, I had other things in my life that weren't about the eating disorder and that is what started to really shift that...Once I made that decision and filled my*



*life with things that were not related to my eating disorder, it just got so much easier. [Lisa]*

*And just confirmation, things to repeat to myself like, is what I'm doing healthy? Is what I'm doing going to bring me a future, a good life? To make me second guess what I'm choosing to do in those moments. That has definitely helped me and trying to look towards the future. [Beatrice]*

### **Pain and Gain 2: Taking Control of Personal Values and Goals**

In this study, young adults with body image and eating concerns frequently used food or exercise as a means to exert control, regulate emotions, or manage stress and anxiety, often when they felt that a situation or aspect of their life was out of their control. The challenge for early interventionists and app developers is that, if these behaviors become regular coping mechanisms, users may fear losing the control that they desire, which may result in a subjective evaluation of change as *risky* and reluctance to engage in intervention content through an app (pain 2). A way to address this is to leverage personal motivations by identifying values and goals that are specific to the user. By concentrating on future potential achievements and outcomes and targeting realistic goals, users may be less frightened and more motivated to initiate their recovery journey, as they know the possible benefits that are personally meaningful to them (gain 2).

### **Theme 3: Social Connection to Facilitate Hope and Support**

Participants discussed a desire for an app to provide a feeling of connection. This was especially pertinent as all participants (7/7, 100%) had experienced some level of loneliness and isolation in the experience of their eating disorder. A number of participants (4/7, 57%) suggested that an app can allow the user to communicate with others, which can facilitate feelings of social support. Although some participants suggested formats within an app including chats, forums, or meetups, most participant (3/4, 75%) also acknowledged that this may be unhelpful or even dangerous if used inappropriately, and thus strict moderation will be needed:

*And just having access to support systems. And also, I don't know, like a chat room or something, where it's just a one-on-one. And if you just feel like you really need to talk to someone there and then, and you're by yourself, just being like, I'll just hop online and talk to someone. [Sarah]*

*I think being able to talk to people who are going through the same thing, or who have, or whatever, is really important, because sometimes you just have questions that no one can answer even if they are a doctor or a therapist, because they've never been through it. So I think having people who can help you in that way is good, but also, you have to be careful as well that they don't say unhelpful things. Which can make it hard, because you have to have someone to moderate it, but then I feel like that might make it harder for people to say something because they know*

*it's being moderated. So, it's, I guess, a fine line between what to do. [Sophia]*

Although existing support, such as professionals or family members, may provide some element of care, it was presumed that they will not understand the issues experienced by participants. Consequently, their opinions seemed less relevant or helpful to participants. In contrast, participants wanted to talk to and hear from others who had first-hand experience. There was some indication that individuals who had been through the same experience will have an overall different and deeper level of understanding and empathy. Participants felt that they will just “get it.” A number of positives of this lived experience perspective were identified by participants, including sharing useful strategies of how individuals dealt with certain situations and providing hope and encouragement for recovery and the future:

*I think I just needed support. And I just needed someone, or people...whether it was a health professional or my friends, just to like, there is going to be an end, and at the end it's going to be okay...So I think definitely hearing from past experiences on how people who have recovered dealt with that, would be really helpful. [Sarah; when asked what they needed when they were in their most vulnerable state]*

*It makes you feel less alone and that you get all of these health professionals saying that it's possible and it's going to be hard, but you don't actually have anyone sitting there saying I've done it, and it's better. In my last admission where the nurses really got stuck into me, one of them actually had lived experience and she was also around my age and she was also in hospital for a little bit. And she was one of the people that definitely helped me get to where I was, where I'm seeing someone who has that future that everyone's making you picture and work for. [Beatrice]*

*And it's proof when you've got so much proof on social media, and I guess in your own head that what you're doing is fine. When you've got someone saying, no, I was in that mindset, it wasn't fine, this is better. That is one of the most important things to me. [Beatrice]*

It appears that this shared lived experience may be an especially powerful aspect of recovery. The idea of hearing from others who had experienced the same concerns also facilitated some level of hope for participants. Participants felt that being able to hear from these people will provide some sense of trust in recovery—that others have done it and that there is light at the end of the tunnel. Having this alternative route and point of view can keep users motivated, knowing that, although it is hard, it is both possible and worthwhile. Seeing a role model may encourage users to achieve this for themselves.

### **Pain and Gain 3: Moving From Loneliness to a Sense of Belonging**

Although intense feelings of loneliness meant that participants desired the ability to communicate with others, regarding an app, this would be challenging to implement in such a way that

it would not cause harm (pain 3). The opportunity from this is that, if app developers can provide helpful content and facilitate a feeling of connection and belonging through an app, this may help users to turn off their inner voice and move away from disordered patterns of thoughts and behaviors. Ultimately, the content, approach, and language used in an app can promote a feeling of belonging and an element of safe social connection, while providing evidence-based and accurate information. This may help users to begin to feel more motivated to engage in the content and develop a sense of hope that they can and will recover (gain 3).

#### **Theme 4: Usability and Language Are Key**

The usability of the app appeared to be very important for participants, including aspects such as content readability and app accessibility. Participants expressed a desire for an app that is intuitive and easy to use. Relatedly, the look and feel of the app are critical factors that likely affect uptake and retention over time. A participant described their experience of using a similar mental health app:

*Well, I know there's one called [eating disorder app name removed for anonymity], which I used with the dietician for a little while. I think its functionality was clunky...it was hard to use and didn't really look that great, so...functionality and look and stuff is probably something that would make a difference to whether I would download it or not if it popped up. [Sophia]*

*But I also think just in terms of access, making things really easy to get through on the app. Rather than a lot of words, because that's really off putting and I definitely wouldn't have like that as a 16 year old. Something really easy to use and not too wordy, or theory based, or jargon based. [Lisa]*

Participants also had the expectation that an eating disorder app should be developed by experts and professionals to ensure its credibility, reliability, and effectiveness and that this would make them more inclined to spend time on an app. This could be achieved by the use of research, connections to universities, and having health professionals directly refer people to the app. The language used within apps also seems to play an important role here, specifically ensuring that users both trust and understand the information being presented and do not feel overwhelmed by it. Achieving this professional yet informal balance will likely encourage continued use and engagement:

*I would be using it as a support and as a friend, so I would want it...the language and stuff to be used, very informal, chatty...Just because then it's not as scary. It's all based on studies and everything, [but has] some sort of informality to it. [Sarah]*

*I think it being suggested to me by health professionals would have been something that I would have been like, okay yeah, I'll give it a whirl. I think I'd have more trust, and I'd feel safer using it. [Sarah]*

*I think if it was promoted in a very compassionate way like, eating is allowed to be enjoyable...You're allowed to feel this way, your body isn't just what you look like. But having it promoted with language in*

*that sort I feel like if that was the case then I might've looked into it. I just think it has to have a lot of sort of compassion focus. [Rachel]*

In comparison with face-to-face support, a key advantage that an app provides is the flexibility and freedom for the user to engage at any time. Participants described how an app should have the ability to be used fluidly to suit the needs of the user:

*I think it has the potential to be really useful and be a success. But it's down to the individual, that's the only thing. I think having an open module with no time limits or anything, that would be preferred by me because there were some days where I didn't have the mental strength to focus on anything. [Sarah]*

*So it's not so much of a structured application or process within the app. I think that would be really appealing for people who were worried about having to commit to something and knowing that they can always, not opt out, but they can always just come back into it. [Alex]*

Participants also indicated that it was important for apps to be personalized or tailored to users' needs. Participants perceived that an app of this nature will require hard work and effort. They described fear of commitment to a long program and concerns regarding the mental capability needed to work through the app content. The emphasis on flexibility suggests that an app will be used more sporadically, rather than, for example, scheduling a certain time to engage each week. The expressed preference for fluidity within an app may also indicate that some users may have difficulty in maintaining motivations or that they only feel moderately committed to working through an app. Aspects that may help to maintain motivation and retention are personalization and relevance for the user within the app and providing some features throughout the app that allow for participant input, including the ability to write ideas, comments, or responses directly into the app and check-in questions for the user:

*I feel like when you can input, like the little faces to check Very Good. I like that. It's also, you can record audio messages, put in photos. I think that's great. And then down the bottom it says like, how much sleep did you get, how much exercise did you do, and how well did you eat? You can just input. I think that's a really good tool. [Claire]*

*I struggle, like [to] be consistent. That's my personal thing. I might do it for a few days and then forget to do it. So like, whether I remind the system or something that would be useful, but that kind of thing, or say you could find a friend to be like, "Oh, have you checked in on the app today?" That kind of accountability thing. [Claire]*

It appears that users may find continued use of an app to be challenging, including maintaining motivation over a prolonged period. Given these concerns, participants welcomed and seemed open to the idea of reminders or check-in notifications used within the app to prompt use. This may also make users more accountable for their engagement.

### ***Pain and Gain 4: Addressing Fear of Failure Through Encouragement and Compassion***

Participants perceived that an app will require significant effort, meaning that some will prefer to use it more sporadically (pain 4). From the empathy mapping process, it appears that this sample typically demonstrated some perfectionistic and competitive qualities and tended to be high achievers, meaning that they often strived to be the best. Therefore, it is likely that they fear feelings of failure when using the app and may be extremely self-critical. In contrast, participants craved and thrived on acknowledgment and praise for their efforts, which some felt that they did not receive during their illness. These types of prompts and rewards can be easily incorporated into apps to foster consistent use and combat lapses, thus ensuring progressive recovery. Given these perfectionistic traits, it may also be important to model a self-compassionate approach and remind users that recovery is neither linear nor straightforward (gain 4). If end users are reminded that setbacks are normal and encouraged to return after periods of less engagement, they may be less likely to drop out or lose interest in an app.

## ***Discussion***

### **Principal Findings**

Web-based and app-based psychological interventions offer some advantages to delivering early interventions and treatment by overcoming barriers related to cost, time, and location and concerns regarding the lack of anonymity and perceived stigma of treatment seeking. However, difficulties such as low uptake and retention remain. Using a novel design thinking approach, this study aimed to provide insights into the lived experience perspectives of potential end users regarding app-based interventions for eating disorders, including identifying key challenges and opportunities. Several issues arose from these insights. First, people with eating disorder will need to recognize that they have a problem that is sufficiently serious to warrant treatment and want to seek help for them to engage with an app. Second, specific content for the app, including processes that enable users to record their experiences, was suggested to facilitate recovery. Third, there was a strong desire for an app to facilitate social connection and sense of belonging. Finally, app usability was identified as important, including the use of compassionate and supportive language and setting a tone that is both credible and friendly.

### **Comparison With Previous Studies**

The first issue that became clear in reaching end users was that many may not know that they have a problem or want help. Individuals with lived experience reported that they would likely not have engaged with this type of app when their eating disorder was most severe or established. As found in previous studies, participants did not believe that their problem was sufficiently serious to warrant treatment [37,38] or they feared losing control [9]. A key challenge for the implementation of an early intervention app is identifying and engaging with those in need, encouraging them to recognize that need, and informing them that they can benefit from using an eating disorder app. To address this, when marketing and disseminating this resource, researchers should aim to increase awareness and understanding

and challenge inaccurate perceptions of eating disorders through psychoeducation, including reference to the broad range of conditions and symptoms. In support of this, a recent Australian study also found that 87% of participants reported a preference for psychoeducation and screening scales to assess symptoms within app-based interventions [16]. Another critical consideration is the framing and *sales pitch* of an app. Researchers developing apps need to think carefully about who the app is intended for and how to engage potential users. For example, is an app intended to be used regularly as a form of treatment to replace or complement therapy, or does it aim to provide in-the-moment support that can be used sporadically? [39] Understanding this and marketing an app accordingly will provide clear expectations to the user, likely resulting in higher retention over time.

Participants identified some particularly helpful activities or perspectives that they experienced, which facilitated their recovery, and recommended that these be incorporated into app content, including verbalizing, writing, and visualizing thoughts, attitudes, and behaviors. These are common in practices such as cognitive behavioral therapy and in other mental health apps, which may theoretically translate well into an app context. A recent systematic review of evidence-based treatment elements in eating disorder apps found that cognitive restructuring was included in only 21% of apps, despite 56% of active monthly users using this function when it was included, making it one of the most frequently used features [40]. In addition, consistent with studies on positive body image, participants in this study also reported content that encouraged more holistic views of beauty and worth as being helpful [41]. Furthermore, the fear of losing control that disordered eating behaviors can cause may be overcome by focusing on the future and identifying personal goals and more meaningful achievements than those associated with weight and appearance. Other studies also suggest that combining goal setting with daily reminders is associated with great app use [42]. This appears to be a helpful strategy to disrupt mechanisms that may maintain eating disorder symptoms, thus making users more open and willing to change.

Given the loneliness and isolation reported by all participants (7/7, 100%) at some stage of their eating disorder, it is perhaps unsurprising that there was a strong desire for an app to facilitate social connection and hope. This may also be useful in helping users to overcome their critical inner voice and allow them to hear alternative views or perspectives. Although a key obstacle of app-based interventions is the lack of human contact or reciprocity [43], participants in this study provided some suggestions to combat this, including communication with other app users through a chat or forum. However, given that this may be accompanied by a range of challenges and potential dangers for users, researchers can instead focus on using innovative and creative thinking to identify alternative ways to reduce feelings of loneliness and enhance a sense of belonging through an app, without direct communication. For example, this can be achieved with inclusive and welcoming language and may include providing notifications or messages to users to indicate that other users are also using the app currently (eg, “[number] other users have been online this week/are now online”). It is possible that an e-therapeutic alliance can also be

achieved in digital health interventions through user engagement and careful design choices [44]. In addition, using recovery stories from individuals with lived experience may facilitate feelings of hope for users. Future technology may also find a way to facilitate social connection, including chatbots or moderation for a chat room–type function [15], as has been done in other areas of mental health research [45]. Although these features can be incorporated into an app, they can also be offered as adjunct supports for people with eating disorder.

The final theme highlights the key role of app usability in participant uptake and engagement. Given that poor usability has been found to predict nonadherence to web-based interventions [46], this finding is an important reminder for researchers to focus considerable efforts on app accessibility, including the need for continued and rigorous testing with end users and design approaches to identify and address issues. The language used within an app itself is also important. End users may fear failure and be especially self-critical. Therefore, ensuring compassionate language that emphasizes the nonlinear journey of recovery and praise when completing activities and progressing through an app may be helpful. Studies suggest that self-compassion may be an effective intervention target, more so than alternatives such as mindfulness [47]. Relatedly, the language and tone of an app need to strike a careful balance between appearing credible (ie, created and recommended by professionals) and being personable and informal (ie, similar to talking to a friend and using language that is understandable and relatable). In addition, it appears that users want an app that can be used flexibly to fit around themselves and their lives. These findings highlight how the look and feel, functionality, features, and usability of an app are extremely important, thus indicating that the time invested in making an app appealing and intuitive during its development is critical [21].

App features that emphasize personalization were mentioned favorably by participants, including allowing users to enter their own input (eg, writings and audio messages) and track their mood. Consistent with this, qualitative studies exploring the usability of an existing eating disorder app found that participants value being able to monitor their progress and felt that it improved their self-efficacy and motivation to use the app [43]. Although some personalized elements may be easy to implement (eg, notifications, prompts, or branching of content and feedback based on early user inputs), others will be more complex and require expertise to be incorporated into an app (eg, use of chatbots to enable an ongoing, personalized dialogue with users). It is also important to emphasize that users and designers may have different interpretations of *personalization*. For users, this may be about feeling that the app is responsive to their inputs and not simply providing generic scripts and videos that fail to pick up on information that may make subsequent content unnecessary. In contrast, programmers may think of personalization in terms of more complex design elements, such as chatbots. Given the rapid evolution of technology, it is important for researchers to be up to date with newly emerging programming that may produce advancements for app personalization. However, it is also important to note that this high-level personalization may be beyond the scope and capacity of most academic-led projects and may not always

be desired or necessary for establishing an engaging and efficacious product. This reinforces the importance of end user testing, co-design elements, and ensuring that researchers are clear in their interpretation of user feedback.

### Limitations

This study used a novel, design thinking approach to provide better understanding of the needs and preferences of potential end users of app-based interventions for eating disorders. These findings can be used to inform the development, framing, and dissemination of eating disorder apps. However, the study has a number of limitations. First, although this study identified several insights into strategies to enhance uptake and retention, the retrospective nature of these data means that suggestions are informed by hindsight and may be different from what participants may have said if interviewed in the early stages of their illness or during their treatment. Second, the present sample was limited to women and may have been moderately biased in that they self-selected to participate, indicating an openness to discussing eating disorders and engaging in research. However, this study was exploratory in nature and the sample was not intended to be representative. Nonetheless, the authors encourage the use of these methodologies among large samples. Finally, even in this small sample, there were some slight inconsistencies in participants' preferences, which highlighted individual differences. For example, some participants placed more emphasis on social support or journal writing within an app than others. Therefore, further studies will be helpful to unpick some of these nuances to better understand individual preferences among users.

### Implications

There are a number of important practical and clinical implications of this study. Researchers developing theoretically informed and evidence-based apps must consider the needs of the target group when developing resources. This study may be used to inform app developers or researchers who may not be familiar with the lived experience of an eating disorder and those trying to put themselves back into an early intervention mindset if they have not experienced this recently. In addition, the pain and gain points can be used to directly inform the design and development of such apps. Relatedly, it is critical that researchers explore and identify key drivers of uptake, which may inform users' willingness to change. Consideration and implementation of such practices will likely enhance app uptake, engagement, and retention, with the potential to support individuals who may not otherwise seek or access eating disorder information elsewhere. Importantly, cost evaluations that examine the impact of various design and implementation features should be performed to ensure that they are appropriate and effective.

### Conclusions

This study extended previous research by using a novel design thinking approach to understand and empathize with individuals with lived experience of eating disorders to identify their needs and preferences for app-based early interventions. Although a number of challenges are experienced by this population, including feelings of loneliness and ambivalent attitudes toward

change, this study has identified some opportunities that may be considered in the design and development of an eating disorder app. These include helping users to develop insights into eating disorders, identifying personal values and goals to encourage behavior change, providing elements of social

connection to facilitate hope, and ensuring that the language and framing of an app is appropriate. Researchers and app developers should consider these aspects to develop future early intervention apps for women in ways that enhance app uptake and retention and improve eating disorder symptoms.

---

## Acknowledgments

The authors would like to thank Melissa Keller-Tuberg for her contribution to this study. This study was funded by the Medical Research Future Fund (APP1179321).

---

## Conflicts of Interest

None declared.

---

### Multimedia Appendix 1

Semi-structured interview schedule.

[[DOCX File, 15 KB - formative\\_v6i10e38387\\_app1.docx](#)]

---

### Multimedia Appendix 2

Overview of common elements across the empathy maps.

[[DOCX File, 16 KB - formative\\_v6i10e38387\\_app2.docx](#)]

---

## References

1. Galmiche M, Déchelotte P, Lambert G, Tavolacci MP. Prevalence of eating disorders over the 2000-2018 period: a systematic literature review. *Am J Clin Nutr* 2019 May 01;109(5):1402-1413. [doi: [10.1093/ajcn/nqy342](#)] [Medline: [31051507](#)]
2. Cooper M, Reilly EE, Siegel JA, Coniglio K, Sadeh-Sharvit S, Pisetsky EM, et al. Eating disorders during the COVID-19 pandemic and quarantine: an overview of risks and recommendations for treatment and early intervention. *Eat Disord* 2022;30(1):54-76. [doi: [10.1080/10640266.2020.1790271](#)] [Medline: [32644868](#)]
3. Griffiths S, Hay P, Mitchison D, Mond JM, McLean SA, Rodgers B, et al. Sex differences in the relationships between body dissatisfaction, quality of life and psychological distress. *Aust N Z J Public Health* 2016 Dec;40(6):518-522. [doi: [10.1111/1753-6405.12538](#)] [Medline: [27372301](#)]
4. Arcelus J, Mitchell AJ, Wales J, Nielsen S. Mortality rates in patients with anorexia nervosa and other eating disorders. A meta-analysis of 36 studies. *Arch Gen Psychiatry* 2011 Jul;68(7):724-731. [doi: [10.1001/archgenpsychiatry.2011.74](#)] [Medline: [21727255](#)]
5. van Hoeken D, Hoek HW. Review of the burden of eating disorders: mortality, disability, costs, quality of life, and family burden. *Curr Opin Psychiatry* 2020 Nov;33(6):521-527 [FREE Full text] [doi: [10.1097/YCO.0000000000000641](#)] [Medline: [32796186](#)]
6. Butterfly report: paying the price the impact of eating disorders in Australia. Deloitte. 2012 Nov 30. URL: <https://www2.deloitte.com/au/en/pages/economics/articles/butterfly-report-paying-price-eating-disorders.html> [accessed 2021-08-21]
7. Streatfeild J, Hickson J, Austin SB, Hutcheson R, Kandel JS, Lampert JG, et al. Social and economic cost of eating disorders in the United States: evidence to inform policy action. *Int J Eat Disord* 2021 May;54(5):851-868. [doi: [10.1002/eat.23486](#)] [Medline: [33655603](#)]
8. Chang PG, Delgadillo J, Waller G. Early response to psychological treatment for eating disorders: a systematic review and meta-analysis. *Clin Psychol Rev* 2021 Jun;86:102032. [doi: [10.1016/j.cpr.2021.102032](#)] [Medline: [33915335](#)]
9. Griffiths S, Rossell SL, Mitchison D, Murray SB, Mond JM. Pathways into treatment for eating disorders: a quantitative examination of treatment barriers and treatment attitudes. *Eat Disord* 2018;26(6):556-574. [doi: [10.1080/10640266.2018.1518086](#)] [Medline: [30222057](#)]
10. Seah XY, Tham XC, Kamaruzaman NR, Yobas PK. Knowledge, attitudes and challenges of healthcare professionals managing people with eating disorders: a literature review. *Arch Psychiatr Nurs* 2017 Feb;31(1):125-136. [doi: [10.1016/j.apnu.2016.09.002](#)] [Medline: [28104050](#)]
11. Hepworth N, Paxton SJ. Pathways to help-seeking in bulimia nervosa and binge eating problems: a concept mapping approach. *Int J Eat Disord* 2007 Sep;40(6):493-504. [doi: [10.1002/eat.20402](#)] [Medline: [17573682](#)]
12. Weissman RS, Bauer S, Thomas JJ. Access to evidence-based care for eating disorders during the COVID-19 crisis. *Int J Eat Disord* 2020 May;53(5):369-376 [FREE Full text] [doi: [10.1002/eat.23279](#)] [Medline: [32338400](#)]
13. Hart LM, Granillo MT, Jorm AF, Paxton SJ. Unmet need for treatment in the eating disorders: a systematic review of eating disorder specific treatment seeking among community cases. *Clin Psychol Rev* 2011 Jul;31(5):727-735. [doi: [10.1016/j.cpr.2011.03.004](#)] [Medline: [21501580](#)]

14. Lee YC, Chatterton ML, Magnus A, Mohebbi M, Le LK, Mihalopoulos C. Cost of high prevalence mental disorders: findings from the 2007 Australian National Survey of Mental Health and Wellbeing. *Aust N Z J Psychiatry* 2017 Dec;51(12):1198-1211. [doi: [10.1177/0004867417710730](https://doi.org/10.1177/0004867417710730)] [Medline: [28565923](https://pubmed.ncbi.nlm.nih.gov/28565923/)]
15. Torous J, Bucci S, Bell IH, Kessing LV, Faurholt-Jepsen M, Whelan P, et al. The growing field of digital psychiatry: current evidence and the future of apps, social media, chatbots, and virtual reality. *World Psychiatry* 2021 Oct;20(3):318-335 [FREE Full text] [doi: [10.1002/wps.20883](https://doi.org/10.1002/wps.20883)] [Medline: [34505369](https://pubmed.ncbi.nlm.nih.gov/34505369/)]
16. Linardon J, Messer M, Lee S, Rosato J. Perspectives of e-health interventions for treating and preventing eating disorders: descriptive study of perceived advantages and barriers, help-seeking intentions, and preferred functionality. *Eat Weight Disord* 2021 May;26(4):1097-1109. [doi: [10.1007/s40519-020-01005-3](https://doi.org/10.1007/s40519-020-01005-3)] [Medline: [32959274](https://pubmed.ncbi.nlm.nih.gov/32959274/)]
17. Cooper Z, Bailey-Straebler S. Disseminating evidence-based psychological treatments for eating disorders. *Curr Psychiatry Rep* 2015 Mar;17(3):551 [FREE Full text] [doi: [10.1007/s11920-015-0551-7](https://doi.org/10.1007/s11920-015-0551-7)] [Medline: [25663154](https://pubmed.ncbi.nlm.nih.gov/25663154/)]
18. Ahmadiankalati M, Steins-Loeber S, Paslakis G. Review of randomized controlled trials using e-health interventions for patients with eating disorders. *Front Psychiatry* 2020 Jun 12;11:568 [FREE Full text] [doi: [10.3389/fpsy.2020.00568](https://doi.org/10.3389/fpsy.2020.00568)] [Medline: [32595546](https://pubmed.ncbi.nlm.nih.gov/32595546/)]
19. Linardon J, Shatte A, Messer M, Firth J, Fuller-Tyszkiewicz M. E-mental health interventions for the treatment and prevention of eating disorders: an updated systematic review and meta-analysis. *J Consult Clin Psychol* 2020 Nov;88(11):994-1007. [doi: [10.1037/ccp0000575](https://doi.org/10.1037/ccp0000575)] [Medline: [32852971](https://pubmed.ncbi.nlm.nih.gov/32852971/)]
20. Melioli T, Bauer S, Franko DL, Moessner M, Ozer F, Chabrol H, et al. Reducing eating disorder symptoms and risk factors using the internet: a meta-analytic review. *Int J Eat Disord* 2016 Jan;49(1):19-31. [doi: [10.1002/eat.22477](https://doi.org/10.1002/eat.22477)] [Medline: [26607683](https://pubmed.ncbi.nlm.nih.gov/26607683/)]
21. Torous J, Nicholas J, Larsen ME, Firth J, Christensen H. Clinical review of user engagement with mental health smartphone apps: evidence, theory and improvements. *Evid Based Ment Health* 2018 Aug;21(3):116-119. [doi: [10.1136/eb-2018-102891](https://doi.org/10.1136/eb-2018-102891)] [Medline: [29871870](https://pubmed.ncbi.nlm.nih.gov/29871870/)]
22. Devakumar A, Modh J, Saket B, Baumer EP, De Choudhury M. A review on strategies for data collection, reflection, and communication in eating disorder apps. In: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. 2021 Presented at: CHI '21; May 8-13, 2021; Yokohama, Japan p. 1-19. [doi: [10.1145/3411764.3445670](https://doi.org/10.1145/3411764.3445670)]
23. Mummah SA, Robinson TN, King AC, Gardner CD, Sutton S. IDEAS (Integrate, Design, Assess, and Share): a framework and toolkit of strategies for the development of more effective digital interventions to change health behavior. *J Med Internet Res* 2016 Dec 16;18(12):e317 [FREE Full text] [doi: [10.2196/jmir.5927](https://doi.org/10.2196/jmir.5927)] [Medline: [27986647](https://pubmed.ncbi.nlm.nih.gov/27986647/)]
24. Baker FW, Moukhless S. Concretising design thinking: a content analysis of systematic and extended literature reviews on design thinking and human - centred design. *Rev Educ* 2019 Dec 11;8(1):305-333. [doi: [10.1002/rev3.3186](https://doi.org/10.1002/rev3.3186)]
25. Dam RF. The 5 stages in the design thinking process. Interaction Design Foundation. 2022. URL: <https://www.interaction-design.org/literature/article/5-stages-in-the-design-thinking-process> [accessed 2022-07-22]
26. Ferreira B, Silva W, Oliveira E, Conte T. Designing personas with empathy map. In: Proceedings of the 27th International Conference on Software Engineering and Knowledge Engineering. 2015 Presented at: 27th International Conference on Software Engineering and Knowledge Engineering; Jul 6-8, 2015; Pittsburgh, PA. [doi: [10.18293/seke2015-152](https://doi.org/10.18293/seke2015-152)]
27. 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](https://doi.org/10.1002/eat.23130)] [Medline: [31313370](https://pubmed.ncbi.nlm.nih.gov/31313370/)]
28. Milton AC, Hambleton A, Dowling M, Roberts AE, Davenport T, Hickie I. Technology-enabled reform in a nontraditional mental health service for eating disorders: participatory design study. *J Med Internet Res* 2021 Feb 16;23(2):e19532 [FREE Full text] [doi: [10.2196/19532](https://doi.org/10.2196/19532)] [Medline: [33591283](https://pubmed.ncbi.nlm.nih.gov/33591283/)]
29. Graham AK, Neubert SW, Chang A, Liu J, Fu E, Green EA, et al. Applying user-centered design methods to understand users' day-to-day experiences can inform a mobile intervention for binge eating and weight management. *Front Digit Health* 2021 Jun 4;3:651749 [FREE Full text] [doi: [10.3389/fdgh.2021.651749](https://doi.org/10.3389/fdgh.2021.651749)] [Medline: [34713124](https://pubmed.ncbi.nlm.nih.gov/34713124/)]
30. Doley JR, McLean SA, Griffiths S, Yager Z. Designing body image and eating disorder prevention programs for boys and men: theoretical, practical, and logistical considerations from boys, parents, teachers, and experts. *Psychol Men Masculinities* 2021 Jan;22(1):124-134. [doi: [10.1037/men0000263](https://doi.org/10.1037/men0000263)]
31. Scholten H, Granic I. Use of the principles of design thinking to address limitations of digital mental health interventions for youth: viewpoint. *J Med Internet Res* 2019 Jan 14;21(1):e11528 [FREE Full text] [doi: [10.2196/11528](https://doi.org/10.2196/11528)] [Medline: [31344671](https://pubmed.ncbi.nlm.nih.gov/31344671/)]
32. Boddy CR. Sample size for qualitative research. *Qual Market Res* 2016 Sep 12;19(4):426-432. [doi: [10.1108/qmr-06-2016-0053](https://doi.org/10.1108/qmr-06-2016-0053)]
33. How to conduct user interviews. Interaction Design Foundation. 2022. URL: <https://www.interaction-design.org/literature/article/how-to-conduct-user-interviews> [accessed 2022-08-05]
34. 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)]

35. Musić S, Elwyn R, Fountas G, Gnat I, Jenkins ZM, Malcolm A, et al. Valuing the voice of lived experience of eating disorders in the research process: benefits and considerations. *Aust N Z J Psychiatry* 2022 Mar;56(3):216-218. [doi: [10.1177/0004867421998794](https://doi.org/10.1177/0004867421998794)] [Medline: [33715452](https://pubmed.ncbi.nlm.nih.gov/33715452/)]
36. Wade TD, Hart LM, Mitchison D, Hay P. Driving better intervention outcomes in eating disorders: a systematic synthesis of research priority setting and the involvement of consumer input. *Eur Eat Disord Rev* 2021 May;29(3):346-354. [doi: [10.1002/erv.2759](https://doi.org/10.1002/erv.2759)] [Medline: [32706169](https://pubmed.ncbi.nlm.nih.gov/32706169/)]
37. Fitzsimmons-Craft EE, Krauss MJ, Costello SJ, Floyd GM, Wilfley DE, Cavazos-Rehg PA. Adolescents and young adults engaged with pro-eating disorder social media: eating disorder and comorbid psychopathology, health care utilization, treatment barriers, and opinions on harnessing technology for treatment. *Eat Weight Disord* 2020 Dec;25(6):1681-1692 [FREE Full text] [doi: [10.1007/s40519-019-00808-3](https://doi.org/10.1007/s40519-019-00808-3)] [Medline: [31679144](https://pubmed.ncbi.nlm.nih.gov/31679144/)]
38. Lipson SK, Jones JM, Taylor CB, Wilfley DE, Eichen DM, Fitzsimmons-Craft EE, et al. Understanding and promoting treatment-seeking for eating disorders and body image concerns on college campuses through online screening, prevention and intervention. *Eat Behav* 2017 Apr;25:68-73 [FREE Full text] [doi: [10.1016/j.eatbeh.2016.03.020](https://doi.org/10.1016/j.eatbeh.2016.03.020)] [Medline: [27117825](https://pubmed.ncbi.nlm.nih.gov/27117825/)]
39. Fuller-Tyszkiewicz M, Richardson B, Lewis V, Linardon J, Mills J, Juknaitis K, et al. A randomized trial exploring mindfulness and gratitude exercises as eHealth-based micro-interventions for improving body satisfaction. *Comput Human Behav* 2019 Jun;95:58-65. [doi: [10.1016/j.chb.2019.01.028](https://doi.org/10.1016/j.chb.2019.01.028)]
40. Wasil AR, Patel R, Cho JY, Shingleton RM, Weisz JR, DeRubeis RJ. Smartphone apps for eating disorders: a systematic review of evidence-based content and application of user-adjusted analyses. *Int J Eat Disord* 2021 May;54(5):690-700. [doi: [10.1002/eat.23478](https://doi.org/10.1002/eat.23478)] [Medline: [33534176](https://pubmed.ncbi.nlm.nih.gov/33534176/)]
41. Wood-Barcalow NL, Tylka TL, Augustus-Horvath CL. "But I Like My Body": positive body image characteristics and a holistic model for young-adult women. *Body Image* 2010 Mar;7(2):106-116. [doi: [10.1016/j.bodyim.2010.01.001](https://doi.org/10.1016/j.bodyim.2010.01.001)] [Medline: [20153990](https://pubmed.ncbi.nlm.nih.gov/20153990/)]
42. Kim JP, Sadeh-Sharvit S, Darcy AM, Neri E, Vierhile M, Robinson A, et al. The utility and acceptability of a self-help smartphone application for eating disorder behaviors. *J Technol Behav Sci* 2018 Sep;3(3):161-164 [FREE Full text] [doi: [10.1007/s41347-018-0048-4](https://doi.org/10.1007/s41347-018-0048-4)] [Medline: [30374444](https://pubmed.ncbi.nlm.nih.gov/30374444/)]
43. Linardon J, King T, Shatte A, Fuller-Tyszkiewicz M. Usability evaluation of a cognitive-behavioral app-based intervention for binge eating and related psychopathology: a qualitative study. *Behav Modif* 2022 Sep;46(5):1002-1020. [doi: [10.1177/01454455211021764](https://doi.org/10.1177/01454455211021764)] [Medline: [34075803](https://pubmed.ncbi.nlm.nih.gov/34075803/)]
44. Baumel A, Faber K, Mathur N, Kane JM, Muench F. Enlight: a comprehensive quality and therapeutic potential evaluation tool for mobile and web-based eHealth interventions. *J Med Internet Res* 2017 Mar 21;19(3):e82 [FREE Full text] [doi: [10.2196/jmir.7270](https://doi.org/10.2196/jmir.7270)] [Medline: [28325712](https://pubmed.ncbi.nlm.nih.gov/28325712/)]
45. Mental health support when and where you need it. MOST. URL: <https://www.most.org.au/> [accessed 2022-07-29]
46. Beatty L, Binnion C. A systematic review of predictors of, and reasons for, adherence to online psychological interventions. *Int J Behav Med* 2016 Dec;23(6):776-794. [doi: [10.1007/s12529-016-9556-9](https://doi.org/10.1007/s12529-016-9556-9)] [Medline: [26957109](https://pubmed.ncbi.nlm.nih.gov/26957109/)]
47. Messer M, Anderson C, Linardon J. Self-compassion explains substantially more variance in eating disorder psychopathology and associated impairment than mindfulness. *Body Image* 2021 Mar;36:27-33. [doi: [10.1016/j.bodyim.2020.10.002](https://doi.org/10.1016/j.bodyim.2020.10.002)] [Medline: [33161205](https://pubmed.ncbi.nlm.nih.gov/33161205/)]

## Abbreviations

**WIRED:** Web-Based Interventions to Reduce Eating Disorders

*Edited by A Mavragani; submitted 31.03.22; peer-reviewed by Anonymous, R Marshall; comments to author 18.07.22; revised version received 05.08.22; accepted 19.08.22; published 31.10.22.*

*Please cite as:*

Jarman HK, McLean SA, Rodgers R, Fuller-Tyszkiewicz M, Paxton S, O'Gorman B, Harris E, Shatte A, Bishop K, Baumann T, Mahoney D, Daugelat MC, Yager Z

*Informing mHealth and Web-Based Eating Disorder Interventions: Combining Lived Experience Perspectives With Design Thinking Approaches*

*JMIR Form Res* 2022;6(10):e38387

URL: <https://formative.jmir.org/2022/10/e38387>

doi: [10.2196/38387](https://doi.org/10.2196/38387)

PMID: [36315225](https://pubmed.ncbi.nlm.nih.gov/36315225/)

©Hannah K Jarman, Siân A McLean, Rachel Rodgers, Matthew Fuller-Tyszkiewicz, Susan Paxton, Beth O'Gorman, Emily Harris, Adrian Shatte, Katie Bishop, Tahlia Baumann, Danielle Mahoney, Melissa-Claire Daugelat, Zali Yager. Originally

published in JMIR Formative Research (<https://formative.jmir.org>), 31.10.2022. 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 Video-Delivered Family Therapeutic Intervention for Perinatal Women With Clinically Significant Depressive Symptoms and Family Conflict: Indicators of Feasibility and Acceptability

Fallon Cluxton-Keller<sup>1\*</sup>, MA, PhD; Mark T Hegel<sup>1\*</sup>, PhD

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

\* all authors contributed equally

**Corresponding Author:**

Fallon Cluxton-Keller, MA, PhD

Department of Psychiatry

Geisel School of Medicine

Dartmouth College

One Medical Center Drive

Lebanon, NH, 03756

United States

Phone: 1 603 650 4724

Email: [Fallon.P.Cluxton-Keller@dartmouth.edu](mailto:Fallon.P.Cluxton-Keller@dartmouth.edu)

## Abstract

**Background:** Variation in family therapeutic intervention fidelity has an impact on outcomes. The use of video conferencing technology can strengthen therapist fidelity to family therapeutic interventions.

**Objective:** This article explores indicators of feasibility and acceptability for a video-delivered family therapeutic intervention for perinatal women with depressive symptoms and family conflict. The objectives of this article are to describe indicators of feasibility, including therapist fidelity to the intervention and technological factors that relate to implementation of the intervention, as well as indicators of acceptability for participants of the intervention.

**Methods:** The data included in this article are from an ongoing randomized trial of the Resilience Enhancement Skills Training (REST) video-delivered family therapeutic intervention. Participant recruitment and data collection are still underway for this clinical trial. Of the 106 participants who are currently enrolled in this study, 54 (51%) have been randomized to receive REST from May 2021 through July 2022. Currently, 2 therapists are delivering the intervention, and the training procedures for therapists are summarized herein. Therapist fidelity to the family therapeutic intervention was assessed in 67 audio recorded sessions. The training procedures were summarized for use of video conferencing technology by therapists and the 54 study participants. Knowledge of the video conferencing technology features was assessed in therapists and study participants by the number of attempts required to use the features. Participant responsiveness to the intervention was assessed by the percentage of attended sessions and percentage of complete homework assignments.

**Results:** To date, both therapists have demonstrated high fidelity to the family therapeutic intervention and used all video conferencing technology features on their first attempt. The current participants required 1 to 3 attempts to use 1 or more of the video conferencing technology features. About 59% (n=32) of the current participants immediately accessed the features on the first attempt. Our results show that perinatal women attended all sessions, and their family members attended 80% of the sessions. To date, participants have completed 80% of the homework assignments.

**Conclusions:** These early findings describe indicators of the feasibility and acceptability of the video-delivered family therapeutic intervention for use with this high priority population. Upon completion of recruitment and data collection, a subsequent article will include a mixed methods process evaluation of the feasibility and acceptability of the video-delivered family therapeutic intervention.

**Trial Registration:** ClinicalTrials.gov NCT04741776; <https://clinicaltrials.gov/ct2/show/NCT04741776>

(*JMIR Form Res* 2022;6(10):e41697) doi:[10.2196/41697](https://doi.org/10.2196/41697)

## KEYWORDS

family intervention; perinatal; postnatal; depression; conflict; telehealth; family conflict; family therapy; family therapist; video conferencing; teleconference; teleconferencing; telemedicine; virtual care; mental health; psychological health; digital health intervention; parenting

## Introduction

### Background

The federal home visiting program is a voluntary program that provides parenting education to pregnant and postpartum women and their families [1]. Several studies have shown that over a third of home visited mothers (pregnant and postpartum) report clinically significant depressive symptoms [2-4]. Home visited mothers infrequently obtain treatment due to barriers (eg, lack of childcare, lack of transportation, stigma [4,5]) that are significantly greater in rural regions. Nonviolent family conflict, defined as expressed anger, criticism, and arguments [6] worsens depressive symptoms in perinatal women [7-10]. Yet, the research is limited on family therapy interventions that address depressive symptoms in perinatal populations with family conflict [11-13].

The use of HIPAA (Health Insurance Portability and Accountability Act)-compliant video conferencing technology (VCT) to deliver family therapeutic interventions that address perinatal depressive symptoms and family conflict may be a feasible strategy to increase access to treatment. However, research is limited on the feasibility of using VCT to deliver family therapeutic interventions to this population [11-13]. While the use of VCT is not new to the field of family therapy, most studies have examined the feasibility and effectiveness of interventions that primarily target parenting skills for youth with specific mental health problems (eg, [14-18]).

The rapid shift from in-person to VCT sessions during the COVID-19 pandemic created fidelity challenges in the implementation of many family therapy models that address family conflict [eg, 17]. Variation in intervention fidelity can have unintended impacts on outcomes [19,20]. Implementation science research shows that several factors (eg, therapist training, frequency of supervision, etc) contribute to intervention fidelity [19,20]. Technology influences how the therapist and family interact in sessions, which can impact adherence to the model and family outcomes [16]. Therapists who lack specific guidance on delivering models using VCT may experience problems with delivering the model as intended. For example, low bandwidth can lead to frustration in therapists and families due to having to repeat themselves, as well as omission of some session content due to reduced audio quality [18]. Thus, model developers that aim to use VCT for session delivery need to include specific steps for therapists to manage technological challenges that could interfere with session quality.

Concordantly, families need to be trained in how to use VCT to participate in the sessions. A recent systematic review [16] on the use of VCT for couple and family therapies showed that of the 37 included studies, only 3 [12,21,22] reported that families were specifically trained in the use of the VCT. Given that many families seek family therapy to address conflict, lack of clarity on how to use VCT to join the session could lead to

arguments before the session even starts or cause them to miss sessions.

This study builds on prior research of a video-delivered family therapeutic intervention called Resilience Enhancement Skills Training (REST) [12,23], which aims to reduce perinatal depressive symptoms and family conflict in home visited families in rural areas. REST is based on general systems theory [24] and informed by family-oriented Dialectical Behavior Therapy Skills Training [25]. REST is not a parenting skills program. It was developed prior to the COVID-19 pandemic and leverages VCT to increase access to treatment [12]. Implementation of REST requires training of the therapist and families to ensure correct use of VCT in sessions [12].

### Objectives

This article explores indicators of feasibility and acceptability from an ongoing randomized trial of REST. The article has 2 objectives: the first is to describe 2 indicators of feasibility that include therapist fidelity to REST and technology factors relating to the implementation of REST, and the second is to describe 2 indicators of acceptability that include session attendance and homework completion.

## Methods

### Study Design

This ongoing study uses an effectiveness-implementation hybrid type 1 design with a pilot randomized trial of REST compared to standard of care video-delivered problem solving therapy (V-PST) (eg, [26]). Mothers are randomly assigned to receive REST or V-PST. Both interventions are delivered using VCT and include a total of 10 weekly 45-minute sessions. The mother participates in REST sessions with a family member, whereas the mother participates alone in the V-PST sessions. This article focuses on indicators of feasibility and acceptability of REST and only uses the current data from that arm of the trial.

### Ethics Approval

This trial was approved by an institutional review board (STUDY02000691) at an academic medical center in New England. Study participants provided consent for the eligibility screen and study enrollment, including consent for audio recorded sessions.

### Participant Eligibility

Each family consists of the mother and her family member, defined as her adult relative or current intimate partner. Based on information from the participating home visiting agencies that most mothers had 1 eligible family member, we limited the number of family members to just 1. A research team member schedules separate phone calls with the mother and her family member to obtain electronic consent for the eligibility screen interview. The eligibility screen interview phone call is done separately with each mother and her family member because

the interview for mothers includes the Abuse Assessment Screen [27] to screen for domestic violence, and family conflict is assessed separately using the Perceived Hostility Survey (PHS) [28].

Mothers are eligible for participation in this study if they meet the following inclusion criteria: (1) enrolled in home visiting at a participating agency; (2) in any trimester of pregnancy and up to 18 months postpartum; (3) at least 15 years old; (4) fluent in English with at least an eighth-grade education, as intervention materials are written in English for this level of education; (5) experiencing moderate-to-severe depressive symptoms, with scores of at least 20 without suicidal ideation on the Beck Depression Inventory-Second Edition (BDI-II, [29]); (6) experiencing moderate to high conflict (PHS ages 18 + raw scores of at least 16; PHS ages 15-17 years old raw scores of at least 14) [30] with the selected family member with whom they live in the same home or have at least weekly contact; and (7) have consistent internet access on a cell phone, tablet, or computer with a working camera and microphone. This study includes a detailed protocol for mothers who report suicidal ideation on the BDI-II (rating of 2 or 3 on item 9), and they are provided with emergency assistance. Mothers who report current domestic violence in their homes or histories of domestic violence with the selected family member on the Abuse Assessment Screen are ineligible for participation in this study. This study includes a detailed protocol for mothers who report domestic violence, and they are provided with emergency assistance by local service providers who address domestic violence.

Family members are eligible for participation if they meet the following inclusion criteria: (1) are the mother's adult relative or current intimate partner; (2) at least 15 years old; (3) fluent in English with at least an eighth-grade education, as intervention materials are written in English for this level of education; (4) experiencing moderate to high family conflict with the mother (PHS ages 18 + raw scores of at least 16; PHS ages 15-17 years old raw scores of at least 14 [30]); and (5) have consistent internet access on a cell phone, tablet, or computer with a working camera and microphone.

## Recruitment

Participants are recruited from participating home visiting agencies that serve low-income families in New England. Home visitors use depression screening and referral procedures to refer mothers to the study [12]. Participant recruitment began in April 2021 and is still underway. The goal is to achieve a sample size of 160 individuals (80 families). As of July 2022, 106 participants have enrolled in the study, and 54 (51%) of them are assigned to REST.

## Overview of REST

REST includes a total of 10 weekly 45-minute sessions delivered by a licensed mental health professional using VCT. In this study, 2 masters-level licensed social workers are currently delivering REST to participants. REST is based on general systems theory [24] and informed by family-oriented Dialectical Behavior Therapy (DBT) skills training [25]. REST uses interventions informed by general systems theory [24] to change

dynamics that produce family conflict. REST is informed by DBT core skills to improve regulation in cognition (mindfulness), emotions (behavioral activation and cognitive restructuring), and behavior (boundaries) [25]. Similar to DBT skills training [25], REST uses a psychoeducational format to teach skills to families. Most families consider REST sessions to be educational classes.

REST requires the therapist's use of 2 documents that pertain to session content and quality: (1) the REST therapist manual (including safety protocols for depression and family conflict) and (2) the skills book for families. The therapist asks participants to select up to 3 goals that pertain to their relationship and mood in the first session. Similar to DBT skills training [25], REST includes sessions that are devoted to teaching participants new skills and some that are devoted to reviewing learned skills. The skill introduction sessions require the therapist to use the screen share feature in the VCT (which allows participants to view a document on the therapist's computer screen during the session) to read content from the skills book on the purpose and use of the skill to participants. All sessions include a mindfulness exercise, a session goal that includes a summary of the activities, and an explanation of the ways in which the skills relate to the participants' goals. Nearly all sessions include skill practice exercises. Therapists verbally communicate participant strengths in at least 8 sessions. Nine sessions include planned skill application questions to help participants prepare for homework assignments. The skill application elements of REST are flexible and allow therapists to work with families to apply the skills to aspects of their relationships based on each family's preferences. Participants receive homework assignments in 9 sessions. Although REST was originally designed for families, we anticipated that some family members would miss sessions. For this reason, practice exercises that involve role-plays and some of the skill application questions were modified for use in sessions only attended by the mother.

## Training Procedures

### *REST Training*

The primary aim of REST training is to teach therapists how to deliver session content to participants. The first author (FC-K) trained both therapists in the implementation of REST during a 6-hour virtual training session. The training session included 2 parts: (1) 2 hours of didactic instruction on the REST therapist manual and practice exercises for the skills, and (2) 4 hours of role-plays with feedback to allow therapists to demonstrate use of the REST therapist manual for skill application for cognitive, emotion, and behavior regulation in different family scenarios. Role-plays were conducted so that therapists could practice delivering REST using VCT. The first author then provided feedback to therapists during role-plays. Supervision sessions are devoted to reinforcing content from the 6-hour training session. The first author supervises each REST therapist separately on a weekly basis using VCT. Results from ongoing fidelity assessments are used to continuously monitor their adherence to the model, and corrective feedback is provided as needed.

## Technology Training

REST requires therapists and participants to use VCT for sessions. WebEx (Cisco) [31] is the VCT used to deliver REST in this study. The technology requirements for therapists include a computer with a working microphone and camera, a word processing program to open the skills book, and consistent, secure internet access (eg, ethernet connection) with a WebEx account. Technology requirements for participants include consistent internet access (eg, subscription to an internet service provider) on an electronic device (eg, cell phone, tablet, or computer) with a working camera and microphone, along with the WebEx app to participate in sessions.

Each therapist underwent a separate training session to learn how to download and install the VCT software on the computer and use all features, including scheduling sessions, audio, video, screen share, and audio recording. One therapist required a single 15-minute training session and the other required a single 30-minute training session on the use of VCT. The second therapist's training was longer because she had recently purchased a new laptop computer and was still in the process of learning how to use it. In this study, 1 therapist used the split screen feature on her computer to view the REST therapist manual on one side of the screen and the skills book for the screen share for participants on the other side. The other therapist did not want to use the split screen feature and chose to place a hard copy of the REST therapist manual on a stand beside her computer and the skills book content on her computer for the screen share element in sessions.

Each participant received an individualized training session on use of the VCT for sessions. In this study, the first author trained the participants in use of the VCT. Each participant was emailed a link to join the VCT demonstration meeting and instructed to click the link to download and install the VCT app on the selected device (cell phone, tablet, or computer). Next, the participant followed the prompts to log in to join the meeting. When the participant joined the meeting, instructions were provided to click the microphone icon for audio and the camera icon for video.

Therapists received ongoing guidance on managing technology challenges that could occur in sessions. Problems with low bandwidth can result in poor audio quality. In these instances, therapists were directed to turn off the video feature and to ask participants to turn off theirs as well. Participants could still see the therapist's screen shared document when the videos were turned off. Turning off the video worked adequately in nearly all instances when low bandwidth was the problem. In the limited instances that turning off the video did not improve the audio quality, therapists were directed to use the phone call feature in the VCT and continue to use the video feature for only the screen shares in sessions. In the rare instance of significant internet service interruptions (eg, power outage due to a storm), therapists were instructed to continue the sessions using only the phone call feature in the VCT. Additional guidance was provided to therapists after we discovered that some participants with specific types of cell phones experienced significant reduced audio quality (low volume and choppy audio) when they used the ignore feature to avoid incoming

phone calls during sessions. In these instances, therapists instructed participants to exit the session and rejoin it.

## Data Collection

Research staff verbally administered the baseline questionnaires to mothers and family members, separately, prior to the first session and entered the data in REDCap (Research Electronic Data Capture) [32,33]. In this article, there are 2 indicators of feasibility, which include (1) REST fidelity, defined as the therapist's knowledge of REST and adherence to session content [34], and (2) knowledge of VCT in therapists and participants. The 2 indicators of acceptability center on participant responsiveness and include (1) session attendance and (2) homework completion [34].

Fidelity was assessed by therapist knowledge of and adherence to the REST therapist manual. For therapist knowledge of REST, the first author (FC-K) rated therapists' knowledge of the cognitive, emotion, and behavior regulation skills on the REST knowledge test during the role-plays during the REST training. Therapists were not allowed to deliver REST until they earned 100% on the test. To monitor therapist fidelity to REST session content in the ongoing trial, sessions are audio recorded. The first author developed the REST fidelity measure to rate therapist adherence to REST session content in the audio recorded sessions, and both authors (FC-K and MTH) use the measure to monitor fidelity in this study. Therapists are rated on the quality of content delivery for each session. For all sessions, therapists are rated in content categories on a scale that ranges from 0 to 2 (0: no content delivered, 1: partial content delivered, and 2: all content delivered). The maximum number of points for each session varies depending on the skill content. The level of fidelity is assessed to ensure that each therapist achieves at least 80% on the REST fidelity measure for each session, and data from these ongoing assessments are used in weekly supervision.

The current trial includes a 2-phase fidelity monitoring process for REST. In phase 1, both authors assessed fidelity in audio recorded sessions for the first 2 families assigned to each REST therapist. Fidelity was assessed following each session and before the next scheduled session so that the first author could immediately provide corrective feedback to REST therapists as needed. Both authors discussed and resolved any discrepancies in these fidelity assessments.

In phase 2, the second author (MTH) assesses fidelity in 10% of randomly selected audio recorded REST sessions that were originally assessed by the first author. Phase 2 is ongoing, as this study is still underway. The level of fidelity is continuously assessed to ensure that each therapist achieves at least 80% adherence to REST with each family. Both authors discuss and resolve any discrepancies in the fidelity assessments as needed.

The first author assessed therapist knowledge of VCT by observing their use of the audio, video, screen share, and recording features on their computers during the training. The first author also assessed participant knowledge of VCT by observing whether they joined the demonstration meeting and used the audio and video features on their electronic device (cell phone, tablet, or computer).

For participant responsiveness, data were collected on participant session attendance and homework completion from the therapists' session notes. Therapists were instructed to document attendance and homework completion in the session notes. The session notes do not contain identifying information about the participants, and all notes are stored on a password-protected, secure site within the medical center network.

### Analytic Plan

Univariate statistics are used to characterize REST participants' demographic and psychosocial information. This article includes demographic characteristics of participants who enrolled in the trial and were randomized to REST from May 2021 to July 2022.

For therapist knowledge of REST, the first author calculated the number of times it took each therapist to demonstrate correct use of the REST skills during role-plays. For adherence, Cohen Kappa was used to assess inter-rater reliability of the REST fidelity measure. A  $\kappa$  value of at least 0.80 suggests sufficient inter-rater reliability [35]. The mean adherence score (percent) for delivery of session content was calculated using the average adherence score divided by the total possible score. The item mean was calculated for adherence to REST session components, and the percentage for all content delivered (rating of 2 on the REST fidelity measure) was calculated for each component. The number of sessions that included lack of adherence to REST and deviations were calculated, and the reasons were recorded.

For knowledge of VCT, univariate statistics were used to summarize the number of attempts needed by each therapist and study participant to access and demonstrate use of the audio and video features in the VCT. For therapists only, the number of attempts needed to correctly use the screen share and audio recording features in the VCT were calculated. For study participants only, the number of attempts needed to join the VCT demonstration meeting were calculated.

For participant responsiveness, the total number of attended sessions were calculated separately for mothers and their family members. The proportion of attended sessions to expected sessions (total of 10) was calculated. The proportion of completed homework assignments to expected homework assignments (total of 9) was calculated.

## Results

### Therapist Characteristics

Both therapists are female licensed social workers with over 10 years of experience working with home visited families. One therapist identifies as more than 1 race, and the other identifies as Caucasian. Both therapists have knowledge of general systems theory and cognitive-behavioral family therapies.

### Participant Characteristics

This trial is ongoing, and recruitment and data collection are still underway. For the purpose of this article, this section

describes the demographic and psychosocial characteristics at baseline for the 54 participants who enrolled in the study and were randomly assigned to REST from May 2021 to July 2022. The race representation of the 54 participants included 72% (n=39) White non-Hispanic, 11% (n=6) more than 1 race, 9% (n=5) Asian, 4% (n=2) American Indian, and 4% (n=2) Black or African American. For highest level of education, 20% (n=11) of participants reported they did not graduate from high school, 43% (n=23) had high school degrees, 17% (n=9) had some college but no degree, and 20% (n=11) had higher education degrees. About 89% (n=48) of the family members were the mother's current intimate partner. On average, mothers were 26.35 (SD 4.90) years old, and their family members were 32.15 (SD 13.20) years old at baseline. Of the 27 mothers assigned to REST, about 15% (n=4) were pregnant at baseline. The baseline mean BDI-II score for mothers was 29.85 (SD 5.69), indicative of severe depressive symptoms. The baseline mean PHS uncorrected *T*-score for mothers was 58.7, indicative of moderately high conflict with family members. The baseline mean PHS uncorrected *T*-score for family members was 60.1, indicative of high conflict with mothers.

### Feasibility

#### Fidelity

Each therapist earned 100% on the REST knowledge test on the first attempt in demonstration of cognitive, emotion, and behavior regulation skills in role-plays. Both authors assessed therapist adherence to REST session content using the REST fidelity measure. The authors established sufficient inter-rater reliability ( $\kappa=0.91$ ) on the REST fidelity measure for the 67 audio recorded sessions. The average session length was 43 (SD 17.07) minutes. Three families experienced significant childcare interruptions, which extended the length of some sessions by 15 to 30 minutes.

One therapist achieved 90% adherence to REST in the sessions with the first 2 families to which she was assigned. She did not deliver some content in a few sessions (described below), but on average, she has adhered to 90% of REST content in the 32 audio recorded sessions assessed by the authors. As previously mentioned, the other therapist experienced some technological problems with her computer in sessions with the first family to which she was assigned, which resulted an adherence score that fell below 80% in the first 4 sessions. In these instances, she delivered content to the family that was missed in subsequent sessions. After her computer was fixed, she achieved 90% adherence to REST in sessions with the second and third families to which she was assigned. She did not deliver some content in a few sessions (described below), but on average, she has adhered to 88% of REST content in 35 audio recorded sessions assessed by the authors. Table 1 includes the average adherence for all content by session number and number of audio recordings assessed by the authors.

Table 2 includes the average adherence to REST session components for both therapists combined.

**Table 1.** Therapist adherence to Resilience Enhancement Skills Training (REST) by session number.

Session by audio recordings, n	Overall adherence to session content, %
Session 1: 10	99
Session 2: 7	94
Session 3: 8	92
Session 4: 5	100
Session 5: 7	89
Session 6: 6	92
Session 7: 4	92
Session 8: 7	92
Session 9: 7	99
Session 10: 6	94

**Table 2.** Therapist adherence to Resilience Enhancement Skills Training (REST) by session content.

Component by audio recorded sessions, n	Delivery of all content, %
Mindfulness exercise: 67	93
Homework review: 57	97
Session goal: 67	90
Purpose of skill <sup>a</sup> : 50	92
Explanation of how skill relates to participant goal: 67	87
Practice exercise: 61	94
Planned skill application: 61	82
Homework assignment: 61	95
Verbalize family strength: 57	88

<sup>a</sup>No variation by skill type.

Both therapists did not deliver some content in 7 sessions (1 session per family), which resulted in session adherence scores below 80%. Lack of adherence to content occurred in 4 sessions when participants arrived over 15 minutes late, and the therapists skipped the explanation of how the skills related to a participant goal, did not verbalize a participant strength, or partially delivered the planned skill application content due to time limitations. To prevent this from occurring again, the authors informed both therapists to reschedule the session if they have back-to-back sessions scheduled and a family does not arrive within 15 minutes of the set time. Lack of adherence to the planned skill application content occurred in 3 sessions with mothers of older children who frequently interrupted sessions with requests for privileges (eg, requests to play video games).

Other minor deviations occurred in sessions, but therapists still achieved at least 80% adherence to the session content. For example, a total of 5 sessions included participants who were diagnosed with significant medical (nonpsychiatric) problems, and they processed their thoughts and feelings at the beginning of the sessions. In these instances, the therapists validated the participants' thoughts and feelings, reinforced the importance of talking with their medical providers about medical treatment options, and redirected to them to the session content.

### **Knowledge of VCT**

Both REST therapists demonstrated use of the audio, video, screen share, and audio recording features in the VCT on the first attempt during the training. One therapist initially had difficulty with consistently accessing the screen share feature in 4 sessions with the first family she served because the mouse pad on her laptop worked inconsistently. The COVID-19 pandemic caused a shipping delay for the external mouse that was compatible with her computer, but when she transitioned to the use of an external mouse, she did not experience any other technological problems in sessions. The therapists initially had some difficulties with the audio recording feature in the first session with a total of 3 families, in that they realized they were not audio recording about 10 minutes into the session.

On average, the training in use of VCT for participants was 10 minutes. All participants were able to download and install the VCT app on their devices on the first attempt. About 87% (n=47) of participants chose to install it on their cell phones, while 2% (n=1) chose to install it on a desktop computer, 4% (n=2) on a tablet, and 7% (n=4) on a laptop computer. About 70% (n=38) of participants immediately joined the demonstration meeting on the first attempt. Moreover, 59% (n=32) of participants immediately accessed the audio and video features on the first attempt. About 6% (n=3) participants

required 2 attempts to access the audio feature, and 4% (n=2) participants required 2 attempts to access the video feature. Approximately 2% (n=1) of participants accessed the audio feature on the third attempt, and 24% (n=13) joined the demonstration meeting and accessed the audio and video features on the second attempt. About 6% (n=3) of participants joined the meeting and accessed the audio and video features on the third attempt.

To date, 9% (n=5) of participants who used cell phones to participate in the sessions experienced instances of low volume when they ignored incoming phone calls during sessions. The therapists asked the participants to exit the session and rejoin it, which resolved the issue in nearly all instances. In a total of 3 sessions, 2 participants used the phone call feature due to low bandwidth.

## Acceptability

### *Participant Responsiveness*

To date, 31 participants have completed REST, and it is currently in progress for 18 participants. Of the 54 participants assigned to REST, 2 dropped out prior to the second session due to busy work schedules and 2 dropped out prior to the fourth session because they moved to a different state. One family member discontinued REST prior to the second session due to a need for intensive medical treatment.

To date, all 16 mothers have attended all 10 sessions. On average, family members attended 80% of the sessions. However, 33% (5 out of 15) of family members had to miss at least 1 session due to variable work schedules or increased work hours. The participants who completed REST also completed 80% of the homework assignments. COVID-19 and other physical illnesses in participants and their children were primary reasons why they did not complete homework assignments.

## Discussion

### Principal Results

Our findings suggest that we have identified feasible strategies to facilitate REST's fidelity. REST is a highly structured psychoeducational intervention that requires the therapist to use basic technological skills (knowledge of VCT software features, average typing speed) to deliver it to families. These preliminary results suggest that the therapists have achieved high fidelity to REST. To date, the findings also suggest that REST is acceptable to families, in that mothers have attended all sessions and family members have attended 80% of sessions. Given the psychoeducational format of REST, participants consider sessions to be classes to learn new skills. To date, the participants who have completed REST have also completed 80% of the homework assignments.

Both therapists mastered use of the audio, video, screen share, and audio recording features in the VCT on the first attempt. About 70% of REST participants joined the VCT demonstration meeting on the first attempt. Over half (59%) of REST

participants accessed the audio and video features in the VCT on the first attempt. To date, REST participants that required 2 or more attempts to access audio or video features in the training have been able to access these features on the first attempt to join the sessions. Therapists have been provided with specific guidelines to prevent and manage potential technology interruptions, which has been helpful in maintaining participant engagement in the sessions. The guideline that is most often used by therapists pertains to poor audio quality that occurs when participants mute incoming phone calls during sessions. To resolve this problem, therapists ask the participants to exit the session and rejoin it.

### Comparison With Prior Work

These early findings on the indicators of feasibility and acceptability of REST align with those of previous research on REST [12,23]. REST has been delivered using 2 types of VCT: Vidy [36] in the preliminary study [12] and WebEx in this study. The early findings of this study suggest that masters-level therapists can deliver REST with high fidelity. To date, findings that pertain to session attendance are consistent with those previously reported in the preliminary pilot study of REST [12], in that mothers have attended all 10 sessions and family members have attended 80% of sessions. In the preliminary pilot study of REST [12,23], the therapist and home visitors reinforced participants' use of the skills, but only therapists reinforced participants' use of the skills in this trial. The decision to eliminate home visitor reinforcement of REST skills in this study was based on home visiting agency staff reports that they did not have enough time to complete this task during home visits. We do not believe this change has impacted the homework completion rates. Unlike the preliminary pilot study of REST [12], this study is being conducted during a pandemic, and some participants were unable to complete homework assignments due to COVID-19 and other physical illnesses in themselves or their children.

### Limitations

This study is ongoing, and more data are needed to gain a comprehensive knowledge of the feasibility and acceptability of REST. The study is currently in phase 2 of the fidelity monitoring plan, and a subsequent article will include the complete findings on therapist adherence to REST. Since recruitment and data collection are currently underway, outcome data that compare REST to the standard of care (V-PST) will be included in a future publication.

### Conclusions

This article explored indicators of feasibility and acceptability of REST. Evidence of feasibility and acceptability is important in justifying REST's potential for scalability. Upon completion of data collection, a mixed methods process evaluation will be conducted to fully explore the feasibility and acceptability of REST. The findings on therapist fidelity and family perceptions of REST will be used to guide interpretation of REST's preliminary effectiveness to inform a larger trial.

## Acknowledgments

This study is supported by funding from the National Institute of Mental Health under grant number 5R34MH124951. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

## Conflicts of Interest

None declared.

## References

1. The maternal, infant, and early childhood home visiting program. Health Resources and Services Administration. 2021. URL: <https://mchb.hrsa.gov/sites/default/files/mchb/about-us/program-brief.pdf> [accessed 2022-07-27]
2. Michalopoulos C, Faucetta K, Hill CJ, Portilla XA, Burrell L, Lee H, et al. Impacts on family outcomes of evidence-based early childhood home visiting: results from the mother and infant home visiting program evaluation. Office of the Administration for Children & Families. 2019. URL: <https://www.acf.hhs.gov/opre/report/impacts-family-outcomes-evidence-based-early-childhood-home-visiting-results-mother-and> [accessed 2022-07-27]
3. Easterbrooks MA, Kotake C, Raskin M, Bumgarner E. Patterns of depression among adolescent mothers: Resilience related to father support and home visiting program. *Am J Orthopsychiatry* 2016 Jan;86(1):61-68. [doi: [10.1037/ort0000093](https://doi.org/10.1037/ort0000093)] [Medline: [26460697](https://pubmed.ncbi.nlm.nih.gov/26460697/)]
4. Ammerman RT, Putnam FW, Altaye M, Chen L, Holleb LJ, Stevens J, et al. Changes in depressive symptoms in first time mothers in home visitation. *Child Abuse Negl* 2009 Mar;33(3):127-138 [FREE Full text] [doi: [10.1016/j.chiabu.2008.09.005](https://doi.org/10.1016/j.chiabu.2008.09.005)] [Medline: [19328548](https://pubmed.ncbi.nlm.nih.gov/19328548/)]
5. Teeters AR, Ammerman RT, Shenk CE, Goyal NK, Folger AT, Putnam FW, et al. Predictors of maternal depressive symptom trajectories over the first 18 months in home visiting. *Am J Orthopsychiatry* 2016;86(4):415-424 [FREE Full text] [doi: [10.1037/ort0000159](https://doi.org/10.1037/ort0000159)] [Medline: [26881983](https://pubmed.ncbi.nlm.nih.gov/26881983/)]
6. Moos RH, Moos BS. Family Environment Scale Manual. Palo Alto, CA: Consulting Psychologists Press; 1994.
7. Hodgkinson S, Beers L, Southammakosane C, Lewin A. Addressing the mental health needs of pregnant and parenting adolescents. *Pediatrics* 2014 Jan;133(1):114-122 [FREE Full text] [doi: [10.1542/peds.2013-0927](https://doi.org/10.1542/peds.2013-0927)] [Medline: [24298010](https://pubmed.ncbi.nlm.nih.gov/24298010/)]
8. Westdahl C, Milan S, Magriples U, Kershaw TS, Rising SS, Ickovics JR. Social support and social conflict as predictors of prenatal depression. *Obstet Gynecol* 2007 Jul;110(1):134-140 [FREE Full text] [doi: [10.1097/01.AOG.0000265352.61822.1b](https://doi.org/10.1097/01.AOG.0000265352.61822.1b)] [Medline: [17601908](https://pubmed.ncbi.nlm.nih.gov/17601908/)]
9. Rollans M, Kohlhoff J, Meade T, Kemp L, Schmied V. PARTNER INVOLVEMENT: NEGOTIATING THE PRESENCE OF PARTNERS IN PSYCHOSOCIAL ASSESSMENT AS CONDUCTED BY MIDWIVES AND CHILD AND FAMILY HEALTH NURSES. *Infant Ment Health J* 2016 May;37(3):302-312. [doi: [10.1002/imhj.21562](https://doi.org/10.1002/imhj.21562)] [Medline: [27111339](https://pubmed.ncbi.nlm.nih.gov/27111339/)]
10. Guintivano J, Sullivan PF, Stuebe AM, Penders T, Thorp J, Rubinow DR, et al. Adverse life events, psychiatric history, and biological predictors of postpartum depression in an ethnically diverse sample of postpartum women. *Psychol Med* 2017 Sep 27;48(7):1190-1200. [doi: [10.1017/s0033291717002641](https://doi.org/10.1017/s0033291717002641)]
11. Cluxton-Keller F, Bruce ML. Clinical effectiveness of family therapeutic interventions in the prevention and treatment of perinatal depression: A systematic review and meta-analysis. *PLoS One* 2018 Jun 14;13(6):e0198730 [FREE Full text] [doi: [10.1371/journal.pone.0198730](https://doi.org/10.1371/journal.pone.0198730)] [Medline: [29902211](https://pubmed.ncbi.nlm.nih.gov/29902211/)]
12. Cluxton-Keller F, Williams M, Buteau J, Donnelly CL, Stolte P, Monroe-Cassel M, et al. Video-delivered family therapy for home visited young mothers with perinatal depressive symptoms: quasi-experimental implementation-effectiveness hybrid trial. *JMIR Ment Health* 2018 Dec 10;5(4):e11513 [FREE Full text] [doi: [10.2196/11513](https://doi.org/10.2196/11513)]
13. Noonan M, Jomeen J, Doody O. A review of the involvement of partners and family members in psychosocial interventions for supporting women at risk of or experiencing perinatal depression and anxiety. *Int J Environ Res Public Health* 2021 May 18;18(10):5396 [FREE Full text] [doi: [10.3390/ijerph18105396](https://doi.org/10.3390/ijerph18105396)] [Medline: [34070148](https://pubmed.ncbi.nlm.nih.gov/34070148/)]
14. Kuulasmaa A, Wahlberg KE, Kuusimäki ML. Videoconferencing in family therapy: a review. *J Telemed Telecare* 2004;10(3):125-129. [doi: [10.1258/135763304323070742](https://doi.org/10.1258/135763304323070742)] [Medline: [15165436](https://pubmed.ncbi.nlm.nih.gov/15165436/)]
15. McLean SA, Booth AT, Schnabel A, Wright BJ, Painter FL, McIntosh JE. Exploring the Efficacy of Telehealth for Family Therapy Through Systematic, Meta-analytic, and Qualitative Evidence. *Clin Child Fam Psychol Rev* 2021 Jun;24(2):244-266 [FREE Full text] [doi: [10.1007/s10567-020-00340-2](https://doi.org/10.1007/s10567-020-00340-2)] [Medline: [33492545](https://pubmed.ncbi.nlm.nih.gov/33492545/)]
16. de Boer K, Muir SD, Silva SSM, Nedeljkovic M, Seabrook E, Thomas N, et al. Videoconferencing psychotherapy for couples and families: A systematic review. *J Marital Fam Ther* 2021 Apr 10;47(2):259-288. [doi: [10.1111/jmft.12518](https://doi.org/10.1111/jmft.12518)] [Medline: [33837968](https://pubmed.ncbi.nlm.nih.gov/33837968/)]
17. Robbins MS, Midouhas H. Adapting the delivery of functional family therapy around the world during a global pandemic. *Glob Implement Res Appl* 2021 Apr 26;1(2):109-121 [FREE Full text] [doi: [10.1007/s43477-021-00009-0](https://doi.org/10.1007/s43477-021-00009-0)] [Medline: [34622211](https://pubmed.ncbi.nlm.nih.gov/34622211/)]
18. Lange AMC, Humayun S, Jefford T. The feasibility of providing remote functional family therapy with adolescents during the COVID-19 pandemic: a mixed-method study. *Child Youth Care Forum* 2022 May 02:1-26 [FREE Full text] [doi: [10.1007/s10566-022-09692-y](https://doi.org/10.1007/s10566-022-09692-y)] [Medline: [35531310](https://pubmed.ncbi.nlm.nih.gov/35531310/)]



19. Fixsen DL, Naoom SF, Blase KA, Friedman RM, Wallace F. Implementation Research: A Synthesis of the Literature. Tampa, FL: University of South Florida Press; 2005.
20. Carroll C, Patterson M, Wood S, Booth A, Rick J, Balain S. A conceptual framework for implementation fidelity. *Implement Sci* 2007 Nov 30;2:40 [FREE Full text] [doi: [10.1186/1748-5908-2-40](https://doi.org/10.1186/1748-5908-2-40)] [Medline: [18053122](https://pubmed.ncbi.nlm.nih.gov/18053122/)]
21. Wade SL, Oberjohn K, Burkhardt A, Greenberg I. Feasibility and preliminary efficacy of a web-based parenting skills program for young children with traumatic brain injury. *J Head Trauma Rehabil* 2009;24(4):239-247. [doi: [10.1097/HTR.0b013e3181ad6680](https://doi.org/10.1097/HTR.0b013e3181ad6680)] [Medline: [19625863](https://pubmed.ncbi.nlm.nih.gov/19625863/)]
22. Cooper-Vince CE, Chou T, Furr JM, Puliafico AC, Comer JS. Videoteleconferencing early child anxiety treatment: a case study of the internet-delivered PCIT CALM (I-CALM) program. *Evid Based Pract Child Adolesc Ment Health* 2016 May;1(1):24-39 [FREE Full text] [doi: [10.1080/23794925.2016.1191976](https://doi.org/10.1080/23794925.2016.1191976)] [Medline: [29104931](https://pubmed.ncbi.nlm.nih.gov/29104931/)]
23. Cluxton-Keller F, Buteau J, Williams M, Stolte P, Monroe-Cassel M, Bruce ML. Engaging rural young mothers in a technology-based intervention for depression. *Child Youth Serv* 2019 Feb 09;40(2):158-183. [doi: [10.1080/0145935x.2018.1561264](https://doi.org/10.1080/0145935x.2018.1561264)]
24. von Bertalanffy BL. *General Systems Theory: Foundations, Development, Applications*. New York, NY: George Braziller; 1968.
25. Rathus JH, Miller AL. *DBT Skills Manual for Adolescents*. New York, NY: Guilford Press; 2015.
26. Hegel MT, Areán PA. *Problem Solving Treatment for Primary Care: A Treatment Manual for Depression*. Lebanon, NH: Project IMPACT, Dartmouth College; 2003.
27. McFarlane J, Parker B, Soeken K, Bullock L. Assessing for abuse during pregnancy. Severity and frequency of injuries and associated entry into prenatal care. *JAMA* 1992 Jun 17;267(23):3176-3178. [doi: [10.1001/jama.267.23.3176](https://doi.org/10.1001/jama.267.23.3176)] [Medline: [1593739](https://pubmed.ncbi.nlm.nih.gov/1593739/)]
28. NIH Toolbox, scoring and interpretation guide for the iPad. National Institutes of Health. URL: <https://nihtoolbox.my.salesforce.com/sfc/p/#2E000001H4ee/a/2E000000Ub7C/9UzNbxkwc9khhqKaYkXmAYUVZaniThvUkA16Ddkp6eY> [accessed 2022-07-27]
29. Beck AT, Steer RA, Brown GK. *Manual for the Beck Depression Inventory-II*. San Antonio, TX: Psychological Corporation; 1996.
30. Gershon R. NIH toolbox norming study. Harvard Dataverse. Boston, MA; 2016. URL: <https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/FF4DI7> [accessed 2020-09-09]
31. WebEx. Cisco Systems Inc. San Jose, CA; 2007. URL: <https://www.webex.com/industries/healthcare.html?dtid=ossdc000283> [accessed 2021-01-09]
32. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform* 2009 Apr;42(2):377-381 [FREE Full text] [doi: [10.1016/j.jbi.2008.08.010](https://doi.org/10.1016/j.jbi.2008.08.010)] [Medline: [18929686](https://pubmed.ncbi.nlm.nih.gov/18929686/)]
33. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, REDCap Consortium. The REDCap consortium: Building an international community of software platform partners. *J Biomed Inform* 2019 Jul;95:103208 [FREE Full text] [doi: [10.1016/j.jbi.2019.103208](https://doi.org/10.1016/j.jbi.2019.103208)] [Medline: [31078660](https://pubmed.ncbi.nlm.nih.gov/31078660/)]
34. Mowbray CT, Holter MC, Teague GB, Bybee D. Fidelity criteria: development, measurement, and validation. *Am J Eval* 2016 Jun 30;24(3):315-340. [doi: [10.1177/109821400302400303](https://doi.org/10.1177/109821400302400303)]
35. Cohen J. A Coefficient of agreement for nominal scales. *Educ Psychol Meas* 2016 Jul 02;20(1):37-46 [FREE Full text] [doi: [10.1177/001316446002000104](https://doi.org/10.1177/001316446002000104)]
36. Vidyo in health care. Vidyo Inc. 2018. URL: <https://www.vidyo.com/video-conferencing-solutions/health> [accessed 2017-02-17]

## Abbreviations

- BDI-II:** Beck Depression Inventory-Second Edition
- DBT:** Dialectical Behavior Therapy
- HIPAA:** Health Insurance Portability and Accountability Act
- PHS:** Perceived Hostility Survey
- REDCap:** Research Electronic Data Capture
- REST:** Resilience Enhancement Skills Training
- VCT:** video conferencing technology
- V-PST:** video-delivered problem solving therapy

*Edited by A Mavragani; submitted 05.08.22; peer-reviewed by K De Boer, V Newman-Morris; comments to author 30.08.22; revised version received 02.09.22; accepted 19.09.22; published 04.10.22.*

*Please cite as:*

*Cluxton-Keller F, Hegel MT*

*A Video-Delivered Family Therapeutic Intervention for Perinatal Women With Clinically Significant Depressive Symptoms and Family Conflict: Indicators of Feasibility and Acceptability*

*JMIR Form Res 2022;6(10):e41697*

*URL: <https://formative.jmir.org/2022/10/e41697>*

*doi: [10.2196/41697](https://doi.org/10.2196/41697)*

*PMID: [36194458](https://pubmed.ncbi.nlm.nih.gov/36194458/)*

©Fallon Cluxton-Keller, Mark T Hegel. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 04.10.2022. 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

# Development of a Web-Based System to Report Medication-Related Adverse Effects: Design and Usability Study

Renly Lim<sup>1</sup>, PhD; Christopher Thornton<sup>2</sup>, PhD; Jan Stanek<sup>3</sup>, MD, PhD; Lisa Kalisch Ellett<sup>1</sup>, PhD; Myra Thiessen<sup>4</sup>, PhD

<sup>1</sup>Quality Use of Medicines and Pharmacy Research Centre, UniSA Clinical and Health Sciences, University of South Australia, Adelaide, Australia

<sup>2</sup>UniSA Creative, University of South Australia, Adelaide, Australia

<sup>3</sup>UniSA STEM, University of South Australia, Adelaide, Australia

<sup>4</sup>Art, Design and Architecture, Monash University, Melbourne, Australia

**Corresponding Author:**

Renly Lim, PhD

Quality Use of Medicines and Pharmacy Research Centre

UniSA Clinical and Health Sciences

University of South Australia

Level 6, Playford Building

Frome Road

Adelaide, 5000

Australia

Phone: 61 883022307

Email: [renly.lim@unisa.edu.au](mailto:renly.lim@unisa.edu.au)

## Abstract

**Background:** Medicine use is the most common intervention in health care. The frequency with which medicines are used means medication-related problems are very common. One common type of medication-related problems is adverse drug events, which are unintended and harmful effects associated with use of medicines. Reporting of adverse drug events to regulatory authorities is important for evaluation of safety of medicines; however, these adverse effects are frequently unreported due to various factors, including lack of consumer-friendly reporting tools.

**Objective:** The aim of this study was to develop a user-friendly digital tool for consumers to report medication-related adverse effects.

**Methods:** The project consisted of 3 parts: (1) content development, including a systematic literature search; (2) iterative system development; and (3) usability testing. The project was guided by participatory design principles, which suggest involving key stakeholders throughout the design process. The first 2 versions were developed as a mobile app and were tested with end users in 2 workshops. The third version was developed as a web application and was tested with consumers who were taking regular medicines. Consumers were asked to complete a modified version of the mHealth app usability questionnaire (MAUQ), an 18-item questionnaire with each item scored using a 7-point Likert scale ranging from 0 (strongly disagree) to 7 (strongly agree). The MAUQ assessed 3 subscales including ease of use (5 items), interface and satisfaction (7 items), and usefulness (6 items). Continuous variables were reported as mean (SD) values, whereas categorical variables were presented as frequencies (percentages). Data analysis was conducted in Microsoft Excel.

**Results:** The content for the system was based on a systematic literature search and short-listing of questions, followed by feedback from project team members and consumers. Feedback from consumers in the 2 workshops were incorporated to improve the functionality, visual design, and stability of the third (current) version. The third version of the system was tested with 26 consumers. A total of 79% (N=307/390) of all responses on the MAUQ were scored 6 or 7, indicating that users generally strongly agree with the usability of the system. When looking at the individual domains, the system had an average score of 6.3 (SD 0.9) for “ease of use,” 6.3 (SD 0.8) for “interface and satisfaction,” and 5.2 (SD 1.4) for “usefulness.”

**Conclusions:** The web-based system for medicine adverse effects reporting is a user-friendly tool developed using an iterative participatory design approach. Future research includes further improving the system, particularly the usefulness of the system, as well as testing the scalability and performance of the system in practice.

(*JMIR Form Res* 2022;6(10):e37605) doi:[10.2196/37605](https://doi.org/10.2196/37605)

## KEYWORDS

adverse drug reaction; adverse drug event; digital health; eHealth; medication safety; mHealth; participatory design; patient reported outcomes; telehealth

## Introduction

Adverse drug events refer to the unintended and harmful effects associated with the use of medicines. Adverse drug reaction (ADR) is a subset of adverse drug events, where there is a causal relationship between the medicines and the adverse effects [1]. Adverse drug events add a significant burden to the health care system due to increased hospital admissions or emergency department visits; prolonged hospital stays; more complex patient management; complications, including disability and death; and potential prescribing cascade, where another medication is prescribed to ‘treat’ the adverse effects [2-6]. In Australia, we estimate that approximately 250,000 hospital admissions annually are medicine-related, costing AUD \$1.4 billion annually [7-9], and that at least one million people have experienced an ADR in the past 6 months [7].

Although some medication-related adverse effects such as fatigue and dizziness may not be considered serious, these adverse effects can have a profound impact on patients’ quality of life [5,10]. Further, these seemingly mild adverse effects have the potential to lead to more serious adverse events, such as falls, fractures, and hospitalizations. Recognizing and preventing medication-related adverse effects is therefore important, so that patients do not suffer from avoidable harms, with more serious adverse events also being preventable, resulting in fewer complications and a reduction in health care costs.

At the population level, reporting of ADRs to regulatory authorities is crucial for effective safety monitoring of medicines. Reporting rates of ADRs are however very low; up to 95% of ADRs are not reported to regulatory authorities [11]. In Australia, reporting of serious ADRs and adverse events to

regulatory authorities is mandatory for pharmaceutical companies but voluntary for health care professionals and patients. Most of ADR reports are made by pharmaceutical companies, about 20% are from health care professionals, and less than 5% are from patients [12].

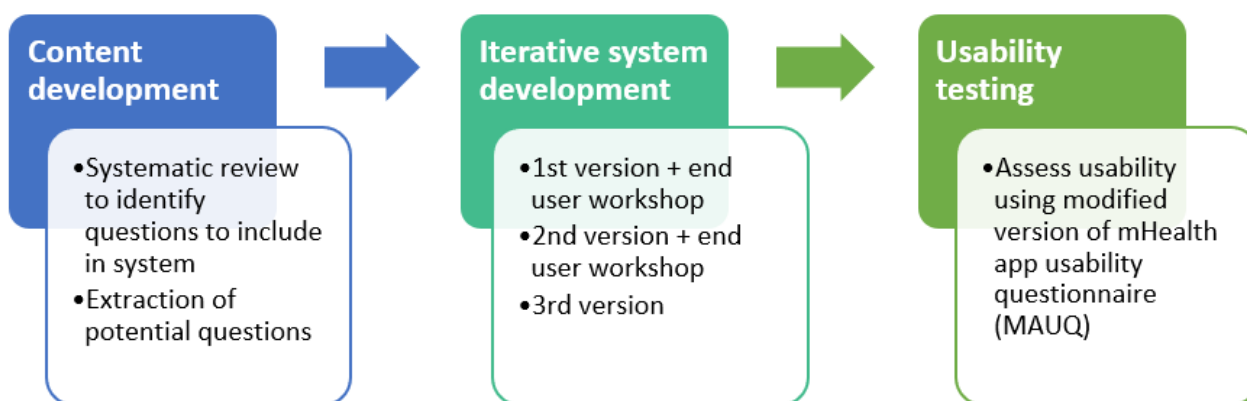
Many interventions have been developed to improve ADR reporting rates but have shown mixed results [13]. Most interventions targeted only health care professionals [13]. Reports from patients are increasingly recognized as important sources of information about ADRs [14]. Previous studies have shown that patients are able to detect ADRs first [15,16], and that patient ADR reports alert regulatory authorities to new and previously unknown ADRs [17]. However, barriers such as lack of awareness of the importance of ADR reporting, difficulty using existing reporting systems due to complicated language and cumbersome interfaces, and negative reporting experiences mean that ADR reporting rates from patients remain very low [11,18,19].

The aim of our study was to develop a user-friendly digital tool (SideRep) to report medication-related adverse effects, primarily for consumer use. This paper describes the development, design, and usability of the SideRep system.

## Methods

The SideRep project consisted of 3 parts: (1) content development, (2) iterative system development, and (3) usability testing (Figure 1). The 3 phases were conducted in South Australia between January and August 2020. The project was guided by participatory design principles [20], which suggest involving key stakeholders throughout the design process. Our project team consisted of researchers with backgrounds in communication design, pharmacy, and software development.

**Figure 1.** Processes involved during content development, system development, and usability testing.



### Content Development

The purpose of the project was to develop a simple and robust system for consumers to report any medication-related side effects that they felt. The first step of the project was therefore to determine relevant questions that should be included in the

system. A systematic literature review was carried out to identify available consumer-reported medicine side effects questionnaires [21]. Variables considered relevant to our app content development were extracted, including the number and types of questions asked, the use of scoring system, the presence of open-ended questions, and the time taken to complete the

questionnaire. One study investigator (RL, a clinical pharmacist) reviewed all short-listed questionnaires to extract possible questions for inclusion in our system, with the understanding that the questions may change following feedback from project team members and consumers during the second part of the study (ie, iterative system development).

### Iterative System Development

An iterative approach to the development of SideRep was taken, drawing on principles of participatory design [20] and human-centered design [22]. This was in recognition that research of this kind is both socially situated and socially constructed, where end users, as experts of their own experience,

should be directly involved in the design decision-making that affects their experience. In the first 2 versions, the systems were developed to run on desktops and various mobile platforms (tested on iPad and Samsung Galaxy 7). The layered architecture of SideRep system allowed some flexibility—the first and second versions were mainly desktop oriented, making use of bespoke JavaScript and designed as a mobile app and desktop application. This dichotomy became a hindrance when moving more heavily to the mobile platform. The development team moved to React library, React Redux toolkit, React Hook Form, Material-UI, Yarn, and Axios for the front end (Figure 2); the third (current) system was hosted and supported by Amazon Web Services (Figure 3).

Figure 2. Front end of the SideRep system.

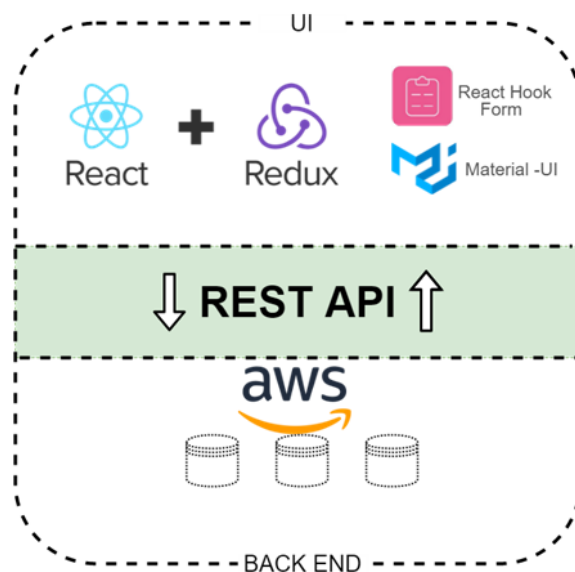
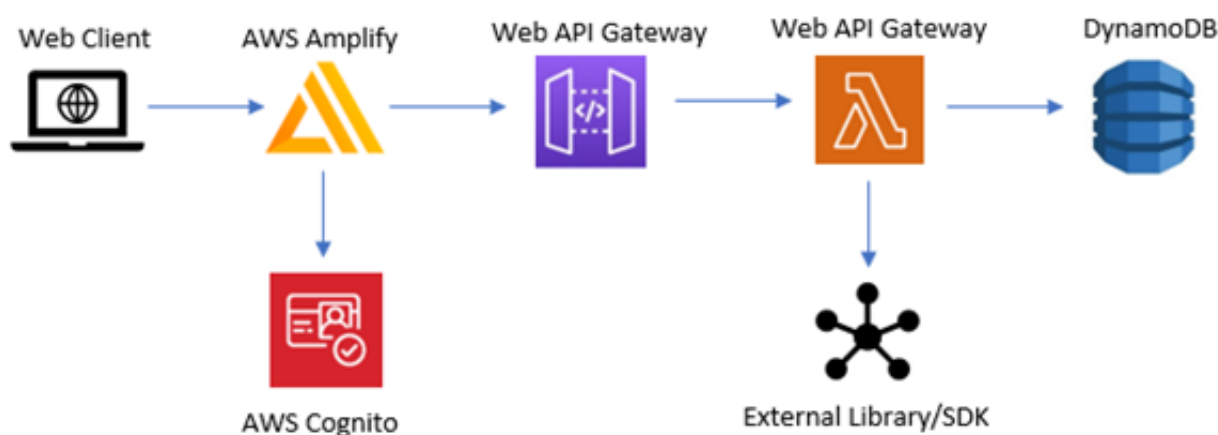


Figure 3. Back end of the SideRep system on Amazon Web Services.



### User Testing

The first 2 SideRep versions were tested in 2 separate workshops each with 3 end user participants. Participants were recruited from members of the University of Third Age in South Australia, a volunteer organization for people over the age of 50 years. Participants were included if they were taking any regular medications, had access to internet, and agreed to download the app on their phones. The workshops consisted of several stepped

activities, where the participants first downloaded the app from the Apple App Store or the Google Play Store, created an account, signed in, and ‘reported’ 3 different and detailed but fictitious adverse events. Each participant was asked to ‘report’ the same adverse effects (ie, headache, skin rash, and indigestion), and therefore, complete the same sequence of events. Participants were observed completing the tasks and were asked a range of semistructured interview questions that aimed to gather data about their experience using the app (eg,

ease of use, design, functionality, and language or terminology). A further set of more probing, unstructured, or informal discussion questions were then asked, giving participants the opportunity to share their opinions and ideas for future design development at greater length (eg, the functions they thought might be useful to include or unnecessary features that might otherwise be removed). Observing how participants interacted with the app and talking to them during and after testing helped the researchers understand any perceived difficulties in use or comprehension, so as to incorporate their critical feedback into the third version of the system.

The third (current) version of the SideRep system was tested with consumers who were taking regular medicines. Participants were recruited from people attending public health talks presented by one of the project team members (RL). Participants were invited to trial the SideRep system and complete a modified version of the mHealth app usability questionnaire (MAUQ) [23] after using the system. The MAUQ is an 18-item questionnaire that evaluates the usability of the mobile health app [23]. The MAUQ assessed 3 subscales including ease of use (5 items), interface and satisfaction (7 items), and usefulness (6 items). Each item is scored with a 7-point Likert scale ranging from 0 (strongly disagree) to 7 (strongly agree), with an additional “not applicable” option if the question does not apply to the participants’ experience. Since the third version of the SideRep system was a multiplatform web-based application (ie, not a mobile app), we replaced the word “app” in the MAUQ with “system.”

### Data Analysis

Continuous variables were reported as mean (SD), whereas categorical variables were presented as frequencies (percentages). To aid interpretation of the usability testing data, results were presented both as mean (SD) and as numbers (percentages) of people who answered “strongly agree” (ie, scores of 6 or 7) to each MAUQ item. Data analysis was conducted in Microsoft Excel.

### Ethics Approval

The project was approved by the University of South Australia Human Research Ethics Committee (202532).

## Results

### Content Development

The initial content for the SideRep system was primarily based on findings from our systematic review. Detailed findings of the review has been published elsewhere [21]. Of the 19 questionnaires identified, 15 were for a specific condition or medication, and 4 were general questionnaires applicable to any medication [21]. Two of the generic questionnaires, developed by Jarensiripornkul et al [24] and de Vries et al [25], were considered the most relevant for our study objective. Both questionnaires had a comprehensive list of symptoms categorized in body categories but were lengthy questionnaires. For example, the questionnaire by de Vries [25] took a median of 30 minutes to complete for patients reporting at least one adverse drug event. The length of time needed to complete the questionnaire was considered too long by the project team. The questionnaires were therefore reviewed to extract only questions deemed important and relevant for our study purpose.

After a review by the project team members, the following 5 main questions were extracted for inclusion in the SideRep system: (1) medication(s) recently started, including the date it started (and the date it stopped, if applicable); (2) whether the participants had been taking the medication as prescribed; (3) symptom change experienced in the last 4 weeks, with options classified by body categories; (4) the degree of bothersome symptoms (eg, not at all bothersome, and minimally, moderately, or severely bothersome); and (5) whether the participants had or planned to inform their health care professionals about the symptoms.

### Iterative System Development

The second part of the project involved iterative system development and workshops with small groups of 3 end users each. Participants from the first workshop thought that the SideRep app was easy to use, and they had no problem reporting medicine adverse effects. They provided the investigators with a list of desired features and highlighted design issues that could be improved. This led to the second version, which focused mainly on the functionality and visual design of the app and improving the stability of the app (Figure 4). Feedback was again that the “app was easy to use” and that the longest time taken was for the app download, registration, and log-in. Participants also gave suggestions on the kind of questions to include.

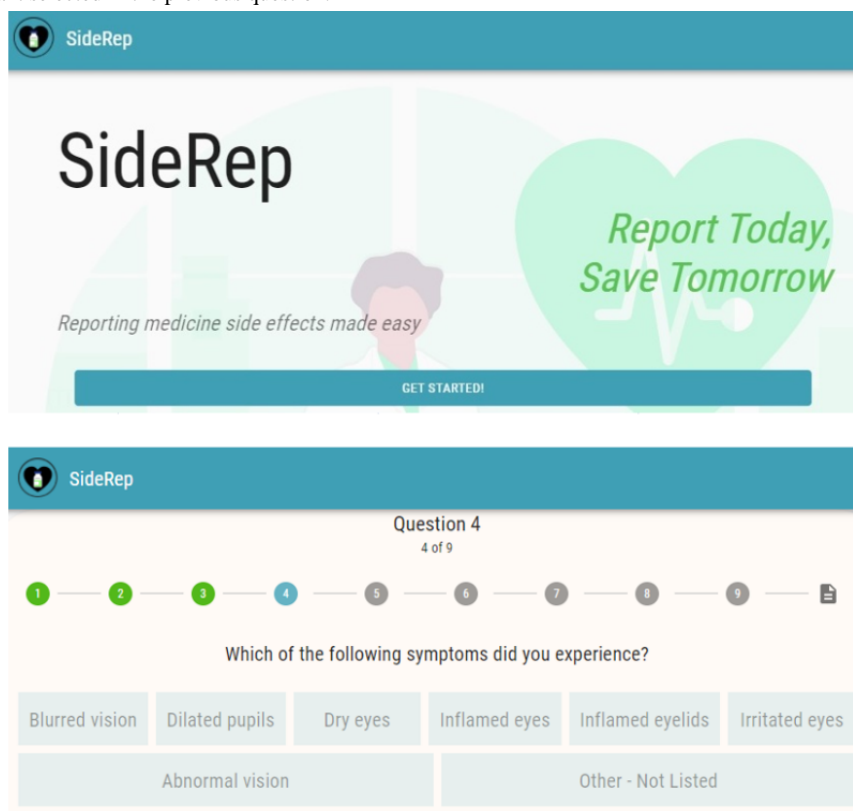
**Figure 4.** Example screenshots of the second version of SideRep system. Left: main screen; middle: list of medications; right: questions related to reporting of medicine side effects.



The third (current) version was developed as a cloud-hosted web-based application with no requirement to download and install an app or to register an account (Figure 5). This immediately mitigated the set of procedures new users found most complex and time-consuming and minimized further risk of error or frustration in user experience. User feedback and consultation within the project team resulted in several minor changes to the final structure and sequence of the questions. There were 8 main questions in the third version of SideRep web system, as follows: (1) what medication(s) was recently

used, including the date it started (and the date it stopped, if applicable); (2) symptom change experienced in the last 4 weeks, with options classified by body categories; (3) the degree of discomfort caused by the symptoms; (4) whether the participants had or planned to inform their health care professionals about the symptoms; (5) whether the participants had been taking the medication as prescribed; (6) whether the problem resolved if the participants indicated they stopped taking the medicine; (7) participant characteristics, including age and sex; and (8) free text to include additional information.

**Figure 5.** Example screenshots of the third (current) iteration of the SideRep system. Top: main screen; bottom: symptoms listed based on body categories that the participant selected in the previous question.



### Usability Testing

The third (current) version of the SideRep system was tested with 26 participants. The average age was 45 (SD 11.8) years. A total of 16 (61%) participants were female, 8 (31%) were male, and 2 (8%) preferred not to specify. The SideRep system received an average MAUQ score of 6.1 (SD 1.1), and 79% of all responses were scored 6 or 7, indicating that users generally strongly agree with the usability of the system. When looking

at the individual domains, the SideRep system had an average score of 6.3 (SD 0.9) for “ease of use,” with 90% (n=109) of the responses being “strongly agree.” In terms of “interface and satisfaction,” the average score was 6.3 (SD 0.8), and 85% (n=155) of the responses were “strongly agree.” In the “usefulness” domain, the average score was 5.2 (SD 1.4), with 50% (n=43) of the responses being “strongly agree.” The mean (SD) scores and number (%) of “strongly agree” responses for each statement can be seen in [Table 1](#).



**Table 1.** Usability of the SideRep system as assessed using the modified version of the mHealth app usability questionnaire (N=26).

Domain and statements	Mean (SD)	Strongly agree, n/N (%) <sup>a</sup>	Not applicable <sup>b</sup>
<b>Ease of use</b>			
The system was easy to use.	6.5 (0.6)	24/26 (92)	— <sup>c</sup>
It was easy for me to learn to use the system.	6.5 (0.6)	25/26 (96)	—
The navigation was consistent when moving between screens.	6.4 (1.1)	23/26 (88)	—
The interface of the system allowed me to use all the functions (such as entering information, responding to reminders, and viewing information) offered by the system.	6.2 (0.7)	23/26 (88)	—
Whenever I made a mistake using the system, I could recover easily and quickly.	6.1 (1.2)	14/17 (82)	9
Overall score	6.3 (0.9)	109/121 (90)	—
<b>Interface and satisfaction</b>			
I like the interface of the system.	6.1 (0.8)	23/26 (88)	—
The information in the system was well organized, so I could easily find the information I needed.	6.1 (0.9)	21/26 (81)	—
The system adequately acknowledged and provided information to let me know the progress of my action.	6.0 (1.0)	18/26 (69)	—
I feel comfortable using this system in social settings.	6.6 (0.7)	23/26 (88)	—
The amount of time involved in using this system has been fitting for me.	6.5 (0.8)	24/26 (92)	—
I would use this system again.	6.4 (0.7)	22/26 (85)	—
Overall, I am satisfied with this system.	6.4 (0.6)	24/26 (92)	—
Overall score	6.3 (0.8)	155/182 (85)	—
<b>Usefulness</b>			
The system would be useful for my health and well-being.	5.2 (1.2)	7/17 (41)	9
The system improved my access to health care services.	5.1 (1.6)	6/11 (55)	15
The system helped me manage my health effectively.	5.1 (1.5)	9/17 (53)	9
This system has all the functions and capabilities I expected it to have.	4.9 (1.3)	9/26 (35)	—
I could use the system even when the Internet connection was poor or not available.	5.5 (1.6)	5/6 (83)	20
This system provides an acceptable way to receive health care services, such as accessing educational materials, tracking my own activities, and performing self-assessment.	5.7 (0.8)	7/10 (70)	16
Overall score	5.2 (1.4)	43/87 (49)	—

<sup>a</sup>N=26 participants except where participants answered “not applicable” to the question.

<sup>b</sup>Number of participants who answered “not applicable” for that statement.

<sup>c</sup>No participants answered “not applicable” to the question.

## Discussion

### Principal Findings

In this paper, we described the development and design of a consumer-friendly digital tool to report medication-related adverse effects and tested the usability of the system in participants who were taking medicines. We drew on principles of participatory design and were guided by a human-centered design methodology, including an iterative design process, and the inclusion of stakeholders, including consumers at key stages of SideRep’s development. Consumer feedback was used to improve the system design and content. The third (current) version of the system was tested in a group of participants. Most

users (>85%) strongly agreed that the system was easy to use and were satisfied with the system, with an average score of 6.3 for both “ease of use” (SD 0.9) and “interface and satisfaction” (SD 0.8). About half of the participants strongly agreed that the system was useful, with an average score of 5.2 (SD 1.4) for the “usefulness” domain.

From the earliest applications of user-centered design, the concern has been to learn about and prioritize the needs and preferences of people who will interact with the designed object [22]. This deep understanding is essential to designing digital health systems, like SideRep, that will improve their ease of use and comprehension. One key insight gained through this iterative design process was that an app platform was too

limiting and rigid to accommodate the SideRep system. By the nature of smart device apps, users must download software updates to keep the platform active. They also require user log-ins; that can make users reluctant to create a personalized account, and they may also fear for the safety of their data and personal information. These are all factors that may de incentivize people from reporting ADRs. The consumer workshops also revealed that individuals may lack the level of digital literacy needed to complete prerequisite tasks such as finding, downloading, and setting up the app on their smart device, and that this may inadvertently exclude consumers from using the system. In response to these issues and to avoid problematic, confusing, and time-consuming activities that may deter people from reporting ADRs [11,18], SideRep was transitioned from a smart device app to a web-based platform. Importantly, a web-based platform is still accessible on mobile devices as well as laptop and desktop computers; however, websites allow for more fluid, seamless, and responsive updating of the content (eg, adding new medicines, improving clarity of language, and site structure) at a lower cost to developers. This effectively removed the procedural barriers of finding, downloading, and updating an app and enabled an anonymous method of reporting without the burden of individual accounts or log-ins. Furthermore, as web technologies change, a server-based deployment enables the updating and improvement of system performance without the need for end users to be involved. This provides an effective means of future-proofing at a lower overall cost to developers and without the risk of deterring existing or new users. As a result, the usability survey showed high ease of use and participation satisfaction with the reporting system.

Many interventions have been developed to improve the rates of spontaneous ADR reporting to regulatory authorities [13,26,27]. These interventions aim to increase use of spontaneous reporting systems but do not address underlying reasons for low system uptake such as lack of time, difficulty accessing or using the system, negative reporting experience, and lack of feedback on submitted reports [11,18,28]. Many factors may influence acceptability and use of an ADR reporting system, including the number of questions, the way the questions are formulated, the options provided (eg, free text vs menu), and the language used (eg, the use of jargon when the target audience is consumers) [29,30]. In mid-2015, the European Union's WEB-RADR (or Recognizing Adverse Drug Reactions) project developed and launched a new mobile app to improve reporting of ADRs from patients and health care professionals [31]. The app was developed following focus groups and interviews with potential users. Several desired functionalities were incorporated including 2-way exchange of safety

information (ie, users could receive safety alerts and news in addition to reporting ADRs) [31]. Despite this, the uptake was low; between 2015 and 2017, only 838 ADR reports were submitted through the app in countries where the app was launched (ie, United Kingdom, Netherlands, and Croatia) [31]. The effects of making access to the SideRep system more straightforward, that is, without needing to install an app, on system uptake will need to be tested in practice.

### Limitations

We tested the first and second versions of the SideRep system with only a small group of end users (ie, 2 workshops each with 3 end user participants). We had originally intended to conduct face-to-face workshops in the second quarter of 2020, with 10 participants at each workshop. However, due to restrictions following the COVID-19 pandemic, we submitted an ethics amendment to conduct the workshops via web-based videoconferencing. As a result, we had to limit the number of participants at each workshop because it would be too difficult to manage a large number of participants in a web-based workshop. This meant that we received feedback from only a small number of participants for the first and second versions. However, feedback from all consumers during the workshops were generally positive. We did not perform causality assessment to determine whether the adverse effects reported by the consumers were attributed to the medicines, that is, whether or not the effects were ADRs. Causality assessment is an important component for medicine safety signal detection; however, the purpose of our study was to first determine whether an alternative system for reporting medication-related adverse effects was feasible. Thus, causality assessment was considered beyond the scope of our project. Finally, we do not yet know how this system will be implemented in practice and whether it will lead to increased consumer-led ADR reporting to regulatory authorities. It will be important to test how the information reported by consumers can be used by regulatory authorities and health professionals. The next step would include testing the SideRep system in practice using real-world situations.

### Conclusions

The SideRep web-based system for medicine adverse effects reporting is a user-friendly tool developed using an iterative participatory design approach. Future research should include further improving the system, particularly the usefulness of the system, as well as testing the scalability and performance of the system in practice. Successful implementation of the system has the potential to allow for early detection and prevention of medicine-induced harms, and it could increase consumer ADR reporting to regulatory authorities.

### Acknowledgments

We would like to thank Carl Green, Ethan Rickaby, Robert Dicker, and Joshua Johnson for their assistance with the development of the SideRep system. The study was funded by a National Health and Medical Research Council (NHMRC) fellowship to RL (APP1156368).

## Authors' Contributions

All authors contributed to the conception, design, and development of the study. CT, LKE, and MT facilitated the workshops. RL performed the data analysis and wrote the first draft of the manuscript. All authors participated in the review and final approval of the manuscript.

## Conflicts of Interest

None declared.

## References

1. Edwards IR, Aronson JK. Adverse drug reactions: definitions, diagnosis, and management. *The Lancet* 2000 Oct;356(9237):1255-1259. [doi: [10.1016/s0140-6736\(00\)02799-9](https://doi.org/10.1016/s0140-6736(00)02799-9)]
2. Giardina C, Cutroneo PM, Mocciaro E, Russo GT, Mandraffino G, Basile G, et al. Adverse drug reactions in hospitalized patients: results of the FORWARD (Facilitation of Reporting in Hospital Ward) study. *Front Pharmacol* 2018 Apr 11;9:350 [FREE Full text] [doi: [10.3389/fphar.2018.00350](https://doi.org/10.3389/fphar.2018.00350)] [Medline: [29695966](https://pubmed.ncbi.nlm.nih.gov/29695966/)]
3. Sultana J, Cutroneo P, Trifirò G. Clinical and economic burden of adverse drug reactions. *J Pharmacol Pharmacother* 2013 Dec 01;4(1\_suppl):S73-S77. [doi: [10.4103/0976-500x.120957](https://doi.org/10.4103/0976-500x.120957)]
4. Kalisch LM, Caughey GE, Roughead EE, Gilbert AL. The prescribing cascade. *Aust Prescr* 2011 Dec 01;34(6):162-166. [doi: [10.18773/austprescr.2011.084](https://doi.org/10.18773/austprescr.2011.084)]
5. Khalil H, Huang C. Adverse drug reactions in primary care: a scoping review. *BMC Health Serv Res* 2020 Jan 06;20(1):5 [FREE Full text] [doi: [10.1186/s12913-019-4651-7](https://doi.org/10.1186/s12913-019-4651-7)] [Medline: [31902367](https://pubmed.ncbi.nlm.nih.gov/31902367/)]
6. Roughead E, Pratt NL, Parfitt G, Rowett D, Kalisch-Ellett LM, Bereznicki L, et al. Effect of an ongoing pharmacist service to reduce medicine-induced deterioration and adverse reactions in aged-care facilities (nursing homes): a multicentre, randomised controlled trial (the ReMInDAR trial). *Age Ageing* 2022 Apr 01;51(4):a [FREE Full text] [doi: [10.1093/ageing/afac092](https://doi.org/10.1093/ageing/afac092)] [Medline: [35460410](https://pubmed.ncbi.nlm.nih.gov/35460410/)]
7. Lim R. Medicine safety: take care. Pharmaceutical Society of Australia. 2019 Jan. URL: <https://www.psa.org.au/wp-content/uploads/2019/01/PSA-Medicine-Safety-Report.pdf> [accessed 2022-10-04]
8. Lim R, Ellett LMK, Semple S, Roughead EE. The extent of medication-related hospital admissions in Australia: a review from 1988 to 2021. *Drug Saf* 2022 Mar 28;45(3):249-257 [FREE Full text] [doi: [10.1007/s40264-021-01144-1](https://doi.org/10.1007/s40264-021-01144-1)] [Medline: [35089582](https://pubmed.ncbi.nlm.nih.gov/35089582/)]
9. Lim R. Medicine safety: aged care. Pharmaceutical Society of Australia. 2020 Feb. URL: <https://www.psa.org.au/wp-content/uploads/2020/02/Medicine-Safety-Aged-Care-WEB-RES1.pdf> [accessed 2022-10-04]
10. Chimirri S, Aiello R, Mazzitello C, Mumoli L, Palleria C, Altomonte M, et al. Vertigo/dizziness as a drugs' adverse reaction. *J Pharmacol Pharmacother* 2013 Dec 01;4(1\_suppl):S104-S109. [doi: [10.4103/0976-500x.120969](https://doi.org/10.4103/0976-500x.120969)]
11. Hazell L, Shakir SAW. Under-reporting of adverse drug reactions : a systematic review. *Drug Saf* 2006;29(5):385-396. [doi: [10.2165/00002018-200629050-00003](https://doi.org/10.2165/00002018-200629050-00003)] [Medline: [16689555](https://pubmed.ncbi.nlm.nih.gov/16689555/)]
12. Annual performance statistics report 2019-20. Australian Government Department of Health. URL: <https://www.tga.gov.au/sites/default/files/annual-performance-statistics-report-july-2019-june-2020.pdf> [accessed 2022-10-18]
13. Paudyal V, Al-Hamid A, Bowen M, Hadi MA, Hasan SS, Jalal Z, et al. Interventions to improve spontaneous adverse drug reaction reporting by healthcare professionals and patients: systematic review and meta-analysis. *Expert Opin Drug Saf* 2020 Sep 03;19(9):1173-1191. [doi: [10.1080/14740338.2020.1807003](https://doi.org/10.1080/14740338.2020.1807003)] [Medline: [32755492](https://pubmed.ncbi.nlm.nih.gov/32755492/)]
14. Inácio P, Cavaco A, Airaksinen M. The value of patient reporting to the pharmacovigilance system: a systematic review. *Br J Clin Pharmacol* 2017 Feb 12;83(2):227-246 [FREE Full text] [doi: [10.1111/bcp.13098](https://doi.org/10.1111/bcp.13098)] [Medline: [27558545](https://pubmed.ncbi.nlm.nih.gov/27558545/)]
15. Mitchell AS, Henry DA, Hennrikus D, O'Connell DL. Adverse drug reactions: can consumers provide early warning? *Pharmacoepidem Drug Safe* 1994 Sep;3(5):257-264. [doi: [10.1002/pds.2630030504](https://doi.org/10.1002/pds.2630030504)]
16. Egberts TCG, Smulders M, de Koning FHP, Meyboom RHB, Leufkens HGM. Can adverse drug reactions be detected earlier? A comparison of reports by patients and professionals. *BMJ* 1996 Aug 31;313(7056):530-531 [FREE Full text] [doi: [10.1136/bmj.313.7056.530](https://doi.org/10.1136/bmj.313.7056.530)] [Medline: [8789980](https://pubmed.ncbi.nlm.nih.gov/8789980/)]
17. Hasford J, Bruchmann F, Lutz M, Thürmann P, Schmiedl S. A patient-centred web-based adverse drug reaction reporting system identifies not yet labelled potential safety issues. *Eur J Clin Pharmacol* 2021 Nov 18;77(11):1697-1704 [FREE Full text] [doi: [10.1007/s00228-021-03134-9](https://doi.org/10.1007/s00228-021-03134-9)] [Medline: [34143228](https://pubmed.ncbi.nlm.nih.gov/34143228/)]
18. Al Dweik R, Stacey D, Kohen D, Yaya S. Factors affecting patient reporting of adverse drug reactions: a systematic review. *Br J Clin Pharmacol* 2017 Apr 25;83(4):875-883 [FREE Full text] [doi: [10.1111/bcp.13159](https://doi.org/10.1111/bcp.13159)] [Medline: [27868226](https://pubmed.ncbi.nlm.nih.gov/27868226/)]
19. Hohl CM, Small SS, Peddie D, Badke K, Bailey C, Balka E. Why clinicians don't report adverse drug events: qualitative study. *JMIR Public Health Surveill* 2018 Mar 27;4(1):e21 [FREE Full text] [doi: [10.2196/publichealth.9282](https://doi.org/10.2196/publichealth.9282)] [Medline: [29487041](https://pubmed.ncbi.nlm.nih.gov/29487041/)]
20. Luck R. What is it that makes participation in design participatory design? *Des Stud* 2018 Nov;59:1-8. [doi: [10.1016/j.destud.2018.10.002](https://doi.org/10.1016/j.destud.2018.10.002)]

21. Lim R, Ellett L, Roughead E, Cheah P, Masnoon N. Patient-reported questionnaires to identify adverse drug reactions: a systematic review. *Int J Environ Res Public Health* 2021 Nov 12;18(22):11877 [FREE Full text] [doi: [10.3390/ijerph182211877](https://doi.org/10.3390/ijerph182211877)] [Medline: [34831635](https://pubmed.ncbi.nlm.nih.gov/34831635/)]
22. van der Bijl-Brouwer M, Dorst K. Advancing the strategic impact of human-centred design. *Des Stud* 2017 Nov;53:1-23. [doi: [10.1016/j.destud.2017.06.003](https://doi.org/10.1016/j.destud.2017.06.003)]
23. Zhou L, Bao J, Setiawan IMA, Saptono A, Parmanto B. The mHealth app usability questionnaire (MAUQ): development and validation study. *JMIR Mhealth Uhealth* 2019 Apr 11;7(4):e11500 [FREE Full text] [doi: [10.2196/11500](https://doi.org/10.2196/11500)] [Medline: [30973342](https://pubmed.ncbi.nlm.nih.gov/30973342/)]
24. Jarensiripornkul N, Krska J, Capps PAG, Richards RME, Lee A. Patient reporting of potential adverse drug reactions: a methodological study. *Br J Clin Pharmacol* 2002 Mar;53(3):318-325 [FREE Full text] [doi: [10.1046/j.0306-5251.2001.01547.x](https://doi.org/10.1046/j.0306-5251.2001.01547.x)] [Medline: [11874396](https://pubmed.ncbi.nlm.nih.gov/11874396/)]
25. de Vries ST, Mol PGM, de Zeeuw D, Haaijer-Ruskamp FM, Denig P. Development and initial validation of a patient-reported adverse drug event questionnaire. *Drug Saf* 2013 Sep 4;36(9):765-777. [doi: [10.1007/s40264-013-0036-8](https://doi.org/10.1007/s40264-013-0036-8)] [Medline: [23553447](https://pubmed.ncbi.nlm.nih.gov/23553447/)]
26. Li R, Zaidi STR, Chen T, Castelino R. Effectiveness of interventions to improve adverse drug reaction reporting by healthcare professionals over the last decade: A systematic review. *Pharmacoepidemiol Drug Saf* 2020 Jan 14;29(1):1-8. [doi: [10.1002/pds.4906](https://doi.org/10.1002/pds.4906)] [Medline: [31724270](https://pubmed.ncbi.nlm.nih.gov/31724270/)]
27. Khalili M, Mesgarpour B, Sharifi H, Daneshvar Dehnavi S, Haghdoost AA. Interventions to improve adverse drug reaction reporting: a scoping review. *Pharmacoepidemiol Drug Saf* 2020 Sep 19;29(9):965-992. [doi: [10.1002/pds.4966](https://doi.org/10.1002/pds.4966)] [Medline: [32431069](https://pubmed.ncbi.nlm.nih.gov/32431069/)]
28. Onakpoya IJ, Heneghan CJ, Aronson JK. Post-marketing withdrawal of 462 medicinal products because of adverse drug reactions: a systematic review of the world literature. *BMC Med* 2016 Feb 4;14(1). [doi: [10.1186/s12916-016-0553-2](https://doi.org/10.1186/s12916-016-0553-2)]
29. de Vries ST, Wong L, Sutcliffe A, Houyez F, Ruiz CL, Mol PGM, IMI Web-RADR Work Package 3b Consortium. Factors influencing the use of a mobile app for reporting adverse drug reactions and receiving safety information: a qualitative study. *Drug Saf* 2017 May 29;40(5):443-455 [FREE Full text] [doi: [10.1007/s40264-016-0494-x](https://doi.org/10.1007/s40264-016-0494-x)] [Medline: [28035492](https://pubmed.ncbi.nlm.nih.gov/28035492/)]
30. Peddie D, Small SS, Badke K, Bailey C, Balka E, Hohl CM. Adverse drug event reporting from clinical care: mixed-methods analysis for a minimum required dataset. *JMIR Med Inform* 2018 Jun 28;6(2):e10248 [FREE Full text] [doi: [10.2196/10248](https://doi.org/10.2196/10248)] [Medline: [29954724](https://pubmed.ncbi.nlm.nih.gov/29954724/)]
31. Pierce CE, de Vries ST, Bodin-Parssinen S, Härmark L, Tregunno P, Lewis DJ, et al. Recommendations on the use of mobile applications for the collection and communication of pharmaceutical product safety information: lessons from IMI WEB-RADR. *Drug Saf* 2019 Apr 25;42(4):477-489 [FREE Full text] [doi: [10.1007/s40264-019-00813-6](https://doi.org/10.1007/s40264-019-00813-6)] [Medline: [30911975](https://pubmed.ncbi.nlm.nih.gov/30911975/)]

## Abbreviations

**ADR:** adverse drug reaction

**MAUQ:** mHealth app usability questionnaire.

*Edited by A Mavragani; submitted 27.02.22; peer-reviewed by D Pessoa, H Mehdizadeh, C Hohl; comments to author 20.04.22; revised version received 02.08.22; accepted 17.08.22; published 07.10.22.*

*Please cite as:*

*Lim R, Thornton C, Stanek J, Ellett LK, Thiessen M*

*Development of a Web-Based System to Report Medication-Related Adverse Effects: Design and Usability Study*

*JMIR Form Res* 2022;6(10):e37605

URL: <https://formative.jmir.org/2022/10/e37605>

doi: [10.2196/37605](https://doi.org/10.2196/37605)

PMID: [36206034](https://pubmed.ncbi.nlm.nih.gov/36206034/)

©Renly Lim, Christopher Thornton, Jan Stanek, Lisa Kalisch Ellett, Myra Thiessen. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 07.10.2022. 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

# Information Resources Among Flemish Pregnant Women: Cross-sectional Study

Dorien Lanssens<sup>1</sup>, RM, MSc, PhD; Inge M Thijs<sup>1,2</sup>, MSc, ING; Pauline Dreesen<sup>1,2</sup>, MSc; Ann Van Hecke<sup>3,4</sup>, RN, MSc, PhD; Pascal Coorevits<sup>3</sup>, MSc, PhD; Gitte Gaethofs<sup>1,2</sup>, MSc; Joyce Derycke<sup>3</sup>, MSc; Inge Tency<sup>3,5</sup>, RM, MSc, PhD

<sup>1</sup>Faculty of Medicine and Life Sciences, Limburg Clinical Research Center, University of Hasselt, Hasselt, Belgium

<sup>2</sup>Obstetric Department and Future Health, Limburg Clinical Research Center, Ziekenhuis Oost-Limburg, Genk, Belgium

<sup>3</sup>Department Public Health and Primary Care, Faculty of Medicine and Health Sciences, Ghent University, Ghent, Belgium

<sup>4</sup>Nursing Department, Ghent University Hospital, Ghent, Belgium

<sup>5</sup>Department of Midwifery, Odisee University of Applied Sciences, Sint Niklaas, Belgium

**Corresponding Author:**

Dorien Lanssens, RM, MSc, PhD  
Faculty of Medicine and Life Sciences  
Limburg Clinical Research Center  
University of Hasselt  
Martelarenlaan 42  
Hasselt, 3500  
Belgium  
Phone: 32 494919476  
Email: [dorien.lanssens@uhasselt.be](mailto:dorien.lanssens@uhasselt.be)

## Abstract

**Background:** There has been an exponential growth in the availability of apps, resulting in increased use of pregnancy apps. However, information on resources and use of apps among pregnant women is relatively limited.

**Objective:** The aim of this study is to map the current information resources and the use of pregnancy apps among pregnant women in Flanders.

**Methods:** A cross-sectional study was conducted, using a semistructured survey (April-June 2019) consisting of four different domains: (1) demographics; (2) use of devices; (3) sources of information; and (4) use of pregnancy apps. Women were recruited by social media, flyers, and paper questionnaires at prenatal consultations. Statistical analysis was mainly focused on descriptive statistics. Differences in continuous and categorical variables were tested using independent Student *t* tests and chi-square tests. Correlations were investigated between maternal characteristics and the women's responses.

**Results:** In total, 311 women completed the entire questionnaire. Obstetricians were the primary source of information (268/311, 86.2%) for pregnant women, followed by websites/internet (267/311, 85.9%) and apps (233/311, 74.9%). The information that was most searched for was information about the development of the baby (275/311, 88.5%), discomfort/complaints (251/311, 80.7%) and health during pregnancy (248/311, 79.7%), administrative/practical issues (233/311, 74.9%), and breastfeeding (176/311, 56.6%). About half of the women (172/311, 55.3%) downloaded a pregnancy app, and primarily searched app stores (133/311, 43.0%). Pregnant women who are single asked their mothers (22/30, 73.3%) or other family members (13/30, 43.3%) for significantly more information than did married women (mother [in law]: 82/160, 51.3%,  $P=.02$ ; family members: 35/160, 21.9%,  $P=.01$ ). Pregnant women with lower education were significantly more likely to have a PC or laptop than those with higher education (72/73, 98.6% vs 203/237, 85.5%;  $P=.008$ ), and to consult other family members for pregnancy information (30/73, 41.1% vs 55/237, 23.1%;  $P<.001$ ), but were less likely to consult a gynecologist (70/73, 95.9% vs 198/237, 83.5%;  $P=.001$ ). They also followed more prenatal sessions (59/73, 80.8% vs 77/237, 32.5%;  $P=.04$ ) and were more likely to search for information regarding discomfort/complaints during pregnancy (65/73, 89% vs 188/237, 79.5%;  $P=.02$ ). Compared to multigravida, primigravida were more likely to solicit advice about their pregnancy from other women in their social networks (family members: primigravida 44/109, 40.4% vs multigravida 40/199, 20.1%;  $P<.001$ ; other pregnant women: primigravida 58/109, 53.2% vs multigravida 80/199, 40.2%;  $P<.03$ ).

**Conclusions:** Health care professionals need to be aware that apps are important and are a growing source of information for pregnant women. Concerns rise about the quality and safety of those apps, as only a limited number of apps are subjected to an external quality check. Therefore, it is important that health care providers refer to high-quality digital resources and take the opportunity to discuss digital information with pregnant women.

(*JMIR Form Res* 2022;6(10):e37866) doi:[10.2196/37866](https://doi.org/10.2196/37866)

## KEYWORDS

pregnancy app; mobile app; questionnaire; pregnancy; pregnant; mHealth; mobile health; cross-sectional; user need; user expectation; survey; maternal; maternity; user experience

## Introduction

Access to reliable information is critical to women's experiences and well-being during pregnancy and childbirth [1,2]. Information and education help women understand what is happening and what can happen during their pregnancy [3] and can improve women's satisfaction with the childbirth experience [4]. Pregnant women seek information to feel more confident and comfortable in their communication with health care providers, to make decisions during the perinatal period, and to prepare themselves for their maternal responsibilities [4-8]. Fulfilling a woman's information needs depends on her access to adequate resources and her ability to comprehend what has been presented to her [7]. In the current context of our information-rich society, women are exposed to a wide range of information sources. This includes information sources from the health care system, conventional sources (eg, family, peers, and books), and digital information sources (eg, websites, apps, and social media) or eHealth [9,10].

The term eHealth refers to the application of electronic information and communication techniques in health care, primarily intended to improve patients' health and the quality of care. eHealth is one of the fastest growing domains within health care [11,12]. Mobile health (mHealth) is a component of eHealth and is defined as the use of mobile technologies such as smartphones, tablets, computers, and other wireless devices (eg, pedometers, smartwatches) to support health services and improve the quality and efficiency of care [11,13-15].

mHealth is on the rise in health care, resulting in the exponential growth of mHealth apps [12,14,16]. The largest group of smartphone users are Millennials, those currently aged 18-34 years, which aligns with the time when many first experience pregnancy and parenthood [17,18]. Therefore, it is not surprising that there is a wide range of mobile apps on the topics of pregnancy, birth, and parenthood, with more than 1000 pregnancy apps available in the various app stores [19-23]. In line with the general trend in mHealth apps, an increase in the use of pregnancy apps has been observed [21,24,25]. The vast majority of pregnant women download on average 3 apps during pregnancy [24], and nearly one-quarter use these apps almost daily [22]. However, health care professionals are concerned about the quality, validity, and accuracy of information freely available through mobile apps [22,24-30] and the reliability and safety of these apps [22,24,25,27,28]. Further, the women may not be able to determine the accuracy of the information, as well as the accessibility and readability of online resources [29]. Research has showed a low level of concern about the validity

of the information in the pregnancy apps, since 74% of users did not check the sources of information [22]. Incorrect and contradictory information can introduce unnecessary confusion, worries, anxiety, and uncertainty among pregnant women [29,31].

Considering the widespread use of the internet and smartphones as means to access health care information and as tools for health care management, it is interesting to know why and how pregnant women use online tools and what kind of information and features they are looking for [21]. However, the actual usage patterns and characteristics of women using pregnancy apps are relatively unknown [30]. In addition, information on information sources and the experience of women in Flanders (the northern Dutch-speaking region of Belgium) using digital tools during pregnancy is lacking. Therefore, the objective of this study was to map the current information resources of pregnancy apps and the use of pregnancy apps among pregnant women in Flanders.

## Methods

### Study Design and Population

A cross-sectional study was conducted, using a semistructured questionnaire that was distributed in Flanders from April 2019 to June 2019. This questionnaire was developed by researchers of Odisee University of Applied Sciences (Sint-Niklaas) and the Limburg Clinical Research Center/Mobile Health Unit (University of Hasselt – ZOL), based on literature and pre-existing questionnaires. Four different domains were questioned: (1) demographics of the pregnant women; (2) use of devices; (3) sources of information; and (4) use of pregnancy apps. A convenience sampling method was used to collect the data. Pregnant women were recruited through two different methods. The first method was the use of flyers in the waiting room of prenatal consultations. If pregnant women were interested in participating in this study, they received a paper questionnaire that they could fill in and return to the midwife at the prenatal consultation. The other method used was an online call for participation on the social media accounts of the participating hospital and universities. This online flyer contained a web-based link to an online survey. The same questions were asked in the online and paper questionnaires.

Pregnant women were recruited by researchers of the Limburg Clinical Research Center/Mobile Health Unit (group 1) and Odisee University of Applied Sciences (group 2). Data from group 1 was received via the prenatal ward of the Ziekenhuis Oost-Limburg (ZOL, Genk, Belgium), a tertiary hospital in Limburg. The Limburg Clinical Research Center/Mobile Health

Unit is a part of the University of Hasselt. Data from group 2 were received via the prenatal ward of VITAZ (Sint-Niklaas, Belgium), a secondary hospital in East Flanders. The data collection at this hospital was performed by Odisee University.

### Data Exclusion

A total of 331 answers were received (group 1:  $n=268$ , 81%; group 2:  $n=63$ , 19%), of which 20 (6%) were not completely filled in (all from group 1). Therefore, 311 questionnaires (94%) were analyzed, of which 92.5% were from group 1 and 7.5% from group 2. Responses were compared online to verify and possibly exclude duplicates. No duplicates were retained.

### Statistical Analysis

Statistical analyses were performed with SPSS (version 22.0; IBM Corp). The statistical analyses were mainly focused on descriptive statistics (frequencies, percentages). Normality was tested by the Shapiro-Wilk test. Differences in continuous and categorical variables were tested using independent Student *t* tests and chi-square tests, respectively. Correlations were investigated between the characteristics of the pregnant women (marital status, educational level, occupation, gravidity) and their responses. All statistical analyses were done at nominal level  $P=.05$ .

### Ethical Considerations

The study was approved by the Medical Ethics Committees of the hospital (Ziekenhuis) Oost-Limburg (ZOL; Genk; no. 19/0026U, eudract/B-no. B371201939699) and Ghent University Hospital (EC 2018/0120, B-no. B670201835156). The survey was anonymous. An information letter was added to the survey to explain the context of the study. By completing the questionnaires, the participants automatically agreed to the terms of the study.

## Results

### Participant Demographics

In total, 311 questionnaires were completed and returned (group 1: 248/311, 79.7% vs group 2: 63/311, 20.3%). There were no significant differences in characteristics between the two groups, except for education level; group 2 had a significantly higher prevalence of participants with a high school and/or university education compared to group 1 (92.1% vs 72.2%;  $P=.004$ ). In both groups, the mean age of the women was 30 years. Most of the women were married or in a civil partnership with their partner, were employees, and multigravida. The details of the characteristics are presented in Table 1. The difference in educational level did not influence the results of this study, so the results of the total study population will be discussed (and will not be divided by group).

**Table 1.** Characteristics of the women.

	Group 1 (n=248)	Group 2 (n=63)	P value (2-tailed)	Total (group 1 + group 2), N=311
Age (years), mean (SD)	30.76 (4.08)	30.13 (4.06)	.27	30.63 (4.08)
<b>Marital status, n (%)</b>			.84	
Married	131 (52.8)	31 (49.2)		162 (52.1)
Living together	94 (37.9)	25 (39.7)		119 (38.3)
Single	23 (9.3)	7 (11.1)		30 (9.6)
<b>Educational level, n (%)</b>			.004	
Lower secondary school and/or higher secondary school	68 (27.4)	5 (7.9)		73 (23.5)
High school and/or university	179 (72.2)	58 (92.1)		237 (72.6)
<b>Occupation, n (%)</b>			.40	
Self-employed	17 (6.9)	7 (11.1)		24 (7.7)
Employee	185 (74.6)	49 (77.8)		234 (75.2)
Worker	21 (8.5)	3 (4.8)		24 (7.7)
Housewife	9 (3.6)	0 (0)		9 (2.9)
Unemployed	5 (2)	2 (3.2)		7 (2.3)
Student	5 (2)	0 (0)		5 (1.6)
Other	6 (2.4)	1 (1.6)		7 (2.3)
Primigravida, n (%)	82 (33.1)	28 (44.4)	.09	110 (35.4)

### Use of Electronic Devices

In this manuscript, a computer/laptop is defined as “an electronic device for storing and processing data, typically in binary form, according to instructions given to it in a variable program,”

while a mobile phone is defined as “a telephone with access to a cellular radio system so it can be used over a wide area, without a physical connection to a network.” A smartphone/iPhone is defined as “a mobile phone that performs many of the functions of a computer, typically having a

touchscreen interface, internet access, and an operating system capable of running downloaded apps,” and a tablet PC/iPad/iPod is defined as “a wireless touch screen personal computer (PC) that is smaller than a notebook but larger than a smartphone... modern tablets are built with wireless Internet or local area networks (LAN) and a variety of software applications, including business applications, Web browsers and games.”

The first domain of the questionnaires was about device use as well as the frequency of use (Table 2). The majority of devices were used daily by the women (computer/laptop: 127/311, 40.8%; mobile phone: 251/311, 80.7%; and smartphone/iPhone: 303/311, 97.4%).

**Table 2.** The use of devices and the frequency of use (n=311).

Device, n (%)	Every day	Several times per week	Several times per month	Once per month	Less than once per month	Never
Computer/laptop	127 (40.8)	83 (26.7)	39 (12.6)	21 (6.8)	29 (9.3)	6 (1.9)
Mobile phone	251 (80.7)	2 (0.6)	0 (0)	0 (0)	1 (0.3)	46 (14.8)
Smartphone/iPhone	303 (97.4)	3 (1)	1 (0.3)	0 (0)	0 (0)	1 (0.3)
Tablet PC/iPad/iPod	61 (19.6)	60 (19.3)	34 (10.9)	13 (4.2)	48 (15.4)	78 (25)

### Sources of Information

Of the 311 respondents, 267 women (85.9%) reported that they searched online for information about pregnancy. Table 3 gives an overview of the manner in which they gathered information about pregnancy. The obstetrician was their first source of information (268/311, 86.2%), followed by websites/the internet (267/311, 85.9%), and apps (233/311, 74.9%). The midwife

was in fifth place (284/311, 59.3%), after friends (194/311, 62.4%).

The information that they sought was mostly about the following themes: development of the baby (276/311, 88.5%), discomfort/complaints during pregnancy (251/311, 80.7%), health during pregnancy (248/311, 79.7%), administration and practical matters (233/311, 74.9%), and breastfeeding (176/311, 56.6%).

**Table 3.** Sources of information accessed by participants (n=311).

Source of information	Participants, n (%)
Obstetricians	268 (86.2)
Websites/the internet	267 (85.9)
Apps	233 (74.9)
Friends	194 (62.4)
Midwife	184 (59.2)
Mother (in law)	179 (57.4)
Media	172 (55.3)
Social media	155 (49.8)
General practitioner	142 (45.7)
Other pregnant women	138 (44.4)
Books	131 (42.1)
Infosessions for future parents	111 (35.7)
Partner	103 (33.1)
Child and family	100 (32.2)
Other family members	84 (27)
Sisters	76 (24.4)
No information searched	2 (0.6)

### Use of Pregnancy Apps

Of the 311 women, 55.3% downloaded a pregnancy app (172/311). The mean number of downloaded apps was 1.59 (SD 0.96), with a maximum number of 7 apps. The top 3 reasons for downloading a pregnancy app were (1) to have a calendar to follow the growth and development of the baby (104/149,

69.8%); (2) to receive push messages with information, advice, and tips about the pregnancy on a daily/weekly basis (103/149, 69.1%); and (3) to have checklists (for baby names, baby layettes, etc; 38/189, 25.5%). The top 5 ways that women were informed about the existence of the apps were (1) a search in app stores (74/172, 43.0%); (2) friends (13/172, 7.6%); (3) a



combination of a search in app stores and (social) media (9/172, 5.2%); (4) a combination of a search in app stores and friends (7/172, 4.1%); and (5) social media (7/172, 4.1%). Advice from a health care provider (midwife) was mentioned by only one respondent (1/172, 0.6%). Women downloaded the apps because they were provided by a reliable organization or institution (10/172, 5.8%) or because of the good ratings by other users (10/172, 5.8%). Only 1.7% (3/172) of the women used an app on the advice of their health care provider. The downloaded app was visited daily by 9.3% (16/172) of the women, weekly by 9.3% (16/172), and monthly by 2.9% (5/172). Few women (5/172, 2.9%) reported rare app use. Only 3.5% (6/172) of the women paid for an app (€-€5, US \$2.04-\$5.10). The most frequently downloaded app was the Pregnancy + app (76/172, 44.2%), which is a worldwide pregnancy app developed by Philips.

### Correlations With Maternal Factors

Differences between maternal characteristics (marital status, education, and gravidity) and use of devices, source of information, and use of pregnancy apps were investigated. Only significant results will be discussed below. A detailed overview of all results is provided in [Multimedia Appendices 1-3](#).

#### Marital Status

There is a significant association between the source of information and a woman's marital status: single women were significantly more likely to ask their mothers (73.3%) or other family members (43.3%) for information than were married pregnant women (mother [in law]: 51.3%;  $P=.02$ ).

#### Education

Pregnant women with a lower education level (lower secondary school and/or higher secondary school) were significantly more likely to use a PC or laptop in their occupation than pregnant women who had a higher education level (high school and/or university; 98.6% vs 85.5%;  $P<.01$ ). They also were more likely to consult other family members for information about their pregnancy (41.1% vs 23.1%;  $P<.01$ ) and less likely to consult their gynecologist (95.9% vs 83.5%;  $P=.001$ ). In addition, they followed more prenatal sessions (80.8% vs 32.5%;  $P=.04$ ) and searched more for specific information about their discomfort and complaints during pregnancy (89% vs 79.5%;  $P=.02$ ).

#### Gravidity

Compared to multigravida, primigravida were more likely to ask for advice about their pregnancy from those in their vicinity, such as family members (primigravida: 40.4% vs multigravida: 20.1%;  $P<.001$ ) and other pregnant women (primigravida: 53.2% vs multigravida 40.2%;  $P=.03$ ). Primigravida also searched for more specific information about their pregnancy compared to multigravida. The specific significant results are general health (primigravida: 64.2% vs multigravida: 41.2%;  $P<.001$ ), health during pregnancy (primigravida: 89% vs multigravida: 75.9%;  $P=.01$ ), sexuality (primigravida 37.6% vs multigravida 10.6%;  $P<.001$ ), emotions, experiences and mood (primigravida: 37.6% vs multigravida: 25.1%;  $P=.02$ ), and administrative and practical matters (primigravida: 82.6% vs multigravida: 71.9%;  $P=.04$ ).

## Discussion

### Principal Results and Comparison With Prior Work

There is relatively little known about the sources of information in pregnancy apps and actual use of these apps by pregnant women, including their characteristics. To our knowledge, this is the first study that sought to map the current information resources and use of websites and mobile apps of Flemish women and their needs and expectations regarding digital information.

The use of websites and mHealth is becoming an increasingly important way for women to receive information about their pregnancy [21,24,25]. Our study showed that most devices (computer/laptop, mobile phone, and smartphone/iPhone) were used daily by women. More than half of the women downloaded a pregnancy app (55.3%), with an average of 1 or 2 apps per pregnancy. This is in line with a study by Lee et al [24], who reported that the average number of free apps downloaded was 2.4 (SD 1.57) [24].

The 3 most common reasons that women in our study downloaded an app were (1) follow-up of the growth and development of the baby (69.8%), (2) daily/weekly notifications based on push messages with information, tips, and advice about the pregnancy (69.1%), and (3) use of checklists (25.5%). A possible reason for this is that this kind of information is easy to explain and very accessible for everyone. In addition, women's care providers (eg, obstetrician, midwife) are not available at every moment of the day, but an app is. In studies by Wang et al [23] and Lupton and Pederson [22], the most reported reason for app usage was for monitoring fetal development (81.5% and 86%, respectively). In addition, the need for tailored advice and tracking of pregnancy changes, enabled by notifications, is another important reason for app usage. Further research is needed to gain insight into reasons for app usage, how app functions meet the expectations of women during pregnancy, and to what extent apps can be complementary to and integrated into current prenatal care.

The majority of women in our study (85.9%) searched for information on the internet, particularly information regarding fetal development, health and complaints during pregnancy, and practical and financial issues. This is in line with other studies showing that 65% [20] and 97% [32] of women sought digital health information. We also found that the obstetrician was the main source of information for many women (86.2%), followed by websites/internet (85.9%) and apps (74.9%). The midwife was in fifth place (59.3%). Our results seem to support those of previous studies [23,29,31,33,34], in which it was found that digital resources are a major source of information during pregnancy and childbirth, in combination with information and support from family, friends, and health professionals. We assume that the underlying reason for this is that online searches of apps are available at any time and place, and they are also accessible for questions that pregnant women might not want to ask their care providers about (eg, sensitive questions about finances).

Health care professionals are concerned about the quality, validity, and accuracy of digital information resources [22,24,25,27,28]. Previous research showed that women did not discuss information they found online with their midwife or doctor and between 8%-12% of women were unsure, worried, or confused about this information [29]. Findings were similar in a study by Wang et al [23] where women expressed the need for information about future apps from health care providers as well as the need to discuss contradictory information with health care providers. However, in our study, only 1.7% (3/172) of the women used an app on the advice of their health care provider, which is a lower percentage compared to a study by Mackintosh et al [29], where 30% of the women used websites or apps recommended by their midwife or doctor [29]. It seems that the recommendation of apps is not yet naturalized in the Flemish field of obstetrics, but this also indicates how important it is that health care providers refer women to high-quality digital resources such as websites and apps and take the opportunity to discuss digital information during consultations.

Similar to Buchanan et al (2021) [30] and Vogels-Broeke et al (2022) [34], our study also highlighted the importance of interpersonal resources such as peers, friends, and family for health- and parenting-related information. However, dependency on information from the internet and relatives can be problematic, particularly when this advice conflicts with recommendations from health professionals. Therefore, it may be recommended to disseminate digital health resources such as mobile apps to the social networks and family members of pregnant women to increase the likelihood of positive health outcomes [30].

Further, our study showed differences in sources of information depending on marital status, educational level, and gravidity. Single women were significantly more likely to ask their relatives for information (ie, their mother or other family members) than were married or cohabiting pregnant women, probably due to the fact that they do not have a partner to turn to with their questions and concerns. Pregnant women with a lower education level were significantly more likely to have a PC or laptop than those with higher education; they also were more likely to consult other family members for information about their pregnancy, followed more prenatal sessions, and searched more for online information on discomfort and complaints during pregnancy. This is in contrast with the results of Buchanan et al (2021) [30], who found that women from a lower socioeconomic background had lower rates of pregnancy app uptake, and were less likely to use written or online resources and digital technologies to search for health information [30]. A study by Vogel-Broeke et al [34] found that the use of websites was lower among women who had a low level of education versus those with a middle or high level of education. However, education level appeared to play little part in the online practices of women in the study by Mackintosh et al [29].

Finally, our study showed that primigravida asked for more advice about their pregnancy from those in their families and social circles compared with multigravida. Primigravida consulted their family members or other pregnant women and also searched for more specific information regarding general

health information, health during pregnancy, sexuality, emotions and mental well-being, and administrative and practical issues. This could be because multigravida have already experienced a pregnancy, and learned a lot from that experience. The need for advice and information is lower compared to primipara. These findings are consistent with other literature indicating that women have the greatest need for information during their first pregnancy and are more likely to use health apps as a source of information [30,31,34]. This emphasizes the importance of customized information, adjusted to the needs of pregnant women. Maternal characteristics need to be considered when developing mobile apps to ensure uptake among pregnant women from broader sociodemographic backgrounds [30].

### Strengths and Limitations

To our knowledge, this is the first study that sought to map the current information resources and use of websites and mobile apps of Flemish women and their needs and expectations regarding digital information. Another strength of this study is that the survey was developed based on existing questionnaires and prior, similar investigations among postnatal women. However, the psychometric properties of the questionnaire were not determined.

This study has several limitations that need consideration. Given the small sample size and the fact that pregnant women were recruited in only two Flemish regions (Genk and Sint-Niklaas), findings may not be generalizable to all Belgian women. Second, there were two versions of the questionnaire: a web-based questionnaire and a paper-based questionnaire. We mainly recruited women through social media, which means that we probably reached women who are more digitally skilled and therefore more familiar with digital technologies. It is known that digital (health) literacy is limited among vulnerable people. A recent report on digital inclusion found that 32% of Belgians have weak digital skills. This value increases up to 75% for people with a low income and low educational level [35]. It is likely that vulnerable pregnant women are underrepresented in our study. We tried to overcome this issue by using paper-based questionnaires, but we are aware of the fact that this could lead to bias in the answers (eg, limited time on the prenatal ward, less privacy). Further, we measured the use of pregnancy apps at one time point, regardless of gestational age. It has been demonstrated that sources of information vary over the course of pregnancy and app use declines as pregnancy progresses [23]. Vogels-Broeke et al [34] also found differences in the use of (digital) information resources between early and late pregnancy. Therefore, a longitudinal approach to studying the use of digital resources among pregnant women is recommended. Finally, we only investigated the needs, expectations, and app usage of pregnant women; it would be interesting to investigate the attitudes and experiences of partners as well as health care professionals.

### Recommendations for Further Research

Health care professionals must be aware that women search for pregnancy information online. Further research is needed to establish how health care professionals can support women's digital use during pregnancy and how digital information resources can be integrated into daily practice. Studies on health

care professionals' attitudes, needs, feasibility, and acceptability toward digital resources are recommended, as well as studies investigating how they are engaging and dealing with pregnancy and parenting apps in clinical practice. In addition, it would be interesting to investigate the experiences and actual app usage of partners. Further insights are needed on how pregnant women select apps, use them through the different trimesters, and evaluate their quality and usefulness.

### Conclusions

Health care professionals need to be aware that mobile health apps and the internet are important and growing sources of information for pregnant women—as shown in our study, they

are the second and third most common sources of information. It is likely that this digital trend will continue in the future and will become even more important. Concerns arise about the quality and safety of such apps, as only a limited number of apps are subjected to an external quality check. Therefore, it is important that health care professionals refer patients to high-quality digital resources and take the opportunity to discuss digital information with pregnant women. The availability of high-quality, evidence-based, and customized mobile pregnancy apps represents an important opportunity to optimize maternal and birth outcomes. Efforts should be made by health care professionals, app developers, and policy makers to ensure the quality of health apps and their integration into maternal care.

### Conflicts of Interest

None declared.

#### Multimedia Appendix 1

Correlation between marital status (married vs single) and use of technology, information resources and topics, and use of pregnancy apps.

[\[DOCX File, 15 KB - formative\\_v6i10e37866\\_app1.docx\]](#)

#### Multimedia Appendix 2

Correlation between educational level (secondary school vs higher education) and use of technology, information resources and topics, and use of pregnancy apps.

[\[DOCX File, 15 KB - formative\\_v6i10e37866\\_app2.docx\]](#)

#### Multimedia Appendix 3

Correlation between gravidity (primigravida vs multigravida) and use of technology, information resources and topics, and use of pregnancy apps.

[\[DOCX File, 15 KB - formative\\_v6i10e37866\\_app3.docx\]](#)

### References

1. Deave T, Johnson D, Ingram J. Transition to parenthood: the needs of parents in pregnancy and early parenthood. *BMC Pregnancy Childbirth* 2008 Jul 29;8:30 [[FREE Full text](#)] [doi: [10.1186/1471-2393-8-30](https://doi.org/10.1186/1471-2393-8-30)] [Medline: [18664251](https://pubmed.ncbi.nlm.nih.gov/18664251/)]
2. Grylka-Baeschlin S, van Teijlingen E, Gross MM. Cultural differences in postnatal quality of life among German-speaking women - a prospective survey in two countries. *BMC Pregnancy Childbirth* 2014 Aug 15;14:277 [[FREE Full text](#)] [doi: [10.1186/1471-2393-14-277](https://doi.org/10.1186/1471-2393-14-277)] [Medline: [25128290](https://pubmed.ncbi.nlm.nih.gov/25128290/)]
3. Rance S, McCourt C, Rayment J, Mackintosh N, Carter W, Watson K, et al. Women's safety alerts in maternity care: is speaking up enough? *BMJ Qual Saf* 2013 Apr;22(4):348-355. [doi: [10.1136/bmjqs-2012-001295](https://doi.org/10.1136/bmjqs-2012-001295)] [Medline: [23417732](https://pubmed.ncbi.nlm.nih.gov/23417732/)]
4. Akca A, Corbacioglu Esmer A, Ozyurek ES, Aydin A, Korkmaz N, Gorgen H, et al. The influence of the systematic birth preparation program on childbirth satisfaction. *Arch Gynecol Obstet* 2017 May;295(5):1127-1133. [doi: [10.1007/s00404-017-4345-5](https://doi.org/10.1007/s00404-017-4345-5)] [Medline: [28303340](https://pubmed.ncbi.nlm.nih.gov/28303340/)]
5. Hollins Martin CJ, Robb Y. Women's views about the importance of education in preparation for childbirth. *Nurse Educ Pract* 2013 Nov;13(6):512-518. [doi: [10.1016/j.nepr.2013.02.013](https://doi.org/10.1016/j.nepr.2013.02.013)] [Medline: [23541361](https://pubmed.ncbi.nlm.nih.gov/23541361/)]
6. Kamali S, Ahmadian L, Khajouei R, Bahaadinbeigy K. Health information needs of pregnant women: information sources, motives and barriers. *Health Info Libr J* 2018 Mar;35(1):24-37 [[FREE Full text](#)] [doi: [10.1111/hir.12200](https://doi.org/10.1111/hir.12200)] [Medline: [29131537](https://pubmed.ncbi.nlm.nih.gov/29131537/)]
7. Vamos CA, Merrell L, Detman L, Louis J, Daley E. Exploring women's experiences in accessing, understanding, appraising, and applying health information during pregnancy. *J Midwifery Womens Health* 2019 Jul;64(4):472-480. [doi: [10.1111/jmwh.12965](https://doi.org/10.1111/jmwh.12965)] [Medline: [31050386](https://pubmed.ncbi.nlm.nih.gov/31050386/)]
8. Wright E, Matthai M, Meyer E. The influence of social media on intrapartum decision making. *J Perinat Neonat Nurs* 2019;10. [doi: [10.1097/jpn.0000000000000377](https://doi.org/10.1097/jpn.0000000000000377)]
9. Risica P, Phipps M. Educational preferences in a prenatal clinic. *Int J Childbirth Education* 2006;21(4):4-7.

10. Criss S, Woo Baidal JA, Goldman RE, Perkins M, Cunningham C, Taveras EM. The role of health information sources in decision-making among Hispanic mothers during their children's first 1000 days of life. *Matern Child Health J* 2015 Nov;19(11):2536-2543 [FREE Full text] [doi: [10.1007/s10995-015-1774-2](https://doi.org/10.1007/s10995-015-1774-2)] [Medline: [26122256](https://pubmed.ncbi.nlm.nih.gov/26122256/)]
11. World Health Organization. From innovation to implementation: eHealth in the WHO European Region. Copenhagen: World Health Organization; 2019. URL: [https://www.euro.who.int/\\_data/assets/pdf\\_file/0012/302331/From-Innovation-to-Implementation-eHealth-Report-EU.pdf](https://www.euro.who.int/_data/assets/pdf_file/0012/302331/From-Innovation-to-Implementation-eHealth-Report-EU.pdf) [accessed 2022-08-17]
12. mHealth als sleutel tot kwaliteit en betaalbaarheid van zorg. VOKA Health Community Mobile Health. 2019. URL: [http://www.healthcommunity.be/sites/default/files/u195/Whitepaper%20Health%20Community%20mHealth\\_lowres.pdf](http://www.healthcommunity.be/sites/default/files/u195/Whitepaper%20Health%20Community%20mHealth_lowres.pdf) [accessed 2022-08-17]
13. European Commission. Report of the Working Group on mHealth assessment guidelines. 2017. URL: <https://digital-strategy.ec.europa.eu/en/library/report-working-group-mhealth-assessment-guideline> [accessed 2022-08-17]
14. mHealth: New Horizons for Health through Mobile Technologies: Based on the Findings of the Second Global Survey on eHealth (Global Observatory for eHealth Series, Volume 3). World Health Organization. 2011. URL: <https://apps.who.int/iris/handle/10665/44607> [accessed 2022-08-17]
15. Ali EE, Chew L, Yap KY. Evolution and current status of mhealth research: a systematic review. *BMJ Innov* 2016 Jan 05;2(1):33-40. [doi: [10.1136/bmjinnov-2015-000096](https://doi.org/10.1136/bmjinnov-2015-000096)]
16. Gagnon M, Ngangue P, Payne-Gagnon J, Desmartis M. m-Health adoption by healthcare professionals: a systematic review. *J Am Med Inform Assoc* 2016 Jan;23(1):212-220 [FREE Full text] [doi: [10.1093/jamia/ocv052](https://doi.org/10.1093/jamia/ocv052)] [Medline: [26078410](https://pubmed.ncbi.nlm.nih.gov/26078410/)]
17. Tripp N, Hainey K, Liu A, Poulton A, Peek M, Kim J, et al. An emerging model of maternity care: smartphone, midwife, doctor? *Women Birth* 2014 Mar;27(1):64-67. [doi: [10.1016/j.wombi.2013.11.001](https://doi.org/10.1016/j.wombi.2013.11.001)] [Medline: [24295598](https://pubmed.ncbi.nlm.nih.gov/24295598/)]
18. Millennials are top smartphone users. The Nielsen Company. 2016. URL: <https://www.nielsen.com/us/en/insights/article/2016/millennials-are-top-smartphone-users/> [accessed 2022-08-19]
19. Brown HM, Bucher T, Collins CE, Rollo ME. A review of pregnancy iPhone apps assessing their quality, inclusion of behaviour change techniques, and nutrition information. *Matern Child Nutr* 2019 Jul;15(3):e12768 [FREE Full text] [doi: [10.1111/mcn.12768](https://doi.org/10.1111/mcn.12768)] [Medline: [30569549](https://pubmed.ncbi.nlm.nih.gov/30569549/)]
20. Overdijkink SB, Velu AV, Rosman AN, van Beukering MD, Kok M, Steegers-Theunissen RP. The usability and effectiveness of mobile health technology-based lifestyle and medical intervention apps supporting health care during pregnancy: systematic review. *JMIR mHealth uHealth* 2018 Apr 24;6(4):e109 [FREE Full text] [doi: [10.2196/mhealth.8834](https://doi.org/10.2196/mhealth.8834)] [Medline: [29691216](https://pubmed.ncbi.nlm.nih.gov/29691216/)]
21. Frid G, Bogaert K, Chen KT. Mobile health apps for pregnant women: systematic search, evaluation, and analysis of features. *J Med Internet Res* 2021 Oct 18;23(10):e25667 [FREE Full text] [doi: [10.2196/25667](https://doi.org/10.2196/25667)] [Medline: [34524100](https://pubmed.ncbi.nlm.nih.gov/34524100/)]
22. Lupton D, Pedersen S. An Australian survey of women's use of pregnancy and parenting apps. *Women Birth* 2016 Aug;29(4):368-375. [doi: [10.1016/j.wombi.2016.01.008](https://doi.org/10.1016/j.wombi.2016.01.008)] [Medline: [26874938](https://pubmed.ncbi.nlm.nih.gov/26874938/)]
23. Wang N, Deng Z, Wen LM, Ding Y, He G. Understanding the use of smartphone apps for health information among pregnant Chinese women: mixed methods study. *JMIR mHealth uHealth* 2019 Jun 18;7(6):e12631 [FREE Full text] [doi: [10.2196/12631](https://doi.org/10.2196/12631)] [Medline: [31215516](https://pubmed.ncbi.nlm.nih.gov/31215516/)]
24. Lee Y, Moon M. Utilization and content evaluation of mobile applications for pregnancy, birth, and child care. *Healthc Inform Res* 2016 Apr;22(2):73-80 [FREE Full text] [doi: [10.4258/hir.2016.22.2.73](https://doi.org/10.4258/hir.2016.22.2.73)] [Medline: [27200216](https://pubmed.ncbi.nlm.nih.gov/27200216/)]
25. Lupton D. The use and value of digital media for information about pregnancy and early motherhood: a focus group study. *BMC Pregnancy Childbirth* 2016 Jul 19;16(1):171 [FREE Full text] [doi: [10.1186/s12884-016-0971-3](https://doi.org/10.1186/s12884-016-0971-3)] [Medline: [27435182](https://pubmed.ncbi.nlm.nih.gov/27435182/)]
26. Frazer C, Hussey L, Bosch E, Squire M. Pregnancy apps: A closer look at the implications for childbirth educators. *International Journal of Childbirth Education* 2015;30(3):1-12.
27. Steen M, Kingdon C. Maternity assist: new midwifery-led digital information and communication support system to enhance maternity care. *Pract Midwife* 2014;17(7):7. [Medline: [25109067](https://pubmed.ncbi.nlm.nih.gov/25109067/)]
28. Stoyanov SR, Hides L, Kavanagh DJ, Wilson H. Development and validation of the user version of the Mobile Application Rating Scale (uMARS). *JMIR mHealth uHealth* 2016 Jun 10;4(2):e72 [FREE Full text] [doi: [10.2196/mhealth.5849](https://doi.org/10.2196/mhealth.5849)] [Medline: [27287964](https://pubmed.ncbi.nlm.nih.gov/27287964/)]
29. Mackintosh N, Agarwal S, Adcock K, Armstrong N, Briley A, Patterson M, et al. Online resources and apps to aid self-diagnosis and help seeking in the perinatal period: A descriptive survey of women's experiences. *Midwifery* 2020 Nov;90:102803 [FREE Full text] [doi: [10.1016/j.midw.2020.102803](https://doi.org/10.1016/j.midw.2020.102803)] [Medline: [32717660](https://pubmed.ncbi.nlm.nih.gov/32717660/)]
30. Buchanan L, Anderson E, MBiostat HX, Phongsavan P, Rissel C, Wen LM. Sources of information and the use of mobile applications for health and parenting information during pregnancy: Implications for health promotion. *Health Informatics J* 2021;27(3):14604582211043146 [FREE Full text] [doi: [10.1177/14604582211043146](https://doi.org/10.1177/14604582211043146)] [Medline: [34569327](https://pubmed.ncbi.nlm.nih.gov/34569327/)]
31. Serçekuş P, Değirmenciler B, Özkan S. Internet use by pregnant women seeking childbirth information. *J Gynecol Obstet Hum Reprod* 2021 Oct;50(8):102144. [doi: [10.1016/j.jogoh.2021.102144](https://doi.org/10.1016/j.jogoh.2021.102144)] [Medline: [33848646](https://pubmed.ncbi.nlm.nih.gov/33848646/)]
32. Lagan B, Sinclair M, Kernohan G. Internet use in pregnancy informs women's decision making: a web-based survey. *Birth* 2010 Jun;37(2):106-115. [doi: [10.1111/j.1523-536X.2010.00390.x](https://doi.org/10.1111/j.1523-536X.2010.00390.x)] [Medline: [20557533](https://pubmed.ncbi.nlm.nih.gov/20557533/)]

33. Goetz M, Müller M, Matthies LM, Hansen J, Doster A, Szabo A, et al. Perceptions of patient engagement applications during pregnancy: a qualitative assessment of the patient's perspective. *JMIR mHealth uHealth* 2017 May 26;5(5):e73 [[FREE Full text](#)] [doi: [10.2196/mhealth.7040](https://doi.org/10.2196/mhealth.7040)] [Medline: [28550005](https://pubmed.ncbi.nlm.nih.gov/28550005/)]
34. Vogels-Broeke M, Daemers D, Budé L, de Vries R, Nieuwenhuijze M. Sources of information used by women during pregnancy and the perceived quality. *BMC Pregnancy Childbirth* 2022 Feb 08;22(1):109 [[FREE Full text](#)] [doi: [10.1186/s12884-022-04422-7](https://doi.org/10.1186/s12884-022-04422-7)] [Medline: [35135487](https://pubmed.ncbi.nlm.nih.gov/35135487/)]
35. Brotcorne P, Faure L, Vendramin P. Digitale inclusie. Essentiële diensten online: toegankelijk voor iedereen. Koning Boudewijn Stichting. URL: <https://www.kbs-frb.be/nl/digitale-inclusie-essentiele-diensten-online-toegankelijk-voor-iedereen> [accessed 2022-08-17]

*Edited by A Mavragani; submitted 09.03.22; peer-reviewed by S Choemprayong, J van den Heuvel; comments to author 15.05.22; revised version received 01.06.22; accepted 29.06.22; published 11.10.22.*

*Please cite as:*

*Lanssens D, Thijs IM, Dreesen P, Van Hecke A, Coorevits P, Gaethofs G, Derycke J, Tency I  
Information Resources Among Flemish Pregnant Women: Cross-sectional Study  
JMIR Form Res 2022;6(10):e37866  
URL: <https://formative.jmir.org/2022/10/e37866>  
doi: [10.2196/37866](https://doi.org/10.2196/37866)  
PMID: [36222794](https://pubmed.ncbi.nlm.nih.gov/36222794/)*

©Dorien Lanssens, Inge M Thijs, Pauline Dreesen, Ann Van Hecke, Pascal Coorevits, Gitte Gaethofs, Joyce Derycke, Inge Tency. Originally published in *JMIR Formative Research* (<https://formative.jmir.org>), 11.10.2022. 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

# RWD-Cockpit: Application for Quality Assessment of Real-world Data

Lmar Marie Babrak<sup>1\*</sup>, PhD; Erand Smakaj<sup>1\*</sup>, MSc; Teyfik Agac<sup>1</sup>, BSc; Petra Maria Asprion<sup>2</sup>, PhD; Frank Grimberg<sup>2</sup>, PhD; Daan Van der Werf<sup>3</sup>; Erwin Willem van Ginkel<sup>3</sup>, MSc; Deniz David Tosoni<sup>1</sup>, BSc; Ieuan Clay<sup>4</sup>, PhD; Markus Degen<sup>1</sup>, MSc; Dominique Brodbeck<sup>1</sup>, PhD; Eriberto Noel Natali<sup>1</sup>, PhD; Erik Schkommodau<sup>1</sup>, PhD; Enkelejda Miho<sup>1,5,6\*</sup>, PhD

<sup>1</sup>University of Applied Sciences and Arts Northwestern Switzerland, School of Life Sciences, Muttenz, Switzerland

<sup>2</sup>Fachhochschule Nordwestschweiz University of Applied Sciences and Arts Northwestern Switzerland, School of Business, Olten, Switzerland

<sup>3</sup>Q-Strip, Steenbergen, Netherlands

<sup>4</sup>Evidation Health Inc, San Mateo, CA, United States

<sup>5</sup>aiNET GmbH, Basel, Switzerland

<sup>6</sup>Swiss Institute of Bioinformatics, Lausanne, Switzerland

\*these authors contributed equally

**Corresponding Author:**

Enkelejda Miho, PhD

University of Applied Sciences and Arts Northwestern Switzerland, School of Life Sciences

Hofackerstrasse 30

Muttenz, 4132

Switzerland

Phone: 41 0612285847

Email: [enkelejda.miho@fnw.ch](mailto:enkelejda.miho@fnw.ch)

## Abstract

**Background:** Digital technologies are transforming the health care system. A large part of information is generated as real-world data (RWD). Data from electronic health records and digital biomarkers have the potential to reveal associations between the benefits and adverse events of medicines, establish new patient-stratification principles, expose unknown disease correlations, and inform on preventive measures. The impact for health care payers and providers, the biopharmaceutical industry, and governments is massive in terms of health outcomes, quality of care, and cost. However, a framework to assess the preliminary quality of RWD is missing, thus hindering the conduct of population-based observational studies to support regulatory decision-making and real-world evidence.

**Objective:** To address the need to qualify RWD, we aimed to build a web application as a tool to translate characterization of some quality parameters of RWD into a metric and propose a standard framework for evaluating the quality of the RWD.

**Methods:** The RWD-Cockpit systematically scores data sets based on proposed quality metrics and customizable variables chosen by the user. Sleep RWD generated de novo and publicly available data sets were used to validate the usability and applicability of the web application. The RWD quality score is based on the evaluation of 7 variables: *manageability* specifies access and publication status; *complexity* defines univariate, multivariate, and longitudinal data; *sample size* indicates the size of the sample or samples; *privacy and liability* stipulates privacy rules; *accessibility* specifies how the data set can be accessed and to what granularity; *periodicity* specifies how often the data set is updated; and *standardization* specifies whether the data set adheres to any specific technical or metadata standard. These variables are associated with several descriptors that define specific characteristics of the data set.

**Results:** To address the need to qualify RWD, we built the RWD-Cockpit web application, which proposes a framework and applies a common standard for a preliminary evaluation of RWD quality across data sets—molecular, phenotypical, and social—and proposes a standard that can be further personalized by the community retaining an internal standard. Applied to 2 different case studies—*de novo*—generated sleep data and publicly available data sets—the RWD-Cockpit could identify and provide researchers with variables that might increase quality.

**Conclusions:** The results from the application of the framework of RWD metrics implemented in the RWD-Cockpit application suggests that multiple data sets can be preliminarily evaluated in terms of quality using the proposed metrics. The output

scores—quality identifiers—provide a first quality assessment for the use of RWD. Although extensive challenges remain to be addressed to set RWD quality standards, our proposal can serve as an initial blueprint for community efforts in the characterization of RWD quality for regulated settings.

(*JMIR Form Res* 2022;6(10):e29920) doi:[10.2196/29920](https://doi.org/10.2196/29920)

## KEYWORDS

real-world data; real-world evidence; quality assessment; application; mobile phone

## Introduction

### Background

Real-world data (RWD) is defined as health care data generated outside of randomized controlled trials (RCTs) [1]. Real-world evidence (RWE) regarding the use, benefits, and risks of medications is obtained through comprehensive analyses and validation of RWD. Examples of RWD include electronic health records, prescription and billing data, insurance claims, genetic and molecular biobanks, medical-related products, disease registries, and patient-generated health data collected through a variety of sources and digital devices such as wearables and smartphones [2,3]. RWD emerged through the widespread use of health-related apps, implementation of electronic health records in hospitals, and routine genetic testing. Recently, these data were recognized as a valuable resource for biopharmaceutical companies to reduce research and development expenditures, and this has been primarily implemented by regulatory agencies in postmarket analysis of medical products [4].

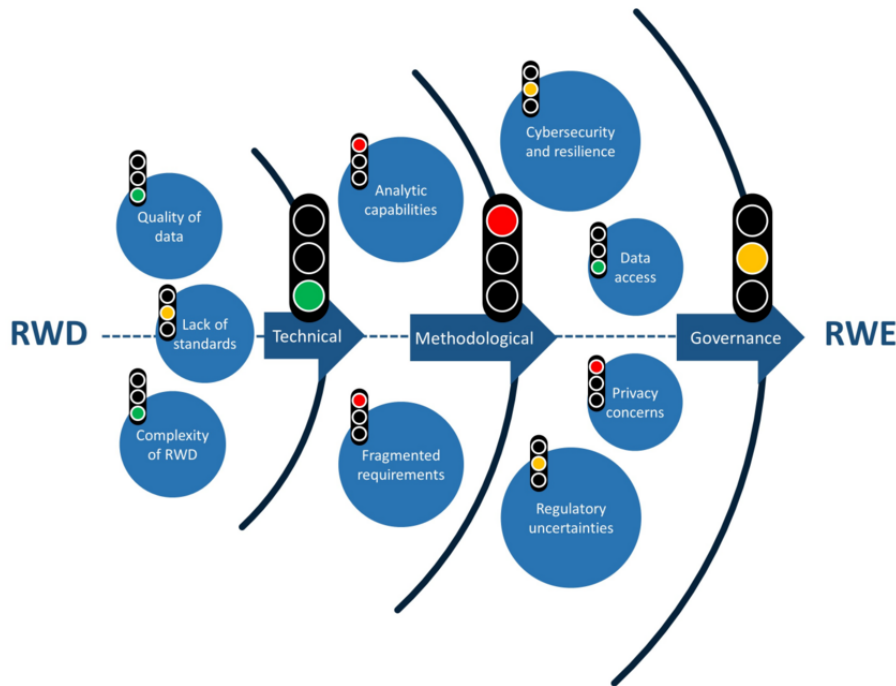
RWE can supplement, and has often served as, primary data to inform on regulatory decisions such as alternative drug indications and is used in orphan and oncological disease studies [5]. In response to this trend, regulatory agencies such as the US Food and Drug Administration (FDA) and European Medicines Agency (EMA) have implemented strategies for the inclusion of RWD and RWE as part of their regulatory approach to digitalization in health care to inform regulatory decisions such as late-term adverse effects or stratifying clinical trial population groups with the US 21st Century Cures Act [6] and the EMA Regulatory Science to 2025 strategy [7]. Several studies have shown that the use of RWD in determining patient health status, especially in cases of progressive or chronic diseases such as Alzheimer disease and Parkinson disease, can greatly affect current diagnosis and prognosis as well as optimize disease management [3,8]. The use of RWE is also crucial for

assessing the safety and effectiveness of processes that cannot be appropriately addressed in an RCT, such as surgical procedures [9]. Ongoing efforts by the regulatory agencies have already seen practical implementations of RWD used to receive regulatory approval as an alternative to RCTs. For example, Prograf (tacrolimus), a drug initially approved to prevent organ rejection in liver transplantations, has received FDA approval for use in kidney and heart transplantations [10] and similar approvals in Europe [11]. These cases and others reflect how well-designed studies relying on fit-for-purpose RWD can be considered adequate under FDA and EMA regulations [12]. To maximize the implementation of RWE, an important challenge currently is to find data that provide the most suitable measurements for biopharmaceutical companies and regulatory agencies [13].

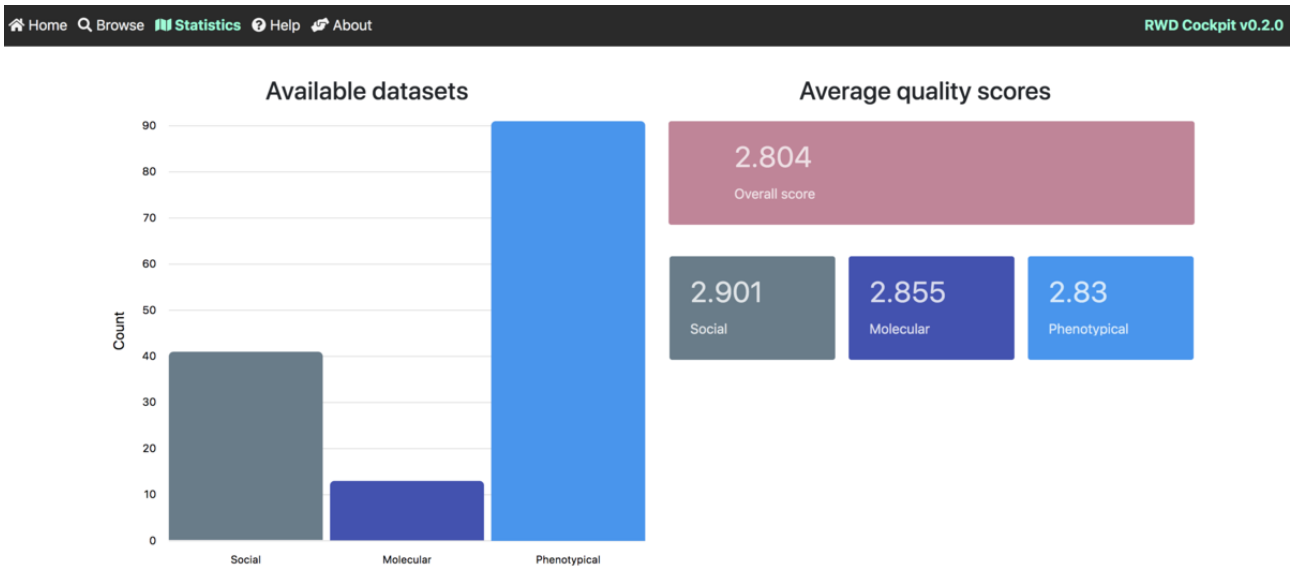
### Objectives

The sources and types of RWD are diverse, ranging from medication orders to patient-generated (eg, PatientsLikeMe and Carenity), digitally collected (fitness trackers), and social media data [14]. However, the criteria used by the biopharmaceutical industry to select appropriate data sets for different applications compared with traditional RCTs are unclear [12]. In addition, data origin, diversity, and complexity make it difficult to consistently rank and assess RWD quality [15]. Lack of standardization and structure among data sets augments and lengthens the process of identifying the right fit-for-purpose RWD and generating meaningful analyses [16]. Carefully curated, validated, standardized, and high-quality data are needed to generate widely accepted RWE that can bridge the knowledge gap between standardized RCTs and the real world. To date, there are neither clear standards nor available tools to assess RWD quality [17]. To address these challenges (Figure 1), we have created an easy-to-use, accessible web application tool that assesses RWD data sets using a customizable selection of proposed standard variables: the *RWD-Cockpit* (Figures 2 and 3).

**Figure 1.** Challenges in the transformation of real-world data (RWD) to real-world evidence (RWE).



**Figure 2.** Screenshot of the Statistics tab of the RWD-Cockpit web-application of the overall score of RWD data sets. A data set tree map allocated under the general statistics bar chart enables the sequential selection of the type of data, its complexity, assessment and other parameters for the identification of the RWD and the quality needed.





**Figure 3.** Search results with the keyword test in data sets filtered by the variable complexity and descriptor longitudinal. The first, highest-score data set entry is shown.

**Search results**

**Parameters:**  
 Keywords: test  
 Type: Molecular Phenotypical Social  
 Filters: Longitudinal

**Key:**

poor	score = 0 - 2.5
average	score = 2.5 - 3.5
good	score = 3.5 - 4.75
superior	score = 4.75 - 5

[Hide Query Parameters & Score Key](#)

**National Health and Nutrition Examination Survey 2013-2014** **3.67**  
 Centers for Disease Control and Prevention

The National Health and Nutrition Examination Survey (NHANES) is a program of studies designed to assess the health and nutritional status of adults and children in the United States. The survey is unique in that it combines interviews and physical examinations. NHANES is a major program of the National Center for Health Statistics (NCHS). NCHS is part of the Centers for Disease Control and Prevention (CDC) and has the responsibility for producing vital and health statistics for the Nation. The NHANES program began in the early 1960s and has been conducted as a series of surveys focusing on different population groups or health topics. In 1999, the survey

## Methods

### Overview

A total of 106 RWD data sets were selected as a target sample group for the development of a scoring method in the RWD-Cockpit to assess RWD data sets. The metadata of these data sets and publications were investigated with regard to data quality to devise a scoring method to assess the quality of RWD. The scoring method takes into account seven variable metrics for data-quality assessment:

1. Manageability specifies the access rights that users may or may not have for a data set as well as whether the data have been published in a peer-reviewed journal.
2. Complexity defines whether the data set is univariate, multivariate, or longitudinal.
3. Sample size defines the sample size of a given data set.
4. Privacy and liability stipulates privacy rules according to the data context of use.
5. Accessibility specifies how the data set can be accessed and to what granularity.

6. Periodicity specifies how often the data set is updated.
7. Standardization specifies whether the data set adheres to any specific technical or metadata standard.

For each variable, there are several associated *descriptors* that define specific characteristics of the data set. The descriptors are explained in detail in [Table 1](#). The RWD scoring formula subsequently averages the performance of variables for a given data set to assign a final score. A specific score (0 to 100) is assigned to each variable's descriptor. Each variable can have >1 descriptor (eg, the data set contains longitudinal and multivariate data), and an average score for each variable is taken. In the *Complexity* variable, multivariate or univariate can be chosen. Subsequently, a cumulative average is calculated for all variables. This cumulative average is normalized to a score from 1 to 5, with 5 being the best quality for a data set. This score is called the quality identifier and is displayed and associated with each data set. The normalization is performed by dividing the cumulative score by 7 (number of descriptors), then dividing by 100 and multiplying by 5. An example of the scoring methodology for 2 data sets is shown in [Table 2](#).

**Table 1.** Variable and descriptor definitions and real-world data examples.

Variable and descriptor (scored highest to lowest)	Definition	Examples
<b>Manageability: describes the level of data management, such as whether the data are protected; have been peer reviewed and published; and require paid access, registration, or are freely available to users</b>		
Protected and peer reviewed	Only selected users have access, or data set has been published in a peer-reviewed journal	SwissRDL [18], Cancer Registry of Norway [19], and EHRs <sup>a</sup>
Attributed	User must register to access the data, and the source is referenced	— <sup>b</sup>
Regulated	User must register to be able to access the data and has no further reference to the generation of the data (eg, scientific publication)	Kaggle [20], Google Dataset Search [21], European Union OpenData.swiss [22], and University of California Irvine Machine Learning Repository [23]
Free	Access is open source, and data are freely available	—
<b>Complexity: describes the extent of complexity within a data set (eg, whether the data set contains single, multiple, or longitudinal measurements)</b>		
Longitudinal	Univariate or multivariate, measured repeatedly over defined time intervals	Panel study of income dynamics [24]
Multivariate	Multiple columns or variables (table containing more information than univariate)	—
Univariate	Only 1 column or variable	Home blood pressure–monitoring pilot: NYU Langone Health EHR [25]
<b>Sample size: describes the number of samples in the data set</b>		
Single	The sample size is 1	Height weight single-variable data [26]
Small	The sample size is 2 to 100	—
Medium	The sample size is 101 to 1,000,000	National Health and Nutrition Examination Survey 2013-2014 [27]
Large	The sample size is >1,000,000	—
<b>Privacy and liability: describes how well the data set addresses privacy concerns, such as use of encryption, anonymization of participants, and other privacy factors</b>		
Encrypted	Data set has been processed through an encryption algorithm and can only be read by authorized parties with the encryption key, and privacy is assured and risk is minimized	Measuring the quality and completeness of medication-related information derived from hospital EHR database (derived data) [28]
Derived	Data have been analytically preprocessed; there are no privacy and anonymization issues, and liability is minimal; and derived data (not the heart rate of all patients but an average)	—
Anonymized	Data do not have any identifying particulars or details that would lead to participant identification and have minimal liability	—
Private	User has private rights or has an established collaboration to access the data (physician with patient data), and liability is reduced because of exclusive rights to the data, but user is responsible for data privacy and safety as determined by law	Twitter data (private) [29]
Open	Data have no protection measures implemented that protect user identity or privacy; thus, users are responsible for the integrity of the data because there is no information on how they were gathered or managed	—
<b>Accessibility: describes how the data set can be accessed, such as from a direct download from the web or as a hard copy document</b>		
API <sup>c</sup>	Data are accessed through API, and specific data sets are queried and requested by the user and acquired	OpenML (download and API) [30] and OpenData.swiss (API) [22]
Download	Data are downloadable from the web, but there is minimal functionality (ie, querying is minimal and usually limited in terms of the number of data sets available)	—
Soft copy	Data are digitally available	USB drive, CD, portable storage, and hard disk
Hard copy	Data are available only as paper documents	Paper documents
<b>Periodicity: describes whether the data set is a single snapshot (collected once) or it is designed to be collected and released continuously or periodically</b>		

Variable and descriptor (scored highest to lowest)	Definition	Examples
Sequential	Data are measured in a specified periodic or continuous manner	Pervasive computing technologies to continuously assess Alzheimer disease progression and intervention efficacy [31]
Ad hoc	Data are measured multiple times when necessary (eg, during sickness)	—
Repeated	Data are generated based on >1 measurement taken at random times	Population census data
Single	Data are generated based on 1 measurement	
<b>Standardization: describes whether the data set adheres to, for example, a common international standard or a specific organization</b>		
Open metadata	Data are organized according to official standards (eg, Health Level Seven and National Council for Prescription Drug Programs)	A validated smartphone-based assessment of gait and gait variability in Parkinson disease [32]
Self-metadata	Data are organized using specific descriptors from data providers that describe the structure of the data in detail	—
Structured	Data are organized in a streamlined and easily interpretable format, but this does not follow any international data set guidelines	Emails, word-processing documents, photographs, and presentations
None	Data have no clear organization or standardization	—

<sup>a</sup>EHR: electronic health record.

<sup>b</sup>Not available.

<sup>c</sup>API: application programming interface.

**Table 2.** Case study of quality-assessed de novo-generated real-world data<sup>a</sup>.

Variable	Descriptor	Descriptor score	Variable score
Manageability	Protected	100	100 / 1 = 100
Complexity	Longitudinal+ multivariate	100+50	(100 + 50) / 2 = 75
Sample size	Small	33	33 / 1 = 33
Privacy and liability	Anonymized	50	50 / 1 = 50
Accessibility	Application programming interface (API)	100	100 / 1 = 100
Periodicity	Sequential	100	100 / 1 = 100
Standardization	Self-metadata	66	66 / 1 = 66

<sup>a</sup>Cumulative score = 100 + 75 + 33 + 50 + 100 + 100 + 66 = 524; quality identifier = (524 / 7) / 100 × 5 = 3.74.

The RWD-Cockpit system is based on a database that manages the available data sets as well as the scoring of data sets. Data sets are assigned descriptors using database object relations. When querying data sets, related flags are automatically fetched for each data set; the data set score is calculated simultaneously and then displayed to users. This simultaneous scoring mechanism allows for assessments to change dynamically over time if there are changes to the variables or the scores. In addition, personal data sets and data sets not provided by the application can also be self-scored through the *Help* link in the application. A user can click on *Enable score calculator* and easily choose the appropriate descriptors for a data set, after which a score is generated automatically.

Furthermore, a statistics page was designed to provide an internet-based chart of the data sets within the application to track global trends in RWD quality assessments from several sources (eg, Kaggle). The global average quality identifier scores of all available data sets and for each type—molecular, phenotypical, and social—are displayed. In terms of technology,

the RWD-Cockpit system is developed as an Angular web application on top of a Loopback server, and the data are stored in a MongoDB database. This architecture is robust and straightforward to deploy and maintain. The RWD-Cockpit web application is free to use, and users are encouraged to input their data set descriptions and add their data sets in the database. For data sets associated with a publication, the publication reference is cited.

### De Novo-Generated RWD as a Case Study

The measurements were performed with a ribbon containing electrodes that is placed underneath a fitted sheet in a bed at chest height. The electrodes quantitatively measure humidity, which is correlated to sweat, during sleep ([Multimedia Appendix 1](#)). In addition, the sensor measures the temperature at the ribbon and in the room using a printed circuit board ([Multimedia Appendix 2](#) and [Multimedia Appendix 3](#)). Resistance and temperature measurements are taken throughout the night by 2 electrodes within the ribbon, and the values collected are

transferred to a moisture sensor connected to a printed circuit board ([Multimedia Appendix 4](#)).

## Results

### A Straightforward Web Tool That Assesses RWD Data Sets

Many challenges [33] and information parameters to guide RWD appropriateness for research questions have been proposed [34]. However, a framework that translates information (eg, quantifying parameters such as the importance of longitudinal data, accessibility, and publication as a metric that indicates quality [34]) into an evaluation framework is still missing. We created the RWD-Cockpit [35], a straightforward web tool that assesses RWD data sets using standard and customizable variables ([Figures 2 and 3](#)). This web application provides a platform to search for, and view, quality-scored RWD data sets. Furthermore, it provides a flexible benchmarking data quality–scoring tool for new user-acquired RWD data sets. Users can search for RWD data sets that fulfill user-selected quality variables and score criteria, as well as set a standard that other users within the same institution or across institutions or regulatory agencies may use. Currently, 106 quality-scored RWD data sets from a variety of sources and areas are available through the RWD-Cockpit, and more are continually added. As new guidelines and laws are passed concerning the use of RWD, novel variables and descriptors can be added as needed by the administrators.

In the RWD-Cockpit, RWD data sets were scored based on 7 variables that were identified to be important metrics in determining data quality to address the challenges in adopting RWD for regulatory decision-making [11,36]. The proposed variables are *manageability*, *complexity*, *sample size*, *privacy and liability*, *accessibility*, *periodicity*, and *standardization*, and they are described in detail in [Table 1](#). Each variable contains 3 to 5 *descriptors* that describe specific characteristics that apply to a data set, such as *multivariate* or *longitudinal* measurements ([Table 1](#)). Each variable has been identified based on its impact on the overall usability of data. Because of the broad landscape of potential use cases for RWD, the identified variables do not consider case-specific suitability or content but create a generalized framework to assess RWD. *Manageability* is an important variable necessary because of the broad diversity and almost nonexistent limitations on what data identify as RWD ([Table 1](#)). The level of data management [36] can be related to the general quality and trustworthiness of the data. A higher score is proposed for either peer-reviewed data sets or data that require additional efforts with regard to data management. The *complexity* of data extends the use-case coverage of the data sets. Univariate data might offer a base to solve single research questions but lack the depth of potential insights. A proposed option to achieve an increased score for *complexity* is to provide or generate diverse data, enabling the data to be integrated into a broader field of use cases. The variable *sample size* is of great importance because RWD are intended to show real-world insights. Real-world behavior can be reflected better in data from large numbers of individuals compared with information on a single individual. The *sample size* of a data set can be

increased at any time, given that the circumstances of the data acquisition, such as used devices, remain the same for each data point. The level of compliance with given data privacy regulations, as represented in the variable *privacy and liability*, can provide further insights on data quality and trustworthiness. Open RWD without data protection measures have a high potential of being simulated data, whereas reliable data sources are compelled to comply with given regulations. The application of data anonymization or encryption measures and compliance with European Union standards (General Data Protection Regulation) or US standards (Health Insurance Portability and Accountability Act) results in a higher assigned score. The state in which data are being transferred and stored takes on a relevant role when intending to extract RWE from RWD; thus, the variable *accessibility* has been identified as a relevant factor. Hard copy of data might provide a wide range of content but increases efforts in preprocessing and requires the transformation from analog to digital. The most efficient method to implement the accessibility variable is to provide the data through an application programming interface. Using an application programming interface allows users to specifically query the data of interest and, furthermore, provides direct digital access to the data. Similar to *complexity* and *sample size*, *periodicity* has been identified as a relevant parameter because of the depth of information. Single snapshots of individuals reflect acute states, whereas data acquired at different points in time of the same individual can generate deeper insights, which indicates reproducibility of the method. Collecting data from individuals repeatedly according to a defined time plan leads to a 2-fold benefit: first, the ability to create an average overview on individuals, and second, the ability to identify time-related patterns or progression. The last variable *standardization* is required to further increase trustworthiness and practicality. Applying state-of-the-art health care data standards to RWD generates a more direct path to use the data, whereas unstructured data or data that follow nonconventional standards require increased efforts with regard to their understanding and use. To achieve a higher score in *standardization*, it is proposed to identify potential community standards or frameworks and apply these to the data set.

The overall average score of all data sets in the RWD-Cockpit application was 2.80, with social data sets scoring 2.90, molecular data sets scoring 2.86, and phenotypical data sets scoring 2.83.

### Case Studies

To provide a practical example on the benefits of using the RWD-Cockpit, quality metrics were calculated in two case studies: (1) a practical case study using temperature data during sleep and (2) two of the publicly available data sets in the database.

To evaluate the applicability and usability of the RWD-Cockpit web application on de novo–generated data, a case study was performed to generate temperature RWD during sleep [37]. The RWD generators were asked to use the RWD-Cockpit application on their data sets, determine the value the application provided, and use it to find quality-scored data sets that were useful for their company. Their preliminary data sets scored

3.74 out of 5, which is determined to be *good* within the application (Table 2). On the basis of this score, the data generators identified variables that can be improved upon within their data set, such as increasing the sample size, publishing the data in a peer-reviewed journal, encrypting the data, and organizing the data according to officially recognized standards, such as Health Level Seven. These developments will increase the quality of their data set, making it more likely to be adopted in a health care environment as a novel digital biomarker for health and to be accepted as a health tool. Furthermore, the data generators identified within the RWD-Cockpit application several scored RWD data sets related to sleep and health that were meaningful and independently collected to further develop their product. This case study validated the purpose and value of the RWD-Cockpit web application for use.

We selected and scored 2 different health care-related data sets to demonstrate how their respective scores reflect the low and high quality of the RWD Table 2 and Table 3. The first data set, named *Height Weight Single Variable data*, studies the relationship between the height and weight of a person, predicting the probabilistic weight for a given height from a list of heights and weights [26]. The second RWD data set, *National Health and Nutrition Examination Survey 2013-2014*, consists of health measurements and surveys (eg, demographics and laboratory measurements) of approximately 5000 individuals across the country over a 2-year period conducted by officials at the US Centers for Disease Control and Prevention [27]. These data include measurements conducted by physicians and laboratories, as well as self-reported measurements. When quality identifiers are compared between the 2 data sets, the second data set scores higher at 3.67 versus the score of 1.47 of the first data set Table 3. The major differences between the 2 data sets are primarily visualized in 5 of the 7 variables: *manageability*, *complexity*, *sample size*, *periodicity*, and *standardization*. Whereas the first data set contains only 35 different samples, the second data set has collected data from 5000 different samples, resulting in a higher score. Regarding *complexity*, the first data set is *univariate*, whereas the second data set is both *longitudinal* and *multivariate*, thus scoring higher than the first data set. Additional differences in descriptors and scores are detailed in Table 1.

The score of the Height Weight Single Variable data set can be improved on various aspects. The data set can be extended with additional individuals to increase the *sample size* variable from *small* to *medium*, and further information related to these individuals can also be used as an extension of the data set.

Furthermore, a structured plan can be applied to measure the required information for the data set at specific time points, further increasing the *periodicity* score of the data set. The *periodicity* variable provides the ability to data providers to define the time ranges and number of independent measurements without constraints, allowing measurements on the same day to be graded *repeatedly* or even *ad hoc*. In addition, when repeated measurements of the same individuals are performed multiple times, the *complexity* score is affected as well, moving from *univariate* to *longitudinal*. On the basis of the additions to the data set, an overall data dictionary can be created to improve the *standardization* score from the current *structured* to *self-metadata*. The extended data set can further be used to conduct research on, granting the option to receive peer reviews on the data set or to move the *manageability* score from *free* to either *attributed* or *peer reviewed*. With regard to *privacy and liability* and *accessibility*, *anonymized* and *download* already provide high scores while covering the variables appropriately for this kind of data set; thus, further improvements are not necessarily required.

The described adjustments and additions to the data set can cause a significant impact on the overall score of the data set. The Height Weight Single Variable data set can increase its score from 1.47, considered a *poor* data set by the platform, to 3.60, reaching the classification *good*. Besides raising the score, the proposed adjustments strongly improve the usability and reliability of the data set. The data set can cover a much wider range of use cases and gain trust of stakeholders. A complete overview of potential improvements to the data set is shown in Table 3.

This simple scoring of RWD enables investigators and health care stakeholders to get a general overview of the suitability of the data sets in relation to the decisions the data will affect. It also allows users to apply the same standards for assessing RWD quality. For high-impact health-related decisions, RWD should be of high quality and scrutinized for validity. The ability to add new potential descriptors to any variable or add completely new variables makes the RWD-Cockpit dynamic and adaptable. In addition, new governmental regulations can be implemented easily by adding new variables. Scoring can be adjusted according to the regulations for each country or specific regulatory or industry-specific requirements. By using the RWD-Cockpit, a fast evaluation of RWD quality can be achieved and the data sets can be scored depending on the fulfilled requirements for further consideration.

**Table 3.** Potential improvements to a case study of publicly available real-world data<sup>a</sup>.

Variable	Value	Score	Improvement	Value after improvement	Score after improvement
Manageability	Free	0	Move from providing the data freely to attributed access or publish the data in a peer-reviewed journal	Attributed	66
Complexity	Univariate	0	Extend the data set by collecting more information on participants at different time points	Longitudinal	100
Sample size	Small	33	Increase the number of individuals included in the data set to >100 participants	Medium	66
Privacy and liability	Anonymized	75	Privacy and liability are already handled appropriately	Anonymized	75
Accessibility	Download	66	Access through download is appropriate for data such as these	Download	66
Periodicity	Single	0	Develop a plan for when the participants will be examined again and extend the data set in defined time steps	Ad hoc	66
Standardization	Structured	33	Create a data dictionary clearly stating the structure of the data set	Self-metadata	66

<sup>a</sup>Score before changes: 1.47; score after changes: 3.60.

## Discussion

### Setting Standards for RWD: Opportunities and Challenges

Digital technological advancements such as measurement of digital biomarkers [13], wide implementation and use of electronic health records, and social media have generated a wealth of health-related data that can potentially be leveraged to generate valuable RWE. Biopharmaceutical companies have already begun to harvest the plethora of data and to integrate the data for new drug applications and postmarket analysis of various therapeutics [4]. Integration of RWD is valuable and has the potential to reduce the huge health care expenditure costs without lowering the standards for evidence [38]. Regulatory authorities such as the FDA or the EMA face challenges when it comes to consideration of RWE generated from RWD in regulatory decision-making and drug approval [2,11]. Some of the challenges [33] the agencies must overcome are related to ensuring the quality of data and providing frameworks for consideration. Without guiding regulations, the currently expanding use of RWD in studies [26,39] fails to follow industry standards. In addition, a robust standard for data sets must be implemented following the Findability, Accessibility, Interoperability, and Reusability (FAIR) principles [40]. However, these principles are related to data upstream the RWD quality assessment, and there are no guidelines from federal agencies on how to standardize data sets or individual data points. Several attempts are being made by groups such as the Observational Medical Outcomes Partnership, the Sentinel System, and the National Patient-Centered Clinical Research Network to partially address standardization by using systems such as the common data models that standardize terminology and transform databases into a similar format and representation [41]. Another strategy is to evaluate apps that generate the RWD [42,43]; however, this would translate into a discrepancy in the

quality evaluation process that would reflect on, and differ in, the steps specific to the application areas of disease and usability. In contrast, information regarding RWD quality in the community is mainly reported in the form of characterization [34], which lacks the formality of a metric.

The complexity of RWD could be greatly reduced by providing a standard for the industry and other health care stakeholders. For RWD to be accepted into mainstay biopharmaceutical pipelines, regulatory agencies must first begin defining and setting standards on what data can be considered valid to be used for health-related decisions. Up to this point, no tools have been provided for the quality assessment of RWD. In addition, generation of RWD is still relatively siloed, with industry-sponsored studies being the main contributors [39]. To harness the creative power of the broader community (eg, academic centers that do not have the resources to generate such data themselves), the process of identifying and assessing the quality of RWD must be streamlined [13]. Nevertheless, considering RWD as a source of evidence in clinical or regulatory decision-making is a process under development. Different variables for the consideration of RWD in making health-related decisions are important and need to be identified appropriately for a variety of end-use analyses. In addition, participants and patients involved in studies that generate RWD must provide complete and trustworthy information. When collecting RWD data sets outside of RCTs, assurance must be given on, for example, data reliability, integrity, availability, and, not least, completeness. The RWD-Cockpit provides an easy-to-use and traceable first general assessment of RWD quality, which is applicable to a wide variety of data sets. Current regulations do not provide a sufficient framework for inclusion of RWD in studies investigating diseases other than orphan diseases or oncological diseases [44-46]. However, the value and possibilities when considering RWD during the whole product lifecycle instead of only postmarket authorization are

recognized across disease areas. The RWD-Cockpit enables health care stakeholders to obtain preliminary quality information across RWD data sets. Users can use this preliminary criterion when searching for, and selecting, specific data sets for consideration in investigations. The web-based RWD-Cockpit application might provide an initial standard blueprint for regulatory authorities considering preliminary approval of the use of RWD in different settings.

The RWD-Cockpit allows users to score data sets independently, where they can assess the scores according to the RWD formula and further choose whether they wish to publish these data sets in the RWD-Cockpit. In addition, the results of the case studies demonstrate the applicability and usability of the RWD-Cockpit application to the wider community. Future versions of the RWD-Cockpit or frameworks for the assessment of RWD quality before further selection and analyses might also consider additional information on the data-generation process, including variables focusing on used devices or firmware. Furthermore, information can be included on whether the data have been centralized by a single institution or person or distributed and combined from multiple sources. Another potential improvement to the RWD-Cockpit could result from the automation of the scoring process. An automated machine-performed grading

mechanism could potentially lead to an increase in consistency compared with the current manual grading by an individual. Another potential future useful feature could enable users to score their own RWD data sets based on their own criteria and similarly apply these quality criteria across a single institution or through multiple institutions.

## Conclusions

The RWD-Cockpit web application is designed to enable a fast and reliable scoring system for evaluating the multi-metric quality of RWD data sets. It aims to reduce preliminary issues related to quality assessment of RWD and streamline the discovery of valuable RWD data sets, and it has the potential to be used in clinical settings. The application of this tool in the context of RWD is diverse and expandable, as demonstrated through the case studies. With the advent of digital medicine and the increasing challenges in data and metadata standards of RWD, there is a pressing need to develop frameworks and tools that represent RWD quality in a metric, comprehensible, and traceable manner and can serve as a standard across data sources and disease areas. The RWD-Cockpit represents a first metric proposal in this direction; however, further community efforts are urgently needed.

---

## Acknowledgments

The authors are grateful to the Stiftung FHNW and the FHNW University of Applied Sciences and Arts Northwestern Switzerland, School of Life Sciences, and the Institute of Medical Engineering and Medical Informatics for supporting this work.

---

## Authors' Contributions

EM, PA, FG, MD, and DB designed the concept. EM, ES, LMB, TA, and DDT developed the concept for the web application. DW and EWG provided data and validated the application method. IC provided data. All authors wrote the manuscript.

---

## Conflicts of Interest

EM owns shares in aiNET GmbH. DW and EWG own shares and work for Q-Strip. IC works for Evidation Health.

---

### Multimedia Appendix 1

The Q-Strip is connected to a printed circuit board where the resistance between 2 electrodes is measured. The data collected are sent to a secure database.

[[PNG File , 306 KB - formative\\_v6i10e29920\\_app1.png](#) ]

---

### Multimedia Appendix 2

The Q-Strip is placed at chest height in a bed to measure nocturnal sweat best accurately.

[[PNG File , 54 KB - formative\\_v6i10e29920\\_app2.png](#) ]

---

### Multimedia Appendix 3

The longitudinal data are collected and then visualized in the Q-Strip software. The x-axis represents date and time and the y-axis represents the resistance value corresponding to amount of sweat detected.

[[PNG File , 222 KB - formative\\_v6i10e29920\\_app3.png](#) ]

---

### Multimedia Appendix 4

Example of quality-assessed real-world data.

[[PDF File \(Adobe PDF File\), 113 KB - formative\\_v6i10e29920\\_app4.pdf](#) ]

---

### Multimedia Appendix 5

Q-strip product is a ribbon that contains electrodes which measure humidity. The resistance between the 2 electrodes within the strip measure humidity which is then correlated to sweat.

[PNG File , 226 KB - formative\_v6i10e29920\_app5.png ]

## References

1. Makady A, de Boer A, Hillege H, Klungel O, Goettsch W, (on behalf of GetReal Work Package 1). What is real-world data? A review of definitions based on literature and stakeholder interviews. *Value Health* 2017;20(7):858-865 [FREE Full text] [doi: [10.1016/j.jval.2017.03.008](https://doi.org/10.1016/j.jval.2017.03.008)] [Medline: [28712614](https://pubmed.ncbi.nlm.nih.gov/28712614/)]
2. U.S. Department of Health and Human Services Food and Drug Administration, Center for Devices and Radiological Health, Center for Biologics Evaluation and Research. Use of real-world evidence to support regulatory decision-making for medical devices: guidance for industry and food and drug administration staff. U.S. Food & Drug Administration. 2017. URL: <https://www.fda.gov/media/99447/download> [accessed 2022-03-24]
3. Tanguy A, Jönsson L, Ishihara L. Inventory of real world data sources in Parkinson's disease. *BMC Neurol* 2017 Dec 08;17(1):213 [FREE Full text] [doi: [10.1186/s12883-017-0985-0](https://doi.org/10.1186/s12883-017-0985-0)] [Medline: [29216834](https://pubmed.ncbi.nlm.nih.gov/29216834/)]
4. Khosla S, White R, Medina J, Ouwens M, Emmas C, Koder T, et al. Real world evidence (RWE) - a disruptive innovation or the quiet evolution of medical evidence generation? *F1000Res* 2018 Jan 25;7:111 [FREE Full text] [doi: [10.12688/f1000research.13585.2](https://doi.org/10.12688/f1000research.13585.2)] [Medline: [30026923](https://pubmed.ncbi.nlm.nih.gov/30026923/)]
5. Katkade VB, Sanders KN, Zou KH. Real world data: an opportunity to supplement existing evidence for the use of long-established medicines in health care decision making. *J Multidiscip Healthc* 2018 Jul 2;11:295-304 [FREE Full text] [doi: [10.2147/JMDH.S160029](https://doi.org/10.2147/JMDH.S160029)] [Medline: [29997436](https://pubmed.ncbi.nlm.nih.gov/29997436/)]
6. 21st century cures act. U.S. Food & Drug Administration. 2020. URL: <https://www.fda.gov/regulatory-information/selected-amendments-fdc-act/21st-century-cures-act> [accessed 2021-10-30]
7. Analysis and summaries of public consultation results: EMA regulatory science to 2025. European Medicines Agency. 2020. URL: [https://www.ema.europa.eu/en/documents/comments/analysis-summaries-public-consultation-results-ema-regulatory-science-2025-strategic-reflection\\_en.pdf](https://www.ema.europa.eu/en/documents/comments/analysis-summaries-public-consultation-results-ema-regulatory-science-2025-strategic-reflection_en.pdf) Please provide the access date for this reference [accessed 2022-03-24]
8. Gallacher J, de Reydet de Vulpillieres F, Amzal B, Angehrn Z, Bexelius C, Bintener C, ROADMAP Consortium. Challenges for optimizing real-world evidence in Alzheimer's disease: the ROADMAP project. *J Alzheimers Dis* 2019;67(2):495-501 [FREE Full text] [doi: [10.3233/JAD-180370](https://doi.org/10.3233/JAD-180370)] [Medline: [30584137](https://pubmed.ncbi.nlm.nih.gov/30584137/)]
9. National Academies of Sciences, Engineering, and Medicine, Health and Medicine Division, Board on Health Sciences Policy, Forum on Drug Discovery, Development, and Translation. In: Shore C, Gee AW, Kahn B, editors. When Is a Real-World Data Element Fit for Assessment of Eligibility, Treatment Exposure, or Outcomes?. Washington, DC, USA: National Academies Press (US); 2019.
10. FDA approves new use of transplant drug based on real-world evidence. U.S. Food & Drug Administration. 2021. URL: <https://www.fda.gov/drugs/news-events-human-drugs/fda-approves-new-use-transplant-drug-based-real-world-evidence> [accessed 2021-10-31]
11. Cave A, Kurz X, Arlett P. Real-world data for regulatory decision making: challenges and possible solutions for Europe. *Clin Pharmacol Ther* 2019 Jul;106(1):36-39 [FREE Full text] [doi: [10.1002/cpt.1426](https://doi.org/10.1002/cpt.1426)] [Medline: [30970161](https://pubmed.ncbi.nlm.nih.gov/30970161/)]
12. Bolislis WR, Fay M, Kühler TC. Use of real-world data for new drug applications and line extensions. *Clin Ther* 2020 May;42(5):926-938 [FREE Full text] [doi: [10.1016/j.clinthera.2020.03.006](https://doi.org/10.1016/j.clinthera.2020.03.006)] [Medline: [32340916](https://pubmed.ncbi.nlm.nih.gov/32340916/)]
13. Babrak LM, Menetski J, Rebhan M, Nisato G, Zinggeler M, Brasier N, et al. Traditional and digital biomarkers: two worlds apart? *Digit Biomark* 2019 Aug 16;3(2):92-102 [FREE Full text] [doi: [10.1159/000502000](https://doi.org/10.1159/000502000)] [Medline: [32095769](https://pubmed.ncbi.nlm.nih.gov/32095769/)]
14. Sources of real-world data. RWE Navigator. URL: <https://rwe-navigator.eu/use-real-world-evidence/sources-of-real-world-data/> [accessed 2021-10-31]
15. National Academies of Sciences, Engineering, and Medicine, Health and Medicine Division, Board on Health Sciences Policy, Forum on Drug Discovery, Development, and Translation. In: Shore C, Gee AW, Kahn B, editors. Barriers and Disincentives to the Use of Real-World Evidence and Real-World Data. Washington, DC, USA: National Academies Press (US); 2019.
16. Lipworth W. Real-world data to generate evidence about healthcare interventions: the application of an ethics framework for big data in health and research. *Asian Bioeth Rev* 2019 Sep;11(3):289-298 [FREE Full text] [doi: [10.1007/s41649-019-00095-1](https://doi.org/10.1007/s41649-019-00095-1)] [Medline: [33717317](https://pubmed.ncbi.nlm.nih.gov/33717317/)]
17. Morton SC, Costlow MR, Graff JS, Dubois RW. Standards and guidelines for observational studies: quality is in the eye of the beholder. *J Clin Epidemiol* 2016 Mar;71:3-10 [FREE Full text] [doi: [10.1016/j.jclinepi.2015.10.014](https://doi.org/10.1016/j.jclinepi.2015.10.014)] [Medline: [26548541](https://pubmed.ncbi.nlm.nih.gov/26548541/)]
18. SwissRDL - medical registries and data linkage. SwissRDL. 2021 May 17. URL: [https://www.swissrdl.unibe.ch/index\\_eng.html](https://www.swissrdl.unibe.ch/index_eng.html) [accessed 2021-10-31]
19. Cancer Registry of Norway. URL: <https://www.kreftregisteret.no/en> [accessed 2021-10-31]
20. Your machine learning and data science community. Kaggle. URL: <https://www.kaggle.com/> [accessed 2021-10-31]
21. Google dataset search. URL: <https://datasetsearch.research.google.com/> [accessed 2021-10-31]



22. Opendata Swiss. URL: <https://opendata.swiss/en/> [accessed 2021-10-31]
23. University of California Irvine machine learning repository. URL: <https://archive.ics.uci.edu/ml/index.php> [accessed 2021-10-31]
24. Panel study of income dynamics. Institute for Social Research, University of Michigan. URL: <https://psidonline.isr.umich.edu/> [accessed 2021-10-31]
25. Home blood pressure monitoring pilot: NYU Langone health EHR. New York University Health Scenes Library. URL: <https://datacatalog.med.nyu.edu/dataset/10354> [accessed 2021-10-31]
26. Height\_Weight\_single\_variable\_data\_101\_series\_1.0: predicting weight for a given height. kaggle. URL: <https://kaggle.com/jamesbasker/height-weight-single-variable-data-101-series-10> [accessed 2021-10-31]
27. National health and nutrition examination survey: NHANES datasets from 2013-2014. Kaggle. URL: <https://kaggle.com/cdc/national-health-and-nutrition-examination-survey> [accessed 2021-10-31]
28. Alwhaibi M, Balkhi B, Alshammari TM, AlQahtani N, Mahmoud MA, Almetwazi M, et al. Measuring the quality and completeness of medication-related information derived from hospital electronic health records database. *Saudi Pharm J* 2019 May;27(4):502-506 [FREE Full text] [doi: [10.1016/j.jsps.2019.01.013](https://doi.org/10.1016/j.jsps.2019.01.013)] [Medline: [31061618](https://pubmed.ncbi.nlm.nih.gov/31061618/)]
29. Twitter - It's what's happening. Twitter. URL: <https://twitter.com/> [accessed 2021-10-31]
30. OpenML: machine learning better, together. OpenML. URL: <https://www.openml.org> [accessed 2021-10-31]
31. Lyons BE, Austin D, Seelye A, Petersen J, Yeagers J, Riley T, et al. Pervasive computing technologies to continuously assess Alzheimer's disease progression and intervention efficacy. *Front Aging Neurosci* 2015 Jun 10;7:102 [FREE Full text] [doi: [10.3389/fnagi.2015.00102](https://doi.org/10.3389/fnagi.2015.00102)] [Medline: [26113819](https://pubmed.ncbi.nlm.nih.gov/26113819/)]
32. Ellis RJ, Ng YS, Zhu S, Tan DM, Anderson B, Schlaug G, et al. A validated smartphone-based assessment of gait and gait variability in Parkinson's disease. *PLoS One* 2015 Oct 30;10(10):e0141694 [FREE Full text] [doi: [10.1371/journal.pone.0141694](https://doi.org/10.1371/journal.pone.0141694)] [Medline: [26517720](https://pubmed.ncbi.nlm.nih.gov/26517720/)]
33. Grimberg F, Asprion PM, Schneider B, Miho E, Babrak L, Habbabeh A. The real-world data challenges radar: a review on the challenges and risks regarding the use of real-world data. *Digit Biomark* 2021 Jun 24;5(2):148-157 [FREE Full text] [doi: [10.1159/000516178](https://doi.org/10.1159/000516178)] [Medline: [34414352](https://pubmed.ncbi.nlm.nih.gov/34414352/)]
34. Reynolds MW, Bourke A, Dreyer NA. Considerations when evaluating real-world data quality in the context of fitness for purpose. *Pharmacoepidemiol Drug Saf* 2020 Oct;29(10):1316-1318 [FREE Full text] [doi: [10.1002/pds.5010](https://doi.org/10.1002/pds.5010)] [Medline: [32374042](https://pubmed.ncbi.nlm.nih.gov/32374042/)]
35. Welcome to the Real-World Data Cockpit Application. aiHealthLab, FHNW University of Applied Sciences Northwestern Switzerland. URL: <http://rwd.aihealth.ch/> [accessed 2022-04-27]
36. Ismail L, Materwala H, Karduck AP, Adem A. Requirements of health data management systems for biomedical care and research: scoping review. *J Med Internet Res* 2020 Jul 7;22(7):e17508 [FREE Full text] [doi: [10.2196/17508](https://doi.org/10.2196/17508)] [Medline: [32348265](https://pubmed.ncbi.nlm.nih.gov/32348265/)]
37. Jolink K, Bedrosian L. Introducing the Q-strip to quantitatively measure nocturnal perspiration as a parameter and predictor: a review of literature. *Med Pharmacol (forthcoming)* 2020:2020040418. [doi: [10.20944/preprints202004.0418.v1](https://doi.org/10.20944/preprints202004.0418.v1)]
38. Swift B, Jain L, White C, Chandrasekaran V, Bhandari A, Hughes DA, et al. Innovation at the intersection of clinical trials and real-world data science to advance patient care. *Clin Transl Sci* 2018 Sep;11(5):450-460 [FREE Full text] [doi: [10.1111/cts.12559](https://doi.org/10.1111/cts.12559)] [Medline: [29768712](https://pubmed.ncbi.nlm.nih.gov/29768712/)]
39. Marra C, Chen JL, Coravos A, Stern AD. Quantifying the use of connected digital products in clinical research. *NPJ Digit Med* 2020 Apr 3;3:3:50 [FREE Full text] [doi: [10.1038/s41746-020-0259-x](https://doi.org/10.1038/s41746-020-0259-x)] [Medline: [32285011](https://pubmed.ncbi.nlm.nih.gov/32285011/)]
40. Wilkinson MD, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A, et al. The FAIR guiding principles for scientific data management and stewardship. *Sci Data* 2016 Mar 15;3:160018 [FREE Full text] [doi: [10.1038/sdata.2016.18](https://doi.org/10.1038/sdata.2016.18)] [Medline: [26978244](https://pubmed.ncbi.nlm.nih.gov/26978244/)]
41. Smith S, Lumsden S, Rocca M, Gordon S, Chen W, Hasan R, et al. Common data model harmonization (CDMH) and open standards for evidence generation. U.S. Food & Drug Administration. 2020. URL: <https://aspe.hhs.gov/sites/default/files/private/pdf/259016/CDMH-Final-Report-14August2020.pdf> [accessed 2022-03-24]
42. Larbi D, Randine P, Årsand E, Antypas K, Bradway M, Gabarron E. Methods and evaluation criteria for apps and digital interventions for diabetes self-management: systematic review. *J Med Internet Res* 2020 Jul 06;22(7):e18480 [FREE Full text] [doi: [10.2196/18480](https://doi.org/10.2196/18480)] [Medline: [32628125](https://pubmed.ncbi.nlm.nih.gov/32628125/)]
43. Robles N, Puigdomènech Puig E, Gómez-Calderón C, Saigí-Rubió F, Cuatrecasas Cambra G, Zamora A, et al. Evaluation criteria for weight management apps: validation using a modified Delphi process. *JMIR Mhealth Uhealth* 2020 Jul 22;8(7):e16899 [FREE Full text] [doi: [10.2196/16899](https://doi.org/10.2196/16899)] [Medline: [32706689](https://pubmed.ncbi.nlm.nih.gov/32706689/)]
44. Breckenridge AM, Breckenridge RA, Peck CC. Report on the current status of the use of real-world data (RWD) and real-world evidence (RWE) in drug development and regulation. *Br J Clin Pharmacol* 2019 Sep;85(9):1874-1877 [FREE Full text] [doi: [10.1111/bcp.14026](https://doi.org/10.1111/bcp.14026)] [Medline: [31290181](https://pubmed.ncbi.nlm.nih.gov/31290181/)]
45. National Academies of Sciences, Engineering, and Medicine, Health and Medicine Division, Board on Health Sciences Policy, Forum on Drug Discovery, Development, and Translation. Looking ahead. In: Shore C, Gee AW, Kahn B, editors. *Examining the Impact of Real-World Evidence on Medical Product Development: Proceedings of a Workshop Series*. Washington, DC, USA: National Academies Press (US); 2019.

46. Khozin S, Blumenthal GM, Pazdur R. Real-world data for clinical evidence generation in oncology. *J Natl Cancer Inst* 2017 Nov 01;109(11):djsx187. [doi: [10.1093/jnci/djsx187](https://doi.org/10.1093/jnci/djsx187)] [Medline: [29059439](https://pubmed.ncbi.nlm.nih.gov/29059439/)]

## Abbreviations

**EMA:** European Medicines Agency  
**FDA:** Food and Drug Administration  
**RCT:** randomized controlled trial  
**RWD:** real-world data  
**RWE:** real-world evidence

*Edited by A Mavragani; submitted 25.04.21; peer-reviewed by AM Auvinen, C Schäfer; comments to author 21.07.21; revised version received 31.10.21; accepted 19.02.22; published 18.10.22.*

*Please cite as:*

*Babrak LM, Smakaj E, Agac T, Asprion PM, Grimberg F, der Werf DV, van Ginkel EW, Tosoni DD, Clay I, Degen M, Brodbeck D, Natali EN, Schkommodau E, Miho E*

*RWD-Cockpit: Application for Quality Assessment of Real-world Data*

*JMIR Form Res* 2022;6(10):e29920

URL: <https://formative.jmir.org/2022/10/e29920>

doi: [10.2196/29920](https://doi.org/10.2196/29920)

PMID: [35266872](https://pubmed.ncbi.nlm.nih.gov/35266872/)

©Lmar Marie Babrak, Erand Smakaj, Teyfik Agac, Petra Maria Asprion, Frank Grimberg, Daan Van der Werf, Erwin Willem van Ginkel, Deniz David Tosoni, Ieuan Clay, Markus Degen, Dominique Brodbeck, Eriberto Noel Natali, Erik Schkommodau, Enkelejda Miho. Originally published in *JMIR Formative Research* (<https://formative.jmir.org>), 18.10.2022. 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

# Understanding Mental Health Apps for Youth: Focus Group Study With Latinx Youth

Elena Agapie<sup>1</sup>, PhD; Katherine Chang<sup>2</sup>, BA; Sneha Patrachari<sup>1</sup>, MHCID; Martha Neary<sup>2</sup>, MSc; Stephen M Schueller<sup>1,2</sup>, PhD

<sup>1</sup>Department of Informatics, University of California, Irvine, Irvine, CA, United States

<sup>2</sup>Department of Psychological Science, University of California, Irvine, Irvine, CA, United States

**Corresponding Author:**

Stephen M Schueller, PhD

Department of Psychological Science

University of California, Irvine

4201 Social and Behavioral Sciences Gateway

Irvine, CA, 92697

United States

Phone: 1 949 824 3850

Email: [s.schueller@uci.edu](mailto:s.schueller@uci.edu)

## Abstract

**Background:** An increasing number of mental health apps (MHapps) are being developed for youth. In addition, youth are high users of both technologies and MHapps. However, little is known about their perspectives on MHapps. MHapps might be particularly well suited to reach the youth underserved by traditional mental health resources, and incorporating their perspectives is especially critical to ensure such tools are useful to them.

**Objective:** The goal of this study was to develop and pilot a process for eliciting youth perspectives on MHapps in a structured and collaborative way. We also sought to generate learnings on the perspectives of Latinx youth on MHapps and their use in ways that might facilitate discovery, activation, or engagement in MHapps, especially in Latinx populations.

**Methods:** We created a series of focus groups consisting of 5 sessions. The groups introduced different categories of MHapps (cognitive behavioral therapy apps, mindfulness apps, and miscellaneous apps). Within each category, we selected 4 MHapps that participants chose to use for a week and provided feedback through both between-session and in-session activities. We recruited 5 youths ranging in age from 15 to 21 (mean 18, SD 2.2) years. All the participants identified as Hispanic or Latinx. After completing all 5 focus groups, the participants completed a brief questionnaire to gather their impressions of the apps they had used.

**Results:** Our focus group methodology collected detailed and diverse information about youth perspectives on MHapps. However, we did identify some aspects of our methods that were less successful at engaging the youth, such as our between-session activities. The Latinx youth in our study wanted apps that were accessible, relatable, youth centric, and simple and could be integrated with their offline lives. We also found that the mindfulness apps were viewed most favorably but that the miscellaneous and cognitive behavioral therapy apps were viewed as more impactful.

**Conclusions:** Eliciting youth feedback on MHapps is critical if these apps are going to serve a role in supporting their mental health and well-being. We refined a process for collecting feedback from the youth and identified factors that were important to a set of Latinx youth. Future work could be broader, that is, recruit larger samples of more diverse youth, or deeper, that is, collect more information from each youth around interests, needs, barriers, or facilitators or better understand the various impacts of MHapps by using qualitative and quantitative measures. Nevertheless, this study advances the formative understanding of how the youth, particularly Latinx youth, might be viewing these tools.

(*JMIR Form Res* 2022;6(10):e40726) doi:[10.2196/40726](https://doi.org/10.2196/40726)

**KEYWORDS**

mental health; mental health apps; youth; child; teenager; focus group; human-centered design; mobile health; mHealth; health app; cognitive behavioral therapy; CBT; perspective; qualitative; mindfulness; health app; digital health tool; Latino; Latinx; mobile phone

## Introduction

### Background

Numerous health apps exist, with estimates of just over 350,000 health apps in 2020, with 18% of those being for stress and wellness and 5% for mental health conditions [1]. Despite this large number of available mental health apps (MHapps), few are downloaded, and those that are downloaded rarely have long-term, sustained use. A review of MHapps targeting depression and anxiety found that the top 3 most downloaded apps accounted for 90% of all downloads and that 63% of depression apps and 56% of anxiety apps had no regular users over a month period [2]. Another study found that the median retention rate for MHapps 15 days after downloading was 3.9%, and by 30 days, it dropped to 3.3% [3]. Thus, app discovery (ie, finding useful MHapps), app activation (ie, initiating the use of an MHapp), and app engagement (ie, sustained and effective use of an MHapp over time) are all major barriers to ensuring that MHapps can provide the most impact on consumers' lives.

It is often proposed that MHapps could have a significant impact on youth [4-6]. Most mental health disorders emerge in adolescence or early adulthood, suggesting that youth is a critical transitory period for mental health [7]. Youth face many barriers to care that include structural and perceptual barriers, including perceptions of both mental health problems and services [8]. Youth of color experience even more barriers to care. They are less likely to live in communities with mental health providers [9] and to receive traditional mental health care than their White counterparts [10,11]. Even when providers and treatments are available, providers are less likely to match their racial and ethnic backgrounds, and culturally relevant care is rarely available [12,13]. For youth with mental health needs, parents and other adults are often the gatekeepers to care. As such, youth often seek care through informal sources first, which includes family and friends but increasingly includes internet-based information and interventions [14]. Seeking informal sources of support before engaging in traditional care might be even more common among Black and Latinx youth than among their White counterparts [15].

MHapps are a potentially useful way to reach youth and provide mental health support and interventions. Smartphone ownership increases substantially from age 11 to 18 years. With over half of youth aged 11 years and 91% of those aged 18 years owning smartphones [16]. As such, youth are quite comfortable using technology, and most report using them almost constantly to access the internet and communicate with peers [17]. Many youth also report that they are using or have used health apps. A recent representative survey of those aged 14 to 22 years found that 69% report having used a health app, a number that climbs to 75% among the youth with moderate to severe symptoms of depression [18]. In this survey, many of the categories of apps youth reported using were related to mental health, including sleep (27%), meditation or mindfulness (17%), stress reduction (14%), mood tracking (10%), depression (9%), and alcohol or substance use (5%). Thus, MHapp use among youth seems common.

MHapps might be particularly useful to provide mental health resources for traditionally underserved groups, such as racial and ethnic minorities [19,20]. First, as mentioned earlier, a sufficient workforce that mirrors the racial and ethnic diversity of the population in need does not exist. As an example, in California, Latinxs represent 38% of the state's population but only 4% of its psychiatrists and 8% of its psychologists [21]. Investment in developing a better pipeline to train and retain diverse mental health providers should be a priority, although this process will take time and resources are needed today to help those in need. Second, MHapps can overcome geographic barriers and can be deployed more effectively where gaps in service provision exist. More than 80% of counties in the United States are designated as mental health professional shortage areas [22], and again, youth of color are less likely to live in areas with access to care [9]. Third, technology might be seen as a more desirable resource either because of the ease of access, overcoming barriers such as transportation or stigma [20], or just because some people might be more inclined to use an MHapp rather than traditional care. Calls to use technology for mental health service delivery have emphasized the need to broaden the portfolio of available mental health resources to promote market segmentation; that is, identifying targeted groups of consumers to better tailor products, such as mental health resources, to those populations. Even when racial and ethnic minorities receive traditional mental health services, they tend to receive fewer sessions [23]. Therefore, in addition to considering how to better engage racial and ethnic minorities in care, we must also consider how to design care that will better serve them when they get there. MHapps might be one such solution to help address this challenge.

Although robust literature exists to suggest that MHapps are effective [24,25], reviews specifically focused on apps for youth have shown that research in this space is sparse [26,27] but have demonstrated some indications of positive benefits [28]. In addition to there being fewer studies that address the effectiveness of MHapps for youth, most of the studies conducted focus on the early feasibility or acceptability of apps that never make it into the hands of consumers. For example, Grist et al [27] found in their systematic review that of the 15 apps described in 24 eligible studies, only 2 were available to download. Studies on MHapps for youth have been characterized by poor uptake and engagement, lack of specification of procedures such as human support, and a lack of research on younger children or traditionally marginalized populations [29]. Given the lack of empirical evidence to speak of the effectiveness of MHapps for racial and ethnic minority youth, processes to better elicit their perspectives and views could be useful to inform the adoption and deployment of MHapps among such populations.

Thus, a gap exists between the use and enthusiasm of MHapps for youth and our understanding of the actual impact on this group. It is worthwhile to better understand youth's impressions, especially youth of color, of MHapps to understand what they would use and why, potentially leading to better design and dissemination of MHapps for youth and youth of color.

## Including Youth Perspectives in the Research Process

One way to address this gap and provide information to better design studies to investigate the potential effectiveness of MHapps for youth is to incorporate youth's perspectives on MHapps, especially those currently available to youth, into research [30]. Indeed, research has taken different approaches to consider the youth perspective, including expert opinions, app reviews, and working with youth themselves. One study conducted a workshop with psychologists with experience working with youth to identify different youth media preferences and then analyzed app features to identify the presence of these media preferences [31]. Media preferences included strategies of social connectivity, use of videos, tailorability, and rich interactions. However, they found that the use of these strategies was limited in apps available for youth, with social connectivity being the most used but with few examples of interactive multimedia experiences. Several studies have worked with youth to co-design digital mental health interventions [32-35], including MHapps, but these studies have often resulted in early-stage ideas or functional prototypes that, although sometimes are evaluated, rarely make it to app stores or other places where youth could access them. Finally, many studies have focused on trying to understand MHapps from app store user reviews [36,37]. Although these studies have not focused specifically on MHapps for youth people or on reviews coming only from youth, given that youth are using these tools, it is possible that some of these reviews come from youth. Overall, these studies demonstrate findings similar to other evaluations that leverage expert opinions or content analysis. Users appreciate apps that are esthetically pleasing and functionally diverse while allowing opportunities for personalization and tailoring. In addition, users dislike apps with poor usability and pricing models that focus on freemium models or emergent costs. However, a limitation of these reviews is that these prompts are general and may not elicit feedback on the aspects of MHapps that might best support their design and implementation.

Diverse methods exist to gather people's impressions of digital products and to team with potential users in ways to better create technologies that meet their needs. These methods, leveraged from the field of human-centered design, have been used in various projects working to develop MHapps for the youth or to understand the needs and interests of the youth when interacting with technologies. Applying such techniques to better understand the impressions of existing apps is also useful to understand what are the affordances of these technologies as well as where such technologies are lacking. In our study, we drew on synchronous methods, used focus groups, and adapted asynchronous remote community (ARC) methods to inform the aspects of our study design. ARC includes a collection of methods intended to facilitate group-based research at a given distance [38]. ARC has been used in various apps, especially for working with populations in which in-person coordination might be challenging, such as populations with rare diseases [39] or stigmatized populations [40]. ARC methods have also been used with teens to leverage digital communication using platforms familiar to teens (eg, social media and asynchronous communication) and to balance their busy schedules [41].

Despite the creation of ARC methods before the pandemic, these methods are also useful for conducting remote design work that is necessary, given the need to follow physical distancing protocols.

In this study, we attempted to develop a process for incorporating youth perspectives on MHapps in a structured and collaborative way. We sought to pilot our methods with a small group of youth while also attempting to learn their perspectives on available MHapps. We were particularly interested in the views of youth from traditionally underserved and marginalized populations (given the potential of MHapps to overcome barriers they face and the relative lack of knowledge of MHapps for these populations in comparison with their White counterparts) and therefore worked with a sample of youth with a Latinx background. We conducted a 5-week focus group with a set of youth aged 15 to 21 years to refine a process for incorporating app feedback into their views on MHapps and to identify some characteristics of MHapps that might serve as barriers to or facilitators of discovery, activation, or engagement. As such, we framed our work in contributing both our methodologies and some formative learnings from our small pilot of these methods.

## Methods

### Recruitment

We recruited participants through a local nonprofit team that works with youth from underrepresented groups, especially Latinx youth, to accelerate readiness in Science, Technology, Engineering, Arts, and Math. Recruitment occurred via email and recruitment messages posted on the nonprofit team's Discord channel. We compensated participants with a US \$25 Amazon gift card for each of the 5 focus group sessions that a participant attended, with a maximum compensation of US \$125 for the entire study.

### Ethical Considerations

Informed consent was obtained from all participants aged  $\geq 18$  years. For participants  $< 18$  years of age, informed consent was obtained from their parent or guardian and assent was obtained from the participants. Researchers had an existing partnership with this nonprofit team, which facilitated consenting participants and their parents or guardians for participants  $< 18$  years. A member of our research team provided a digital consent or assent form to the participants and a digital consent form to parents for participants  $< 18$  years of age. A member of the nonprofit team helped facilitate contact with parents or guardians when necessary. Institutional review board approval was obtained from the University of California, Irvine (#2019-5609).

### In-Session Activities

We conducted 5 focus group sessions, 1 per week, totaling 5 weeks. For the first 3 weeks, each week, participants used a new category of MHapps and then discussed the app they used during the following session. In the fourth week, participants were asked to choose 1 app to use from all the apps we introduced across all the weeks. In each session, we collected feedback about the apps participants used in the past week. In

the fifth session, we discussed feedback across all the apps participants used.

The app categories were cognitive behavioral therapy (CBT) apps (week 1), mindfulness apps (week 2), and miscellaneous apps, which did not pertain to 1 category, including a coping app, journaling app, mood-tracking app, and peer support app (week 3). Each category contained 4 apps. These apps were selected to represent the general categories of MHapps that the youth use and the features present in popular MHapps. Furthermore, our research team used our experience of identifying and evaluating apps and considered apps that were used by youth based on previous analyses [42]. Each of the 4 apps discussed in the week was assigned to a different participant (2 participants used the same app).

We selected focus groups for our method because focus groups enabled us to contrast between the different apps by engaging in group discussions. We used this approach to enable discussions about the different types of functionalities of related apps. Focus groups also enabled us to address known challenges in working with teens [43] by making it easier to balance power dynamics between researchers and teens. We drew on design

theory that has established that discussion around different prototypes can result in increased participant rapport and sharing about the prototypes they each engaged with [44]. We also built on prior work through which exposing participants to different types of designs and functionalities can elicit feedback on functionalities that are most important to users [45].

### Focus Group Session Structure

Each focus group session was approximately 90 minutes and was organized around a similar schedule (Table 1): icebreakers to build rapport between participants and researchers, discussion and activities about the apps assigned to participants in the past week (week 2-5), a brief presentation of a core concept for the apps in the upcoming week (eg, CBT or mindfulness), and assignment of apps that participants would use in the following week. To ensure that all the apps were evaluated, each participant chose a different app until all the apps were chosen by at least one person. We then asked the participants to use the app of their choice throughout the next week. In the week 4 session, the participants could choose an app that they had already used in the past few weeks or they could choose a completely new app that they were interested in.

**Table 1.** Overview of session topics and apps assigned.

Week	Activities in focus group sessions	MHapps <sup>a</sup> assigned
Week 1	<ul style="list-style-type: none"> <li>• Overview of the study</li> <li>• Getting to know each other (icebreaker)</li> <li>• Introduction to CBT<sup>b</sup></li> </ul>	<ul style="list-style-type: none"> <li>• Mindshift</li> <li>• Sanvello</li> <li>• Woebot</li> <li>• Wysa</li> </ul>
Week 2	<ul style="list-style-type: none"> <li>• Getting to know each other (icebreaker)</li> <li>• Feedback on CBT apps</li> <li>• Introduction to mindfulness</li> </ul>	<ul style="list-style-type: none"> <li>• Headspace</li> <li>• Insight Timer</li> <li>• Shine</li> <li>• Smiling Mind</li> </ul>
Week 3	<ul style="list-style-type: none"> <li>• Getting to know each other (icebreaker)</li> <li>• Feedback on mindfulness apps</li> <li>• Introduction to “wellness hacks”</li> </ul>	<ul style="list-style-type: none"> <li>• Covid Coach</li> <li>• Daylio</li> <li>• Moodflow</li> <li>• Talk Life</li> </ul>
Week 4	<ul style="list-style-type: none"> <li>• Getting to know each other (icebreaker)</li> <li>• Feedback on miscellaneous apps</li> <li>• Overview of last week’s assignments</li> <li>• Review and selection of previous weeks’ apps</li> </ul>	<ul style="list-style-type: none"> <li>• An app from a previous week</li> </ul>
Week 5	<ul style="list-style-type: none"> <li>• Getting to know each other (icebreaker)</li> <li>• Reflecting on all apps</li> <li>• App design activity</li> </ul>	<ul style="list-style-type: none"> <li>• N/A<sup>c</sup></li> </ul>

<sup>a</sup>MHapp: mental health app.

<sup>b</sup>CBT: cognitive behavioral therapy.

<sup>c</sup>N/A: not applicable.

### Focus Group Activities for Collecting Feedback

We collected feedback about participants’ experiences using the assigned apps in the past week for approximately 50 of the total 90 minutes of sessions in weeks 2 to 5. Participants filled out sticky notes on an interactive remote whiteboard named Miro [46] in response to guided questions. Participants then shared their experiences with the app they used orally with the

rest of the group. We used the following questions to guide the discussion: “What did you like most about the app?”, “What did you like least?”, “How do you think apps like this could be useful for youth like you?”, “What do you think apps like this are missing for youth like you?”, and “If you could design a [category, i.e., CBT, mindfulness, mental health]-app for youth like you, what would you do differently?”

## Between-Session Activities

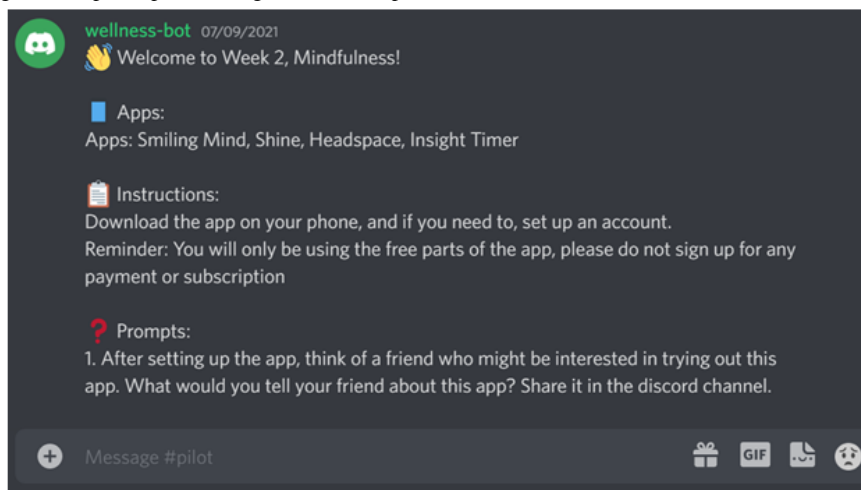
We used a web-based community platform to communicate with participants and elicit further information from participants about their experiences using the app they chose to interact with for the week. We chose Discord [47] because the participants were already using it as part of the community from which we recruited. The goal of these activities was to capture participant perspectives as they were using the apps in their lives, rather than reflecting on their use in the focus group sessions. We planned to deliver 3 questions every week. We used a generic user profile called “wellness-bot,” which would post the questions in Discord (Figure 1). The profile was controlled by a research team member. We named the account wellness-bot to encourage users to interact with the account as if it was a

chatbot and feel more comfortable with the anonymity of sharing with a chatbot instead of a person. Some questions were delivered in a channel created for this study that all participants shared, while others were delivered through direct messages. The questions prompted the participants to use the app and provide feedback on their experiences using the app.

Engagement in Discord started immediately after the first focus group. We continued that week with the second question:

*Now that you've set up the app, try out an activity! Some examples include logging how you're feeling today, reviewing past entries, listening to a video, etc. What is something about the app that you find useful or not useful? Share in the channel.*

**Figure 1.** Example message sent to participants through the Discord platform.



## Follow-up Survey

At the end of participation, the participants received a follow-up survey about the apps they used. They were asked to indicate the app they used within each category. For each app they indicated, they were asked the following questions (response options) shown in Textbox 1.

This survey was created by adapting the One Mind PsyberGuide Consumer Review Questionnaire [48] and questions from the

Mobile App Rating Scale [49], which have been recommended by service users for consumer evaluation in another study [48]. This included star ratings, satisfaction as indicated by likelihood to recommend the product, and open-ended questions about what they liked the most and least about the app. These questions also align with those used to evaluate technologies and apps in other settings, such as star ratings from app stores or the Net Promoter Score.

**Textbox 1.** Follow-up survey about the apps they used.

1. During the week you used the app, how often did you use it? (1, One day of the week; 2, A few days of the week; 3, Once a day; 4, Multiples times per day)
2. What did you like most about the app? (Free-text response)
3. What did you like least about the app? (Free-text response)
4. What impact, if any, did this app have on your wellness or how you felt? (1, Negative impact; 2, No impact; 3, Small positive impact; 4, Moderate to large positive impact)
5. Did you continue to use this app after the week you were asked to use it? (1-No; 2-Yes)
6. Would you recommend this app to people who might benefit from it? (1, Not at all--I would not recommend this app to anyone; 2, There are very few people I would recommend this app to; 3, Maybe -- There are several people whom I would recommend it to; 4, There are many people I would recommend this app to; 5, Definitely -- I would recommend this app to everyone)
7. What is your overall star rating of the app? (1, ☆ -- One of the worst apps I've used; 2, ☆☆; 3, ☆☆☆ -- Average; 4, ☆☆☆☆; 5, ☆☆☆☆☆ -- One of the best apps I've used)

## Analytic Strategy

### *Qualitative Analysis of Feedback From the Youth*

We used thematic analysis to analyze the focus group data, using an inductive approach [50] to conceptualize themes. A senior (EA) and junior (SP) researcher open coded the focus group transcripts using a descriptive coding approach [51]. The codes described the topics mentioned by the participants on a line-by-line basis. The codes were identified inductively based on the data. Given the small sample of data, we did not follow a fixed codebook; rather, we coded for the different topics that participants discussed and then conceptualized themes. Examples of codes identified include easy-to-consume content, personalization, social interaction, and relatability of content. The 2 researchers worked closely together, discussing data from all the focus groups. On the basis of the resultant codes, the 2 researchers conceptualized the data into themes related to each other through brief memos summarizing the coded data. During weekly meetings, the researchers discussed codes and themes in the data. The 2 researchers grouped subthemes together into themes such as personalization, connecting with others, and reliability of content. The other 4 researchers in the team assessed the subthemes and provided feedback, which led to a slight reorganization of the themes as presented in the results. Because the data collected centered on feedback about the functionalities of the apps, the themes that were conceptualized are also connected to app functionalities (eg, content type, features, and design of the app). We did not conduct interrater reliability, as any misalignments in coding supported our process of synthesizing data, understanding themes, and revising memos [52]. The themes presented are comprehensive of our sample, but given the small sample size and formative nature of this research, we did not intend to reach data saturation.

### *Analysis of Research Approach*

The research team reflected on the stages of the research design and assessed which ones were successful at engaging the youth: the way we engaged and built rapport with youth, how we engaged with the group of participants, and the activities we designed during and between the focus group sessions. Because this was a preliminary and formative study, our research team met on a weekly basis to discuss that week's session, make any modifications to the next week's activities, and discuss lessons learned. We note that our team conducting the research procedures did not include any members that identified as Latinx. However, this study was part of an existing research partnership with a nonprofit team from which participants were recruited, and members of the nonprofit team shaped the research design and recruitment considerations. For example, together we agreed that focus groups were suitable and to use the nonprofit team's existing Discord infrastructure for between-session interaction. The nonprofit team's members also directed our recruitment to students who were part of the program, influencing who was willing to participate. The research team included 1 researcher who had been conducting qualitative studies for >10 years (EA) and another for >5 years (SMS). In addition to collaboration with this nonprofit team for this project, our team also had multiple years of research and

clinical training with Latinx communities and collaboration with Latinx investigators.

## Results

### Participants

Five participants attended each of the focus groups. The participants ranged in age from 15 to 21 (mean 18, SD 2.2) years. Of 5 participants, 4 (80%) identified as women and 1 (20%) as men. All the participants identified as Hispanic or Latinx, and 1 (20%) participant also identified as Black.

### Focus Group Insights

#### *Overview*

Participants desired a range of features that would tailor the apps to the needs and practices of the youth, such as topically relevant content to youth issues, lightweight and playful content, and brief interactive exercises or content. The youth noted that the content of the apps came from authoritative sources. The youth also emphasized the importance of authentic social connections with others.

#### *Personalizing Content With Youth's Needs*

Participants shared that they often found the content of these apps not specific enough to the problems that youth encountered. Users mentioned wanting apps specifically designed for their demographics, age, challenges, and stage of life. One of the participants currently enrolled in high school felt "different timelines or milestones that youth deal with, [specifically] graduating from high school, and navigating adult life" could be helpful for others in the youth community.

#### *Customizing Content and App Display*

The participants found that tailoring the amount of content to their particular interests would help them engage with the apps. Such useful tailoring occurred when participants could customize the topics shown in the app so that they were better aligned with their interests (P4). The participants also wanted to customize the look and feel of the app by changing the colors used in the app (P4 and P1). They also preferred topics customized to the user, for example, by avoiding sensitive topics:

*I took into consideration that some people might have trauma or something. And it was like, oh, would you rather not see posts related to certain topics. So they pretty much asked you that, before you even get to see anything, which I really liked. [P4]*

The participants valued activities that could capture their personal situations. For example, P4 valued having the opportunity to create a plan to manage anxiety that was personal to themselves, not just to listen to or read generic content about the topic. When the content fit the participant's life, they felt the app was more relatable:

*I clicked on one of the manage stress [short audio clips] and they all let me see, oh, that's my story too. [P1]*



### **Content Aligned to Youth Use of Technology: “Lighten the Mood”**

The participants hoped that MHapps would complement the types of activities they engaged in when using technology. They wanted the duration of the content to be aligned with their use of technology for entertainment:

*Nowadays, people want something entertaining over something that takes more time. So in order for...youth to want to use a wellness app, they would have to bring in entertainment...maybe...including games, like something to make it more fun. [P5]*

Another participant mentioned that it was important to have content to “lighten the mood” and make it “light on emotion,” such as comedic content, or visual and interactive elements, such as emojis and stickers (P2). They thought this could be more aligned with a positive interaction with technology:

*I feel like this type of app should feel more like, I want to do this [as] part of my daily routine that’s actually going to bring me joy. [P2]*

They thought that content related to mental health could be alternated with other types of engaging content:

*Maybe I don’t want to be asked how I’m doing. Maybe I just want to play a game. [P1]*

A different way of making content more lightweight was through encouraging messages:

*As soon as I woke up the notification with words of encouragement was already there. And I would just have to click on it, and it would take me to a meditation. [P5]*

### **Integration With Offline Activities**

#### **Overview**

Several participants mentioned that they did not carry their phones with them all the time or that they preferred to do certain activities supported by the apps in a nondigital context; for example, journaling on paper. Some participants (P1 and P2) did not foresee using the digital tool for journaling:

*I don’t really use my phone for any other apps, mainly just internet browsing and YouTube. [P1]*

One participant thought it could be useful to get complementary support in their journaling activities; for example, by getting journaling templates that can guide the activity (P2).

#### **Valuing Diversity in Content Over Single-Focus Apps**

Participants found it valuable to have access to a range of content because it could give them control over what features to use:

*There was a lot of different categories, and I just chose the ones that I was, more like I needed help with. [P4]*

The type of content that interested participants included informational articles, audio content, meditation videos, or social content to help them relate to other people’s experiences (P4 and P5). The participants valued seeing content on different

topics, such as work, exercise, sleep, nutrition, mindfulness, emotions, or managing the pandemic (P3, P4, and P5) and expressed reluctance in using an app that was focused on only 1 topic, such as anxiety (P3 and P4). They preferred entertaining features in apps. They wanted access to music playlists to boost their mood (P4).

#### **Connecting With Others: Valuing Receiving and Providing Authentic Support**

The participants valued apps that had a social component, making them feel less alone (P2, P4, and P5). They were more interested in fostering a community with other users than in the general content of the apps (P5). They identified 4 different needs for feeling connected to others: being able to share challenging content, seeing positive or motivational content, receiving support from others, and offering help to others.

Participants wanted to share their experiences and struggles with others who have similar experiences:

*A safe space in which people who are experiencing the same feelings can talk to one another and just know that they’re not alone. [Participant not identifiable from the transcript]*

Participants valued having a space on the web to talk about things they would not be able to speak about with other people:

*I feel like it’s for people that don’t have someone to talk to so they come on this app to like, I guess vent about their lives. [P5]*

P1 felt it was useful for them to receive support from a peer who was going through similar challenges than through generic content in the app:

*Links to people who could help you out... who know what they’re doing... [P1]*

The participants also thought that offering support to others can bring them satisfaction as well. One participant posted a positive message on a community forum in the hope that it might help someone else who needed support (P5).

The participants valued authentic connections. They were reluctant to connect with others whose identities they could not recognize. For example, participants could not recognize if some of the social profiles of other people on the app were actual people or some type of automated account not associated with a person (P4).

#### **Tracking Progress**

Participants valued seeing their growth through tracking features, such as mood tracking or saved journal entries (P2). P2 noted that they felt they could not personally assess their emotional growth, and the app allowed them to reflect on their progress:

*When you don’t track your progress, you don’t realize how far you come. But when you can physically see it, it can either be a positive thing or a negative, but sometimes is usually helps you adjust your mindset. [P3]*

### **Quick Engagement With Content**

Some participants found it important to get value out of the apps quickly through short periods of engagement and lightweight interactions (P1 and P4). P4 mentioned that they valued audio content when they wanted to relax because they were able to get the support they needed to unwind quickly, or to complete an activity quickly, such as recognizing anxiety (P4). They did not like engaging with conversational apps because it took too much time to arrive at a helpful solution. They found it valuable when the apps were interactive; for example, conversational interfaces that felt like talking to a therapist (P4) or to another person (P3).

### **Ease of Use**

Participants valued simplicity in apps: “simple and easy to use...you just know where to go, almost like you’re walking in your house” (P2) reflected in simple layouts, esthetics, and functionality. Participants wanted content to be easy to navigate and for them to quickly know how to use the app, “you just know...where to find everything” (P2). Participants thought that too many features were overwhelming:

*There’s a lot of different features that pop up in the app...might feel overwhelmed. [P4]*

When apps contained too many features, it can make it difficult to find the desired content or even the content the user previously visited, which led to frustration:

*Everything was all over the place in that app. It was hard to find a topic that I had seen earlier...couldn’t find it anymore. [P5]*

The participants felt meditation apps did this particularly well:

*Especially meditation. And I feel like they’re simple enough...someone wouldn’t feel like they’re going out of their way. [P4]*

### **Match App Content With the User’s Context of Use**

The participants “were more inclined to actually use the app” if the content was relevant to the need they had at the moment. In 1 instance, a participant used their app to wind down because the app had categories relevant to their intended use, to wind down at the end of the day. Because the user’s interests aligned with the app, they were able to engage with the app more meaningfully:

*I was seeing an article about difficulty falling asleep sometimes. So, I was really interested in the article about sleep. [P4]*

### **Reliability and Accessibility of Content**

#### **Accessibility of Apps and of Free Content**

The participants identified certain aspects that made it hard for them to engage with the app. For example, 1 app included a community feature where people posted in foreign languages. P5 was concerned that some people might be expressing a need for help through their posts but that they might not be able to provide help because they do not understand the language.

The participants valued the availability of free content in the app and found frequent requests for upgrading to a paid version disruptive:

*It was really pushing...need to buy...the premium....every time I go on the app. [P2]*

The participants preferred using apps that were richer in free content:

*This app was definitely way better. I feel like there was more content that was free and available to public use. [P3 and P5]*

The trial version of apps also made participants reluctant about the benefits of the app after the trial period.

### **Reliability of Content**

Some participants found it important to know that the content of the apps was validated by professionals and that it was not biased (P1). They wanted to ensure that people could connect with professions if they wish to:

*I kind of feel like at that stage, it should bring up links to professionals, because we’re dealing with mental health problems here. [P1]*

### **Between-Session Activities**

Once we began the study, the number of student responses to between-session activities over Discord was disproportionately low, yielding only 2 user responses. During this time, we first posted messages on Discord using an account that presented as a bot sending updates. This did not receive responses, so we changed our updates to be posted by one of our undergraduate researchers. We also asked about Discord engagement during our focus group session and encouraged participants to respond to prompts. This prompted responses from another participant. However, the responses were consistent and overlapped with what we were learning in the focus groups. Therefore, after the second week, we discontinued the between-session activities because engagement was low, and we were not receiving additional information in the feedback sessions. We also thought it would be more beneficial to keep participant feedback confined to focus groups where it could be discussed more in depth.

### **Follow-up Survey**

In the follow-up survey, we identified 2 themes related to overall feedback on the apps. *Organization* of the app was identified as an important aspect across all app categories (CBT, mindfulness, and miscellaneous). The participants reported that they preferred apps with a simple interface such that they were “to the point,” organized, and easy to navigate. They disliked the presence of too many tabs, features, or steps to work through to get to the support they wanted. *Payment* was the second theme. In general, the participants reported that they disliked when most features were only available in a premium, paid version of the app, and they also disliked seeing promotions of paid content. This aligned with comments made throughout the sessions that participants thought that these types of apps should be free or indicated that they would be unlikely to download an app that required payment when free versions exist.

On average, the participants reported that the apps had a small positive impact on their mental well-being across all apps. No one indicated that the apps had a negative impact, although “no impact” was reported for 1 CBT app, 2 mindfulness apps, and 1 miscellaneous app. Miscellaneous apps were ranked as the most impactful, followed by CBT apps and then mindfulness apps.

When asked how often they used their app during the week they were assigned to use it, the participants were most likely to report using the app for “a few days of the week.” Participants were slightly more likely to continue using mindfulness apps after the study compared with CBT or miscellaneous apps. Of the 5 participants, 2 (40%) reported that they continued to use their CBT and miscellaneous apps after the week they were asked to use it, and 3 (60%) participants reported continuing to use their mindfulness app after they were asked to do so. The average star ratings assigned to the app categories by the participants were 3 for CBT apps and 3.6 for both mindfulness and miscellaneous apps out of a total of 5 stars.

## Discussion

### Principal Findings

In this study, we developed and refined a process for obtaining feedback on various MHapps. Our resultant series of focus groups allowed us to gain initial insights about Latinx youth’s preferences for and within our selected set of MHapps and their views on the benefits and use of the apps, albeit from a small sample of participants. As we gained experience with our procedure, we made a few changes; for example, removing the between-session activities and adding a final session to get impressions across all the MHapps. Despite refining the methods over the course of the 5 focus groups, we still identified some consistent preliminary themes across these sessions that provide some impressions of Latinx youth’s views on MHapps.

Latinx youth wanted apps that were youth centric in terms of content and functionality. However, youth-centric content meant not just making age-appropriate content in terms of language, visuals, or examples, but it also meant making the interactions brief and fun. Our participants also emphasized the importance of MHapps being able to transcend their digital components and facilitate meaningful offline interactions. Our follow-up survey also showed that ratings of MHapps might differ based on the types of questions asked. Although the CBT apps had lower star ratings than the other apps (3.0 vs 3.6), they were rated as being more impactful than the mindfulness apps. We discuss our findings in subsequent sections, both insights gained from what we learned from the participants and reflections on our methods.

The youth wanted MHapps that were reflective of youth problems, interests, and technology use. Many existing MHapps, even those that claim to be focused on the youth, often contain minimal surface-level tailoring, such as changing examples or esthetics [53]. Other approaches include using techniques such as gamification that might appeal to some youth. On the other hand, others might find that games trivialize mental health issues and would appreciate a more serious approach [33]. Indeed,

another study that used a workshop approach to ideate multiple ideas for well-being technologies with the youth identified interests in diverse engagement strategies and types of technologies [32]. Thus, it is worth noting that “youth” is not a homogenous population, and differences exist among different subpopulations (eg, specific age ranges, gender or gender identity, race, or ethnic background).

Our study focused on the subpopulation of Latinx youth. Although we cannot comment on which themes identified were specific to Latinx youth, as we did not have a comparison group, it is worth reflecting on some things noted. Our participants noted the importance of social connectedness through MHapps. The importance of social connection has been noted in other mobile health studies comparing Spanish and English speakers, finding that Spanish speakers emphasized feelings of social support, whereas English speakers emphasized introspection and self-awareness [54]. It is hard to separate how much of these findings are related to youth generally or Latinx youth specifically, and further work could replicate these methods across different subgroups to help lead to more specific and definitive design recommendations.

Youth also go through different developmental processes. Therefore, adopting a developmental science lens might introduce affordances for MHapps and other youth-focused technologies [55]. One proposed solution to better reflect youth preferences in MHapps aimed at youth is to include youth and various youth stakeholders in the design process [56]. Indeed, various co-design methods (eg, questionnaires, interviews, focus groups, interactive workshops, and meetings) might be helpful to solicit input from youth and can be selected based on the stage of development and type of information needed [34]. We used some co-design methods in this project to solicit information regarding the developed MHapps, but of course, these methods should precede fully developed products.

It is also worth reflecting on considerations for evaluation methods for MHapps. Several evaluation frameworks or systems exist, such as the American Psychiatric Association framework [57], One Mind PsyberGuide, and Framework to Assist Stakeholders in Technology Evaluation for Recovery to Mental Health and Wellness [58]. Although these frameworks are intended to assist decision-making by diverse stakeholders, including consumers, providers, advocacy organizations, payers, and health systems, they often require specialized knowledge. This results in evaluations being conducted by individuals with expertise to apply such evaluation frameworks. Therefore, although consumer input might have been solicited in the development of these frameworks [48,59], methods for incorporating bottom-up input from consumers, especially consumers from diverse backgrounds, are needed. Some measures are relevant for consumers, such as the user version of the Mobile App Rating Scale [60] or the System Usability Scale [61], but these are mostly questions, rather than processes, for soliciting feedback. Our approach had both processes and questions, and many of the questions we used for our follow-up survey were derived from these measures [48,60].

We also found discrepancies between the youth feedback depending on how questions were asked. For example, we found

differences between apps that were viewed as most liked (ie, mindfulness), using a 5-star rating scale, and those seen as more impactful (ie, CBT and miscellaneous). Given the vast amount of information available through app stores, many researchers have used ratings to understand the user perspectives of MHapps [36,37]. However, this information might obscure more nuanced perspectives on apps. The specific discrepancies identified in our study, for example, for mindfulness apps and CBT, might reflect how this content is delivered to youth. It is worth noting that we did not collect formal outcome measures regarding the impact of these apps; therefore, our findings are entirely based on the youth's subjective perspective. Research on traditional mental health services has found that youth preferences do not always align with better outcomes [62]. As such, it is possible that these apps might have meaningful impacts on wellness and clinical targets or that satisfaction and clinical outcomes are not aligned, but additional research would be needed to verify this. Consistent with our previous points about the importance of separating app discovery (identifying apps), activation (initiating use), and engagement (repeated and sustained use), it is possible that elements that contribute to satisfaction, such as esthetics and learnability, might lead to people starting with an MHapp but that clinical effectiveness (ie, benefiting) might lead to sustained use. Other work has identified that people do discuss different factors related to barriers and facilitators to use across early initiation and long-term use [63], and studies that further explore the contributions of satisfaction and benefits could be useful here. For example, people who experience early symptom change might be more likely to stop using an MHapp (ie, "happy abandonment") but might be more likely to return to it if symptoms reoccur or to recommend it to others struggling with similar issues.

### **Reflections on the Apps That the Youth Designed in the Last Session**

In our final focus group session, we had the participants design their own apps using a web-based collaboration platform. Participants received a series of prompts that covered different common features they identified (ie, symptom tracking, distraction tools, games, information, and links to resources). The participants consistently identified simplicity as a major design feature, but other common features included information, meditation or mindfulness tools, and links to other resources. This is extremely similar to popular features identified by unemployed individuals and essential workers for MHapps intended to support distress during the COVID-19 pandemic [64]. A few of our participants also noted distraction tools, which were also identified in the study [64], especially among unemployed workers. However, these interests are somewhat inconsistent with common features within MHapps, with the finding that the most common features both alone and in combination are journaling and mood tracking [65]. In fact, MHapps containing only journaling and mood tracking combined accounted for 16.5% of the reviewed 278 MHapps [65]. Many MHapps tend to be complex and multifeatured; however, our participants came up with suggestions that were simple and targeted. The IntelliCare app suite was designed to include simple, single-featured apps [66], and an analysis of users of this suite found that participants tended to use focused

subsets of the various apps available to them [67]. Especially for the youth, who tend to be tech-savvy and technologically engaged, if a given MHapp does not meet their needs or is overly complex, they might use other technologies, perhaps not specifically designed for mental health, to meet their needs. Efforts should be made to better understand how youth use MHapps as part of their digital ecologies, and what needs and opportunities are lacking from how they are currently using technology to support their mental health and well-being.

### **Lessons on Methodology to Engage With Youth**

#### ***Recruitment and Community Partnership***

We engaged with youth who had already participated in an educational community. Our research group had also engaged with the community before. These youth were receptive to participating in the research, which might be because of the norms of the organization and partnership with the research team. They were also insightful and engaged in providing feedback, which might not have been true for youth recruited through other means. Nevertheless, working with youth from this organization helped build trust and collaboration more quickly. It also provided opportunities to present our findings back to the leadership in the organization to consider how these lessons could shape other efforts to involve their youth in activities or consider opportunities for technology to support their wellness.

#### ***Familiarity of the Participant Group and Research Team***

The participants knew each other before the focus groups started. However, the research team did not and was not aware of any of their relationships. This can create imbalances in the group because the research team is an outsider to the group. For future work, we recommend considering a group that has less familiarity so that everyone, including the research team, gets to know each other in a balanced way. Alternatively, our research team could have spent more time getting to know the youth and building a rapport before shifting to data collection. Another approach would be switching to remote activities only after a synchronous focus group where participants met each other [68]. Other work engaging children and adolescents in participatory design [69,70], especially using remote activities made necessary by the COVID-19 pandemic, has similarly demonstrated the need to build rapport with children and adolescents both at the beginning and repeatedly throughout extended design activities as we conducted.

#### ***Engagement With the Youth and Data Collection***

We found it beneficial to conduct icebreaker activities as part of the focus group sessions because it got the participants started with sharing information about themselves with the rest of the participants and with the research team. We recommend keeping the icebreaker activities short so that they provide a warm-up to the activities but do not take a significant amount of time. One of the primary facilitators of the focus group was an undergraduate student (KC), which we found beneficial because this student was closer to the age of the participants than our other facilitators (EA, SMS, and MN) and could relate to them more in the discussion. We also found it useful to use informal language, including emojis, in communication with the youth.

Although we initially attempted to send updates from an account that simulated a bot, we found that youth had slightly more updates when one of our undergraduate team members (Ashley McKinnon) messaged the Discord channel. We attribute this to participants relating more to a human being than to a generic bot.

### *Using a Social Network for Data Collection*

We used the Discord channel to communicate with youth between sessions. This was beneficial because they were already using Discord for other activities, and the Discord channel we created for this study was part of the wider Discord of the community. However, the research team was not familiar with the norms of the community and how it was used by the group. Our interactions and questions did not result in responses from participants on the web. We recommend that in the future, researchers use a social platform that is familiar to the participants but is separate from their existing communities (eg, a separate Discord server, a separate Slack channel, or a private group on a social network). This allows researchers to define the norms of the community and participation. Similar to guidelines in prior work [69], researchers could design activities that encourage using the platform early on, or during focus groups, to familiarize participants with how they are expected to use the platform.

### **Limitations**

This work was limited by its formative nature; that is, we worked with a small sample of youth who were recruited through 1 specific community partnership. However, we would also like to note other limitations in this work that might be helpful to other researchers who conduct work in similar areas. As previously noted, the youth population is not homogenous. The participants we worked with were in the age range of 15 to 21 years, which captures mid- to late adolescence. Researchers should be mindful about what age ranges of youth populations they are most interested in working with and also appreciate that in this period of rapid development, a developmental science lens to understand the interplay of youth and technologies is necessary. We also worked with only youth who identified as Latinx, and our results may not be representative of the views of MHapps in other populations. We might also have gained different insights if we used different categories of MHapps. Our last category was a miscellaneous category, and we could have either replaced the categories we used (ie, CBT or mindfulness) or extended the number of focus group sessions to focus on different categories of apps, such as tracking or peer support apps. Not every aspect of our project went as planned, and we had to adjust methodologies along the way; for example, removing between-session activities. We expect that other projects likely change their initial plans and encourage other researchers to share these deviations as an opportunity to help others conducting similar work. Our team did not include members with a Latinx personal background. It is possible that participants would be less likely to connect and open up with our team members if they shared a Latinx personal background or that our analyses did not capture some elements of the Latinx experience because of the positionality of our research team. Finally, it is worth calling out some

specific things that we did not hear from the participants we worked with. For example, recent work has called out several aspects related to the digital divide that might be barriers to use among marginalized groups, such as cost and internet and device access [71]. It is hard to disentangle whether these concerns are not present among the youth generally, perhaps because they are the most tech-savvy, these youth specifically, perhaps because we recruited them from a nonprofit team focused on Science, Technology, Engineering, Arts, and Math, or related to our selection of apps, which focused on free products. Nevertheless, it is important to note that assumptions related to digital literacy and technology access are worth checking and rechecking as knowledge and access continue to evolve.

### **Future Directions and Work**

We identified a range of needs for personalization of MHapps for the needs of youth. The content of the app should have some amount of topical relevance to youth's interests and situations. Future work is needed to understand the topics that are most relevant to the youth and in what way they should be presented. For example, how specific content can be so that it is relatable to many but still feels meaningful to those for whom it is intended. The youth valued playful and lightweight interactions as part of the apps. More work is needed to understand how to implement such designs as part of existing MHapps; for example, MHapp designs could include lightweight audio content, activities designed in a way that feels less serious, or pairing activities that are perceived as demanding with activities that make the user feel good at the moment.

Building on this work, we see various potential avenues for future studies. One would be to conduct similar workshops with more youth, additionally drawing in youth from different racial and ethnic backgrounds or for different age ranges. We also see multiple ways in which this work could go deeper—individual interviews with the youth, refined between-session activities, and additional quantitative measures to look at the impact of MHapps both while using the apps (ie, process measures) and at the end of participation (ie, outcome measures). We believe that interviews and between-session activities could help better understand youth's early experiences with MHapps, or app activation as we described earlier, as well as continued and sustained use (ie, app engagement). Additional quantitative measures could help better understand the impact of MHapps on youth. Finally, although we noted that app discovery, activation, and engagement were all challenges, given that we provided a select list of MHapps for participants, we did not learn much about discovery beyond how they navigated the choices we gave them. Future work could also work with the youth who elect to use MHapps on their own to better understand why and how they use them. It is likely that not all mental health and wellness needs that the youth are attempting to address with technologies are being approached using MHapps. The youth might be using other technologies, such as social media and games, and further work could help better understand how the youth approach those technologies in addition to MHapps.

We conducted a co-design activity with the participants in the last focus group. Future work could use structured design

activities with the youth to understand specific preferences related to MHapps, that is, the types of functionalities, content, and visual look and feel of MHapps as well as that interactions align with their technology use practices. Our work provides some high-level insights—it should be youth centric, simple, and focused; translating these insights into design guidelines and app features would be a useful next step.

## Conclusions

Given the growing number of MHapps and youth's tendency to use these resources, understanding their perspectives on MHapps can help guide the development, evaluation, and implementation of these tools for the youth. Our study developed and used a multisession focus group design to introduce MHapps

to youth participants, elicit feedback on those apps, and learn more about what youth might want in MHapps. This work adds to a growing body of research focused on understanding youth's needs and interests with regard to mental health technologies while also attempting to identify affordances that such technologies might offer, especially through a developmental science lens. We highlighted emerging themes, such as the need for simple and tailored content and the intersection of MHapps with youth's offline lives. Overall, the participants we worked with expressed enthusiasm for this space, but more work needs to support building MHapps that are effective for the youth and determining how best to make these tools available to youth and integrate them into their lives.

## Acknowledgments

The authors thank Listo America and specifically Gustavo Madrigal for help with recruiting youth and the support of their members. Ashley McKinnon assisted with the study procedures. Funding for this study was supported by Pivotal Ventures and One Mind with funding supporting One Mind PsyberGuide.

## Conflicts of Interest

SMS is a member of Scientific Advisory Board for Headspace, for which he receives compensation, and has received consulting payments from Otsuka Pharmaceuticals and K Health (Trusst) for unrelated work.

## References

1. Institute Report. Digital Health Trends 2021: Innovation, Evidence, Regulation, and Adoption. IQVIA Institute for Human Data Sciences. 2021 Jul 22. URL: <https://www.iqvia.com/insights/the-iqvia-institute/reports/digital-health-trends-2021> [accessed 2022-06-03]
2. Wasil AR, Gillespie S, Shingleton R, Wilks CR, Weisz JR. Examining the reach of smartphone apps for depression and anxiety. *Am J Psychiatry* 2020 May 01;177(5):464-465. [doi: [10.1176/appi.ajp.2019.19090905](https://doi.org/10.1176/appi.ajp.2019.19090905)] [Medline: [32354266](https://pubmed.ncbi.nlm.nih.gov/32354266/)]
3. Baumel A, Muench F, Edan S, Kane JM. Objective user engagement with mental health apps: systematic search and panel-based usage analysis. *J Med Internet Res* 2019 Sep 25;21(9):e14567 [FREE Full text] [doi: [10.2196/14567](https://doi.org/10.2196/14567)] [Medline: [31573916](https://pubmed.ncbi.nlm.nih.gov/31573916/)]
4. Schueller SM, Stiles-Shields C, Yarosh L. Online treatment and virtual therapists in child and adolescent psychiatry. *Child Adolesc Psychiatr Clin N Am* 2017 Jan;26(1):1-12 [FREE Full text] [doi: [10.1016/j.chc.2016.07.011](https://doi.org/10.1016/j.chc.2016.07.011)] [Medline: [27837935](https://pubmed.ncbi.nlm.nih.gov/27837935/)]
5. Uhlhaas P, Torous J. Digital tools for youth mental health. *NPJ Digit Med* 2019 Oct 18;2:104 [FREE Full text] [doi: [10.1038/s41746-019-0181-2](https://doi.org/10.1038/s41746-019-0181-2)] [Medline: [31646184](https://pubmed.ncbi.nlm.nih.gov/31646184/)]
6. Baghaei N, Naslund JA, Hach S, Liang HN. Editorial: designing technologies for youth mental health. *Front Public Health* 2020 Feb 25;8:45 [FREE Full text] [doi: [10.3389/fpubh.2020.00045](https://doi.org/10.3389/fpubh.2020.00045)] [Medline: [32161745](https://pubmed.ncbi.nlm.nih.gov/32161745/)]
7. Kessler RC, Berglund P, Demler O, Jin R, Merikangas KR, Walters EE. Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Arch Gen Psychiatry* 2005 Jun;62(6):593-602. [doi: [10.1001/archpsyc.62.6.593](https://doi.org/10.1001/archpsyc.62.6.593)] [Medline: [15939837](https://pubmed.ncbi.nlm.nih.gov/15939837/)]
8. Owens PL, Hoagwood K, Horwitz SM, Leaf PJ, Poduska JM, Kellam SG, et al. Barriers to children's mental health services. *J Am Acad Child Adolesc Psychiatry* 2002 Jun;41(6):731-738. [doi: [10.1097/00004583-200206000-00013](https://doi.org/10.1097/00004583-200206000-00013)] [Medline: [12049448](https://pubmed.ncbi.nlm.nih.gov/12049448/)]
9. Cummings JR, Allen L, Clennon J, Ji X, Druss BG. Geographic access to specialty mental health care across high- and low-income US communities. *JAMA Psychiatry* 2017 May 01;74(5):476-484 [FREE Full text] [doi: [10.1001/jamapsychiatry.2017.0303](https://doi.org/10.1001/jamapsychiatry.2017.0303)] [Medline: [28384733](https://pubmed.ncbi.nlm.nih.gov/28384733/)]
10. Kataoka SH, Zhang L, Wells KB. Unmet need for mental health care among U.S. children: variation by ethnicity and insurance status. *Am J Psychiatry* 2002 Sep;159(9):1548-1555. [doi: [10.1176/appi.ajp.159.9.1548](https://doi.org/10.1176/appi.ajp.159.9.1548)] [Medline: [12202276](https://pubmed.ncbi.nlm.nih.gov/12202276/)]
11. Marrast L, Himmelstein DU, Woolhandler S. Racial and ethnic disparities in mental health care for children and young adults: a national study. *Int J Health Serv* 2016 Oct;46(4):810-824. [doi: [10.1177/0020731416662736](https://doi.org/10.1177/0020731416662736)] [Medline: [27520100](https://pubmed.ncbi.nlm.nih.gov/27520100/)]
12. Lokko HN, Chen JA, Parekh RI, Stern TA. Racial and ethnic diversity in the US psychiatric workforce: a perspective and recommendations. *Acad Psychiatry* 2016 Dec;40(6):898-904. [doi: [10.1007/s40596-016-0591-2](https://doi.org/10.1007/s40596-016-0591-2)] [Medline: [27421839](https://pubmed.ncbi.nlm.nih.gov/27421839/)]
13. McGuire TG, Miranda J. New evidence regarding racial and ethnic disparities in mental health: policy implications. *Health Aff (Millwood)* 2008;27(2):393-403 [FREE Full text] [doi: [10.1377/hlthaff.27.2.393](https://doi.org/10.1377/hlthaff.27.2.393)] [Medline: [18332495](https://pubmed.ncbi.nlm.nih.gov/18332495/)]

14. Rickwood DJ, Deane FP, Wilson CJ. When and how do young people seek professional help for mental health problems? *Med J Aust* 2007 Oct 01;187(S7):S35-S39. [doi: [10.5694/j.1326-5377.2007.tb01334.x](https://doi.org/10.5694/j.1326-5377.2007.tb01334.x)] [Medline: [17908023](https://pubmed.ncbi.nlm.nih.gov/17908023/)]
15. McMiller WP, Weisz JR. Help-seeking preceding mental health clinic intake among African-American, Latino, and Caucasian youths. *J Am Acad Child Adolesc Psychiatry* 1996 Aug;35(8):1086-1094. [doi: [10.1097/00004583-199608000-00020](https://doi.org/10.1097/00004583-199608000-00020)] [Medline: [8755806](https://pubmed.ncbi.nlm.nih.gov/8755806/)]
16. Rideout V, Robb MB. The Common Sense census: media use by teens and teens. Common Sense Media. 2019. URL: <https://www.commonsensemedia.org/sites/default/files/research/report/2019-census-8-to-18-full-report-updated.pdf> [accessed 2022-06-03]
17. Teens, social media and technology 2018. Pew Research Center. 2018 May 31. URL: [https://www.pewresearch.org/internet/wp-content/uploads/sites/9/2018/05/PI\\_2018.05.31\\_TeensTech\\_FINAL.pdf](https://www.pewresearch.org/internet/wp-content/uploads/sites/9/2018/05/PI_2018.05.31_TeensTech_FINAL.pdf) [accessed 2022-05-30]
18. Rideout V, Fox S, Peebles A, Robb MB. Coping with COVID-19: how young people use digital media to manage their mental health. Common Sense Media. 2021. URL: <https://amerispeak.norc.org/content/dam/amerispeak/research/pdf/Coping%20with%20COVID-19%20-%20How%20Young%20People%20Use%20Digital%20Media%20to%20Manage%20Their%20Mental%20Health.pdf> [accessed 2022-06-02]
19. Schueller SM, Hunter JF, Figueroa C, Aguilera A. Use of digital mental health for marginalized and underserved populations. *Curr Treat Options Psych* 2019 Jul 5;6(3):243-255. [doi: [10.1007/s40501-019-00181-z](https://doi.org/10.1007/s40501-019-00181-z)]
20. Ramos G, Chavira DA. Use of technology to provide mental health care for racial and ethnic minorities: evidence, promise, and challenges. *Cogn Behavi Pract* 2022 Feb;29(1):15-40. [doi: [10.1016/j.cbpra.2019.10.004](https://doi.org/10.1016/j.cbpra.2019.10.004)]
21. Coffman J, Bates T, Geyn I, Spetz J. California's current and future behavioral health workforce. Healthforce Center at UCSF. 2018 Feb 12. URL: <https://healthforce.ucsf.edu/sites/healthforce.ucsf.edu/files/publication-pdf/California%E2%80%99s%20Current%20and%20Future%20Behavioral%20Health%20Workforce.pdf> [accessed 2022-06-02]
22. Ku BS, Li J, Lally C, Compton MT, Druss BG. Associations between mental health shortage areas and county-level suicide rates among adults aged 25 and older in the USA, 2010 to 2018. *Gen Hosp Psychiatry* 2021;70:44-50 [FREE Full text] [doi: [10.1016/j.genhosppsy.2021.02.001](https://doi.org/10.1016/j.genhosppsy.2021.02.001)] [Medline: [33714795](https://pubmed.ncbi.nlm.nih.gov/33714795/)]
23. Alexandre PK, Martins SS, Richard P. Disparities in adequate mental health care for past-year major depressive episodes among Caucasian and Hispanic youths. *Psychiatr Serv* 2009 Oct;60(10):1365-1371. [doi: [10.1176/ps.2009.60.10.1365](https://doi.org/10.1176/ps.2009.60.10.1365)] [Medline: [19797377](https://pubmed.ncbi.nlm.nih.gov/19797377/)]
24. Linardon J, Cuijpers P, Carlbring P, Messer M, Fuller-Tyszkiewicz M. The efficacy of app-supported smartphone interventions for mental health problems: a meta-analysis of randomized controlled trials. *World Psychiatry* 2019 Oct;18(3):325-336 [FREE Full text] [doi: [10.1002/wps.20673](https://doi.org/10.1002/wps.20673)] [Medline: [31496095](https://pubmed.ncbi.nlm.nih.gov/31496095/)]
25. Lecomte T, Potvin S, Corbière M, Guay S, Samson C, Cloutier B, et al. Mobile apps for mental health issues: meta-review of meta-analyses. *JMIR Mhealth Uhealth* 2020 May 29;8(5):e17458 [FREE Full text] [doi: [10.2196/17458](https://doi.org/10.2196/17458)] [Medline: [32348289](https://pubmed.ncbi.nlm.nih.gov/32348289/)]
26. Hollis C, Falconer CJ, Martin JL, Whittington C, Stockton S, Glazebrook C, et al. Annual Research Review: digital health interventions for children and young people with mental health problems - a systematic and meta-review. *J Child Psychol Psychiatry* 2017 Apr;58(4):474-503. [doi: [10.1111/jcpp.12663](https://doi.org/10.1111/jcpp.12663)] [Medline: [27943285](https://pubmed.ncbi.nlm.nih.gov/27943285/)]
27. Grist R, Porter J, Stallard P. Mental health mobile apps for preadolescents and adolescents: a systematic review. *J Med Internet Res* 2017 May 25;19(5):e176 [FREE Full text] [doi: [10.2196/jmir.7332](https://doi.org/10.2196/jmir.7332)] [Medline: [28546138](https://pubmed.ncbi.nlm.nih.gov/28546138/)]
28. Punukollu M, Marques M. Use of mobile apps and technologies in child and adolescent mental health: a systematic review. *Evid Based Ment Health* 2019 Nov;22(4):161-166. [doi: [10.1136/ebmental-2019-300093](https://doi.org/10.1136/ebmental-2019-300093)] [Medline: [31358537](https://pubmed.ncbi.nlm.nih.gov/31358537/)]
29. Bergin AD, Vallejos EP, Davies EB, Daley D, Ford T, Harold G, et al. Preventive digital mental health interventions for children and young people: a review of the design and reporting of research. *NPJ Digit Med* 2020 Oct 15;3:133 [FREE Full text] [doi: [10.1038/s41746-020-00339-7](https://doi.org/10.1038/s41746-020-00339-7)] [Medline: [33083568](https://pubmed.ncbi.nlm.nih.gov/33083568/)]
30. Stiles-Shields C, Cummings C, Montague E, Plevinsky JM, Psihogios AM, Williams KD. A call to action: using and extending human-centered design methodologies to improve mental and behavioral health equity. *Front Digit Health* 2022 Apr 25;4:848052 [FREE Full text] [doi: [10.3389/fdgth.2022.848052](https://doi.org/10.3389/fdgth.2022.848052)] [Medline: [35547091](https://pubmed.ncbi.nlm.nih.gov/35547091/)]
31. Michel T, Tachtler F, Slovak P, Fitzpatrick G. A review of youth mental health promotion apps towards their fit with youth media preferences. *EAI Endorsed Trans Pervasive Health Technol* 2019 Feb 12;19(17):e2 [FREE Full text] [doi: [10.4108/eai.13-7-2018.161419](https://doi.org/10.4108/eai.13-7-2018.161419)]
32. Yarosh S, Schueller SM. "Happiness inventors": informing positive computing technologies through participatory design with children. *J Med Internet Res* 2017 Jan 17;19(1):e14 [FREE Full text] [doi: [10.2196/jmir.6822](https://doi.org/10.2196/jmir.6822)] [Medline: [28096066](https://pubmed.ncbi.nlm.nih.gov/28096066/)]
33. Thabrew H, Fleming T, Hetrick S, Merry S. Co-design of eHealth interventions with children and young people. *Front Psychiatry* 2018 Oct 18;9:481 [FREE Full text] [doi: [10.3389/fpsy.2018.00481](https://doi.org/10.3389/fpsy.2018.00481)] [Medline: [30405450](https://pubmed.ncbi.nlm.nih.gov/30405450/)]
34. Bevan Jones R, Stallard P, Agha SS, Rice S, Werner-Seidler A, Stasiak K, et al. Practitioner review: co-design of digital mental health technologies with children and young people. *J Child Psychol Psychiatry* 2020 Aug;61(8):928-940 [FREE Full text] [doi: [10.1111/jcpp.13258](https://doi.org/10.1111/jcpp.13258)] [Medline: [32572961](https://pubmed.ncbi.nlm.nih.gov/32572961/)]
35. Stoyanov SR, Zelenko O, Staneva A, Kavanagh DJ, Smith C, Sade G, et al. Development of the niggle app for supporting young people on their dynamic journey to well-being: co-design and qualitative research study. *JMIR Mhealth Uhealth* 2021 Apr 20;9(4):e21085 [FREE Full text] [doi: [10.2196/21085](https://doi.org/10.2196/21085)] [Medline: [33877050](https://pubmed.ncbi.nlm.nih.gov/33877050/)]

36. Alqahtani F, Orji R. Insights from user reviews to improve mental health apps. *Health Informatics J* 2020 Sep;26(3):2042-2066 [FREE Full text] [doi: [10.1177/1460458219896492](https://doi.org/10.1177/1460458219896492)] [Medline: [31920160](https://pubmed.ncbi.nlm.nih.gov/31920160/)]
37. Oyebo O, Alqahtani F, Orji R. Using machine learning and thematic analysis methods to evaluate mental health apps based on user reviews. *IEEE Access* 2020 Jun 12;8:111141-111158. [doi: [10.1109/ACCESS.2020.3002176](https://doi.org/10.1109/ACCESS.2020.3002176)]
38. MacLeod H, Jelen B, Prabhakar A, Oehlberg L, Siek K, Connelly K. A guide to using asynchronous remote communities (ARC) for researching distributed populations. *EAI Endorsed Trans Pervasive Health Technol* 2017 Jul 18;17(11):e4. [doi: [10.4108/eai.18-7-2017.152898](https://doi.org/10.4108/eai.18-7-2017.152898)]
39. MacLeod H, Jelen B, Prabhakar A, Oehlberg L, Siek K, Connelly K. Asynchronous remote communities (ARC) for researching distributed populations. In: *Proceedings of the 10th EAI International Conference on Pervasive Computing Technologies for Healthcare*. 2016 May Presented at: PervasiveHealth '16; May 16-19, 2016; Cancun, Mexico p. 1-8. [doi: [10.4108/eai.16-5-2016.2263322](https://doi.org/10.4108/eai.16-5-2016.2263322)]
40. Maestre JF, MacLeod H, Connelly CL, Dunbar JC, Beck J, Siek KA, et al. Defining through expansion: conducting asynchronous remote communities (arc) research with stigmatized groups. In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 2018 Presented at: CHI '18; April 21-26, 2018; Montreal, Canada p. 1-13. [doi: [10.1145/3173574.3174131](https://doi.org/10.1145/3173574.3174131)]
41. Bhattacharya A, Liang C, Zeng EY, Shukla K, Wong ME, Munson SA, et al. Engaging teenagers in asynchronous online groups to design for stress management. In: *Proceedings of the 18th ACM International Conference on Interaction Design and Children*. 2019 Jun Presented at: IDC '19; June 12-15, 2019; Boise, ID, USA p. 26-37. [doi: [10.1145/3311927.3323140](https://doi.org/10.1145/3311927.3323140)]
42. Ito M, Odgers C, Schueller S, Cabrera J, Conaway E, Cross R, et al. *Social Media and Youth Wellbeing: What We Know and Where We Could Go*. Connected Learning Alliance. 2020 Jun. URL: <https://clalliance.org/wp-content/uploads/2020/06/Social-Media-and-Youth-Wellbeing-Report.pdf> [accessed 2022-06-02]
43. Poole ES, Peyton T. Interaction design research with adolescents: methodological challenges and best practices. In: *Proceedings of the 12th International Conference on Interaction Design and Children*. 2013 Presented at: IDC '13; June 24-27, 2013; New York, NY, USA p. 211-217. [doi: [10.1145/2485760.2485766](https://doi.org/10.1145/2485760.2485766)]
44. Dow S, Fortuna J, Schwartz D, Altringer B, Schwartz DL, Klemmer SR. Prototyping dynamics: sharing multiple designs improves exploration, group rapport, and results. In: *Proceedings of the 2011 SIGCHI Conference on Human Factors in Computing Systems*. 2011 Presented at: CHI '11; May 7-12, 2011; Vancouver, Canada p. 2807-2816. [doi: [10.1145/1978942.1979359](https://doi.org/10.1145/1978942.1979359)]
45. Haldar S, Mishra SR, Khelifi M, Pollack AH, Pratt W. Beyond the patient portal: supporting needs of hospitalized patients. In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 2019 Presented at: CHI '19; May 4-9, 2019; Glasgow, UK p. 1-14. [doi: [10.1145/3290605.3300596](https://doi.org/10.1145/3290605.3300596)]
46. The Visual Collaboration Platform for Every Team. miro. URL: <https://miro.com/> [accessed 2022-06-02]
47. Discord. URL: <https://discord.com/> [accessed 2022-06-02]
48. Hudson G, Negbenose E, Neary M, Jansli SM, Schueller SM, Wykes T, et al. Comparing professional and consumer ratings of mental health apps: mixed methods study. *JMIR Form Res* 2022 Sep 23;6(9):e39813 [FREE Full text] [doi: [10.2196/39813](https://doi.org/10.2196/39813)] [Medline: [36149733](https://pubmed.ncbi.nlm.nih.gov/36149733/)]
49. 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/)]
50. Charmaz K. *Constructing Grounded Theory*. 2nd edition. Thousand Oaks, CA, USA: Sage Publications; 2006.
51. Saldaña J. *The Coding Manual for Qualitative Researchers*. 4th edition. Thousand Oaks, CA, USA: Sage Publications; 2021.
52. McDonald N, Schoenebeck S, Forte A. Reliability and inter-rater reliability in qualitative research: norms and guidelines for CSCW and HCI practice. *Proc ACM Hum Comput Interact* 2019 Nov 07;3(CSCW):1-23. [doi: [10.1145/3359174](https://doi.org/10.1145/3359174)]
53. Odgers CL, Schueller SM, Ito M. Screen time, social media use, and adolescent development. *Annu Rev Dev Psychol* 2020 Dec 15;2(1):485-502. [doi: [10.1146/annurev-devpsych-121318-084815](https://doi.org/10.1146/annurev-devpsych-121318-084815)]
54. Aguilera A, Berridge C. Qualitative feedback from a text messaging intervention for depression: benefits, drawbacks, and cultural differences. *JMIR Mhealth Uhealth* 2014 Nov 05;2(4):e46 [FREE Full text] [doi: [10.2196/mhealth.3660](https://doi.org/10.2196/mhealth.3660)] [Medline: [25373390](https://pubmed.ncbi.nlm.nih.gov/25373390/)]
55. Moreno MA, Uhls YT. Applying an affordances approach and a developmental lens to approach adolescent social media use. *Digit Health* 2019 Jan 31;5:2055207619826678 [FREE Full text] [doi: [10.1177/2055207619826678](https://doi.org/10.1177/2055207619826678)] [Medline: [30746154](https://pubmed.ncbi.nlm.nih.gov/30746154/)]
56. Psihogios AM, Lane-Fall MB, Graham AK. Adolescents are still waiting on a digital health revolution: accelerating research-to-practice translation through design for implementation. *JAMA Pediatr* 2022 Jun 01;176(6):545-546. [doi: [10.1001/jamapediatrics.2022.0500](https://doi.org/10.1001/jamapediatrics.2022.0500)] [Medline: [35404381](https://pubmed.ncbi.nlm.nih.gov/35404381/)]
57. Torous JB, Chan SR, Gipson SY, Kim JW, Nguyen TQ, Luo J, et al. A hierarchical framework for evaluation and informed decision making regarding smartphone apps for clinical care. *Psychiatr Serv* 2018 May 01;69(5):498-500. [doi: [10.1176/appi.ps.201700423](https://doi.org/10.1176/appi.ps.201700423)] [Medline: [29446337](https://pubmed.ncbi.nlm.nih.gov/29446337/)]



58. Agarwal S, Jalan M, Wilcox HC, Sharma R, Hill R, Pantalone E, et al. Evaluation of Mental Health Mobile Applications. Technical Brief No. 41. Rockville, MD, USA: Agency for Healthcare Research and Quality; 2022.
59. Lagan S, Emerson MR, King D, Matwin S, Chan SR, Proctor S, et al. Mental health app evaluation: updating the American Psychiatric Association's framework through a stakeholder-engaged workshop. *Psychiatr Serv* 2021 Sep 01;72(9):1095-1098. [doi: [10.1176/appi.ps.202000663](https://doi.org/10.1176/appi.ps.202000663)] [Medline: [33882716](https://pubmed.ncbi.nlm.nih.gov/33882716/)]
60. Stoyanov SR, Hides L, Kavanagh DJ, Wilson H. Development and validation of the user version of the Mobile Application Rating Scale (uMARS). *JMIR Mhealth Uhealth* 2016 Jun 10;4(2):e72 [FREE Full text] [doi: [10.2196/mhealth.5849](https://doi.org/10.2196/mhealth.5849)] [Medline: [27287964](https://pubmed.ncbi.nlm.nih.gov/27287964/)]
61. Lewis JR. The system usability scale: past, present, and future. *Int J Human Comput Interact* 2018 Mar 30;34(7):577-590. [doi: [10.1080/10447318.2018.1455307](https://doi.org/10.1080/10447318.2018.1455307)]
62. Lau AS, Kim JJ, Nguyen DJ, Nguyen HT, Kodish T, Weiss B. Effects of preference on outcomes of preventive interventions among ethnically diverse adolescents at-risk of depression. *J Clin Child Adolesc Psychol* 2020;49(6):820-836 [FREE Full text] [doi: [10.1080/15374416.2019.1639514](https://doi.org/10.1080/15374416.2019.1639514)] [Medline: [31407937](https://pubmed.ncbi.nlm.nih.gov/31407937/)]
63. Balaskas A, Schueller SM, Cox AL, Doherty G. Understanding users' perspectives on mobile apps for anxiety management. *Front Digit Health* 2022 Sep 1;4:854263 [FREE Full text] [doi: [10.3389/fdgth.2022.854263](https://doi.org/10.3389/fdgth.2022.854263)] [Medline: [36120712](https://pubmed.ncbi.nlm.nih.gov/36120712/)]
64. Mata-Greve F, Johnson M, Pullmann MD, Friedman EC, Griffith Fillipo I, Comtois KA, et al. Mental health and the perceived usability of digital mental health tools among essential workers and people unemployed due to COVID-19: cross-sectional survey study. *JMIR Ment Health* 2021 Aug 05;8(8):e28360 [FREE Full text] [doi: [10.2196/28360](https://doi.org/10.2196/28360)] [Medline: [34081592](https://pubmed.ncbi.nlm.nih.gov/34081592/)]
65. Lagan S, D'Mello R, Vaidyam A, Bilden R, Torous J. Assessing mental health apps marketplaces with objective metrics from 29,190 data points from 278 apps. *Acta Psychiatr Scand* 2021 Aug;144(2):201-210. [doi: [10.1111/acps.13306](https://doi.org/10.1111/acps.13306)] [Medline: [33835483](https://pubmed.ncbi.nlm.nih.gov/33835483/)]
66. Lattie EG, Schueller SM, Sargent E, Stiles-Shields C, Tomasino KN, Corden ME, et al. Uptake and usage of IntelliCare: a publicly available suite of mental health and well-being apps. *Internet Interv* 2016 May;4(2):152-158 [FREE Full text] [doi: [10.1016/j.invent.2016.06.003](https://doi.org/10.1016/j.invent.2016.06.003)] [Medline: [27398319](https://pubmed.ncbi.nlm.nih.gov/27398319/)]
67. Kwasny MJ, Schueller SM, Lattie E, Gray EL, Mohr DC. Exploring the use of multiple mental health apps within a platform: secondary analysis of the IntelliCare field trial. *JMIR Ment Health* 2019 Mar 21;6(3):e11572 [FREE Full text] [doi: [10.2196/11572](https://doi.org/10.2196/11572)] [Medline: [30896433](https://pubmed.ncbi.nlm.nih.gov/30896433/)]
68. Liang CA, Albertson K, Williams F, Inwards-Breland D, Munson SA, Kientz JA, et al. Designing an online sex education resource for gender-diverse youth. In: Proceedings of the Interaction Design and Children Conference. 2020 Presented at: IDC '20; June 21-24, 2020; London, UK p. 108-120. [doi: [10.1145/3392063.3394404](https://doi.org/10.1145/3392063.3394404)]
69. Fails JA, Ratakonda DK, Koren N, Elsayed-Ali S, Bonsignore E, Yip J. Pushing boundaries of co-design by going online: lessons learned and reflections from three perspectives. *Int J Child Comput Interact* 2022 Sep;33:100476 [FREE Full text] [doi: [10.1016/j.ijcci.2022.100476](https://doi.org/10.1016/j.ijcci.2022.100476)] [Medline: [36034969](https://pubmed.ncbi.nlm.nih.gov/36034969/)]
70. Lee KJ, Roldan W, Zhu TQ, Saluja HK, Na S, Chin B, et al. The show must go on: a conceptual model of conducting synchronous participatory design with children online. In: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. 2021 May Presented at: CHI '21; May 8-13, 2021; Yokohama, Japan p. 1-16. [doi: [10.1145/3411764.3445715](https://doi.org/10.1145/3411764.3445715)]
71. Ramos G, Ponting C, Labao JP, Sobowale K. Considerations of diversity, equity, and inclusion in mental health apps: a scoping review of evaluation frameworks. *Behav Res Ther* 2021 Dec;147:103990 [FREE Full text] [doi: [10.1016/j.brat.2021.103990](https://doi.org/10.1016/j.brat.2021.103990)] [Medline: [34715396](https://pubmed.ncbi.nlm.nih.gov/34715396/)]

## Abbreviations

**ARC:** asynchronous remote community

**CBT:** cognitive behavioral therapy

**MHapp:** mental health app

*Edited by A Mavragani; submitted 02.07.22; peer-reviewed by KT Win, G Ramos; comments to author 24.07.22; revised version received 12.09.22; accepted 13.09.22; published 18.10.22.*

*Please cite as:*

Agapie E, Chang K, Patrachari S, Neary M, Schueller SM

Understanding Mental Health Apps for Youth: Focus Group Study With Latinx Youth

*JMIR Form Res* 2022;6(10):e40726

URL: <https://formative.jmir.org/2022/10/e40726>

doi: [10.2196/40726](https://doi.org/10.2196/40726)

PMID: [36256835](https://pubmed.ncbi.nlm.nih.gov/36256835/)

©Elena Agapie, Katherine Chang, Sneha Patrachari, Martha Neary, Stephen M Schueller. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 18.10.2022. 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

# Long-term Follow-up of Patients With Hernia Using the Hernia-Specific Quality-of-Life Mobile App: Feasibility Questionnaire Study

Ching-Shui Huang<sup>1,2</sup>, MD; Feng-Chuan Tai<sup>1</sup>, MD; Heng-Hui Lien<sup>1,3</sup>, MD; Jia-Uei Wong<sup>1</sup>, MD; Chi-Cheng Huang<sup>4,5,6</sup>, MD, PhD

<sup>1</sup>Division of General Surgery, Department of Surgery, Cathay General Hospital, Taipei, Taiwan

<sup>2</sup>School of Medicine, College of Medicine, Taipei Medical University, Taipei, Taiwan

<sup>3</sup>School of Medicine, College of Medicine, Fu-Jen Catholic University, New Taipei, Taiwan

<sup>4</sup>Department of Surgery, Taipei Veterans General Hospital, Taipei, Taiwan

<sup>5</sup>Comprehensive Breast Health Center, Taipei Veterans General Hospital, Taipei, Taiwan

<sup>6</sup>Institute of Epidemiology and Preventive Medicine, National Taiwan University, Taipei, Taiwan

**Corresponding Author:**

Chi-Cheng Huang, MD, PhD

Department of Surgery

Taipei Veterans General Hospital

No 201, Sec 2, Shipai Rd

Beitou District

Taipei, 106

Taiwan

Phone: 886 2 28757808

Email: [chishenh74@gmail.com](mailto:chishenh74@gmail.com)

## Abstract

**Background:** Hernia repair is one of the most common surgical procedures; however, the long-term outcomes are seldom reported due to incomplete follow-up.

**Objective:** The aim of this study was to examine the use of a mobile app for the long-term follow-up of hernia recurrence, complication, and quality-of-life perception.

**Methods:** A cloud-based corroborative system drove a mobile app with the HERQL (Hernia-Specific Quality-of-Life) questionnaire built in. Patients who underwent hernia repair were identified from medical records, and an invitation to participate in this study was sent through the post.

**Results:** The response rate was 11.89% (311/2615) during the 1-year study period, whereas the recurrence rate was 1.0% (3/311). Causal relationships between symptomatic and functional domains of the HERQL questionnaire were indicated by satisfactory model fit indices and significant regression coefficients derived from structural equation modeling. Regarding patients' last hernia surgeries, 88.7% (276/311) of the patients reported them to be satisfactory or very satisfactory, 68.5% (213/311) of patients reported no discomfort, and 61.1% (190/311) of patients never experienced mesh foreign body sensation. Subgroup analysis for the most commonly used mesh repairs found that mesh plug repair inevitably resulted in worse symptoms and quality-of-life perception from the group with groin hernias.

**Conclusions:** The mobile app has the potential to enhance the quality of care for patients with hernia and facilitate outcomes research with more complete follow-up.

(*JMIR Form Res* 2022;6(10):e39759) doi:[10.2196/39759](https://doi.org/10.2196/39759)

**KEYWORDS**

hernia; mobile app; quality of life; Hernia-Specific Quality-of-Life (HERQL); mobile health; mHealth; app; self-management

## Introduction

Hernia repair is one of the most common surgical procedures performed each day worldwide, which can be traced back to the era of ancient Egypt [1]. With the development of prosthetic mesh and tension-free techniques, the recurrence rate following hernia repair has been reduced drastically [2-8]. For abdominal wall (ie, ventral or incisional) hernias, the recurrence rate can be reduced from 50% of primary repair to 10% to 23% with a prosthetic mesh [9]. On the other hand, the open anterior approach with mesh repair has replaced the Shouldice procedure as the standard operation with the benefits of shorter hospital stays, lower recurrences, and decreased postoperative pain; the recurrence rate was even lower (<1%) for groin hernia repair [10].

With reduced herniorrhaphy failures, outcomes research of hernia surgery should concentrate on postoperative quality of life and complications such as chronic pain [11,12]. The long-term outcomes of herniorrhaphy, however, have not been thoroughly evaluated. Loss of follow-up and poor compliance from patients result in a biased evaluation of recurrence, complication, and patients' subjective quality of life [13,14].

To understand treatment outcomes, several quality-of-life instruments specific to hernia disease, such as the Carolinas Comfort Scale (CCS), the Hernia-Related Quality-of-Life Survey, the European Abdominal Wall Hernia Quality-of-Life Scores, the Core Outcome Measures Index adapted for patients with hernia, the Inguinal Pain Questionnaire, and the Brief Pain Inventory, have been developed and reported [15-20]. In the past few years, we have developed and validated an instrument—the Hernia-Specific Quality-of-Life (HERQL) questionnaire—for both groin and abdominal wall hernias. The questionnaire comprises a 4-item summative pain score measuring pain and discomfort resulting from various strenuous activities. Both symptomatic and functional domains, as well as postoperative satisfaction, are assessed with additional evaluations of hernia-related complications [21,22]. The validation study was conducted among 183 Taiwanese patients with groin hernia and 386 assessments; the internal reliability of the multi-item summative pain score was satisfactory (Cronbach  $\alpha=.85$ ). Criterion validity was evidenced by substantial to moderate correlations of the HERQL questionnaire with the five-level EQ-5D in pain/discomfort and health impact subscales [23]. Clinical validity was ascertained from worse hernia protrusion, pain during mild to heavy exercise, activity restriction, and health impairment scores reported from preoperative compared to postoperative patients. Clinical responsiveness was indicated by the time effect of  $-1.63$  in the summative pain score from repeated measures [21].

The HERQL questionnaire targets both abdominal wall and inguinal hernias, traditional open and minimally invasive surgeries, and various mesh materials [21,22]. One merit of using the HERQL questionnaire for hernia outcomes research is the determination of the causal relationship between formative symptomatic scales and reflective functional indicators, which is elaborated through the pathway analysis of structural equation modeling (SEM) [24-26].

As previously mentioned, low compliance and high loss-of-follow-up rates among patients with hernia heavily compromised outcome evaluation of herniorrhaphy, especially when long-term outcomes were pursued. Indeed, there remains an unmet need to understand the true recurrence rate of hernia surgery, as well as the associated complications and subjective well-being [27]. To overcome these limitations, we purposed a novel mobile app to enhance the follow-up and outcomes research of patients with hernia. Indeed, mobile devices have been advocated as an effective tool for administration of screening tests among populations, from school-age children to adults; there is comprehensive evidence regarding the effectiveness of smartphone-based mobile apps for follow-up among surgical patients, especially for those who have undergone hernia repair [28-30].

## Methods

### Study Design and Subjects

We invited patients who completed hernia repair at our hospital to participate in this study. Index cases were identified from medical records. Both groin and abdominal hernias were eligible, with the latter comprising primary ventral and incisional hernias. The latest hernia surgery should have been performed at least 1 year prior to the starting date of the study. The enrollment period was between April 1, 2016, and March 31, 2017. Identified cases were contacted through the post using addresses from the medical charts.

A preset combination of a unique ID and password was sent to each invitee concurrently, and electronically signed informed consent was obtained with the mobile app through the built-in signature module (see Mobile App section). A copy of the informed consent document was sent to the email address provided by each invitee for reference.

### The HERQL Questionnaire

The HERQL questionnaire has been described elsewhere [21]. In brief, the HERQL questionnaire comprises a 4-item summative pain score measuring pain and discomfort resulting from various strenuous activities (ie, rest or mild, moderate, or heavy activities). In the meantime, symptomatic burden and functional domains, as well as postoperative satisfaction and potential complications, were assessed concurrently.

Pain and activity restriction due to pain or discomfort were rated on an 11-point Likert-type scale, ranging from 0 to 10, for each item, while symptomatic and functional domains (ie, hernia protrusion, analgesic usage, hernia's impact on health, economic burden, and subjective quality of life/global health) were evaluated using a 5-point Likert-type scale. An auxiliary postoperative module, also equipped with 5-point Likert-type scales, was designed for potential complications following hernia repairs; these items included mesh foreign body sensation, severity of complications, overall satisfaction with hernia repair, confidence that hernia will not recur, and quality-of-life improvement by hernia repair. All scales were arranged with higher values representing compromised functionality or worse symptoms. The causal and indicator variables model proposed by Fayers et al [24,25] and Boehmer et al [26] formed the basis

of HERQL questionnaire structure [21,22]. Elaboration on causal-indicator duality recognized one-way causal effects of symptomatic scales on functional domains, but not vice versa

[16]. The content of the HERQL instrument is displayed in [Textbox 1](#).

**Textbox 1.** Content of the Hernia-Specific Quality-of-Life instrument.

<p><b>Summative pain score measures pain and discomfort resulting from rest or mild, moderate, or heavy activities:</b></p> <ul style="list-style-type: none"><li>• 11-point Likert-type scale</li><li>• Q01, Q03, Q04, and Q05 (Q: question)</li></ul> <p><b>Activity restriction due to pain or discomfort:</b></p> <ul style="list-style-type: none"><li>• 11-point Likert-type scale</li><li>• Q09</li></ul> <p><b>Symptomatic domains:</b></p> <ul style="list-style-type: none"><li>• 5-point Likert-type scale</li><li>• Hernia protrusion: Q02</li><li>• Analgesic use: Q08</li></ul> <p><b>Functional domains:</b></p> <ul style="list-style-type: none"><li>• 5-point Likert-type scale</li><li>• Hernia's impact on health: Q11</li><li>• Economic burden: Q12</li><li>• Subjective quality-of-life/global health perception: Q13</li></ul> <p><b>Postoperative module:</b></p> <ul style="list-style-type: none"><li>• 5-point Likert-type scale</li><li>• Mesh foreign body sensation: Q15</li><li>• Severity of complications: Q17</li><li>• Overall satisfaction with hernia repair: Q18</li><li>• Confidence that hernia will never recur: Q19</li><li>• Quality-of-life improvement by hernia repair: Q20</li></ul>
--

## Mobile App

The mobile app version of the HERQL questionnaire assessing patients' quality of life was ready for log-on for index cases identified from medical records, and an invitation was sent by post to those who had undergone hernia repairs at our institute at least 1 year before the study began. Both Android and iOS platforms were supported; a URL that linked to an online Google Docs-based questionnaire was provided as an alternative for those not equipped with a smartphone but who had internet access [31]. [Figure 1](#) shows the QR code for the HERQL questionnaire mobile app, and [Figure S1 in Multimedia Appendix 1](#) shows screenshots from an iOS-based device capturing all steps from log-on to the end of the survey. A corresponding Google website was established for

communication and educational purposes ([Figure S2 in Multimedia Appendix 1 \[32\]](#)). The mobile app system was developed in cooperation with SynerFUN Technology Corporation, based in Hsinchu City, Taiwan.

The mobile app was driven by a cloud-based corroborative system. The system comprised a data management and storage subunit, as well as a security information subunit. Patients with hernia could administer HERQL, to report their quality-of-life and outcomes following hernia repair. The platform provided an easy and efficient way for patients to report any discomfort to their surgeons, which was designed to enhance the long-term follow-up and compliance of patients with hernia. A built-in signature module was developed to facilitate acquisition of electronically signed informed consent ([Figure S3 in Multimedia Appendix 1](#)).

**Figure 1.** QR code for the HERQL questionnaire mobile app. HERQL: Hernia-Specific Quality-of-Life.

## Statistical Methods

Subgroup comparisons were conducted between the most commonly used mesh materials from the group with groin hernias. Quality-of-life scores were treated as continuous variables; the Student *t* test was used for between-group comparisons. P values less than .05 were considered statistically significant.

The SEM concept was evaluated with the following model fit indices. Goodness of fit was evaluated by the ratio of chi-square to the degrees of freedom, and a ratio of less than 3 indicated a good fit of the hypothesized construct to the experimental data. Additional fit indices included the goodness-of-fit index (GFI; >0.90), the adjusted GFI (>0.80), the standardized root mean square residual (<0.1), the comparative fit index (>0.9), and the root mean square error of approximation (RMSEA; <0.08). All variances for latent factors were determined to be the ones used for model identification purposes.

## Ethics Approval

Human subject research ethics review and approval was in accordance with the regulation of the Institutional Review Board of Cathay General Hospital (protocol number: CGH-P102069).

Written informed consent using electronic signatures was obtained from all participants, and all study subjects were compensated by post with a remuneration of 200 New Taiwan Dollars (approximately US \$7 in 2016). Analyses were conducted after all data were anonymized for privacy and confidentiality protection.

## Results

### Study Population

During the 1-year study period, 2615 patients who had their hernia repaired at our institute were invited to participate in the study via the post. Among them, 2245 (85.85%) were male and 370 (14.15%) were female. The mean age was 60 (SD 15) years (median 62; range 18-95 years). The response rate was 11.89%: 311 patients followed the instructions, with successful log-on, and completed the HERQL survey. There were 93 (29.9%) abdominal wall (ie, incisional and ventral) hernias, 202 (65.0%) groin hernias, and 16 (5.14%) patients had both. The earliest hernia repair took place more than 13 years ago (mean 5.5, SD 2.7 years; median 5.4, range 1-13.6 years). Most responders were within 5 years of hernia repairs. [Table 1](#) shows the types of prosthetic mesh adopted during herniorrhaphy.

**Table 1.** Types of prosthetic mesh used for hernia repairs.

Mesh type	Groin hernia (n=218 <sup>a</sup> ), n (%)	Abdominal wall hernia (n=109 <sup>a</sup> ), n (%)
Composix or Ventrio	0 (0)	38 (34.9)
Kugel or modified Kugel	60 (27.5)	11 (10.1)
Prolene Hernia System or Ultrapro Hernia System	57 (26.1)	11 (10.1)
Parietex	2 (0.9)	0 (0)
Mesh plug	77 (35.3)	9 (8.3)
Laparoscopy	13 (6.0)	2 (1.8)
Others	9 (4.1)	38 (34.9)

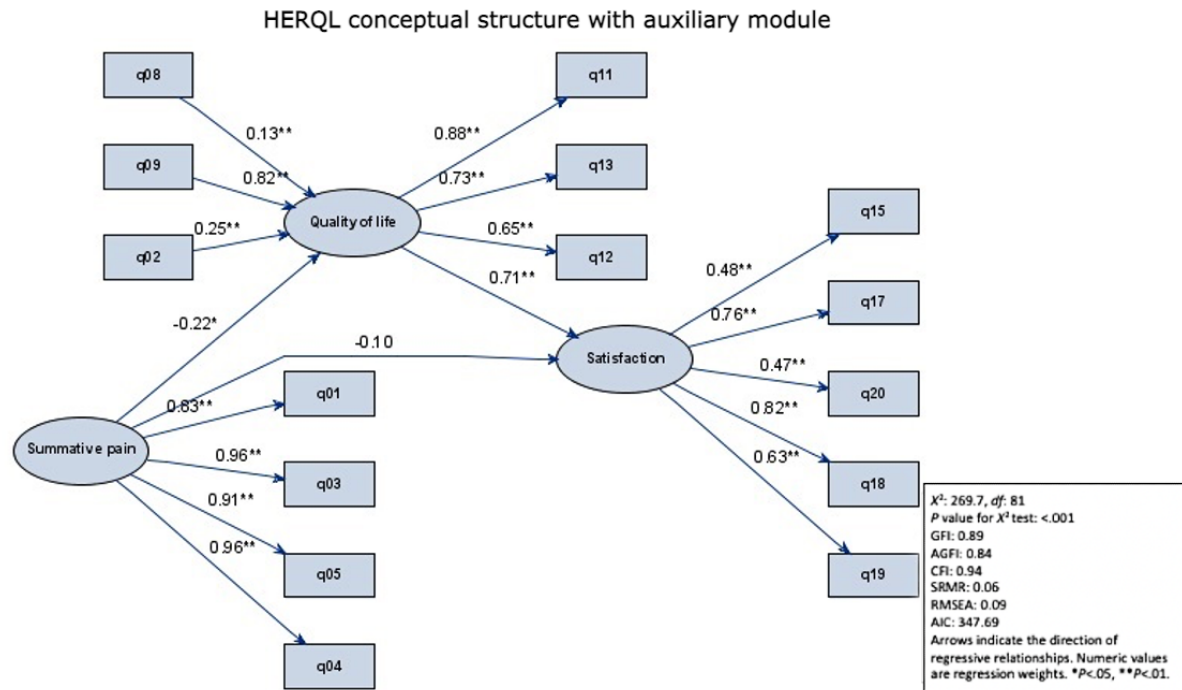
<sup>a</sup>This number includes 16 patients with both groin and abdominal wall hernias.

## SEM Concept of the HERQL Questionnaire

[Figure 2](#) shows the conceptual structure of the HERQL questionnaire with the postoperative module. Goodness of fit ( $\chi^2/df=3.3$ ) was slightly deviated from a good fit of the hypothesized construct to the experimental data. In large data sets with hundreds of samples, this deviation is not uncommon and was acceptable. Other indices supported the concept of the

HERQL questionnaire for long-term follow-up of patients with hernia, with slight deviations of GFI (0.89) and RMSEA (0.09). Most importantly, the causal relationship between the summative pain score and the quality-of-life latent factor was indicated by a significant -0.22 regression coefficient, whereas the significant and positive 0.71 regression coefficient was reported for the causal relationship between the quality-of-life and postoperative satisfactory latent factors.

**Figure 2.** Conceptual structure of HERQL (Hernia-Specific Quality-of-Life) with the auxiliary post-operative module. Circles: latent factors, rectangles: measured variables (questionnaire items). AGFI: adjusted goodness-of-fit index; AIC: Akaike's information criterion; CFI: comparative fit index; GFI: goodness-of-fit index; Q01: pain at rest, Q02: hernia protrusion, Q03: pain from mild activity, Q04: pain from moderate activity, Q05: pain from heavy activity, Q08: analgesic usage, Q09: activity restriction, Q11: hernia's impact on health, Q12: economic burden, Q13: quality-of-life/global health, Q15: foreign body sensation, Q17: complication severity, Q18: overall satisfaction, Q19: confidence in hernia repair, Q20: quality-of-life improvement by hernia repair; RMSEA: root mean square error of approximation; SRMR: standardized root mean square residual. Arrows indicate the direction of regressive relationships. Numeric values are regression weights. \* $P < .05$ , \*\* $P < .01$ .



## Long-term Follow-Up of Hernia Repairs

The overall recurrence rate among 311 participants was 1.0% ( $n=3$ ; Q19), and 88.7% ( $n=276$ ) of participants rated their last hernia repair as satisfactory or very satisfactory (Q18). Approximately 70% of invitees ( $n=213$ , 68.5%) reported no discomfort with their hernia repair (Q16), and 61.1% ( $n=190$ ) never experienced mesh foreign body sensation (Q15). Quality-of-life improvement following hernia repairs was ascertained in 90.4% ( $n=281$ ) of enrolled subjects (Q20).

## Subgroup Analysis

Comparisons were conducted between 77 patients with mesh plugs and 57 patients with the Prolene Hernia System (PHS) or the Ultrapro Hernia System (UHS) for groin hernia repair. Patients with mesh plug hernia repairs reported higher analgesic usage than those with PHS or UHS (Q08 score: 1.2 vs 1,  $P=.009$ ), worse impact on health (Q11 score: 1.9 vs 1.5,  $P=.03$ ), economic burden (Q12 score: 1.4 vs 1.2,  $P=.04$ ), foreign body sensation (Q15 score: 1.8 vs 1.4,  $P=.002$ ), discomfort severity (Q17 score: 1.5 vs 1.2,  $P=.005$ ), less satisfaction in herniorrhaphy (Q18 score: 1.7 vs 1.4,  $P=.03$ ), less confidence in hernia repair (Q19 score: 2.2 vs 1.9,  $P=.009$ ), and compromised quality-of-life improvement (Q20 score: 1.4 vs 1.1,  $P=.03$ ).

## Discussion

### Principal Findings

In this study, we reported the development of a mobile app to facilitate the long-term follow-up of patients with hernia and

ascertained the feasibility of the app. The cloud-based system eliminated the need for return visits by subjects who had undergone hernia repairs as early as a decade ago, which, in turn, enhanced the long-term follow-up and outcomes research. Hernia is a type of surgically treated disease with compromised long-term follow-up, as there is neither periodic surveillance nor medication prescription once the defect has been repaired. As well, patients with recurrent disease may seek a second opinion and receive further operations from surgeons in addition to the one resulting in failed repair [13,14]. Therefore, an easy-to-assess reporting system will be of great value for patients to present their immediate abdominal or groin conditions and for surgeons to update treatment outcomes.

For these reasons, we sent invitation letters by post to subjects who had their hernia repaired at our institute more than 1 year ago, with index cases identified from medical records. With an enclosed preset ID and password, invitees could easily download the iOS or Android version of the mobile app, complete the quality-of-life survey, and provide their electrical signature for informed consent within a few minutes. For those not familiar with mobile apps or not equipped with a smartphone, a Google Forms survey provided an online alternative. The response rate was 11.9%, or slightly more than one-tenth of the identified candidates. The majority of the 311 responders were diagnosed with groin hernias, reflecting the clinical scenario of the hernia population.

From our study, the long-term recurrence rate was less than 1%, with 3 patients reporting hernia recurrence. Most of the 311 patients ( $n=276$ , 88.7%) reported that they were satisfied or very satisfied with their last hernia repair, 68.5% of patients

(n=213) reported no hernia-related discomfort, and 61.1% (n=190) never experienced mesh foreign body sensation. We also compared patients with groin hernias receiving either mesh plug repairs or PHS or UHS repairs and found that those with mesh plugs inevitably experienced more analgesic usage, worse health impact, economic burden, foreign body sensation, discomfort severity, less satisfaction, less confidence in hernia repair, and compromised quality-of-life improvement. Most importantly, 90.3% of the participants experienced an improvement in quality of life following hernia surgery, indicating that elimination of hernia-related symptoms might be the main contributor to such improvement.

The conceptual structure of the HERQL questionnaire with the postoperative module displayed satisfactory model fit indices (Figure 2), further augmenting the superiority of SEM. Fayers et al [24,25] initiated the efforts to use SEM for the conceptual structure of the instrument used to measure quality of life; they aimed to separate causal variables (ie, symptoms) from effect indicators (ie, functional domains). The critical rationale underpinning the causal-indicative duality was that hernia-associated symptoms impaired subjective perception of quality of life, which was subsequently reflected in functional domain indicator variables as well as in patients' satisfaction as measured by the postoperative module.

### Comparison With Prior Work

Our study was not the first to perform outcomes research for hernia outside a hospital. Heniford et al [15] conducted outcomes research using the CCS questionnaire, which was mailed to 1048 patients and had a response rate was 12.9%. We invited patients with hernia who had completed hernia repair more than 1 year before the study began, and our response rate was similar to that of the CCS study; however, there was a much longer time interval between herniorrhaphy and questionnaire administration in this study. One major reason for the low response rate was loss of contact due to incorrect addresses, which resulted in undelivered mail. With longer follow-up times, migration could occur naturally and some invitees could pass away; in these cases, patients would become inaccessible. Although compensation was arranged, lack of incentive, worry about fraud, and reluctance to participate might compromise the uptake of the mobile app, constituting another reason for low response. In addition, recall bias did occur, and it was postulated that patients with a recurrence after hernia repair might be more reluctant to participate in this study; consequently, the recurrence rate might be underestimated.

In our previous study, 192 patients who had groin hernias repaired with mesh plugs were compared with 234 patients who had PHS repairs. Postoperatively, the group who had mesh plug repairs had a higher incidence of chronic nondisabling groin pain [33]. In this study, a subgroup analysis was conducted comparing mesh plug repairs with PHS or UHS repairs. Coinciding with our previous study, mesh plugs hampered hernia surgery outcomes with worse symptoms and compromised functionality. It deserves notice that the median follow-up time was only 26.6 months in our previous hospital-based study, which was much shorter than the median follow-up time of 66 months in this study. On the other hand, the HERQL

questionnaire validation study was designed with repeated measurements up to 1 year following hernia repair, which is also distinct from the long-term follow-up scope of this study [21]. There are different definitions of chronic pain following hernia repair; we used 1 year as the cutoff for chronic pain, as our previous validation study with repeated measurements indicated that 1 year after hernia repair was a reasonable timepoint for a stable long-term condition.

### Limitations

There were some limitations of this study. First, the retrospective design inevitably introduced recall bias, especially for those with longer follow-up periods. This study evaluated the feasibility of mobile app-based outcomes research, and further study is warranted to eliminate this bias with a more uniform follow-up interval. Second, not all clinical and demographic data were available through chart reviews, such as BMI and fascia defect size, which could hamper post hoc and multivariate analysis considerably. Third, some older adult patients might not be able to complete the survey without an assistant, and there was no printed questionnaire available in the event that the mobile app was not properly installed. Fourth, no further reminder letters were sent by mail and no further phone calls were attempted if there was no response from the initial invitation letter, which inevitably compromised the response rate.

### Conclusions

Knowledge gained from this study can translate into the design of future hernia outcomes research. For example, an updated hernia registry could be established that includes a novel mobile app to enhance the follow-up of patients with hernia, and a cloud-based database could be established for surgeons as well as patients with hernia. A corroborative database would be useful for surgeons to collect clinical and operative details from hernia surgeries, and this would provide a platform for real-time communication between patients with hernia and surgeons and to enhance postoperative follow-up and outcomes assessment. Surgeons could enter the clinical and operative data immediately after completion of hernia repairs using mobile devices, whereas sensitive clinical data would be secured and restricted to authorized personnel. In addition, patients with hernia could review their clinical and operative details in a well-designed and self-explanatory manner. Patients with hernia could also record postoperative events, such as results from a visual analog pain scale, wound condition, and complications, and they could complete the HERQL questionnaire periodically in order to assess the outcomes of hernia repair. Finally, the instant message communication feature can be established to provide an easy and efficient way for patients to report any discomfort to their surgeons, and a proper response from the latter could enhance the long-term follow-up compliance rate of patients with hernia.

The establishment of the mobile app could enhance the quality of care for patients with hernia and facilitate outcomes research for hernia disease with a more comprehensive and complete follow-up, and the feasibility is ascertained herein. The knowledge gained from this project could be extended to other common surgical procedures [34,35]. This study will facilitate hernia outcomes research and enhance the quality of care for



this common disease by providing a validated HERQL instrument with enhanced sensitivity.

## Acknowledgments

This work was supported, in part, by the Ministry of Science and Technology, Executive Yuan, Taipei, Taiwan (grant MOST-103-2314-B-281-004-MY2) and Cathay General Hospital, Taipei, Taiwan (grant CGH-MR-A10510P). This study contains materials presented at the 39th European Hernia Society Annual International Congress that took place from May 24 to 27, 2017, in Vienna, Austria. The authors would like to thank Taiwan Clinical Oncology Research Foundation, Melissa Lee Cancer Foundation and Mr Morris Chang for their kind help during the study.

## Conflicts of Interest

None declared.

Multimedia Appendix 1

Supplementary materials.

[DOCX File, 605 KB - [formative\\_v6i10e39759\\_app1.docx](#) ]

## References

- Legutko J, Pach R, Solecki R, Matyja A, Kulig J. [The history of treatment of groin hernia] [Article in Polish]. *Folia Med Cracov* 2008;49(1-2):57-74. [Medline: [19140492](#)]
- Lomanto D, Cheah W, Faylona JM, Huang CS, Lohsiriwat D, Maleachi A, et al. Inguinal hernia repair: Toward Asian guidelines. *Asian J Endosc Surg* 2015 Feb;8(1):16-23. [doi: [10.1111/ases.12141](#)] [Medline: [25598054](#)]
- Simons MP, Aufenacker T, Bay-Nielsen M, Bouillot JL, Campanelli G, Conze J, et al. European Hernia Society guidelines on the treatment of inguinal hernia in adult patients. *Hernia* 2009 Aug;13(4):343-403 [FREE Full text] [doi: [10.1007/s10029-009-0529-7](#)] [Medline: [19636493](#)]
- Huang C. The evolution of groin hernioplasty, an update. *Formosan J Surg* 2010;43:278-278. [doi: [10.29833/FJS.201012.0002](#)]
- Mason R, Moazzez A, Sohn H, Berne T, Katkhouda N. Laparoscopic versus open anterior abdominal wall hernia repair: day morbidity and mortality using the ACS-NSQIP database. *Ann Surg* 2011;254:30. [doi: [10.1097/sla.0b013e31823009e6](#)]
- Nguyen MT, Berger RL, Hicks SC, Davila JA, Li LT, Kao LS, et al. Comparison of outcomes of synthetic mesh vs suture repair of elective primary ventral herniorrhaphy: A systematic review and meta-analysis. *JAMA Surg* 2014 May;149(5):415-421. [doi: [10.1001/jamasurg.2013.5014](#)] [Medline: [24554114](#)]
- Scott N, McCormack K, Graham P, Go P, Ross S, Grant A. Open mesh versus non-mesh for repair of femoral and inguinal hernia. *Cochrane Database Syst Rev* 2002;CD002197. [doi: [10.1002/14651858.cd002197](#)]
- Woods B, Neumayer L. Open repair of inguinal hernia: An evidence-based review. *Surg Clin North Am* 2008 Feb;88(1):139-155, ix. [doi: [10.1016/j.suc.2007.11.005](#)] [Medline: [18267167](#)]
- Luijendijk RW, Hop WC, van den Tol MP, de Lange DC, Braaksma MM, IJzermans JN, et al. A comparison of suture repair with mesh repair for incisional hernia. *N Engl J Med* 2000 Aug 10;343(6):392-398. [doi: [10.1056/nejm200008103430603](#)]
- Fränneby U, Sandblom G, Nordin P, Nyrén O, Gunnarsson U. Risk factors for long-term pain after hernia surgery. *Ann Surg* 2006 Aug;244(2):212-219. [doi: [10.1097/01.sla.0000218081.53940.01](#)] [Medline: [16858183](#)]
- Jensen KK, Henriksen NA, Harling H. Standardized measurement of quality of life after incisional hernia repair: A systematic review. *Am J Surg* 2014 Sep;208(3):485-493. [doi: [10.1016/j.amjsurg.2014.04.004](#)] [Medline: [25017051](#)]
- Sandø A, Rosen MJ, Heniford BT, Bisgaard T. Long-term patient-reported outcomes and quality of the evidence in ventral hernia mesh repair: A systematic review. *Hernia* 2020 Aug;24(4):695-705. [doi: [10.1007/s10029-020-02154-1](#)] [Medline: [32152807](#)]
- Holihan JL, Hannon C, Goodenough C, Flores-Gonzalez JR, Itani KM, Olavarria O, et al. Ventral hernia repair: A meta-analysis of randomized controlled trials. *Surg Infect (Larchmt)* 2017;18(6):647-658. [doi: [10.1089/sur.2017.029](#)] [Medline: [28557648](#)]
- Al Chalabi H, Larkin J, Mehigan B, McCormick P. A systematic review of laparoscopic versus open abdominal incisional hernia repair, with meta-analysis of randomized controlled trials. *Int J Surg* 2015 Aug;20:65-74 [FREE Full text] [doi: [10.1016/j.ijsu.2015.05.050](#)] [Medline: [26074289](#)]
- Heniford B, Walters A, Lincourt A, Novitsky Y, Hope W, Kercher K. Comparison of generic versus specific quality-of-life scales for mesh hernia repairs. *J Am Coll Surg* 2008 Apr;206(4):638-644. [doi: [10.1016/j.jamcollsurg.2007.11.025](#)] [Medline: [18387468](#)]

16. Krpata D, Schmotzer B, Flocke S, Jin J, Blatnik J, Ermlich B, et al. Design and initial implementation of HerQLes: A hernia-related quality-of-life survey to assess abdominal wall function. *J Am Coll Surg* 2012 Nov;215(5):635-642. [doi: [10.1016/j.jamcollsurg.2012.06.412](https://doi.org/10.1016/j.jamcollsurg.2012.06.412)] [Medline: [22867715](https://pubmed.ncbi.nlm.nih.gov/22867715/)]
17. Staerkle R, Villiger P. Simple questionnaire for assessing core outcomes in inguinal hernia repair. *Br J Surg* 2011 Jan;98(1):148-155. [doi: [10.1002/bjs.7236](https://doi.org/10.1002/bjs.7236)] [Medline: [20814965](https://pubmed.ncbi.nlm.nih.gov/20814965/)]
18. Chung L, O'Dwyer PJ. Pain and its effects on physical activity and quality of life before operation in patients undergoing elective inguinal and ventral hernia repair. *Am J Surg* 2014 Sep;208(3):406-411. [doi: [10.1016/j.amjsurg.2014.02.011](https://doi.org/10.1016/j.amjsurg.2014.02.011)] [Medline: [24997490](https://pubmed.ncbi.nlm.nih.gov/24997490/)]
19. Muysoms F, Campanelli G, Champault GG, DeBeaux AC, Dietz UA, Jeekel J, et al. EuraHS: The development of an international online platform for registration and outcome measurement of ventral abdominal wall hernia repair. *Hernia* 2012 Jun;16(3):239-250 [FREE Full text] [doi: [10.1007/s10029-012-0912-7](https://doi.org/10.1007/s10029-012-0912-7)] [Medline: [22527930](https://pubmed.ncbi.nlm.nih.gov/22527930/)]
20. Fränneby U, Gunnarsson U, Andersson M, Heuman R, Nordin P, Nyrén O, et al. Validation of an inguinal pain questionnaire for assessment of chronic pain after groin hernia repair. *Br J Surg* 2008 Apr;95(4):488-493. [doi: [10.1002/bjs.6014](https://doi.org/10.1002/bjs.6014)] [Medline: [18161900](https://pubmed.ncbi.nlm.nih.gov/18161900/)]
21. Huang C, Tai F, Chou T, Lien H, Jeng J, Ho T, et al. Quality of life of inguinal hernia patients in Taiwan: The application of the hernia-specific quality of life assessment instrument. *PLoS One* 2017;12(8):e0183138 [FREE Full text] [doi: [10.1371/journal.pone.0183138](https://doi.org/10.1371/journal.pone.0183138)] [Medline: [28817703](https://pubmed.ncbi.nlm.nih.gov/28817703/)]
22. Huang C, Lien H, Wong J, Ho T, Chang W, Hunag C. Revisiting HERQL, the hernia-specific quality-of-life assessment instrument, to extend the clinical applicability for abdominal wall hernias. *Hernia* 2020 Aug;24(4):771-780. [doi: [10.1007/s10029-019-02066-9](https://doi.org/10.1007/s10029-019-02066-9)] [Medline: [31781964](https://pubmed.ncbi.nlm.nih.gov/31781964/)]
23. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res* 2011 Dec;20(10):1727-1736 [FREE Full text] [doi: [10.1007/s11136-011-9903-x](https://doi.org/10.1007/s11136-011-9903-x)] [Medline: [21479777](https://pubmed.ncbi.nlm.nih.gov/21479777/)]
24. Fayers P, Hand D, Bjordal K, Groenvold M. Causal indicators in quality of life research. *Qual Life Res* 1997;6:393-406. [doi: [10.1023/a:1018491512095](https://doi.org/10.1023/a:1018491512095)]
25. Fayers PM, Hand DJ. Factor analysis, causal indicators and quality of life. *Qual Life Res* 1997 Mar;6(2):139-150. [doi: [10.1023/a:1026490117121](https://doi.org/10.1023/a:1026490117121)] [Medline: [9161114](https://pubmed.ncbi.nlm.nih.gov/9161114/)]
26. Boehmer S, Luszczynska A. Two kinds of items in quality of life instruments: 'Indicator and causal variables' in the EORTC qlq-c30. *Qual Life Res* 2006 Feb;15(1):131-141. [doi: [10.1007/s11136-005-8290-6](https://doi.org/10.1007/s11136-005-8290-6)] [Medline: [16411037](https://pubmed.ncbi.nlm.nih.gov/16411037/)]
27. Harji D, Thomas C, Antoniou S, Chandraratan H, Griffiths B, Henniford B, NoSTRA HarMoNY. A systematic review of outcome reporting in incisional hernia surgery. *BJS Open* 2021 Mar 05;5(2):zrab006 [FREE Full text] [doi: [10.1093/bjsopen/zrab006](https://doi.org/10.1093/bjsopen/zrab006)] [Medline: [33839746](https://pubmed.ncbi.nlm.nih.gov/33839746/)]
28. Chu Y, Cheng Y, Lai Y, Tsao Y, Tu T, Young ST, et al. A mobile phone-based approach for hearing screening of school-age children: Cross-sectional validation study. *JMIR Mhealth Uhealth* 2019 Apr 01;7(4):e12033 [FREE Full text] [doi: [10.2196/12033](https://doi.org/10.2196/12033)] [Medline: [30932870](https://pubmed.ncbi.nlm.nih.gov/30932870/)]
29. Lin HH, Chu Y, Lai Y, Cheng H, Lai F, Cheng Y, et al. A smartphone-based approach to screening for sudden sensorineural hearing loss: Cross-sectional validity study. *JMIR Mhealth Uhealth* 2020 Nov 11;8(11):e23047 [FREE Full text] [doi: [10.2196/23047](https://doi.org/10.2196/23047)] [Medline: [33174845](https://pubmed.ncbi.nlm.nih.gov/33174845/)]
30. Chen C, Lin HH, Wang M, Chu Y, Chang C, Huang C, et al. Diagnostic accuracy of smartphone-based audiometry for hearing loss detection: Meta-analysis. *JMIR Mhealth Uhealth* 2021 Sep 10;9(9):e28378 [FREE Full text] [doi: [10.2196/28378](https://doi.org/10.2196/28378)] [Medline: [34515644](https://pubmed.ncbi.nlm.nih.gov/34515644/)]
31. HERQL. Google Docs. 2014. URL: <http://goo.gl/FMHICt> [accessed 2022-09-18]
32. HERQL. Google Sites. URL: <https://sites.google.com/site/herqlsurvey>
33. Huang CS, Huang CC, Lien HH. Prolene hernia system compared with mesh plug technique: A prospective study of short-to mid-term outcomes in primary groin hernia repair. *Hernia* 2005 May;9(2):167-171. [doi: [10.1007/s10029-005-0318-x](https://doi.org/10.1007/s10029-005-0318-x)] [Medline: [15703858](https://pubmed.ncbi.nlm.nih.gov/15703858/)]
34. Wu J, Yu H, Ho T, Su X, Lin M, Lai F. Tablet PC-enabled application intervention for patients with gastric cancer undergoing gastrectomy. *Comput Methods Programs Biomed* 2015 Apr;119(2):101-109. [doi: [10.1016/j.cmpb.2015.03.004](https://doi.org/10.1016/j.cmpb.2015.03.004)] [Medline: [25819034](https://pubmed.ncbi.nlm.nih.gov/25819034/)]
35. McCarthy MJ, Jonasson O, Chang C, Pickard A, Giobbie-Hurder A, Gibbs J, et al. Assessment of patient functional status after surgery. *J Am Coll Surg* 2005 Aug;201(2):171-178. [doi: [10.1016/j.jamcollsurg.2005.03.035](https://doi.org/10.1016/j.jamcollsurg.2005.03.035)] [Medline: [16038812](https://pubmed.ncbi.nlm.nih.gov/16038812/)]

## Abbreviations

- AGFI:** adjusted goodness-of-fit index
- CCS:** Carolinas Comfort Scale
- GFI:** goodness-of-fit index
- HERQL:** Hernia-Specific Quality-of-Life
- PHS:** Prolene Hernia System

**RMSEA:** root mean square error of approximation

**SEM:** structural equation modeling

**UHS:** Ultrapro Hernia System

*Edited by A Mavragani; submitted 22.05.22; peer-reviewed by S Pit, A Bwanika Naggirinya; comments to author 15.08.22; revised version received 04.09.22; accepted 07.09.22; published 19.10.22.*

*Please cite as:*

*Huang CS, Tai FC, Lien HH, Wong JU, Huang CC*

*Long-term Follow-up of Patients With Hernia Using the Hernia-Specific Quality-of-Life Mobile App: Feasibility Questionnaire Study*  
*JMIR Form Res 2022;6(10):e39759*

*URL: <https://formative.jmir.org/2022/10/e39759>*

*doi: [10.2196/39759](https://doi.org/10.2196/39759)*

*PMID: [36260390](https://pubmed.ncbi.nlm.nih.gov/36260390/)*

©Ching-Shui Huang, Feng-Chuan Tai, Heng-Hui Lien, Jia-Uei Wong, Chi-Cheng Huang. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 19.10.2022. 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 Family Level Assessment of Screen Use—Mobile Approach: Development of an Approach to Measure Children’s Mobile Device Use

Oriana Perez<sup>1</sup>, MSc; Anil Kumar Vadathya<sup>2</sup>, MSc; Alicia Beltran<sup>1</sup>, MSc; R Matthew Barnett<sup>3</sup>, BA; Olivia Hinder<sup>4</sup>, BSc; Tatyana Garza<sup>1</sup>, BSc; Salma M Musaad<sup>1</sup>, MD, PhD; Tom Baranowski<sup>1</sup>, PhD; Sheryl O Hughes<sup>1</sup>, PhD; Jason A Mendoza<sup>5,6</sup>, MPH, MD; Ashutosh Sabharwal<sup>2</sup>, PhD; Ashok Veeraraghavan<sup>2</sup>, PhD; Teresia M O'Connor<sup>1</sup>, MPH, MD

<sup>1</sup>United States Department of Agriculture/Agricultural Research Service Children's Nutrition Research Center, Baylor College of Medicine, Houston, TX, United States

<sup>2</sup>Department of Electrical & Computer Engineering, Rice University, Houston, TX, United States

<sup>3</sup>Center for Research Computing, Rice University, Houston, TX, United States

<sup>4</sup>Baylor College of Medicine, Houston, TX, United States

<sup>5</sup>Public Health Sciences Division, Fred Hutchinson Cancer Center, Seattle, WA, United States

<sup>6</sup>Center for Child Health, Behavior and Development, Seattle Children's Research Institute, Seattle, WA, United States

**Corresponding Author:**

Teresia M O'Connor, MPH, MD

United States Department of Agriculture/Agricultural Research Service Children's Nutrition Research Center

Baylor College of Medicine

1100 Bates St

Houston, TX, 77030

United States

Phone: 1 713 798 6782

Email: [teresiao@bcm.edu](mailto:teresiao@bcm.edu)

## Abstract

**Background:** There is a strong association between increased mobile device use and worse dietary habits, worse sleep outcomes, and poor academic performance in children. Self-report or parent-proxy report of children’s screen time has been the most common method of measuring screen time, which may be imprecise or biased.

**Objective:** The objective of this study was to assess the feasibility of measuring the screen time of children on mobile devices using the Family Level Assessment of Screen Use (FLASH)—mobile approach, an innovative method that leverages the existing features of the Android platform.

**Methods:** This pilot study consisted of 2 laboratory-based observational feasibility studies and 2 home-based feasibility studies in the United States. A total of 48 parent-child dyads consisting of a parent and child aged 6 to 11 years participated in the pilot study. The children had to have their own or shared Android device. The laboratory-based studies included a standardized series of tasks while using the mobile device or watching television, which were video recorded. Video recordings were coded by staff for a gold standard comparison. The home-based studies instructed the parent-child dyads to use their mobile device as they typically use it over 3 days. Parents received a copy of the use logs at the end of the study and completed an exit interview in which they were asked to review their logs and share their perceptions and suggestions for the improvement of the FLASH-mobile approach.

**Results:** The final version of the FLASH-mobile approach resulted in user identification compliance rates of >90% for smartphones and >80% for tablets. For laboratory-based studies, a mean agreement of 73.6% (SD 16.15%) was achieved compared with the gold standard (human coding of video recordings) in capturing the target child’s mobile use. Qualitative feedback from parents and children revealed that parents found the FLASH-mobile approach useful for tracking how much time their child spends using the mobile device as well as tracking the apps they used. Some parents revealed concerns over privacy and provided suggestions for improving the FLASH-mobile approach.

**Conclusions:** The FLASH-mobile approach offers an important new research approach to measure children’s use of mobile devices more accurately across several days, even when the child shares the device with other family members. With additional

enhancement and validation studies, this approach can significantly advance the measurement of mobile device use among young children.

(*JMIR Form Res* 2022;6(10):e40452) doi:[10.2196/40452](https://doi.org/10.2196/40452)

## KEYWORDS

screen time; mobile media apps; children; mobile phone use; tablet use; mobile phone

## Introduction

### Background

Screen use includes mobile device use, television watching, playing video games, and computer use. Strong associations have been demonstrated between increased screen time and worse sleep outcomes, locomotive skills, and dietary habits in children [1-4]. Screen-based activities, particularly with television and video games, may be related to poor academic performance, whereas mobile phone use has been shown to be correlated with poor mental health in adolescents [5-7]. Screen media use has also been associated with less favorable body composition and worse cardiometabolic risk in youth [1,8], but the associations seem less clear in longitudinal studies [9]. Limitations in measurement might explain discrepancies in the association of screen use with health outcomes in the literature, raising concern over the accuracy of screen media exposure assessments in past studies.

The most common approach to measuring screen use is through self-report or parent-proxy report of children's screen use by questionnaires, which introduces errors and biases [10], affecting the conclusions drawn from the studies. Within this context, screen time has continued to rise among children and adolescents and increasingly includes the use of mobile devices, such as tablets and smartphones [11,12]. Approximately, 48% of children aged 0 to 8 years personally owned a mobile device in 2020 [12], and 41% of tweens (aged 8-12 years) and 84% of teens had their own smartphone in 2019 [11]. This increased variety of screen use among youth supports the need for accurate measurements of different types of screen use to inform outcome-based research studies. The gold standard for measuring children's television viewing remains to be direct observation [10], but its use is limited by expense and privacy concerns. Observation can be even more challenging when used to measure mobile screen use. Parent-proxy-reported general estimates are less cumbersome and less costly, but parent report of young children's mobile device use had a high frequency of underestimation and overestimation of use compared with objective assessment (35.7% and 34.8%, respectively) [13]. Mobile screen use by children may be particularly difficult to assess because of its intermittent use, the short duration of the use of some apps [13], and the use of devices away from parents because of the mobile nature of the device.

The difficulties in accurately measuring children's mobile device use demonstrate an acute need for new measurement tools to accomplish this goal. We have already developed a novel approach for measuring children's television viewing [14]. There are also assessment tools to track the screen use of a person on a mobile device used only by that person [13,15,16]. However, the existing methods cannot distinguish the user of

the device, which is an important consideration for children who may access a parent's phone or share a device with a sibling.

### Objectives

In this study, our goal was to develop an innovative approach that leverages the capabilities of smartphone apps to develop a less intrusive yet accurate method for assessing children's mobile screen use. This study describes the Family Level Assessment of Screen use (FLASH) for mobile devices. The FLASH-mobile approach makes use of (1) an app to track device and app use logging, (2) notifications to identify the user of the device, (3) a system to provide feedback on compliance with the user identification notification to improve compliance, and (4) a minimum threshold of user identification to be included in the analysis. In this study, we used the HealthSense platform [17] for implementing the FLASH-mobile approach with device and app use tracking and user identification, but the FLASH-mobile approach could be implemented with other Android tracking apps in the future. This study describes the development of and assesses the feasibility of the FLASH-mobile approach to track children's mobile smartphone or digital tablet use, compliance with the user log-in step, and parents' perceptions of the FLASH-mobile approach.

## Methods

### Overview

A total of 4 feasibility studies tested multiple iterations of the FLASH-mobile approach on mobile phones and tablets. Participants included a parent and at least one child aged 6 to 11 years. Data were collected from December 2019 to August 2021 by installing HealthSense, an app that tracks the use of the device and corresponding apps, on the participants' mobile phone or tablet.

### Ethical Considerations

This research was approved by the Baylor College of Medicine Institutional Review Board (H-40556), with a reciprocal authorized agreement with Rice University. Parents provided written, informed consent for themselves and their child, and the children provided assent to participate.

### Participants

Participants were recruited from the community and had to live in the United States, speak English, and be a parent of a child aged between 6 and 11 years, and the child had to have their own or share another family member's Android device (smartphone or tablet). Participants were excluded if the parent or child had a developmental, medical, mental, or physical condition that would prevent them from following the study

protocol. Families with mobile devices with other operating systems (OSs), such as Apple iOS or Amazon Fire OS, were excluded because other OSs do not support the HealthSense platform. For laboratory-based protocols, a study-owned backup device was available in case of technical difficulties with the participant's device. Families were compensated US \$40 for their participation.

### FLASH-Mobile Approach

The FLASH-mobile approach was developed to estimate the total time a target child spent using a mobile device. Children often share a mobile device with their siblings or a parent. Therefore, 4 steps were developed to assess the mobile device use of a child from a shared device: (1) an app to track the Android OS app use statistics on the device to identify mobile device use time, duration, and which apps were used; (2) identification of the user of the mobile device via device notifications when unlocked; (3) a system to provide feedback on compliance with the user identification notifications to improve compliance; and (4) a minimum threshold of user identification compliance to be included in the analysis. Several iterative versions of the user identification procedure were tested based on feedback from the participants on earlier versions.

The FLASH-mobile approach was implemented using the HealthSense app [17], developed by researchers at Rice University and used for all the validation and testing results presented in this paper. Similar to other apps such as Chronicle [13] and Minuku [18], HealthSense uses Android's UsageStatsManager [19] function to access app use event entries recorded by the phone's OS [20]. Specifically, each time an app is opened (moved to foreground) or closed (moved to background), the Android OS automatically records these as individual time-stamped events in the Android event log, regardless of whether HealthSense or any other app is installed on the device. HealthSense reads this time-stamped event log, resulting in accurate app use data. To monitor phone unlock events for user identification, the app monitors another automatic event log, the ACTION\_USER\_PRESENT [21] event. When the participants installed the HealthSense app, they were explicitly asked to authorize the app to read these event logs that the Android OS had already recorded on their devices.

For the user identification step, multiple approaches were tested: version 1 (V1), a pop-up prompt appeared when the phone was unlocked, which had to be answered to use the device; version 2 (V2), a notification banner appeared when the phone was unlocked and after every 15 minutes of use; and version 3 (V3),

a notification banner that appeared only when the phone was unlocked. In each case, the user had to identify themselves by selecting 1 of 2 options, an orange button with the word *child* or a gray button with the word *other* (Multimedia Appendix 1). There was no option for multiple users. To refine the log-in prompt, three major requirements were considered: (1) the app should be compatible across as many Android versions as possible, (2) newer versions of the Android OS had to support it, and (3) the log-in prompt must be user-friendly to encourage high compliance among children.

We conducted 4 feasibility studies. Parents were instructed to download the HealthSense app. The type of Android device used, model number of the device, OS number, whether the HealthSense app was successfully downloaded, and whether the user identification feature worked as expected were tracked. Two laboratory-based protocols (Table 1: feasibility studies A [in laboratory] and B [in laboratory]) assessed the ability of the FLASH-mobile approach to accurately capture participants' mobile device use, compared with the research staff's assessment of the participants' gaze on their mobile device from video data. Parents and children were observed in a laboratory set up like a living room while being video recorded by 4 high-definition cameras placed in each corner. The protocols included a series of tasks (eg, using the mobile phone or tablet, watching television, playing with the toys in the room, or free play).

Furthermore, 2 home-based studies (Table 1: feasibility studies C [at home] and D [at home]) gathered real-world information about the ability of the FLASH-mobile approach to collect data on the child's smartphone (study C [at home]) and tablet use (study D [at home]) over 3 days. Feasibility was assessed by the (1) number of mobile devices the HealthSense app could be downloaded on for the at-home and laboratory studies, (2) ability of the app to collect data over the 3-day at-home study period, (3) compliance with user notification, and (4) parents' perceptions of the FLASH-mobile approach during the exit interviews. All 3 versions of HealthSense (V1, V2, and V3) were iteratively tested in study C (at home). We observed low participant compliance with V3, which prompted a modification of the protocol to include feedback for compliance. Feedback on compliance was accomplished by sending reminders to respond to the user notification via telephone, SMS text message, or email to participants with low compliance and by sending encouraging messages to participants with high compliance.

**Table 1.** Description of the feasibility studies (N=48).

Feasibility study	Setting	HealthSense app version tested	Parent-child dyads enrolled, n (%)	Device tested	Duration
A	Laboratory based	Version 1 <sup>a</sup>	5 (10)	Mobile phone	Approximately 2 hours (video-recorded tasks)
B	Laboratory based	Version 3 <sup>b</sup>	10 (21)	Mobile phone or tablet	Approximately 2 hours (video-recorded tasks)
C	Home based	Versions 1, 2 <sup>c</sup> , and 3	21 (44)	Mobile phone	3 days (naturalistic)
D	Home based	Version 3	12 (25)	Tablet	3 days (naturalistic)

<sup>a</sup>A pop-up prompt appeared when the device was unlocked, and the user had to identify themselves by selecting 1 of 2 options, *child* or *other*, to use the device. It is no longer supported by the Android operating system.

<sup>b</sup>A notification banner appeared only when the phone was unlocked, and the user had to identify themselves by selecting 1 of 2 options, *child* or *other*. Daily text or email feedback was sent to the family on their compliance.

<sup>c</sup>The notification banner appeared when the phone was unlocked and after every 15 minutes of use. The user had to identify themselves by selecting 1 of 2 options, *child* or *other*.

## Video Coding

Screen use was defined as the duration of time a child spent actively watching the mobile device screen (gaze on the mobile device) or using the screen while multitasking (eg, while playing with toys). This is similar to how others have defined screen use, including as background television [22]. A total of 5 research staff members coded video data from studies A (in laboratory) and B (in laboratory) to identify when the target child was using or watching the mobile device during the observation protocol to provide a gold standard for assessing the accuracy of the FLASH-mobile approach in tracking child device use. The 4 video frames were viewed simultaneously by staff to get 4 different angles of view of the room during coding. The video was coded using duration coding with one of ten codes: (1) child screen use, (2) parent screen use, (3) both screen use, (4) child multitasking, (5) child audio use, (6) parent audio use, (7) both audio use, (8) no mobile device use, (9) uncertain, or (10) out of frame (informed by previous publications [23-25]). Screen use was differentiated from audio use based on whether the person was gazing at the screen or making a call, speaking voice commands, or listening to music. To be considered screen use, gaze could be quickly diverted for <3 seconds but should still mainly be focused on the mobile device, similar to other coding protocols [26]. The *multitasking code* was applied when the child used a mobile device screen concurrently with other tasks, such as watching television, playing with toys, or talking to another person [24]. In all, 10% of each video was double coded by 2 staff members, and the agreement was high [27-29] (feasibility study A [in laboratory], Cohen  $\kappa=0.76$ , SD 0.38; feasibility study B [in laboratory], Cohen  $\kappa=0.79$ , SD 0.37).

## Identification of the Apps Used

HealthSense listed the names of all the opened apps. A log of the apps used was generated when the child was the user. These were categorized by the type of app: educational (eg, Math Jumps and Khan Academy Kids), streaming video services (eg, YouTube and Netflix), gaming, social media, browsing, Android OS, and others (Multimedia Appendix 2 presents the full list of categorized apps). Two staff members reviewed and independently coded the apps into one of these categories by reading the description included in the Google App Store. A

third staff member reviewed the app categorization. Differences in the categorization of apps were reviewed and discussed as a team until a consensus was reached.

## Exit Interviews

All the parents were asked to share their perceptions of using the FLASH-mobile approach and provide suggestions for its improvement. The semistructured interviews followed a standardized script, which included probes and prompts to clarify responses (Multimedia Appendix 3 presents the full interview script). For studies C (at home) and D (at home), parents and children were provided with a copy of their mobile device use over the past 3 days before their interview and were asked to review it. The log included which apps were used, the time of day, the duration of use, and the user identified (*child*, *other*, or none if the user identification prompt was ignored). Parents were asked to verify whether the apps listed were correct; whether any apps were missing or added incorrectly; and whether the timing, duration, and user identified were correct. All the interviews were audio recorded, transcribed, and coded by 3 study staff members (OH, TG, and OP), using NVivo (version 11, QSR International) software. A codebook was created based on an inductive coding of themes [30,31]. Pairs of reviewers independently coded each interview. Discrepancies in the themes were discussed and resolved until a consensus was reached. The authors identified the main themes which arose and categorized the subthemes within them.

## Analyses

Means, SDs, and percentages describe the demographics and dyad's compliance with user identification. User compliance was measured along 2 dimensions: compliance with the log-in identification and unidentified use. Compliance score was defined as the percentage of times the user responded to the log-in notification over the total number of times the mobile device was unlocked. Unidentified use was calculated as the percentage of use for which the user was not known over the total mobile device use.

*Child use* of device was defined by collating staff video codes of (1) child directly using mobile screen alone, (2) using mobile screen with someone else, and (3) multitasking mobile screen

use with other activities (eg, playing). The agreement between the HealthSense output and staff coding for mobile screen use in feasibility studies A (in laboratory) and B (in laboratory) was tested using the percentage of agreement, prevalence- and bias-adjusted Cohen  $\kappa$  statistic, and intraclass correlation coefficient.

## Results

### Participant Demographics

A total of 48 parent-child dyads participated in the study (n=5, 10% for study A [in laboratory]; n=10, 21% for study B [in

laboratory]; n=21, 44% for study C [at home]; and n=12, 25% for study D [at home]). Each dyad participated in only 1 study. Of all the 48 enrolled participants, 33 (68%) were able to download the app on their own device and have all the components function. For the observational laboratory studies, a study device was available for 2 families whose device did not work, resulting in 73% (35/48) of participants with complete data from the HealthSense app. The demographics of the sample are shown in Table 2, and those for each study can be found in Multimedia Appendix 4.

**Table 2.** Demographic characteristics of participants with complete data (N=48).

Characteristic	All tests with incomplete data	All tests with complete data
<b>Target children, n (%)</b>	13 (27)	35 (73)
Age (years), mean (SD)	8.6 (1.9)	8.4 (1.5)
Sex (female), n (%)	6 (46)	18 (51)
<b>Race or ethnicity, n (%)</b>		
Non-Hispanic White	3 (23)	8 (23)
Hispanic White	3 (23)	10 (28)
Non-Hispanic Black	5 (39)	9 (26)
Hispanic Black	0 (0)	1 (3)
Asian	0 (0)	2 (6)
Hispanic, non-Hispanic mix, other, and unknown	2 (15)	5 (14)
<b>Parent, n (%)</b>	13 (27)	35 (73)
Age (years), mean (SD)	40.1 (4)	38.3 (6)
Sex (female), n (%)	13 (100)	33 (94)
<b>Race or ethnicity, n (%)</b>		
Non-Hispanic White	2 (18)	10 (28)
Hispanic White	3 (25)	11 (31)
Non-Hispanic Black	5 (42)	9 (26)
Asian	1 (8)	2 (6)
Hispanic, non-Hispanic mix, other, and unknown	1 (8)	3 (8)
<b>Education, n (%)</b>		
High school graduate	1 (8)	1 (3)
Technical school	1 (8)	4 (11)
Some college	3 (25)	12 (34)
College	4 (33)	9 (26)
Graduate school	3 (25)	9 (26)
<b>Income (US \$), n (%)</b>		
<30,000	0 (0)	6 (17)
<60,000	7 (58)	10 (28)
>60,000	5 (42)	18 (51)
Do not know	0 (0)	1 (3)



## HealthSense Versions Tested and Installation Issues

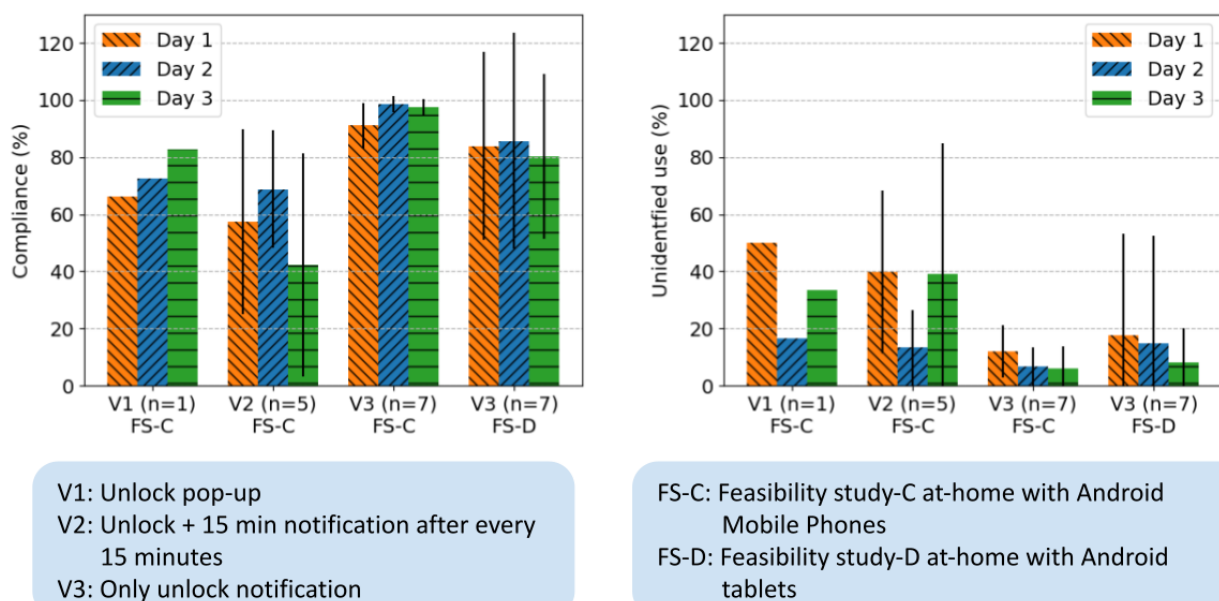
Problems with the installation of the app on participants' mobile devices included the nonappearance of user identification, lack of tracking app use, inability to install HealthSense on the device, problems due to phones and tablets with older Android OS systems (<8), and incompatibility of certain mobile device models with the HealthSense app. In some cases, the reason for malfunction was unknown. A full description of installation issues is provided in [Multimedia Appendix 5](#).

## Usability of the HealthSense App Interface and Participant Compliance

Compliance with the user identification procedure for the at-home feasibility tests increased over the iterations of the

HealthSense app from versions V1 to V3 ([Figure 1](#)). HealthSense V3, which included reminders and encouraging messages sent to participants daily, achieved the highest compliance among users, with 95.7% compliance among smartphone users and 83.3% for tablet users maintained across 3 days of the feasibility study ([Figure 1](#), left graph). [Figure 1](#) also illustrates the percentage of unidentified use in the total mobile device use, which decreased as the compliance approached 100% ([Figure 1](#), right graph). Note that the percentage of unidentified use time could be high even with relatively high compliance, for instance, if someone used the mobile device for a long time without identifying themselves at the start of that use period.

**Figure 1.** Compliance of user identification and percentage of unidentified mobile device use across the different versions of HealthSense. V1: version 1; V2: version 2; V3: version 3.



## Accuracy of the FLASH-Mobile Approach Compared With Gold Standard

The accuracy of the FLASH-mobile approach in capturing children's mobile device use was assessed in the laboratory observation, and it was found that most users (12/15, 80%) had high compliance with the log-in notification, with >80% of use identified, whereas a smaller proportion (3/15, 20%) had very low compliance, with >70% of use unidentified. When the user compliance was low, HealthSense identified app use but could not identify who used the app, resulting in an unidentified use label and disagreement with staff coding between *child* or *other's* use. Hence, a threshold of  $\geq 70\%$  of identified use in the overall use of the mobile device was selected to identify quality assessments for reporting target child's mobile device use. Those with lower identified mobile device use were screened out because their inclusion would lead to discrepancies resulting from user behavior (low compliance) rather than HealthSense app behavior.

For the users with high compliance ( $\geq 70\%$  of app use identified), HealthSense identified the target child's mean use as 21.76 (SD 20.71) minutes, whereas the gold standard was 18.30 (SD 10.77)

minutes. Among the 12 families with high compliance with user identification, HealthSense achieved a 73.6% (range 44.7%-96.14%) mean agreement with human labelers in capturing the target child's mobile use, 78.9% (range 3.55%-97.5%) in capturing other's mobile use, 2.9% (range 0%-17.75%) agreement in capturing unidentified mobile use, and 85.2% (range 38.37%-98.37%) agreement in capturing no mobile use. Among the 3 families with low compliance (<70% of app use identified), mean agreement between human labelers and target child's mobile use was 37% (range 0%-78.29%), 2.1% (range 0%-6.33%) agreement in capturing other's mobile use, 82.7% (range 73.6%-97.17%) agreement in capturing unidentified mobile use, and 95.9% (range 93.82%-98.33%) agreement in capturing no mobile use. The prevalence- and bias-adjusted Cohen  $\kappa$  value for the agreement between HealthSense and the gold standard was 0.711 (SD 0.301), and the intraclass correlation coefficient was 0.714 for the high-compliance users, indicating high agreement between the gold standard and HealthSense estimate [32-35].

The times HealthSense and gold standard disagreed were primarily due to HealthSense either identifying child use as *other* use or no user was identified. The former happened when

either parent or sibling logged in as the user and the child started using the device or the target child started coviewing the mobile device along with their parent or sibling, which staff labeled as *child* use. The latter case happened when the user of the device ignored the log-in prompt, failing to identify the user.

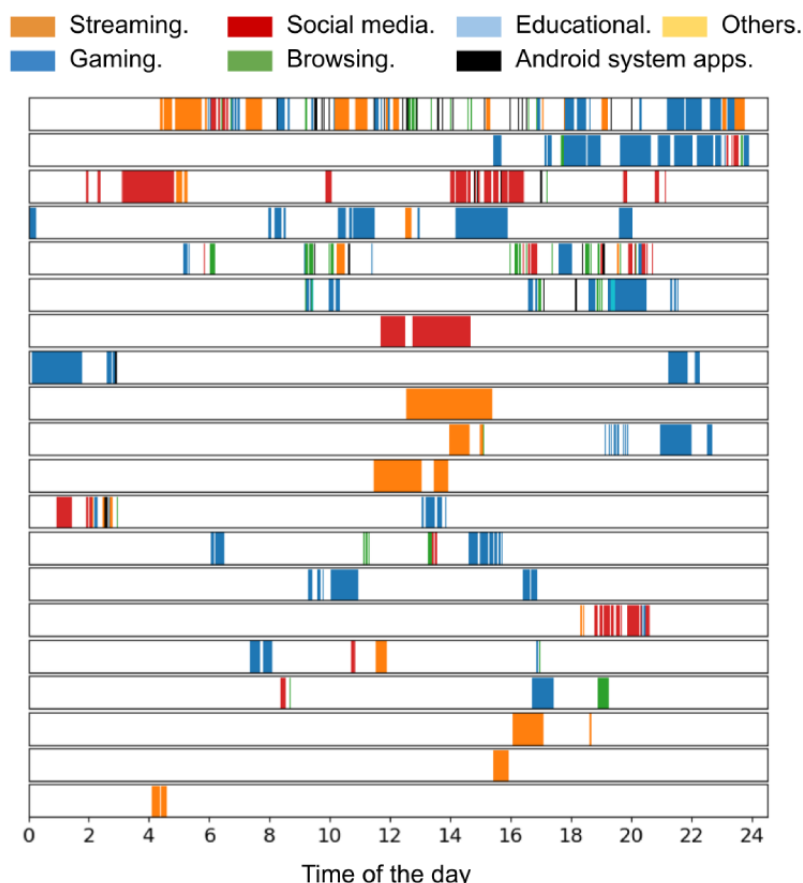
**FLASH-Mobile Approach in Naturalistic Setting**

After qualitatively reviewing the 3-day app and user log, among 23, only 2 (9%) parents reported minor discrepancies in the mobile use reports of their child (one noted a misidentified user

for an app use during the 3-day protocol, and another found 2 apps that they did not recall their child using). The remaining 91% (21/23) of parents agreed with the HealthSense log of their child’s mobile device use.

Figure 2 demonstrates the categories of app use of 20 children from the at-home studies on the second day of the 3-day study; a child (not depicted) had no mobile device use on the second day. The children’s device use duration varied widely, ranging from just half an hour per day to >10 hours a day, with 3 to 7 different apps used by the children during a single day.

**Figure 2.** Mobile device app use by 20 children on day 2 of the 3-day at-home study. This included 20 children from feasibility studies C and D. One child in the full sample (N=21) had no mobile use on day 2 and was, therefore, not depicted.



**Qualitative Feedback From Parents and Children in Exit Interviews**

**Laboratory-Based Studies**

The parents from the observation laboratory studies tended to have positive or neutral feelings toward the FLASH-mobile approach, stating that it seemed fine, did not impact the child’s enjoyment while using the phone, and could allow them to track their child’s screen time. Although most parents had no issues with HealthSense and did not think that it was intrusive, there were a few concerns about privacy, for instance, regarding who would have access to the data and what was actually being recorded.

**At-Home Studies**

Most (18/23, 78%) parents from the at-home studies found the app useful for determining how much time their child spent

using the mobile device and to review the apps they used. More than half (5/8, 63%) of participants who tested the versions with a user identification pop-up and 15-minute reminders (V1 and V2) disliked the pop-up button. They found the pop-up annoying and too frequent. Those who tested V3, with an unlock user identification notification, had mixed reactions. Approximately, half (9/15, 60%) of the participants liked the notification and found no issues with it, but 13% (2/15) of participants stated that the notification was annoying. The remaining parents (4/15, 27%) had a positive reaction to the notification but added that it was difficult to notice. Moreover, 13% (2/15) of participants reported that the app slowed down their phones and did not allow some calls to come through.

Suggestions for the improvement of the FLASH-mobile approach on phones included simplifying the user identification feature; making it available for other devices such as tablets, televisions, and Apple iOS devices; adding a time-restriction

function; adding games or an educational component; having a unique personal identification number or password per user; and blocking phone use until user identification is selected. Others suggested including the user identification requirement only on shared devices. A few parents using mobile phones expressed privacy concerns over their information being sold or stolen and other security risks, and some felt uncomfortable with HealthSense recording not just the child's but also the parent's mobile phone use. None of the parents using tablets expressed similar concerns about privacy.

## Discussion

### Principal Findings

The FLASH-mobile approach, a mobile device tracking app on an Android device, included four features: (1) user identification prompts, (2) app use tracking ability, (3) feedback on compliance with the user identification notification via text or email, and (4) a minimum threshold of user identification compliance to best capture mobile device use by a child. This approach was found to be feasible to measure children's mobile device use. The FLASH-mobile approach addresses differentiating users of shared devices by sending a notification banner for the user to identify themselves every time the mobile device is unlocked along with intermittent communication with participants to praise or remind participants to be compliant. This provided >90% compliance rates with user identification for smartphones and >80% for tablets. In this small feasibility study, the threshold was 70% compliance with user identification, but further studies need to verify whether this threshold continues to correspond to high agreement between the FLASH-mobile approach and gold standard observation. When compliance with user identification reached this threshold, FLASH-mobile approach achieved a 73.6% (SD 16.15%) mean agreement with human labelers in capturing the target child's mobile use (defined as a child actively watching the mobile device's screen alone or with someone else or multitasking with mobile device use and other activities such as playing with toys). There is no agreed-upon definition in the literature for screen use by children. Others have conceptualized screen use of mobile devices as any time a device is used by its owner, regardless of whether the person is viewing the screen [13,36]. The latter definition is difficult to apply to a situation where a phone is shared between a child and another user. However, the definition used for screen use by a study will directly affect the agreement of a tool like FLASH-mobile for actually measuring a child's screen use.

During the course of this study, modifications were made to the FLASH-mobile approach to optimize it and adapt to the continually changing aspects of the Android OS. For example, the full-screen pop-up prompts used in the initial version (V1) of the FLASH-mobile approach are no longer supported by the latest Android OS versions because of new security constraints. Therefore, notification banners were used in the second version (V2), which appeared when the phone was unlocked. In addition, to better identify the user of the device, notification banners were programmed to appear after every 15 minutes of device use. The results from study B (at home) demonstrated that the

15-minute notification was not user-friendly, thus reducing user compliance. Therefore, the final version (V3) of the FLASH-mobile approach included only the unlock notification banner. Such a notification banner is supported by past and current versions of the Android OS. Although a few parents disliked the notification banner, it was well received by most parents and resulted in the lowest proportion of unidentified mobile device use.

Although FLASH-mobile is not publicly available, it could be reproduced using specific Android methods [19-21]. Moreover, similar apps have been developed by others [13,18] and are available for researchers to log device use. The specific features of FLASH-mobile such as user identification notification are straightforward to replicate in any of these Android apps using the methods mentioned [21].

### Comparisons With Prior Work

Only a few other digital tracking approaches have been used to measure children's mobile device use, such as similar tracking apps on Android devices [13], screenshots of the device's battery page in the settings on Apple devices [13], or frequent screenshots of the device [16]. However, all these approaches pose challenges in differentiating a child's use from use by others when the mobile device is shared, which is common among younger children [13]. Only one of these studies assessed children's mobile device use using an approach similar to FLASH-mobile [13]. Radesky et al [13] examined the association between young children's mobile device use and their emotion regulation and executive functioning. Parents who had an Android mobile device (n=126) were instructed to download Chronicle [37], an app similar to HealthSense, for a 9-day study period. Parents with an iPhone (n=220) were asked to take screenshots of the device's battery page in the settings to visualize app use for the past 7 to 10 days. For 71% of the young children who shared the device with someone else, parents completed a form to indicate which apps their child used during the study week. Complete mobile device data were available for 71% of the children, similar to the 69% (33/48) of the participants who could download the HealthSense app with full functionality for the FLASH-mobile approach described here. An important limitation that Radesky et al [13] identified was that they could not identify the user among the 70.6% of the sample who shared the device with someone else; an important issue that the FLASH-mobile approach overcame with the user notification banner. The observational studies (studies A [in laboratory] and B [in laboratory]) for the FLASH-mobile approach additionally demonstrated that when a parent and child were together and used the mobile device, some of that time was shared viewing time. However, most of that occurred appropriately when the child was logged in as the user, ensuring that the use was logged under the child. The duration of time when the parent and child viewed the mobile device together during the observational studies may be artificially high, given that the protocol was performed when the parent and child were together and specifically focused on the screen. It is likely that in free-living conditions, this happens less frequently.

Similar to Radesky et al [13], the FLASH-mobile approach could not be used on all the Android devices tested, and tablets experienced slightly more problems than smartphones. As illustrated in [Multimedia Appendix 5](#), some phones and tablets with older Android OS systems and certain mobile device models did not support the HealthSense app. Inclusion criteria for future studies using background apps that track device and app use should include the use of an Android OS that is newer ( $\geq 7.0$ ), and pretests should be conducted to find which specific Android devices support the tracking app. Future studies also need to further explore the appropriate threshold of compliance with user identification that results in valid and reliable data and how many days of mobile device use are required to capture typical mobile device use by a child. The assessment of physical activity using wearable monitors has adopted similar approaches to develop guidelines for scoring and processing accelerometer data with minimum thresholds of hours and days of wear to be considered valid [38].

### Limitations

This study has several important limitations. The sample was small, recruited through convenience sampling, and thus not representative. However, the sample does reflect a racially and economically diverse group, which is important for demonstrating usefulness for studies targeting diverse children. The small sample with a limited number of days of data collection (3 days) was also not meant to establish typical mobile device use among children. Because the FLASH-mobile approach requires an app, such as HealthSense, Chronicle [37], or Minuku [18], that reads the Android's UsageStatsManager [19] function to track device and app use, it only worked on the Android OS. This may limit the population that can be assessed using this approach. The FLASH-mobile approach also did not work on all Android devices or with all versions of the Android OS. It is important to proactively design studies to screen for mobile devices that support the FLASH-mobile approach. Although the current version of the FLASH-mobile approach leveraged the HealthSense platform for data management, future

versions of FLASH-mobile will need to use alternate platforms because the HealthSense platform will be discontinued in the near future. We are already piloting a version of FLASH-mobile using the Chronicle platform [37].

The low compliance of some participants with the log-in notifications was a challenge, as it resulted in high unidentified app use on the mobile device in earlier versions of testing. In the future, user identification prompts should be made more child-friendly, or they may be turned off when the child does not share the mobile device with others. In addition, future versions of the approach could include automated reminders and encouraging messages for participants to comply with user identification instead of relying on staff. We also found that having a uniform auto screen lock time (eg, 2 minutes) increased the app's ability to accurately identify the user (data not presented). However, the problem of user identification arises only when the mobile device is shared. Many young children now have their own mobile device, including 46% of children aged 2 to 4 years and 67% of children aged 5 to 8 years [12], in which case, apps that read the Android's UsageStatsManager [19] function to track device and app use could more easily track the child's mobile device use.

### Conclusions

The FLASH-mobile approach can be implemented using a mobile device app on an Android device with 4 features: user identification prompts, device and app use tracking ability, feedback on compliance with user identification, and thresholds of compliance to identify valid data. It offers an important new research approach to more accurately measure children's use of mobile devices across one or several days, even when the child shares the device with others. By providing time-stamped data of app use for a specific user, the FLASH-mobile approach will allow for better causal assessments of how the duration, timing, and type of mobile device use by children can affect their developmental and health outcomes. Other researchers can use this approach to further advance the measurement of mobile device use among children.

---

### Acknowledgments

This work was supported by the National Institute of Diabetes and Digestive and Kidney Diseases of the National Institutes of Health (grant R01DK113269) and the United States Department of Agriculture/Agricultural Research Service (cooperative agreement 58-3092-0-001). The contents of this work are solely the responsibility of the authors and do not necessarily represent the official views of the National Institutes of Health or the United States Department of Agriculture.

---

### Conflicts of Interest

None declared.

---

Multimedia Appendix 1

Mobile app interface.

[\[DOCX File, 996 KB - formative\\_v6i10e40452\\_app1.docx \]](#)

---

Multimedia Appendix 2

Classification of apps.

[\[DOCX File, 13 KB - formative\\_v6i10e40452\\_app2.docx \]](#)

## Multimedia Appendix 3

Interview script.

[\[DOCX File, 29 KB - formative\\_v6i10e40452\\_app3.docx\]](#)

## Multimedia Appendix 4

Demographic characteristics of participants with complete and incomplete data.

[\[DOCX File, 21 KB - formative\\_v6i10e40452\\_app4.docx\]](#)

## Multimedia Appendix 5

HealthSense versions tested and problems during study.

[\[DOCX File, 14 KB - formative\\_v6i10e40452\\_app5.docx\]](#)**References**

1. Li C, Cheng G, Sha T, Cheng W, Yan Y. The relationships between screen use and health indicators among infants, toddlers, and preschoolers: a meta-analysis and systematic review. *Int J Environ Res Public Health* 2020 Oct 07;17(19):7324 [FREE Full text] [doi: [10.3390/ijerph17197324](#)] [Medline: [33036443](#)]
2. Hale L, Guan S. Screen time and sleep among school-aged children and adolescents: a systematic literature review. *Sleep Med Rev* 2015 Jun;21:50-58 [FREE Full text] [doi: [10.1016/j.smrv.2014.07.007](#)] [Medline: [25193149](#)]
3. Domingues-Montanari S. Clinical and psychological effects of excessive screen time on children. *J Paediatr Child Health* 2017 Apr;53(4):333-338. [doi: [10.1111/jpc.13462](#)] [Medline: [28168778](#)]
4. Madigan S, Browne D, Racine N, Mori C, Tough S. Association between screen time and children's performance on a developmental screening test. *JAMA Pediatr* 2019 Mar 01;173(3):244-250 [FREE Full text] [doi: [10.1001/jamapediatrics.2018.5056](#)] [Medline: [30688984](#)]
5. Adelantado-Renau M, Moliner-Urdiales D, Caverro-Redondo I, Beltran-Valls MR, Martínez-Vizcaíno V, Álvarez-Bueno C. Association between screen media use and academic performance among children and adolescents: a systematic review and meta-analysis. *JAMA Pediatr* 2019 Nov 01;173(11):1058-1067 [FREE Full text] [doi: [10.1001/jamapediatrics.2019.3176](#)] [Medline: [31545344](#)]
6. Sohn S, Rees P, Wildridge B, Kalk NJ, Carter B. Prevalence of problematic smartphone usage and associated mental health outcomes amongst children and young people: a systematic review, meta-analysis and GRADE of the evidence. *BMC Psychiatry* 2019 Nov 29;19(1):356 [FREE Full text] [doi: [10.1186/s12888-019-2350-x](#)] [Medline: [31779637](#)]
7. Sohn SY, Rees P, Wildridge B, Kalk NJ, Carter B. Correction to: prevalence of problematic smartphone usage and associated mental health outcomes amongst children and young people: a systematic review, meta-analysis and GRADE of the evidence. *BMC Psychiatry* 2021 Jan 22;21(1):52 [FREE Full text] [doi: [10.1186/s12888-020-02986-2](#)] [Medline: [33482751](#)]
8. Carson V, Hunter S, Kuzik N, Gray CE, Poitras VJ, Chaput J, et al. Systematic review of sedentary behaviour and health indicators in school-aged children and youth: an update. *Appl Physiol Nutr Metab* 2016 Jun;41(6 Suppl 3):S240-S265 [FREE Full text] [doi: [10.1139/apnm-2015-0630](#)] [Medline: [27306432](#)]
9. Biddle SJ, García Bengoechea E, Wiesner G. Sedentary behaviour and adiposity in youth: a systematic review of reviews and analysis of causality. *Int J Behav Nutr Phys Act* 2017 Mar 28;14(1):43 [FREE Full text] [doi: [10.1186/s12966-017-0497-8](#)] [Medline: [28351363](#)]
10. Bryant MJ, Lucove JC, Evenson KR, Marshall S. Measurement of television viewing in children and adolescents: a systematic review. *Obes Rev* 2007 May;8(3):197-209. [doi: [10.1111/j.1467-789X.2006.00295.x](#)] [Medline: [17444962](#)]
11. Rideout VJ, Foehr UG, Roberts DF. Generation M2: media in the lives of 8- to 18-year-olds. Kaiser Family Foundation. 2010 Jan 1. URL: <https://www.kff.org/wp-content/uploads/2013/01/8010.pdf> [accessed 2022-10-04]
12. Rideout V, Robb MB. The common sense census: media use by kids age zero to eight. Common Sense Media. 2020. URL: [https://www.commonsensemedia.org/sites/default/files/research/report/2020\\_zero\\_to\\_eight\\_census\\_final\\_web.pdf](https://www.commonsensemedia.org/sites/default/files/research/report/2020_zero_to_eight_census_final_web.pdf) [accessed 2022-10-05]
13. Radesky JS, Weeks HM, Ball R, Schaller A, Yeo S, Durnez J, et al. Young children's use of smartphones and tablets. *Pediatrics* 2020 Jul 01;146(1):e20193518 [FREE Full text] [doi: [10.1542/peds.2019-3518](#)] [Medline: [32482771](#)]
14. Vadathya AK, Musaad S, Beltran A, Perez O, Meister L, Baranowski T, et al. An objective system for quantitative assessment of television viewing among children (family level assessment of screen use in the home-television): system development study. *JMIR Pediatr Parent* 2022 Mar 24;5(1):e33569 [FREE Full text] [doi: [10.2196/33569](#)] [Medline: [35323113](#)]
15. Wilcockson TD, Ellis DA, Shaw H. Determining typical smartphone usage: what data do we need? *Cyberpsychol Behav Soc Netw* 2018 Jun;21(6):395-398. [doi: [10.1089/cyber.2017.0652](#)] [Medline: [29781716](#)]
16. Ram N, Yang X, Cho M, Brinberg M, Muirhead F, Reeves B, et al. Screenomics: a new approach for observing and studying individuals' digital lives. *J Adolesc Res* 2020 Jan;35(1):16-50 [FREE Full text] [doi: [10.1177/0743558419883362](#)] [Medline: [32161431](#)]
17. Curtis A, Pai A, Cao J, Moukaddam N, Sabharwal A. HealthSense: software-defined mobile-based clinical trials. In: *Proceedings of the 25th Annual International Conference on Mobile Computing and Networking*. 2019 Presented at:

- MobiCom '19: The 25th Annual International Conference on Mobile Computing and Networking; Oct 21 - 25, 2019; Los Cabos Mexico. [doi: [10.1145/3300061.3345433](https://doi.org/10.1145/3300061.3345433)]
18. Yuan N, Weeks HM, Ball R, Newman MW, Chang Y, Radesky JS. How much do parents actually use their smartphones? Pilot study comparing self-report to passive sensing. *Pediatr Res* 2019 Oct;86(4):416-418 [FREE Full text] [doi: [10.1038/s41390-019-0452-2](https://doi.org/10.1038/s41390-019-0452-2)] [Medline: [31195401](https://pubmed.ncbi.nlm.nih.gov/31195401/)]
  19. UsageStatsManager. Developers. URL: <https://developer.android.com/reference/android/app/usage/UsageStatsManager> [accessed 2022-10-04]
  20. UsageEvents. Developers. URL: <https://developer.android.google.cn/reference/android/app/usage/UsageEvents> [accessed 2022-10-04]
  21. ACTION\_USER\_PRESENT. Developers. URL: [https://developer.android.com/reference/android/content/Intent#ACTION\\_USER\\_PRESENT](https://developer.android.com/reference/android/content/Intent#ACTION_USER_PRESENT) [accessed 2022-10-04]
  22. Madigan S, McArthur BA, Anhorn C, Eirich R, Christakis DA. Associations between screen use and child language skills: a systematic review and meta-analysis. *JAMA Pediatr* 2020 Jul 01;174(7):665-675 [FREE Full text] [doi: [10.1001/jamapediatrics.2020.0327](https://doi.org/10.1001/jamapediatrics.2020.0327)] [Medline: [32202633](https://pubmed.ncbi.nlm.nih.gov/32202633/)]
  23. Lin L, Cranton P, Lee J. Research methodologies for multitasking studies. In: *Handbook of Research on Scholarly Publishing and Research Methods*. Hershey, PA: IGI Global; 2015.
  24. Song KS, Nam S, Lim H, Kim J. Analysis of youngsters' media multitasking behaviors and effect on learning. *Intl J Multimed Ubiquit Eng* 2013;8(4):191-198.
  25. Hawkins RP, Pingree S, Bruce L, Tapper J. Strategy and style in attention to television. *J Broadcast Electron Media* 2009 May 18;41(2):245-264. [doi: [10.1080/08838159709364404](https://doi.org/10.1080/08838159709364404)]
  26. O'Connor TM, Elias C, Thompson RR, Arlinghaus KR, Vollrath K, Hernandez DC, et al. The association of TV viewing during dinner meals with quality of dietary intake and BMI z-scores among low income, ethnic minority preschool children. *Appetite* 2019 Sep 01;140:231-238. [doi: [10.1016/j.appet.2019.05.023](https://doi.org/10.1016/j.appet.2019.05.023)] [Medline: [31121200](https://pubmed.ncbi.nlm.nih.gov/31121200/)]
  27. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977 Mar;33(1):159-174. [Medline: [843571](https://pubmed.ncbi.nlm.nih.gov/843571/)]
  28. Mandrekar JN. Measures of interrater agreement. *J Thorac Oncol* 2011 Jan;6(1):6-7 [FREE Full text] [doi: [10.1097/JTO.0b013e318200f983](https://doi.org/10.1097/JTO.0b013e318200f983)] [Medline: [21178713](https://pubmed.ncbi.nlm.nih.gov/21178713/)]
  29. Albaum MN, Hill LC, Murphy M, Li YH, Fuhrman CR, Britton CA, et al. Interobserver reliability of the chest radiograph in community-acquired pneumonia. PORT investigators. *Chest* 1996 Aug;110(2):343-350. [doi: [10.1378/chest.110.2.343](https://doi.org/10.1378/chest.110.2.343)] [Medline: [8697831](https://pubmed.ncbi.nlm.nih.gov/8697831/)]
  30. Thomas DR. A general inductive approach for analyzing qualitative evaluation data. *Am J Eval* 2016 Jun 30;27(2):237-246. [doi: [10.1177/1098214005283748](https://doi.org/10.1177/1098214005283748)]
  31. 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)]
  32. Bhatnagar G, Quinn L, Higginson A, Plumb A, Halligan S, Tolan D, METRIC study investigators. Observer agreement for small bowel ultrasound in Crohn's disease: results from the METRIC trial. *Abdom Radiol (NY)* 2020 Oct;45(10):3036-3045 [FREE Full text] [doi: [10.1007/s00261-020-02405-w](https://doi.org/10.1007/s00261-020-02405-w)] [Medline: [32037466](https://pubmed.ncbi.nlm.nih.gov/32037466/)]
  33. Heal C, Cotterill S, Rowland AG, Garratt N, Long T, Brown S, et al. Inter-rater reliability of paediatric emergency assessment: physiological and clinical features. *Arch Dis Child* 2021 Feb;106(2):149-153 [FREE Full text] [doi: [10.1136/archdischild-2019-318664](https://doi.org/10.1136/archdischild-2019-318664)] [Medline: [32988814](https://pubmed.ncbi.nlm.nih.gov/32988814/)]
  34. Xue C, Yuan J, Lo GG, Chang AT, Poon DM, Wong OL, et al. Radiomics feature reliability assessed by intraclass correlation coefficient: a systematic review. *Quant Imaging Med Surg* 2021 Oct;11(10):4431-4460 [FREE Full text] [doi: [10.21037/qims-21-86](https://doi.org/10.21037/qims-21-86)] [Medline: [34603997](https://pubmed.ncbi.nlm.nih.gov/34603997/)]
  35. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropractic Med* 2016 Jun;15(2):155-163 [FREE Full text] [doi: [10.1016/j.jcm.2016.02.012](https://doi.org/10.1016/j.jcm.2016.02.012)] [Medline: [27330520](https://pubmed.ncbi.nlm.nih.gov/27330520/)]
  36. Kristensen PL, Olesen LG, Egebæk HK, Pedersen J, Rasmussen MG, Grøntved A. Criterion validity of a research-based application for tracking screen time on android and iOS smartphones and tablets. *Comput Human Behav Reports* 2022 Mar;5:100164. [doi: [10.1016/j.chbr.2021.100164](https://doi.org/10.1016/j.chbr.2021.100164)]
  37. Chronicle: the end-to-end platform for behavioral research. OpenLattice. URL: <https://openlattice.com/> [accessed 2022-10-03]
  38. Adams MA, Johnson WD, Tudor-Locke C. Steps/day translation of the moderate-to-vigorous physical activity guideline for children and adolescents. *Int J Behav Nutr Phys Act* 2013 Apr 21;10:49 [FREE Full text] [doi: [10.1186/1479-5868-10-49](https://doi.org/10.1186/1479-5868-10-49)] [Medline: [23601369](https://pubmed.ncbi.nlm.nih.gov/23601369/)]

## Abbreviations

- FLASH:** Family Level Assessment of Screen Use  
**OS:** operating system  
**V1:** version 1  
**V2:** version 2

**V3:** version 3

*Edited by A Mavragani; submitted 22.06.22; peer-reviewed by B Armstrong, K Pedersen; comments to author 28.07.22; revised version received 30.08.22; accepted 15.09.22; published 21.10.22.*

*Please cite as:*

*Perez O, Kumar Vadathya A, Beltran A, Barnett RM, Hinderer O, Garza T, Musaad SM, Baranowski T, Hughes SO, Mendoza JA, Sabharwal A, Veeraraghavan A, O'Connor TM*

*The Family Level Assessment of Screen Use–Mobile Approach: Development of an Approach to Measure Children’s Mobile Device Use*

*JMIR Form Res 2022;6(10):e40452*

URL: <https://formative.jmir.org/2022/10/e40452>

doi: [10.2196/40452](https://doi.org/10.2196/40452)

PMID: [36269651](https://pubmed.ncbi.nlm.nih.gov/36269651/)

©Oriana Perez, Anil Kumar Vadathya, Alicia Beltran, R Matthew Barnett, Olivia Hinderer, Tatyana Garza, Salma M Musaad, Tom Baranowski, Sheryl O Hughes, Jason A Mendoza, Ashutosh Sabharwal, Ashok Veeraraghavan, Teresia M O'Connor. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 21.10.2022. 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 Tailored Gender-Sensitive mHealth Weight Loss Intervention (I-GENDO): Development and Process Evaluation

Magdalena Pape<sup>1,2</sup>, MSc; Tanja Färber<sup>3</sup>, MSc; Caroline Seiferth<sup>2</sup>, MSc; Tanja Roth<sup>1</sup>, BSc; Stefanie Schroeder<sup>2,3</sup>, DPhil; Joerg Wolstein<sup>3</sup>, PhD; Stephan Herpertz<sup>1</sup>, PhD; Sabine Steins-Loeber<sup>2</sup>, PhD

<sup>1</sup>Department of Psychosomatic Medicine and Psychotherapy, LWL-University Hospital of the Ruhr-University Bochum, Bochum, Germany

<sup>2</sup>Department of Clinical Psychology and Psychotherapy, University of Bamberg, Bamberg, Germany

<sup>3</sup>Department of Pathopsychology, University of Bamberg, Bamberg, Germany

**Corresponding Author:**

Magdalena Pape, MSc

Department of Psychosomatic Medicine and Psychotherapy

LWL-University Hospital of the Ruhr-University Bochum

Alexandrinstraße 1-3

Bochum, 44791

Germany

Phone: 49 0234 509 ext 59039

Email: [magdalena.pape@rub.de](mailto:magdalena.pape@rub.de)

## Abstract

**Background:** Given the increase in the prevalence of overweight and obesity worldwide, the number of digital weight loss interventions has also risen. However, these interventions often lack theoretical background and data on long-term effectiveness. The consideration of individual and gender differences in weight-related psychological parameters might enhance the efficacy and sustainability of mobile-based weight loss interventions.

**Objective:** This paper presented an introduction to and the process evaluation of a 12-week gender-sensitive mobile health (mHealth) weight loss intervention (I-GENDO) combining computer-based and self-tailoring features.

**Methods:** Between August 2020 and August 2021, individuals with overweight (BMI 25.0-29.9 kg/m<sup>2</sup>), those with obesity class I (BMI 30.0-34.9 kg/m<sup>2</sup>), and those with obesity class II (BMI 35.0-39.9 kg/m<sup>2</sup>) were recruited to the I-GENDO project, a multicenter study in Germany. The mHealth intervention aimed at targeting individual psychological factors associated with the development and persistence of overweight and obesity (eg, emotional eating) using computer-based tailoring. Moreover, the intervention took a gender-sensitive approach by implementing self-tailoring of gender-targeted module versions. The computer-based assignment of the main modules, self-selection of gender-targeted module versions, and use patterns were evaluated while considering gender. Moreover, gender differences in the usability assessment were analyzed.

**Results:** Data from the intervention arm of the study were processed. A total of 116 individuals with overweight and obesity (77/116, 66.4% women; age mean 47.28, SD 11.66 years; BMI mean 33.58, SD 3.79 kg/m<sup>2</sup>) were included in the analyses. Overall, the compliance (90/109, 82.6%) and satisfaction with the app (mean 86% approval) were high and comparable with those of other mobile weight loss interventions. The usability of the intervention was rated with 71% (5.0/7.0 points) satisfaction. More women obtained the main module that focused on emotion regulation skills. Most men and women selected women-targeted versions of the main modules. Women used the app more frequently and longer than men. However, women and men did not differ in the progress of use patterns throughout the intervention.

**Conclusions:** We developed a tailored gender-sensitive mHealth weight loss intervention. The usability of and engagement with the intervention were satisfactory, and the overall satisfaction with the intervention was also high. Gender differences must be considered in the evaluation of the effectiveness and sustainability of the intervention.

(JMIR Form Res 2022;6(10):e38480) doi:[10.2196/38480](https://doi.org/10.2196/38480)

**KEYWORDS**

mobile health; mHealth; eHealth; tailoring; gender; weight loss intervention; mobile phone



## Introduction

Within the last few decades, a vast number of digital health apps have been developed worldwide [1,2]. eHealth interventions (ie, mobile health [mHealth] interventions) are cost-effective and feasible in everyday life and represent a useful addition to analog health care services, not only in times of a worldwide pandemic [3]. In 2021, 87% of German adults and adolescents aged >14 years owned a smartphone, and 27% reported using mHealth interventions regularly [4]. The use of mHealth interventions requires an active and self-determined engagement of the user and therefore facilitates behavioral changes [5]. For example, mHealth lifestyle interventions show good efficacy in promoting healthy behaviors such as dietary intake and physical activity [6-10]. Therefore, they are promising tools that could promote behavioral change in participants wishing to reduce weight [11]. However, most available interventions to date demonstrate only short-term effects of behavioral change, whereas long-term effectiveness, especially regarding weight loss, has either not been investigated or not been demonstrated [12-14]. An explanation for the lack of effects is that most weight loss apps have not been developed from a scientific background and thus lack sufficient consideration of psychological evidence-based strategies [15], which are an important aspect of effective weight loss programs (WLPs) according to international guidelines [16,17]. Moreover, most weight loss apps have been developed on a *one-size-fits-all* approach, despite indications from prior studies that targeted (tailored) interventions are more effective [18-20].

The term “tailoring” refers to the customization of a feature of an intervention based on the individual characteristics of the participants [21]. The participants might customize an intervention based on their own preferences (self-tailoring), or they might receive individualized interventions in which the program tailors the content, usually based on algorithms (computer-based tailoring). In the latter case, tailoring can be based on data from 1 assessment (static tailoring) or adapted to different assessments within an intervention process (dynamic tailoring). Studies have indicated that participants feel more strongly addressed by individualized interventions, are more satisfied with them, and are subsequently more engaged in their use, which enhances the efficacy of the programs [6,11,19,22-25]. Various psychological aspects are involved in the development and maintenance of overweight and obesity, including the experience of weight-related stigmatization [26], maladaptive coping strategies [27], or dysfunctional eating behaviors [28]. Therefore, developing computer-based tailoring features that consider such psychological aspects might be a key element in the optimization of digital WLPs.

Gender differences in the development and treatment of obesity and overweight have also been investigated [29,30]. In Germany, more men (43.3%) develop overweight (BMI 25-29.9 kg/m<sup>2</sup>) compared with women (28.8%), but there are no gender differences in the prevalence of obesity (BMI >30.0 kg/m<sup>2</sup>), with increasing prevalence rates in the past decades among both genders [31,32]. Men with overweight and obesity are less likely to accurately perceive their weight and are less dissatisfied with their overweight status [29]. Moreover, gender differences in

physical activity, eating behavior, and weight-related psychological parameters have been reported. For example, women engage more often in problematic eating behaviors, such as emotional eating (EE) and craving of special foods than do men [33]. Women consistently report higher levels of perceived stress and engage more in emotion-focused coping, such as rumination, whereas men often use problem-focused or avoidant coping strategies [34,35]. On average, men are more physically active [36]. Some biological sex differences have been published; for instance, in males, fat depositions are often in the visceral depot, which increases their risk for cardiovascular disorders [37-39]. More women participate in WLPs, yet the participating men lose more absolute weight [40,41]. Results on the adherence to WLPs are heterogeneous, depending on the intervention type, among other factors [42-44]. On the basis of reviewed studies, investigating the effect of gender on overweight and obesity outcomes to improve the effectiveness of WLPs is an important research agenda. A recently published meta-analysis comparing the effects of gender-targeted and gender-neutral WLPs however revealed no differences in weight-related outcomes, although gender-targeted interventions were more effective in promoting activity and improving nutrition [45]. However, the included *gender-targeted* WLPs were offered either to male or to female participants based on sex. We support the idea that psychological interventions should be gender sensitive instead of gender dichotomous and assume an increase in the effectiveness of the intervention if it is gender sensitive [46]. Therefore, to avoid prejudiced gender-based distinctions between individuals with overweight and obesity, we recommend implementing gender-sensitive self-tailoring features.

Against this background, we aimed at developing a smartphone-based psychological and gender sensitive weight-loss intervention with computer-based and self-tailoring features. In the first part of this paper, we have described the development process of the app with particular focus on the tailoring features of the intervention. The subsequent process evaluation focuses on the evaluation of the app with regard to the psychological and gender-sensitive tailoring features, use patterns, and satisfaction with the app derived from a sample of 116 participants taking part in the I-GENDO project [47].

## Methods

### The I-GENDO Project

The project “Gender-sensitive enhancement of common weight-loss strategies for overweight and obesity: A personalized smartphone app” was proposed by the University of Bamberg, Departments of Clinical Psychology and Psychotherapy and Pathopsychology, in cooperation with LWL-University Hospital of Ruhr-University Bochum, Department of Psychosomatic Medicine and Psychotherapy, and funded by the Federal Ministry of Education and Research of Germany (01GL1719A/B). The project was preregistered (ClinicalTrials.gov identifier: NCT04080193).

### Ethical Considerations

This study was conducted in accordance with the Declaration of Helsinki. The Institutional Review Board of Ruhr-University

Bochum approved the study (number 18-6415). All participants were informed about the study and provided written informed consent.

### Development of the mHealth Intervention I-GENDO

From September 2017 to November 2019, a modular app system was developed at the University of Bamberg in cooperation with an external software provider (groupXS Solutions GmbH).

**Figure 1.** The I-GENDO app interface.

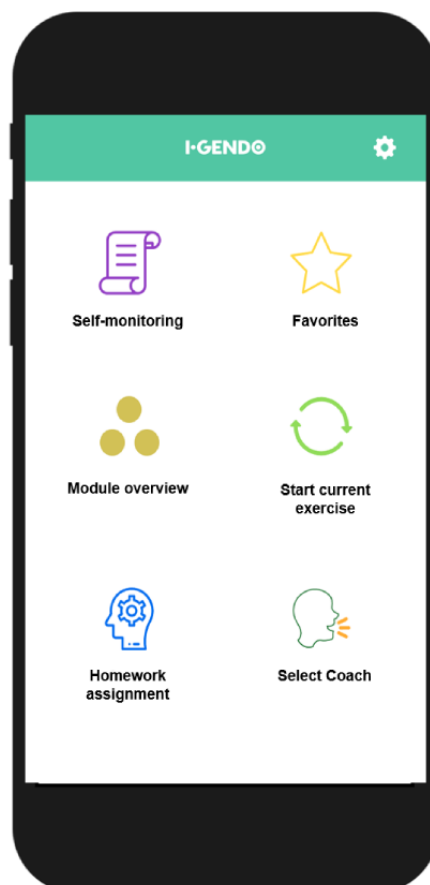


Figure 1 provides an overview of the I-GENDO app interface. The app provided the following elements: module-based psychological intervention; selection of an accompanying coach; and self-monitoring of hunger, appetite, and mood.

The content of the modules was based on the existing evidence-based manuals, qualitative data from focus groups of individuals with overweight and obesity, and interviews with experts in the field of psychological treatment of obesity. To implement a gender-sensitive approach, extensive literature reviews were conducted on the disparities between genders in the psychological and behavioral aspects of obesity treatment. Furthermore, a steering committee consisting of experts in the field of prevention and treatment of overweight and obesity, digital transformation, and qualitative data analyses was formed. All principal decisions regarding app development were made in consensus with the members of the steering committee.

On the basis of this information, 7 modules that served as the heart of the 12-week I-GENDO intervention were constructed. Of the 7 modules, 2 modules addressed the introduction (goal setting) and conclusion (relapse prevention strategies) of the intervention. The remaining 5 modules (main modules) focused on different psychological parameters associated with the development and maintenance of overweight and obesity: stress management skills (*stress module*), emotion regulation skills

(*emotion module*), dealing with the consequences of overweight (*consequences module*), self-regulation skills (*control module*), and self-efficacy (*self-efficacy module*). Each module contained 9 sessions, which included psychoeducational elements delivered through texts and videos, several therapeutic tools from different therapeutic approaches (ie, cognitive behavioral therapy, dialectical behavioral therapy, and mindfulness), and various behavior change techniques [48]. These sessions could be repeated as many times as desired, and users could set a short link to their favorite exercises via the toolbox.

Each module was presented in either a women-targeted version (*version A*) or a men-targeted version (*version B*), which differed in terms of knowledge transfer, communication style, and prioritization of topics. For example, in the *stress module*, this was achieved using appealing case examples in the women-targeted version and fact presenting in the men-targeted version to transfer general knowledge about stress. Another example is that the men-targeted version in the *emotion module* highlighted and trained the recognition and labeling of emotions, whereas in the women-targeted version, the association between

dysfunctional beliefs and eating behavior was prioritized. [Multimedia Appendix 1](#) [48-77] provides an overview of the operationalization of the gender-sensitive modules and the origin of evidence. The versions were briefly introduced, with both introductions presented on 1 screen page. Participants could then freely choose between *version A* or *B* regardless of biological sex (gender-sensitive instead of gender dichotomous tailoring). Participants were blind to the manipulation of the gender-targeted versions.

## Process Evaluation of the mHealth Intervention I-GENDO

From December 2019 to December 2021, the effectiveness of the 12-week I-GENDO intervention was evaluated in a randomized controlled trial conducted at the University of Bamberg and LWL-University Hospital Bochum, Department of Psychosomatic Medicine and Psychotherapy (ClinicalTrials.gov Identifier: NCT04080193). The main results of the randomized controlled trial will be published elsewhere. In this manuscript, the relevant process evaluation data from the intervention arm were analyzed.

## Study Sample

Individuals were informed about the I-GENDO project via newspaper articles, radio features, and oral presentations at rehabilitation centers. Participants interested in the study were screened for eligibility ([Textbox 1](#)) and, if eligible, were invited to participate. According to the guidelines of the German Association for the Study of Obesity and the German Society for General and Visceral Surgery, individuals with obesity class III (BMI >39.9 kg/m<sup>2</sup>) experience a complex multifactorial framework of severe social, mental, and physical problems and are recommended to undergo bariatric surgery. Therefore, we excluded individuals with obesity class III from participation but provided further support. Because the effect of bariatric surgery on weight loss is mainly driven by physical limitations and varies significantly between the types of operative procedure [78], we decided to exclude individuals who underwent or planned to undergo bariatric surgery. The total study sample consisted of 213 individuals with overweight and obesity, of which 116 (n=77, 66.4% women) were randomly assigned to the intervention group for this study and subsequently included in this analysis.

**Textbox 1.** Eligibility criteria of the I-GENDO project.

### Inclusion criteria

- Legal age (≥18 years)
- Obesity class I or II with subjectively experienced weight-related impairment and a current intention to lose weight
- Overweight (ie, BMI between 25 and 29.9 kg/m<sup>2</sup>) with weight-related health problems, visceral adipose tissue, or high psychosocial weight-related distress and a current intention to lose weight

### Exclusion criteria

- Obesity class III (ie, BMI >39.9 kg/m<sup>2</sup>)
- Current (or within the last 12 months) involvement in a structured weight loss intervention
- Insulin-dependent type 1 diabetes
- Previous or intended bariatric surgery
- Current psychotherapeutic treatment of weight-related health problems
- Weight-enhancing drugs
- Drugs that promote weight loss (eg, antiobesity drugs)
- Weight-enhancing health problems that are not yet treated
- Cancerous disease within the last 5 years
- Current substance-use disorders, major depression, psychosis, suicidal tendency, or pregnancy
- Severe cognitive impairments
- Insufficient knowledge of the German language
- Binge-eating disorder or bulimia nervosa

## Intervention Phase

Participants in the intervention group received a 12-week tailored app intervention. In the first week of intervention, the introduction module was unlocked for each participant, followed by 9 weeks of tailored intervention comprising 3 of the 5 main modules. Each session of the 3 main modules was unlocked successively between weeks 2 and 9. The basic, minimal content of the remaining 2 modules was provided in the form of mini

modules, which were unlocked in week 11. Finally, the conclusion module was provided to each participant in week 12.

## Tailoring

[Figure 2](#) displays computer-based and self-tailoring features of the intervention. The introduction and conclusion module were mandatory elements framing the intervention that conveyed general content, whereas the main modules targeted individual

differences in weight-related psychological parameters. The main module assignment was computer-based and depended on the results of the Revised Illness Perception Questionnaire (IPQ-R), a standardized questionnaire adapted to overweight and obesity that measures illness beliefs (eg, “my overweight strongly affects the way others see me”) and causal attribution of overweight (eg, “my emotional state, e.g. feeling down, lonely, anxious, empty”) [79]. Participants completed the IPQ-R at the baseline assessment. Each of the 32 items were rated on a 5-point Likert scale ranging from 1 (“strongly disagree”) to 5 (“strongly agree”). In this study, the internal consistency of the scale was good (Cronbach  $\alpha=.714$ ). Scales were regenerated with higher means representing severe problems on the related psychological parameters associated with overweight and obesity (eg, EE). Of the 5 dimensions, 3 on which the

participants reported the highest impairments were tailored to the participants (computer-based tailoring). In addition to the computer-based tailoring feature, individual adaption of content and functions was enabled (self-tailoring). Each module was presented in either a men-specific (*version B*) or a women-specific version (*version A*; “App features” section and [Multimedia Appendix 1](#)). The app additionally contained customization features to enhance the adherence to the intervention [80]. In particular, the participants could choose between different coaches at the beginning of the 12-week intervention. A total of 4 different coaches were introduced: 2 men and 2 women coaches depicted as being either more friendly (eg, informal and motivating tips) or more professional (eg, formal and directive tips).

**Figure 2.** Tailoring features of the I-GENDO intervention. Of the 5 main modules (in the box), 3 were assigned to the participants based on the results of the revised illness perception questionnaire (computer-based tailoring). Each of the modules was presented in either a women- or men-targeted version (self-tailoring).



## Measurements

### Engagement With the App

Use patterns were retrieved from individual app data and subsequently analyzed. Actions were defined as time slots of active engagement with the app, for example, log-in to the app and processing a session (use frequency). Inactivity for 20 minutes defined the completion of one action. The overall app use time was calculated in minutes (use time). The participants who used the app at least 12 times (actions) and for 120 minutes within the 12-week intervention were defined as being compliant with the I-GENDO app.

### Satisfaction With the App

At the end of the conclusion module, the users could give feedback about their satisfaction with the app and the relevance and daily usefulness of the app on scales ranging from 0 (“not at all”) to 100 (“very much”). In the last session of each module, participants could evaluate how satisfied they were with the corresponding module.

### Usability Rating of the App

After the 12-week intervention, the mHealth App Usability Questionnaire for stand-alone mHealth apps used by patients was administered [81]. The original English questionnaire was translated into German by a member of the research group and retranslated by a native speaker. Deviations were discussed and subsequently adjusted. The self-report questionnaire consisted

of 18 items, which were scored on a scale from 1 (“strongly disagree”) to 7 (“strongly agree”), with higher means reflecting higher usability. Prior research indicated good psychometric properties of the English version of the mHealth App Usability Questionnaire [81]. In this study, the internal consistency of the total scale was excellent (Cronbach  $\alpha=.935$ ).

## Data Analysis

All analyses were conducted using SPSS for Windows (version 26.0; IBM Corp) and Excel (version 16.0; Microsoft Corp). App data were retrieved from Apache CouchDBTM. Descriptive analyses were conducted using percentages and frequencies for categorical variables and means and SDs for continuous variables. *Chi-square* distributions that compared categorical variables between genders were implemented, and Bonferroni-adjusted independent 2-tailed *t* tests were conducted to compare metrically scaled variables. Mann-Whitney *U* tests were conducted to compare results between genders on nonnormally-distributed variables. Friedman tests and Dunn-Bonferroni post hoc tests were implemented to compare app engagement between genders over the 12 weeks of intervention.

## Results

### Participants

We found no significant gender differences in age, BMI, marital status, and education level at baseline ([Table 1](#)).

**Table 1.** Sociodemographic factors (N=116).

Characteristic	Overall	Women (n=77)	Men (n=39)	Women vs men		
				2-tailed <i>t</i> test ( <i>df</i> )	Chi-square ( <i>df</i> )	<i>P</i> value <sup>a</sup>
Age (years), mean (SD)	47.28 (11.66)	46.40 (12.22)	49.00 (10.38)	1.14 (114)	N/A <sup>b</sup>	.26
BMI (kg/m <sup>2</sup> ), mean (SD)	33.58 (3.79)	33.75 (3.69)	33.23 (4.02)	0.70 (114)	N/A	.49
Marital status (yes), n (%) <sup>c</sup>	91 (78.4)	57 (74)	34 (87)	N/A	1.9 (1)	.17
Education (university), n (%) <sup>d</sup>	36 (31)	25 (32)	11 (28)	N/A	0.1 (1)	.80

<sup>a</sup>Bonferroni-adjusted *P* values.

<sup>b</sup>N/A: not applicable.

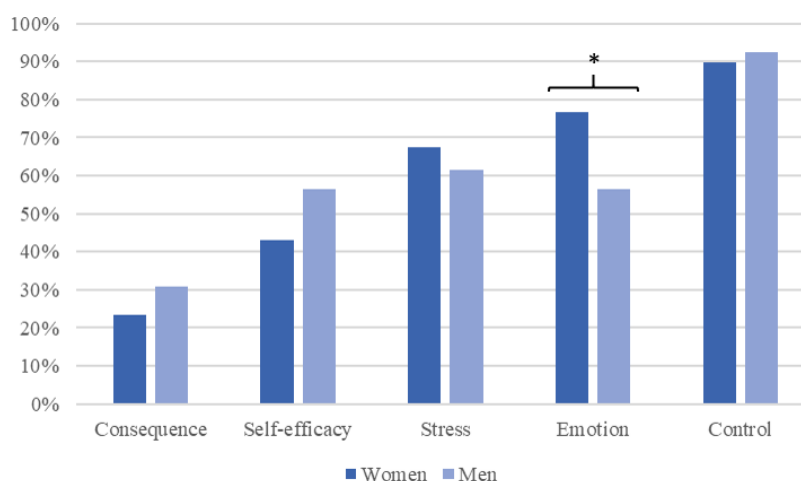
<sup>c</sup>Number of participants in a relationship.

<sup>d</sup>Number of participants with a university degree.

## Tailoring

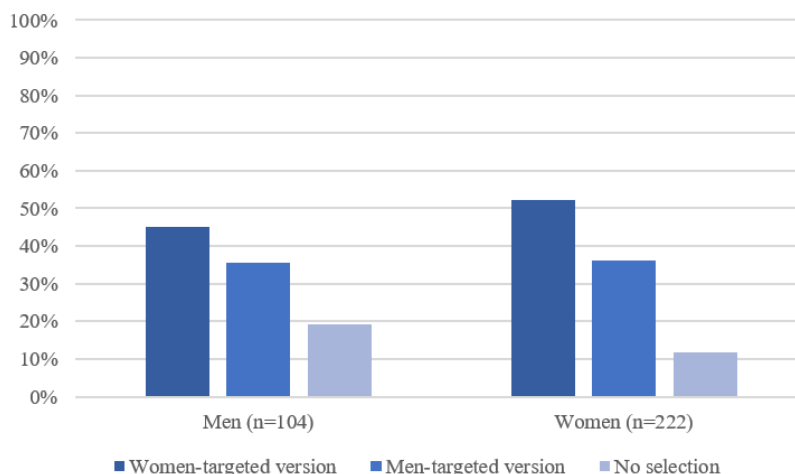
Three main modules were tailored to each of the 116 participants by computer-based tailoring according to their IPQ-R results (see the section *Tailoring*). Most participants (105/116, 90.5%) received the *control module*, followed by the *emotion module* (81/116, 69.8%), *stress module* (76/116, 65.5%), and *self-efficacy module* (55/116, 47.4%). One-quarter of the

participants (30/116, 25.9%) received the *consequence module*. Figure 3 illustrates the module assignments for the participating men and women separately. Significantly more women obtained the *emotion module* than men ( $\chi^2_1=4.1$ ;  $P=.04$ ;  $\phi=0.21$ ). The genders did not differ in the assignment of the *consequence* ( $\chi^2_1=0.4$ ;  $P=.53$ ), *self-efficacy* ( $\chi^2_1=1.6$ ;  $P=.23$ ), *stress* ( $\chi^2_1=0.2$ ;  $P=.66$ ), or *control module* ( $\chi^2_1=0.02$ ;  $P=.89$ ).

**Figure 3.** Assigned full-version modules (computer-based tailoring) in percentage (\* $P<.005$ ).

As described earlier, at the beginning of each module, the participants were instructed to choose between either a women-targeted or a men-targeted version (self-tailoring). In 50% (163/326) of the choices, the women-targeted versions were selected (women: 116/222, 52.3%; men: 47/104, 45.2%). In 35.9% (117/326) of the choices, the men-targeted versions

were selected (women: 80/222, 36%; men: 37/104, 35.6%). In the remaining 14.1% (46/326) of the choices, no selection was made (Figure 4). When the participants did choose a version, they chose version A 58.2% (163/280) of the time (women: 116/196, 59.2%; men: 47/84, 56%).

**Figure 4.** Module version assignments (self-tailoring) in percentage (total choices: N=326).

Another customization feature of the intervention was the selection of an accompanying coach when starting the app for the first time. Most women (35/74, 47%) chose a friendly woman coach, 19% (14/74) chose a professional man coach, 18% (13/74) chose a friendly man coach, and 16% (12/74) chose a professional woman coach. Coach assessment in men was more balanced, with 34% (12/35) choosing a friendly woman coach, 23% (8/35) choosing a friendly man or professional woman coach, and 20% (7/35) choosing a professional man coach. No significant gender differences were found in coach assessment ( $\chi^2_3=1.9$ ;  $P=.60$ ).

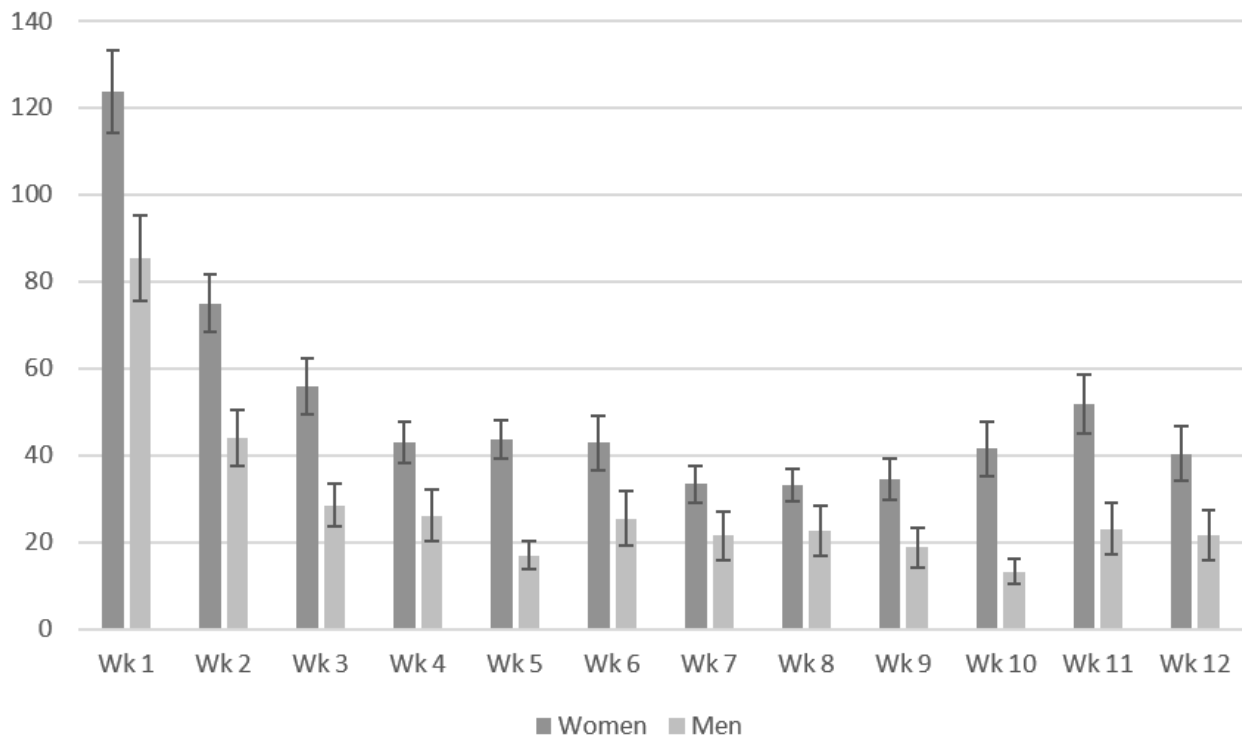
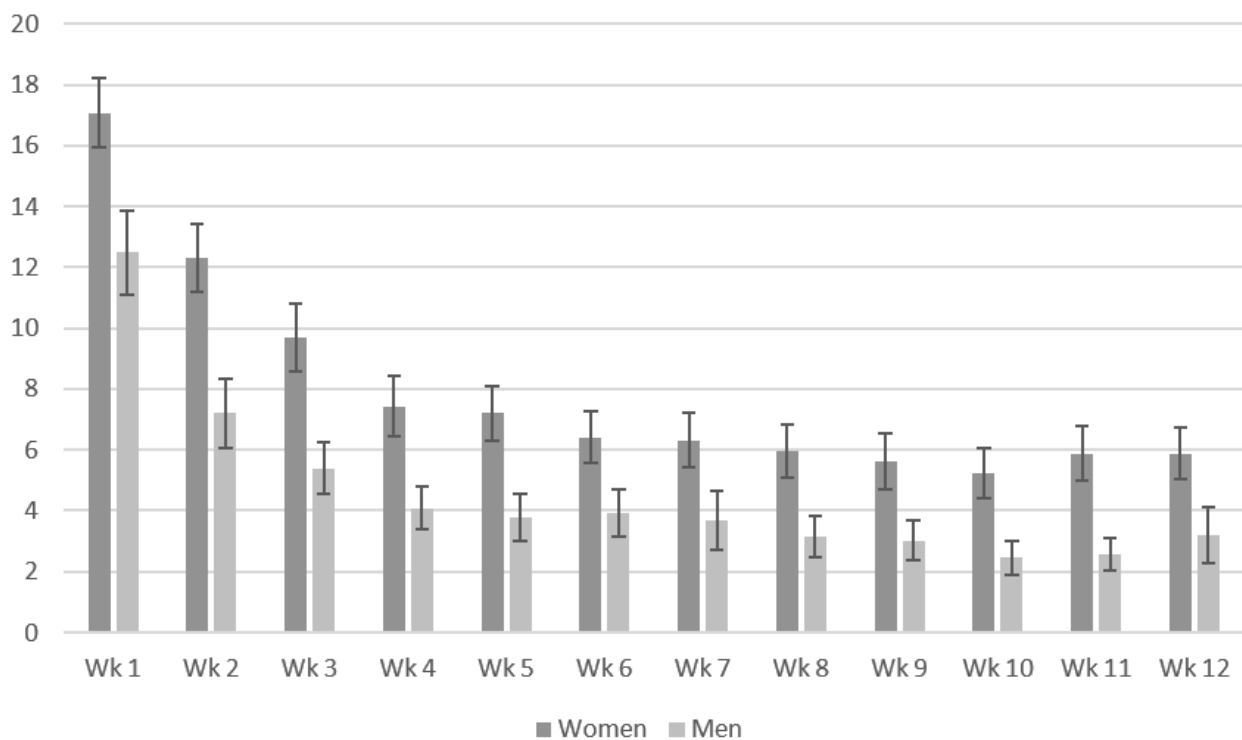
### Engagement With the App

Of the 116 participants in the intervention group, 109 actively participated in the app intervention phase. During the 12-week intervention period, the use frequency and use time were recorded.

We found significant gender difference in use frequency ( $U=908.00$ ;  $z$  score=-2.51;  $P=.01$ ;  $r=-0.24$ ) and use time ( $U=736.00$ ;  $z$  score=-3.63;  $P<.001$ ;  $r=-0.35$ ). The participating women used the app 97 (SD 88.03) times and for 625 (SD

427.94) minutes on average throughout the intervention, whereas the participating men used the app 56 (SD 45.62) times and for 347 (SD 285.68) minutes on average. In total, 82.6% (90/109) of the users were compliant with the app (women: 63/74, 85%; men: 27/35, 77%).

During the 12-week intervention phase, the use time ( $\chi^2_{11}=126.03$ ;  $P<.001$ ) and use frequency ( $\chi^2_{11}=139.51$ ;  $P<.001$ ) of the participating men ( $n=35$ ) decreased (Figures 5 and 6). The use time, ( $\chi^2_{11}=231.34$ ;  $P<.001$ ) and use frequency ( $\chi^2_{11}=309.16$ ;  $P<.001$ ) of the participating women ( $n=74$ ) also decreased. Dunn-Bonferroni post hoc tests revealed a significant decrease in use time within the first 3 weeks of intervention ( $z$  score=3.99;  $P<.001$ ;  $r=0.46$ ). From week 3 to week 12, use time and frequency leveled off at approximately 6.56 (SD 7.21) actions per week and 41.99 (SD 34.03) minutes per week for the participating women and 3.53 actions per week (SD 3.36) and 21.75 minutes per week (SD 21.88) for the participating men. We found no gender differences in use time progress ( $U=1075.00$ ;  $z$  score=-1.43;  $P=.15$ ) and use frequency progress ( $U=1106.00$ ,  $z$  score=-1.23;  $P=.22$ ) during the 12-week intervention period.

**Figure 5.** Use time per week in minutes (means and SEs of means).**Figure 6.** Use frequency per week in actions (means and SEs of means).

### Evaluation of the App

After completion, 41 participants evaluated the I-GENDO app. On average, the overall satisfaction with the app was high (mean 85.54, SD 19.36). In addition, the relevance of the content (mean 83.34, SD 20.03) and daily life usefulness (mean 78.95, SD

22.24) were evaluated as satisfactory. Of the main modules, the *stress module* (n=36) was rated best (mean 82.92, SD 14.05), followed by the *emotion module* (n=50; mean 81.66, SD 16.45), the *control module* (n=60; mean 80.47, SD 18.08), the *self-efficacy module* (n=29; mean 78.48, SD 17.66), and finally the *consequence module* (n=16; mean 67.75, SD 21.68).

In addition to the evaluation, the usability of the app was assessed using a standardized questionnaire (see the section *Usability Rating of the App*). The usability of the app was rated, on average, with 71% satisfaction (mean 5.00, SD 1.08 points; maximum: 7.00 points). No gender differences could be found between the usability ratings of men (mean 4.72, SD 1.07) and women (mean 5.13, SD 1.07;  $t_{99}=-1.76$ ;  $P=.08$ ).

## Discussion

### Overview

We aimed to introduce the I-GENDO app, a tailored gender-sensitive mHealth weight loss intervention, and present results from its process evaluation data. Therefore, data from the intervention arm of the I-GENDO project were analyzed. The sample included 116 ( $n=77$ , 66.4% women) individuals with overweight and obesity.

### Principal Findings

We developed a module-based 12-week intervention combining computer-based and self-tailoring features. Most participants (105/116, 90.5%) received the *control module*, which focused on self-regulation skills of food craving. The *stress module* was assigned to 65.5% (76/116) of the participants, and the *self-efficacy module* to 47.4% (55/116). The *consequence module* was obtained by 25.9% (30/116) of the participants. Significantly more women (59/77, 77%) than men (22/39, 56%) received the *emotion module*. Another tool of the intervention was the implementation of gender-sensitive self-tailoring features. We developed women- and men-targeted versions of the main modules. At the beginning of each module, participants could choose between the 2 versions. Among the participants who chose a version, *version A* was chosen 58.2% (163/280) of the time (women: 116/196, 59.2%; men: 47/84, 56%), which means that among both genders, the women-targeted module versions were predominantly selected.

In total, 82.6% (90/109) of the participants (women: 63/74, 85%; men: 27/35, 77%) were compliant with the I-GENDO app during the intervention phase. Use time and frequency significantly decreased during the 12-week intervention phase for both genders. After the first 3 weeks of intervention, use time leveled off and remained stable at approximately 42 minutes per week for the participating women and 22 minutes per week for the participating men. Similarly, use frequencies were approximately stable as of week 3 for both genders. Compared with the women, the men used the app infrequently and spent less time with the app. Nevertheless, the average use times and frequencies in both genders were satisfactory even in the last weeks.

The overall satisfaction with the app was high, with almost 86% (86/100) approval. In addition, the daily life usefulness and relevance of the content were ranked satisfactory by 79% (79/100) and 83% (83/100) of participants, respectively. The highest-rated main module was the *stress module* (83/100, 83%), but even the satisfaction with the *consequence module* was acceptable (68/100, 68%). In general, the usability ratings indicated that the I-GENDO intervention was good, averaging 5.0 out of 7.0 points (71%).

### Comparison With Prior Work

The heterogeneous computer-based administration of the main modules supports the tailoring feature. The *control module* was assigned to most participants. This is in line with the observation that decreased food-related inhibitory control is regularly associated with overweight and obesity [49,82,83]. Gender differences were found in the computer-based assignments of the *emotion module*, which significantly more women obtained. The module focused on dysfunctional emotion regulation and associations between negative emotions and (eating) behavior. EE refers to problems in the distinction between physiological appetite and eating as a strategy to cope with negative feelings [84]. EE is correlated with higher weight, severe depression symptoms, and the consumption of sweet energy-dense foods [85]. More women report negative emotions as causes for their overweight and engage more often in EE compared with men [50,85,86]. EE is associated with less intuitive eating by women, which could be a barrier to the implementation of healthy eating behaviors [87]. Studies indicate that more women undergo weight loss treatment, whereas participating men lose more absolute weight [29]. Focusing more on EE in treatment might contribute to a close in this gap. In addition, previous studies indicated that a relevant subgroup of individuals with overweight and obesity exhibit addiction-like eating behavior (ie, food addiction [FA]), characterized by an impaired food-related inhibitory control, EE, and food craving [88,89]. The prevalence of FA is higher in women than in men and is among other factors associated with higher BMI, dysfunctional eating behavior, and psychological distress [90,91]. Some studies reported lower adherence to and decreased effectiveness of WLPs in individuals experiencing FA, whereas others found no influence of FA on the success of WLPs [92-95]. As the *control* and *emotion* modules implement the treatment of dysfunctional EE behavior and exercises to improve food-related inhibitory control, participants experiencing FA might especially benefit from the intervention. Thus, the association between FA and the effectiveness of our intervention should be further investigated.

One-quarter of the participants received the *consequences module*, which focused on weight-related discrimination and the improvement of self-esteem and body image, as well as the social competences to deal with discrimination. The extent of this use might explain the prevalence of weight discrimination being higher in our sample than in the results of a representative German study reporting prevalence rates ranging from 5.6% to 18.7% in individuals with overweight and obesity (classes I and II) [96]. We hypothesized that individuals who have experienced discrimination might prefer seeking WLPs based on psychological rather than lifestyle features. Moreover, in our study, the *consequence module* was assigned to more men (12/39, 31%) than women (18/77, 23%), which appears to be in contrast to the results of the previously cited study that reported double the prevalence of weight-based discrimination in women [96]. The anonymity of a smartphone-based intervention combined with the opportunity to receive specialized psychological support targeted to individual needs could have been particularly appealing for men who had experienced weight-related discrimination and were affected by the consequences of their overweight. Nevertheless, the



module generally focused on weight-related emotional and physical consequences, which might be appealing to individuals with overweight and obesity regardless of whether they experienced discrimination.

Gender differences in health care services are an important consideration for the improvement of treatment outcomes [97]. Prior studies have indicated gender differences in eating behavior, as well as the psychological factors associated with weight gain and maintenance, highlighting the need for gender-targeted weight loss interventions [29,40]. As the effectiveness of gender dichotomous tailoring does not significantly differ from that of gender-neutral interventions [45], we implemented gender-sensitive self-tailoring features. The participants could choose between 2 gender-targeted versions at the beginning of the modules. The selection of the versions was heterogeneous, with most participants choosing women-targeted versions. This result supports the idea of gender-sensitive interventions to overcome gender binary [46]. However, its influence on the effectiveness of the intervention needs to be further investigated.

In complex digital interventions, the consideration of relevant process evaluation data (eg, usability testing and use patterns) is crucial before interpreting the effectiveness of the intervention [98]. The compliance with the app was satisfactory (90/109, 82.6%) and comparable with other studies. Signal et al [99] developed an eHealth intervention for prediabetes and diabetes self-management. They reported that 74% of the participants were actively engaged (ie, any use data were detected at any time throughout the 16-week intervention). Ruf et al [100] developed an mHealth intervention that assesses event-contingent dietary intake and physical activity, as well as relevant psychological parameters. Compliance, defined as the percentage of complete prompts within the total number of prompts received, was 80%. Another mHealth intervention focused on the management of food-related impulses to facilitate weight loss [101]. In that study, the completion rate (the number of participants who provided data at the 3-month follow-up) was 76%. These findings suggest that our compliance rate is comparable or even higher, although the differences in operational definitions cloud the interpretation.

Throughout the intervention, the use time and frequency decreased in both genders. Decreases in engagement were also reported in other studies; that is, in those with extended intervention periods [99,102]. Reductions in engagement and high dropouts are typical for internet-based interventions and are caused by a variety of reasons [103]. We hypothesized that the reduction in engagement observed in our study might be associated with the high number of competing commercial digital weight loss interventions, which might be less demanding, compared with psychological interventions. Moreover, the intervention phase of our study fell within the first and second lockdowns of the COVID-19 pandemic in Germany in 2020. During this period, the level of psychological distress increased, and vulnerable people engaged more often in dysfunctional eating patterns (ie, EE) [104]. In addition, many people were affected by short-term work or job losses and subsequent income losses [105]. It is likely that people neglected the intervention during this burdensome period.

The results from previous studies on the adherence to mHealth interventions are heterogeneous, with some reporting higher engagement in men [29,40,106] and others reporting higher engagement in women [99]. In our study, women used the app more frequently and spent more time on it. In the general German population, women report higher smartphone use time (mean 167 min/day) than men (mean 154 min/day), which might at least partially explain these differences [107]. Moreover, women are more interested in body appearance and health-related topics than men and use the internet more frequently for medical and health research [108-110]. Studies have also reported that women are more likely to use mHealth interventions focusing on nutrition and self-care apps, whereas men are more likely to use fitness apps [111-113]. Therefore, the lower engagement of the participating men in this study might be because the app focused on psychological rather than physiological determinants of overweight and obesity.

As reported in a recently published systematic review [114], other studies on mHealth interventions have either failed to report gender differences in the adherence to and usability of these interventions or reported results from biased samples with approximately 90% of women [115-117]. Given that higher engagement in mHealth interventions is usually associated with better outcomes [22,24,118], we propose that the samples in future studies should be more balanced with regard to gender and implement gender-sensitive feasibility and usability testing. Overall, the compliance with the app (90/109, 82.6%) and satisfaction with the app (86/100, 86%) were high and comparable with those of other mHealth interventions [99-101,119]. The usability of the app was rated with 71% (5.0/7.0 points) satisfaction. Other evidence-based mHealth weight loss interventions reported comparable or even lower usability scores, between 61.9% and 69.3% [100,119]. In addition, Ferrara et al [120] reviewed the usability of commercial weight loss apps, which can be downloaded from Google Play and the Apple Store. Scientists ranked the usability of these apps between 47% and 89%.

## Limitations

In our study, men and women differed in the assignment of main modules, which focused on psychological parameters associated with the development and maintenance of overweight and obesity. Interestingly, most men and women selected the women-targeted versions of the main modules. Given that the participants were blind to the gender-targeted manipulation, we suggest that the selections were not influenced by social desirability. Future studies should distinguish between gender differences based on the results from explicit and implicit assessments to adjust for social norms. Moreover, the participants were forced to select one version at the beginning of each module and were not allowed to switch versions. A reasonable approach could be to allow participants to test both versions to enhance their adherence to the app. In addition, it should be verified whether the introductions of the versions sufficiently hint at different module content.

It should be noted that only few participants (41/109, 37.6%) evaluated the app after completion. The evaluation was voluntary and was assessed at the end of the last session of the

intervention. Therefore, results regarding satisfaction with the app and the main modules should be interpreted cautiously.

The results from the process evaluation revealed that men and women differed in their app use. Women used the app more frequently and longer than men. Most of the scientists involved in the development process were women. Therefore, the women-targeted features of the app might have been more salient and thus confounded the selection by both genders. This methodological aspect might subsequently explain the higher use patterns of the participating women. Future studies or revisions of the app intervention should involve men scientists.

## Conclusions

In summary, given the high diversity in module assignment, we hypothesize that tailoring was successfully implemented in the intervention. The heterogeneous selection of the gender-targeted features might underscore the need for gender-sensitive (self-tailoring and blind choice) instead of gender dichotomous (computer-based tailoring) targeting but could also hint at methodological limitations, which need to be considered and further investigated in future studies. Further studies need to clarify whether the reported gender differences in the use and evaluation of the app confound the effectiveness and sustainability of the I-GENDO intervention.

## Acknowledgments

The Federal Ministry of Education and Research, Germany funded this study. The authors would like to acknowledge support from the Open Access Publication Funds of the Ruhr-Universität Bochum. The authors would also like to acknowledge Zoë Blumenwitz for research assistance.

## Authors' Contributions

MP was involved in conceptualization, the acquisition of data, formal analysis, the interpretation of data, and the writing of the original draft. TF was involved in the acquisition and interpretation of data, review, and editing. CS and TR were involved in the acquisition of data, review, and editing. SS contributed to the study design. JW and SH contributed to the study design, supervision, review, and editing. SS-L contributed to the study design, conceptualization, study supervision, review, and editing.

## Conflicts of Interest

None declared.

## Multimedia Appendix 1

Evidence, content, and adaptations of the gender-sensitive main modules.

[[DOCX File, 55 KB - formative\\_v6i10e38480\\_app1.docx](#)]

## References

1. World Health Organization. mHealth: New horizons for health through mobile technologies. Geneva: World Health Organization; 2011.
2. Gesundheit digital. Perspektiven zur Digitalisierung im Gesundheitswesen. Book in German. Cham: Springer; 2019.
3. Webster P. Virtual health care in the era of COVID-19. *Lancet* 2020 Apr;395(10231):1180-1181. [doi: [10.1016/s0140-6736\(20\)30818-7](https://doi.org/10.1016/s0140-6736(20)30818-7)]
4. D21 Digital Index 2020/2021. Initiative D21. 2013. URL: [https://initiated21.de/app/uploads/2021/02/d21-digital-index-2020\\_2021.pdf](https://initiated21.de/app/uploads/2021/02/d21-digital-index-2020_2021.pdf) [accessed 2022-03-28]
5. Farley H. Promoting self-efficacy in patients with chronic disease beyond traditional education: a literature review. *Nurs Open* 2020 Jan 20;7(1):30-41 [FREE Full text] [doi: [10.1002/nop2.382](https://doi.org/10.1002/nop2.382)] [Medline: [31871689](https://pubmed.ncbi.nlm.nih.gov/31871689/)]
6. Krebs P, Prochaska JO, Rossi JS. A meta-analysis of computer-tailored interventions for health behavior change. *Prev Med* 2010;51(3-4):214-221 [FREE Full text] [doi: [10.1016/j.ypmed.2010.06.004](https://doi.org/10.1016/j.ypmed.2010.06.004)] [Medline: [20558196](https://pubmed.ncbi.nlm.nih.gov/20558196/)]
7. Stephens J, Allen J. Mobile phone interventions to increase physical activity and reduce weight: a systematic review. *J Cardiovasc Nurs* 2013;28(4):320-329 [FREE Full text] [doi: [10.1097/JCN.0b013e318250a3e7](https://doi.org/10.1097/JCN.0b013e318250a3e7)] [Medline: [22635061](https://pubmed.ncbi.nlm.nih.gov/22635061/)]
8. Patel B, Thind A. Usability of mobile health apps for postoperative care: systematic review. *JMIR Perioper Med* 2020 Jul 20;3(2):e19099 [FREE Full text] [doi: [10.2196/19099](https://doi.org/10.2196/19099)] [Medline: [33393925](https://pubmed.ncbi.nlm.nih.gov/33393925/)]
9. Turner-McGrievy GM, Beets MW, Moore JB, Kaczynski AT, Barr-Anderson DJ, Tate DF. Comparison of traditional versus mobile app self-monitoring of physical activity and dietary intake among overweight adults participating in an mHealth weight loss program. *J Am Med Inform Assoc* 2013 May 01;20(3):513-518 [FREE Full text] [doi: [10.1136/amiajnl-2012-001510](https://doi.org/10.1136/amiajnl-2012-001510)] [Medline: [23429637](https://pubmed.ncbi.nlm.nih.gov/23429637/)]
10. Alnasser A, Kyle J, Aloumi N, Al-Khalifa A, Marais D. The Twazon Arabic weight loss app: app-based intervention for Saudi women with obesity. *JMIR Mhealth Uhealth* 2019 May 28;7(5):e10923 [FREE Full text] [doi: [10.2196/10923](https://doi.org/10.2196/10923)] [Medline: [31140444](https://pubmed.ncbi.nlm.nih.gov/31140444/)]
11. Cavero-Redondo I, Martinez-Vizcaino V, Fernandez-Rodriguez R, Saz-Lara A, Pascual-Morena C, Álvarez-Bueno C. Effect of behavioral weight management interventions using lifestyle mHealth self-monitoring on weight loss: a systematic

- review and meta-analysis. *Nutrients* 2020 Jul 03;12(7):1977 [FREE Full text] [doi: [10.3390/nu12071977](https://doi.org/10.3390/nu12071977)] [Medline: [32635174](https://pubmed.ncbi.nlm.nih.gov/32635174/)]
12. Jeffery RW, Epstein LH, Wilson GT, Drenowski A, Stunkard AJ, Wing RR. Long-term maintenance of weight loss: current status. *Health Psychol* 2000 Jan;19(1, Suppl):5-16. [doi: [10.1037/0278-6133.19.Supp1.5](https://doi.org/10.1037/0278-6133.19.Supp1.5)]
  13. Beleigoli AM, Andrade AQ, Caçado AG, Paulo MN, Diniz MD, Ribeiro AL. Web-based digital health interventions for weight loss and lifestyle habit changes in overweight and obese adults: systematic review and meta-analysis. *J Med Internet Res* 2019 Jan 08;21(1):e298 [FREE Full text] [doi: [10.2196/jmir.9609](https://doi.org/10.2196/jmir.9609)] [Medline: [30622090](https://pubmed.ncbi.nlm.nih.gov/30622090/)]
  14. Gudzone KA, Doshi RS, Mehta AK, Chaudhry ZW, Jacobs DK, Vakil RM, et al. Efficacy of commercial weight-loss programs. *Ann Intern Med* 2015 Apr 07;162(7):501-512. [doi: [10.7326/m14-2238](https://doi.org/10.7326/m14-2238)]
  15. Pagoto S, Schneider K, Jovic M, DeBiasse M, Mann D. Evidence-based strategies in weight-loss mobile apps. *Am J Prev Med* 2013 Nov;45(5):576-582. [doi: [10.1016/j.amepre.2013.04.025](https://doi.org/10.1016/j.amepre.2013.04.025)] [Medline: [24139770](https://pubmed.ncbi.nlm.nih.gov/24139770/)]
  16. National Clinical Guideline Centre (UK). Obesity: Identification, Assessment and Management of Overweight and Obesity in Children, Young People and Adults. London: National Institute for Health and Care Excellence (NICE); Nov 2014.
  17. Hauner H, Moss A, Berg A, Bischoff SC, Colombo-Benkmann M, Ellrott T, et al. Interdisziplinäre Leitlinie der Qualität S3 zur "Prävention und Therapie der Adipositas". Article in German. *Adipositas Ursachen Folgeerkrankungen Therapie* 2017 Dec 21;08(04):179-221. [doi: [10.1055/s-0037-1618857](https://doi.org/10.1055/s-0037-1618857)]
  18. Davis A, Sweigart R, Ellis R. A systematic review of tailored mHealth interventions for physical activity promotion among adults. *Transl Behav Med* 2020 Oct 12;10(5):1221-1232. [doi: [10.1093/tbm/ibz190](https://doi.org/10.1093/tbm/ibz190)] [Medline: [33044542](https://pubmed.ncbi.nlm.nih.gov/33044542/)]
  19. Ryan K, Dockray S, Linehan C. A systematic review of tailored eHealth interventions for weight loss. *Digit Health* 2019;5:2055207619826685 [FREE Full text] [doi: [10.1177/2055207619826685](https://doi.org/10.1177/2055207619826685)] [Medline: [30783535](https://pubmed.ncbi.nlm.nih.gov/30783535/)]
  20. DiClemente R, Nowara A, Shelton R, Wingood G. Need for innovation in public health research. *Am J Public Health* 2019 Feb;109(S2):S117-S120. [doi: [10.2105/AJPH.2018.304876](https://doi.org/10.2105/AJPH.2018.304876)] [Medline: [30785791](https://pubmed.ncbi.nlm.nih.gov/30785791/)]
  21. Noar SM, Grant Harrington N, Van Stee SK, Shemanski Aldrich R. Tailored health communication to change lifestyle behaviors. *Am J Lifestyle Med* 2010 Nov 19;5(2):112-122. [doi: [10.1177/1559827610387255](https://doi.org/10.1177/1559827610387255)]
  22. Puls H, Schmidt R, Herpertz S, Zipfel S, Tuschen-Caffier B, Friederich H, et al. Adherence as a predictor of dropout in Internet-based guided self-help for adults with binge-eating disorder and overweight or obesity. *Int J Eat Disord* 2020 Apr 31;53(4):555-563. [doi: [10.1002/eat.23220](https://doi.org/10.1002/eat.23220)] [Medline: [31891225](https://pubmed.ncbi.nlm.nih.gov/31891225/)]
  23. Butryn ML, Godfrey KM, Martinelli MK, Roberts SR, Forman EM, Zhang F. Digital self-monitoring: does adherence or association with outcomes differ by self-monitoring target? *Obes Sci Pract* 2020 Apr;6(2):126-133 [FREE Full text] [doi: [10.1002/osp4.391](https://doi.org/10.1002/osp4.391)] [Medline: [32313670](https://pubmed.ncbi.nlm.nih.gov/32313670/)]
  24. Calugi S, Marchesini G, El Ghoch M, Gavasso I, Dalle Grave R. The association between weight maintenance and session-by-session diet adherence, weight loss and weight-loss satisfaction. *Eat Weight Disord* 2020 Feb 21;25(1):127-133. [doi: [10.1007/s40519-018-0528-8](https://doi.org/10.1007/s40519-018-0528-8)] [Medline: [29931447](https://pubmed.ncbi.nlm.nih.gov/29931447/)]
  25. Neve M, Morgan PJ, Jones PR, Collins CE. Effectiveness of web-based interventions in achieving weight loss and weight loss maintenance in overweight and obese adults: a systematic review with meta-analysis. *Obes Rev* 2010 Apr;11(4):306-321. [doi: [10.1111/j.1467-789X.2009.00646.x](https://doi.org/10.1111/j.1467-789X.2009.00646.x)] [Medline: [19754633](https://pubmed.ncbi.nlm.nih.gov/19754633/)]
  26. Schvey NA, Puhl RM, Brownell KD. The impact of weight stigma on caloric consumption. *Obesity (Silver Spring)* 2011 Oct;19(10):1957-1962 [FREE Full text] [doi: [10.1038/oby.2011.204](https://doi.org/10.1038/oby.2011.204)] [Medline: [21760636](https://pubmed.ncbi.nlm.nih.gov/21760636/)]
  27. Tomiyama AJ. Weight stigma is stressful. A review of evidence for the Cyclic Obesity/Weight-Based Stigma model. *Appetite* 2014 Nov;82:8-15. [doi: [10.1016/j.appet.2014.06.108](https://doi.org/10.1016/j.appet.2014.06.108)] [Medline: [24997407](https://pubmed.ncbi.nlm.nih.gov/24997407/)]
  28. van Strien T. Causes of emotional eating and matched treatment of obesity. *Curr Diab Rep* 2018 Apr 25;18(6):35 [FREE Full text] [doi: [10.1007/s11892-018-1000-x](https://doi.org/10.1007/s11892-018-1000-x)] [Medline: [29696418](https://pubmed.ncbi.nlm.nih.gov/29696418/)]
  29. Tsai SA, Lv N, Xiao L, Ma J. Gender differences in weight-related attitudes and behaviors among overweight and obese adults in the United States. *Am J Mens Health* 2016 Sep 22;10(5):389-398 [FREE Full text] [doi: [10.1177/1557988314567223](https://doi.org/10.1177/1557988314567223)] [Medline: [25595019](https://pubmed.ncbi.nlm.nih.gov/25595019/)]
  30. Koritzky G, Yechiam E, Bukay I, Milman U. Obesity and risk taking. A male phenomenon. *Appetite* 2012 Oct;59(2):289-297. [doi: [10.1016/j.appet.2012.05.020](https://doi.org/10.1016/j.appet.2012.05.020)] [Medline: [22634199](https://pubmed.ncbi.nlm.nih.gov/22634199/)]
  31. Mensink G, Schienkiewitz A, Haftenberger M, Lampert T, Ziese T, Scheidt-Nave C. Overweight and obesity in Germany: results of the German Health Interview and Examination Survey for Adults (DEGS1). *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz* 2013 May;56(5-6):786-794 [FREE Full text] [doi: [10.1007/s00103-012-1656-3](https://doi.org/10.1007/s00103-012-1656-3)] [Medline: [23703499](https://pubmed.ncbi.nlm.nih.gov/23703499/)]
  32. Schienkiewitz A, Mensink G, Kuhnert R, Lange C. Overweight and obesity among adults in Germany. *J Health Monitoring* 2017 Jun 15;2:20-27 [FREE Full text] [doi: [10.17886/RKI-GBE-2017-038](https://doi.org/10.17886/RKI-GBE-2017-038)]
  33. Kiefer I, Rathmanner T, Kunze M. Eating and dieting differences in men and women. *J Men's Health Gender* 2005 Jun;2(2):194-201. [doi: [10.1016/j.jmhg.2005.04.010](https://doi.org/10.1016/j.jmhg.2005.04.010)]
  34. Anbumalar C, Dorathy Agines P, Jaswanti VP, Priya D, Reniangelin D. Gender differences in perceived stress levels and coping strategies among college students. *Int J Indian Psychol* 2017 Sep 25;4(4). [doi: [10.25215/0404.103](https://doi.org/10.25215/0404.103)]
  35. Tamres LK, Janicki D, Helgeson VS. Sex differences in coping behavior: a meta-analytic review and an examination of relative coping. *Pers Soc Psychol Rev* 2016 Dec 21;6(1):2-30. [doi: [10.1207/s15327957pspr0601\\_1](https://doi.org/10.1207/s15327957pspr0601_1)]

36. Rütten A, Abu-Omar K. Prevalence of physical activity in the European Union. *Soz Präventivmed* 2004 Aug;49(4):281-289. [doi: [10.1007/s00038-004-3100-4](https://doi.org/10.1007/s00038-004-3100-4)] [Medline: [15357531](https://pubmed.ncbi.nlm.nih.gov/15357531/)]
37. Novelle MG, Diéguez C. Updating gender differences in the control of homeostatic and hedonic food intake: implications for binge eating disorder. *Mol Cell Endocrinol* 2019 Nov 01;497:110508. [doi: [10.1016/j.mce.2019.110508](https://doi.org/10.1016/j.mce.2019.110508)] [Medline: [31319096](https://pubmed.ncbi.nlm.nih.gov/31319096/)]
38. Frank AP, de Souza Santos R, Palmer BF, Clegg DJ. Determinants of body fat distribution in humans may provide insight about obesity-related health risks. *J Lipid Res* 2019 Oct;60(10):1710-1719. [doi: [10.1194/jlr.R086975](https://doi.org/10.1194/jlr.R086975)]
39. Kroll DS, Feldman DE, Biesecker CL, McPherson KL, Manza P, Joseph PV, et al. Neuroimaging of sex/gender differences in obesity: a review of structure, function, and neurotransmission. *Nutrients* 2020 Jun 30;12(7):1942 [FREE Full text] [doi: [10.3390/nu12071942](https://doi.org/10.3390/nu12071942)] [Medline: [32629783](https://pubmed.ncbi.nlm.nih.gov/32629783/)]
40. Cooper AJ, Gupta SR, Moustafa AF, Chao AM. Sex/gender differences in obesity prevalence, comorbidities, and treatment. *Curr Obes Rep* 2021 Dec 02;10(4):458-466. [doi: [10.1007/s13679-021-00453-x](https://doi.org/10.1007/s13679-021-00453-x)] [Medline: [34599745](https://pubmed.ncbi.nlm.nih.gov/34599745/)]
41. Chin SO, Keum C, Woo J, Park J, Choi HJ, Woo J, et al. Successful weight reduction and maintenance by using a smartphone application in those with overweight and obesity. *Sci Rep* 2016 Nov 07;6:34563 [FREE Full text] [doi: [10.1038/srep34563](https://doi.org/10.1038/srep34563)] [Medline: [27819345](https://pubmed.ncbi.nlm.nih.gov/27819345/)]
42. Klenk S, Reifegerste D, Renatus R. Gender differences in gratifications from fitness app use and implications for health interventions. *Mobile Media Commun* 2017 Mar 01;5(2):178-193. [doi: [10.1177/2050157917691557](https://doi.org/10.1177/2050157917691557)]
43. Greenberg I, Stampfer MJ, Schwarzfuchs D, Shai I, DIRECT Group. Adherence and success in long-term weight loss diets: the dietary intervention randomized controlled trial (DIRECT). *J Am Coll Nutr* 2009 Apr;28(2):159-168. [doi: [10.1080/07315724.2009.10719767](https://doi.org/10.1080/07315724.2009.10719767)] [Medline: [19828901](https://pubmed.ncbi.nlm.nih.gov/19828901/)]
44. Bautista-Castaño I, Molina-Cabrillana J, Montoya-Alonso JA, Serra-Majem L. Variables predictive of adherence to diet and physical activity recommendations in the treatment of obesity and overweight, in a group of Spanish subjects. *Int J Obes Relat Metab Disord* 2004 May 2;28(5):697-705. [doi: [10.1038/sj.ijo.0802602](https://doi.org/10.1038/sj.ijo.0802602)] [Medline: [14993911](https://pubmed.ncbi.nlm.nih.gov/14993911/)]
45. Sharkey T, Whatnall MC, Hutchesson MJ, Haslam RL, Bezzina A, Collins CE, et al. Correction to: effectiveness of gender-targeted versus gender-neutral interventions aimed at improving dietary intake, physical activity and/or overweight/obesity in young adults (aged 17-35 years): a systematic review and meta-analysis. *Nutr J* 2020 Aug 26;19(1):90 [FREE Full text] [doi: [10.1186/s12937-020-00605-0](https://doi.org/10.1186/s12937-020-00605-0)] [Medline: [32847592](https://pubmed.ncbi.nlm.nih.gov/32847592/)]
46. Hyde JS, Bigler RS, Joel D, Tate CC, van Anders SM. The future of sex and gender in psychology: five challenges to the gender binary. *Am Psychol* 2019 Feb;74(2):171-193. [doi: [10.1037/amp0000307](https://doi.org/10.1037/amp0000307)] [Medline: [30024214](https://pubmed.ncbi.nlm.nih.gov/30024214/)]
47. I-GENDO study. Universität Bamberg. 2022. URL: <https://www.uni-bamberg.de/klinpsych/forschung/projekte/i-gendo> [accessed 2022-03-29]
48. 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 20;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/)]
49. Houben K, Nederkoorn C, Jansen A. Eating on impulse: the relation between overweight and food-specific inhibitory control. *Obesity (Silver Spring)* 2014 May 09;22(5):E6-E8 [FREE Full text] [doi: [10.1002/oby.20670](https://doi.org/10.1002/oby.20670)] [Medline: [24910860](https://pubmed.ncbi.nlm.nih.gov/24910860/)]
50. Larsen JK, van Strien T, Eisinga R, Engels RC. Gender differences in the association between alexithymia and emotional eating in obese individuals. *J Psychosom Res* 2006 Mar;60(3):237-243. [doi: [10.1016/j.jpsychores.2005.07.006](https://doi.org/10.1016/j.jpsychores.2005.07.006)] [Medline: [16516654](https://pubmed.ncbi.nlm.nih.gov/16516654/)]
51. Dallman MF, Pecoraro N, Akana SF, La Fleur SE, Gomez F, Houshyar H, et al. Chronic stress and obesity: a new view of "comfort food". *Proc Natl Acad Sci U S A* 2003 Sep 30;100(20):11696-11701 [FREE Full text] [doi: [10.1073/pnas.1934666100](https://doi.org/10.1073/pnas.1934666100)] [Medline: [12975524](https://pubmed.ncbi.nlm.nih.gov/12975524/)]
52. Sinha R, Jastreboff AM. Stress as a common risk factor for obesity and addiction. *Biol Psychiatry* 2013 May 01;73(9):827-835 [FREE Full text] [doi: [10.1016/j.biopsych.2013.01.032](https://doi.org/10.1016/j.biopsych.2013.01.032)] [Medline: [23541000](https://pubmed.ncbi.nlm.nih.gov/23541000/)]
53. Barry D, Petry N. Gender differences in associations between stressful life events and body mass index. *Prev Med* 2008 Nov;47(5):498-503 [FREE Full text] [doi: [10.1016/j.ypmed.2008.08.006](https://doi.org/10.1016/j.ypmed.2008.08.006)] [Medline: [18793665](https://pubmed.ncbi.nlm.nih.gov/18793665/)]
54. Bennett E, Gough B. In pursuit of leanness: the management of appearance, affect and masculinities within a men's weight loss forum. *Health (London)* 2013 May;17(3):284-299 [FREE Full text] [doi: [10.1177/1363459312454149](https://doi.org/10.1177/1363459312454149)] [Medline: [22815334](https://pubmed.ncbi.nlm.nih.gov/22815334/)]
55. Galdas P, Darwin Z, Kidd L, Blickem C, McPherson K, Hunt K, et al. The accessibility and acceptability of self-management support interventions for men with long term conditions: a systematic review and meta-synthesis of qualitative studies. *BMC Public Health* 2014 Nov 27;14(1):1230-1296 [FREE Full text] [doi: [10.1186/1471-2458-14-1230](https://doi.org/10.1186/1471-2458-14-1230)] [Medline: [25428230](https://pubmed.ncbi.nlm.nih.gov/25428230/)]
56. Robertson C, Archibald D, Avenell A, Douglas F, Hoddinott P, van Teijlingen E, et al. Systematic reviews of and integrated report on the quantitative, qualitative and economic evidence base for the management of obesity in men. *Health Technol Assess* 2014 May;18(35):v-vi, xxiii [FREE Full text] [doi: [10.3310/hta18350](https://doi.org/10.3310/hta18350)] [Medline: [24857516](https://pubmed.ncbi.nlm.nih.gov/24857516/)]
57. Becker S, Zipfel S, Teufel M. *Psychotherapie der Adipositas: Interdisziplinäre Diagnostik und differenzielle Therapie*. Stuttgart, Germany: Kohlhammer Verlag; 2015.
58. Cooper Z, Fairburn CG. A new cognitive behavioural approach to the treatment of obesity. *Behav Res Ther* 2001 May 01;39(5):499-511. [doi: [10.1016/s0005-7967\(00\)00065-6](https://doi.org/10.1016/s0005-7967(00)00065-6)] [Medline: [28465183](https://pubmed.ncbi.nlm.nih.gov/28465183/)]

59. Kaluza G. Stressbewältigung: Trainingsmanual zur psychologischen Gesundheitsförderung. Berlin, Germany: Springer-Verlag; 2018.
60. Kaluza G. Gelassen und sicher im Stress-Das Stresskompetenz-Buch. Berlin, Heidelberg: Springer; 2018.
61. Bennett J, Greene G, Schwartz-Barcott D. Perceptions of emotional eating behavior. A qualitative study of college students. *Appetite* 2013 Jan;60(1):187-192. [doi: [10.1016/j.appet.2012.09.023](https://doi.org/10.1016/j.appet.2012.09.023)] [Medline: [23046706](https://pubmed.ncbi.nlm.nih.gov/23046706/)]
62. Nolen-Hoeksema S, Aldao A. Gender and age differences in emotion regulation strategies and their relationship to depressive symptoms. *Personal Individ Differ* 2011 Oct;51(6):704-708. [doi: [10.1016/j.paid.2011.06.012](https://doi.org/10.1016/j.paid.2011.06.012)]
63. Munsch S, Hilbert A. Übergewicht und Adipositas, vol 59. Göttingen, Germany: Hogrefe Verlag; 2015.
64. Bohus M, Wolf-Arehult M. Interaktives Skillstraining für Borderline-Patienten. Stuttgart, Germany: Schattauer Verlag; 2013.
65. Eismann E, Lammers CH. Therapie-Tools Emotionsregulation: Mit E-Book inside und Arbeitsmaterial. Weinheim, Baden-Württemberg, Germany: Beltz; 2017.
66. Puhl RM, Brownell KD. Confronting and coping with weight stigma: an investigation of overweight and obese adults. *Obesity (Silver Spring)* 2006 Oct;14(10):1802-1815 [FREE Full text] [doi: [10.1038/oby.2006.208](https://doi.org/10.1038/oby.2006.208)] [Medline: [17062811](https://pubmed.ncbi.nlm.nih.gov/17062811/)]
67. Puhl RM, Heuer CA. The stigma of obesity: a review and update. *Obesity (Silver Spring)* 2009 May;17(5):941-964 [FREE Full text] [doi: [10.1038/oby.2008.636](https://doi.org/10.1038/oby.2008.636)] [Medline: [19165161](https://pubmed.ncbi.nlm.nih.gov/19165161/)]
68. Couch D, Han GS, Robinson P, Komesaroff P. Men's weight loss stories: How personal confession, responsibility and transformation work as social control. *Health (London)* 2019 Jan;23(1):76-96. [doi: [10.1177/1363459317724855](https://doi.org/10.1177/1363459317724855)] [Medline: [28803501](https://pubmed.ncbi.nlm.nih.gov/28803501/)]
69. Hinsch R, Pflingsten U. Gruppentraining sozialer Kompetenzen: GSK; Grundlagen, Durchführung, Anwendungsbeispiele. Weinheim, Baden-Württemberg, Germany: Beltz; 2015.
70. Bartholdy S, Dalton B, O'Daly OG, Campbell IC, Schmidt U. A systematic review of the relationship between eating, weight and inhibitory control using the stop signal task. *Neurosci Biobehav Rev* 2016 May;64:35-62. [doi: [10.1016/j.neubiorev.2016.02.010](https://doi.org/10.1016/j.neubiorev.2016.02.010)] [Medline: [26900651](https://pubmed.ncbi.nlm.nih.gov/26900651/)]
71. Carey JB, Saules KK, Carr MM. A qualitative analysis of men's experiences of binge eating. *Appetite* 2017 Sep 01;116:184-195. [doi: [10.1016/j.appet.2017.04.030](https://doi.org/10.1016/j.appet.2017.04.030)] [Medline: [28465183](https://pubmed.ncbi.nlm.nih.gov/28465183/)]
72. Preuss H, Schnicker K, Legenbauer T. ImpulsE zur Verbesserung der Impuls-und Emotionsregulation: Ein kognitiv-verhaltenstherapeutisches Behandlungsprogramm. Göttingen, Germany: Hogrefe Verlag; 2018.
73. Hofmann W, Friese M, Strack F. Impulse and self-control from a dual-systems perspective. *Perspect Psychol Sci* 2009 Mar;4(2):162-176. [doi: [10.1111/j.1745-6924.2009.01116.x](https://doi.org/10.1111/j.1745-6924.2009.01116.x)] [Medline: [26158943](https://pubmed.ncbi.nlm.nih.gov/26158943/)]
74. Lavagnino L, Arnone D, Cao B, Soares JC, Selvaraj S. Inhibitory control in obesity and binge eating disorder: a systematic review and meta-analysis of neurocognitive and neuroimaging studies. *Neurosci Biobehav Rev* 2016 Sep;68:714-726. [doi: [10.1016/j.neubiorev.2016.06.041](https://doi.org/10.1016/j.neubiorev.2016.06.041)] [Medline: [27381956](https://pubmed.ncbi.nlm.nih.gov/27381956/)]
75. Clark MM, Abrams DB, Niaura RS, Eaton CA, Rossi JS. Self-efficacy in weight management. *J Consult Clin Psychol* 1991;59(5):739-744. [doi: [10.1037/0022-006X.59.5.739](https://doi.org/10.1037/0022-006X.59.5.739)]
76. Richman RM, Loughnan GT, Droulers AM, Steinbeck KS, Caterson ID. Self-efficacy in relation to eating behaviour among obese and non-obese women. *Int J Obes Relat Metab Disord* 2001 Jun;25(6):907-913. [doi: [10.1038/sj.ijo.0801606](https://doi.org/10.1038/sj.ijo.0801606)] [Medline: [11439307](https://pubmed.ncbi.nlm.nih.gov/11439307/)]
77. Linde JA, Rothman AJ, Baldwin AS, Jeffery RW. The impact of self-efficacy on behavior change and weight change among overweight participants in a weight loss trial. *Health Psychol* 2006 May;25(3):282-291. [doi: [10.1037/0278-6133.25.3.282](https://doi.org/10.1037/0278-6133.25.3.282)] [Medline: [16719599](https://pubmed.ncbi.nlm.nih.gov/16719599/)]
78. Wolfe BM, Kvach E, Eckel RH. Treatment of obesity. *Circ Res* 2016 May 27;118(11):1844-1855. [doi: [10.1161/circresaha.116.307591](https://doi.org/10.1161/circresaha.116.307591)]
79. Moss-Morris R, Weinman J, Petrie K, Horne R, Cameron L, Buick D. The revised illness perception questionnaire (IPQ-R). *Psychol Health* 2002 Jan;17(1):1-16. [doi: [10.1080/08870440290001494](https://doi.org/10.1080/08870440290001494)]
80. Jakob R, Harperink S, Rudolf AM, Fleisch E, Haug S, Mair JL, et al. Factors influencing adherence to mHealth apps for prevention or management of noncommunicable diseases: systematic review. *J Med Internet Res* 2022 May 25;24(5):e35371 [FREE Full text] [doi: [10.2196/35371](https://doi.org/10.2196/35371)] [Medline: [35612886](https://pubmed.ncbi.nlm.nih.gov/35612886/)]
81. Zhou L, Bao J, Setiawan IM, Saptono A, Parmanto B. The mHealth app usability questionnaire (MAUQ): development and validation study. *JMIR Mhealth Uhealth* 2019 Apr 11;7(4):e11500 [FREE Full text] [doi: [10.2196/11500](https://doi.org/10.2196/11500)] [Medline: [30973342](https://pubmed.ncbi.nlm.nih.gov/30973342/)]
82. Liu Y, Gao X, Zhao J, Zhang L, Chen H. Neurocognitive correlates of food-related response inhibition in overweight/obese adults. *Brain Topogr* 2020 Jan 28;33(1):101-111. [doi: [10.1007/s10548-019-00730-y](https://doi.org/10.1007/s10548-019-00730-y)] [Medline: [31564028](https://pubmed.ncbi.nlm.nih.gov/31564028/)]
83. Duraccio KM, Zaugg K, Jensen CD. Effects of sleep restriction on food-related inhibitory control and reward in adolescents. *J Pediatr Psychol* 2019 Jul 01;44(6):692-702. [doi: [10.1093/jpepsy/jsz008](https://doi.org/10.1093/jpepsy/jsz008)] [Medline: [30861067](https://pubmed.ncbi.nlm.nih.gov/30861067/)]
84. Bruch H. Psychological aspects of overeating and obesity. *Psychosomatics* 1964 Sep 29;5(5):269-274. [doi: [10.1016/S0033-3182\(64\)72385-7](https://doi.org/10.1016/S0033-3182(64)72385-7)]

85. Konttinen H, Männistö S, Sarlio-Lähteenkorva S, Silventoinen K, Haukkala A. Emotional eating, depressive symptoms and self-reported food consumption. A population-based study. *Appetite* 2010 Jun;54(3):473-479. [doi: [10.1016/j.appet.2010.01.014](https://doi.org/10.1016/j.appet.2010.01.014)] [Medline: [20138944](https://pubmed.ncbi.nlm.nih.gov/20138944/)]
86. Henning C, Schroeder S, Steins-Loeber S, Wolstein J. Gender and emotional representation matter: own illness beliefs and their relationship to obesity. *Front Nutr* 2022;9:799831 [FREE Full text] [doi: [10.3389/fnut.2022.799831](https://doi.org/10.3389/fnut.2022.799831)] [Medline: [35211498](https://pubmed.ncbi.nlm.nih.gov/35211498/)]
87. Smith JM, Serier KN, Belon KE, Sebastian RM, Smith JE. Evaluation of the relationships between dietary restraint, emotional eating, and intuitive eating moderated by sex. *Appetite* 2020 Dec 01;155:104817. [doi: [10.1016/j.appet.2020.104817](https://doi.org/10.1016/j.appet.2020.104817)] [Medline: [32739329](https://pubmed.ncbi.nlm.nih.gov/32739329/)]
88. Davis C, Curtis C, Levitan RD, Carter JC, Kaplan AS, Kennedy JL. Evidence that 'food addiction' is a valid phenotype of obesity. *Appetite* 2011 Dec;57(3):711-717. [doi: [10.1016/j.appet.2011.08.017](https://doi.org/10.1016/j.appet.2011.08.017)] [Medline: [21907742](https://pubmed.ncbi.nlm.nih.gov/21907742/)]
89. Meule A. Back by popular demand: a narrative review on the history of food addiction research. *Yale J Biol Med* 2015 Sep;88(3):295-302 [FREE Full text] [Medline: [26339213](https://pubmed.ncbi.nlm.nih.gov/26339213/)]
90. Pape M, Herpertz S, Schroeder S, Seiferth C, Färber T, Wolstein J, et al. Food addiction and its relationship to weight- and addiction-related psychological parameters in individuals with overweight and obesity. *Front Psychol* 2021;12:736454 [FREE Full text] [doi: [10.3389/fpsyg.2021.736454](https://doi.org/10.3389/fpsyg.2021.736454)] [Medline: [34621227](https://pubmed.ncbi.nlm.nih.gov/34621227/)]
91. Pursey KM, Stanwell P, Gearhardt AN, Collins CE, Burrows TL. The prevalence of food addiction as assessed by the Yale Food Addiction Scale: a systematic review. *Nutrients* 2014 Oct 21;6(10):4552-4590 [FREE Full text] [doi: [10.3390/nu6104552](https://doi.org/10.3390/nu6104552)] [Medline: [25338274](https://pubmed.ncbi.nlm.nih.gov/25338274/)]
92. Burmeister JM, Hinman N, Koball A, Hoffmann DA, Carels RA. Food addiction in adults seeking weight loss treatment. Implications for psychosocial health and weight loss. *Appetite* 2013 Jan;60(1):103-110. [doi: [10.1016/j.appet.2012.09.013](https://doi.org/10.1016/j.appet.2012.09.013)] [Medline: [23017467](https://pubmed.ncbi.nlm.nih.gov/23017467/)]
93. Clark SM, Saules KK. Validation of the Yale Food Addiction Scale among a weight-loss surgery population. *Eat Behav* 2013 Apr;14(2):216-219. [doi: [10.1016/j.eatbeh.2013.01.002](https://doi.org/10.1016/j.eatbeh.2013.01.002)] [Medline: [23557824](https://pubmed.ncbi.nlm.nih.gov/23557824/)]
94. Meule A, Hermann T, Kübler A. Food addiction in overweight and obese adolescents seeking weight-loss treatment. *Eur Eat Disord Rev* 2015 May;23(3):193-198. [doi: [10.1002/erv.2355](https://doi.org/10.1002/erv.2355)] [Medline: [25778000](https://pubmed.ncbi.nlm.nih.gov/25778000/)]
95. Fielding-Singh P, Patel ML, King AC, Gardner CD. Baseline psychosocial and demographic factors associated with study attrition and 12-month weight gain in the DIETFITS trial. *Obesity (Silver Spring)* 2019 Dec;27(12):1997-2004 [FREE Full text] [doi: [10.1002/oby.22650](https://doi.org/10.1002/oby.22650)] [Medline: [31633313](https://pubmed.ncbi.nlm.nih.gov/31633313/)]
96. Sikorski C, Spahlholz J, Hartlev M, Riedel-Heller SG. Weight-based discrimination: an ubiquitous phenomenon? *Int J Obes (Lond)* 2016 Feb;40(2):333-337. [doi: [10.1038/ijo.2015.165](https://doi.org/10.1038/ijo.2015.165)] [Medline: [26311336](https://pubmed.ncbi.nlm.nih.gov/26311336/)]
97. Lippi D, Bianucci R, Donell S. Gender medicine: its historical roots. *Postgrad Med J* 2020 Aug;96(1138):480-486. [doi: [10.1136/postgradmedj-2019-137452](https://doi.org/10.1136/postgradmedj-2019-137452)] [Medline: [32471879](https://pubmed.ncbi.nlm.nih.gov/32471879/)]
98. Jake-Schoffman DE, Silfee VJ, Waring ME, Boudreaux ED, Sadasivam RS, Mullen SP, et al. Methods for evaluating the content, usability, and efficacy of commercial mobile health apps. *JMIR Mhealth Uhealth* 2017 Dec 18;5(12):e190 [FREE Full text] [doi: [10.2196/mhealth.8758](https://doi.org/10.2196/mhealth.8758)] [Medline: [29254914](https://pubmed.ncbi.nlm.nih.gov/29254914/)]
99. Signal V, McLeod M, Stanley J, Stairmand J, Sukumaran N, Thompson DM, et al. A mobile- and web-based health intervention program for diabetes and prediabetes self-management (BetaMe/Melon): process evaluation following a randomized controlled trial. *J Med Internet Res* 2020 Dec 01;22(12):e19150 [FREE Full text] [doi: [10.2196/19150](https://doi.org/10.2196/19150)] [Medline: [33258776](https://pubmed.ncbi.nlm.nih.gov/33258776/)]
100. Ruf A, Koch ED, Ebner-Priemer U, Knopf M, Reif A, Matura S. Studying microtemporal, within-person processes of diet, physical activity, and related factors using the APPetite-Mobile-App: feasibility, usability, and validation study. *J Med Internet Res* 2021 Jul 05;23(7):e25850 [FREE Full text] [doi: [10.2196/25850](https://doi.org/10.2196/25850)] [Medline: [34342268](https://pubmed.ncbi.nlm.nih.gov/34342268/)]
101. van Beurden SB, Smith JR, Lawrence NS, Abraham C, Greaves CJ. Feasibility randomized controlled trial of ImpulsePal: smartphone app-based weight management intervention to reduce impulsive eating in overweight adults. *JMIR Form Res* 2019 Apr 30;3(2):e11586 [FREE Full text] [doi: [10.2196/11586](https://doi.org/10.2196/11586)] [Medline: [31038464](https://pubmed.ncbi.nlm.nih.gov/31038464/)]
102. Butryn ML, Martinelli MK, Crane NT, Godfrey K, Roberts SR, Zhang F, et al. Counselor surveillance of digital self-monitoring data: a pilot randomized controlled trial. *Obesity (Silver Spring)* 2020 Dec;28(12):2339-2346 [FREE Full text] [doi: [10.1002/oby.23015](https://doi.org/10.1002/oby.23015)] [Medline: [33098278](https://pubmed.ncbi.nlm.nih.gov/33098278/)]
103. Eysenbach G. The law of attrition. *J Med Internet Res* 2005 Mar 31;7(1):e11-e24 [FREE Full text] [doi: [10.2196/jmir.7.1.e11](https://doi.org/10.2196/jmir.7.1.e11)] [Medline: [15829473](https://pubmed.ncbi.nlm.nih.gov/15829473/)]
104. Bühlmeier J, Frölich S, Ludwig C, Knoll-Pientka N, Schmidt B, Föcker M, et al. Changes in patterns of eating habits and food intake during the first German COVID-19 lockdown: results of a cross-sectional online survey. *Eur J Nutr* 2022 Sep 03;61(6):3293-3306 [FREE Full text] [doi: [10.1007/s00394-022-02919-7](https://doi.org/10.1007/s00394-022-02919-7)] [Medline: [35759031](https://pubmed.ncbi.nlm.nih.gov/35759031/)]
105. Moehring K, Weiland A, Reifenscheid M, Naumann E, Wenz A, Rettig T. Inequality in employment trajectories and their socio-economic consequences during the early phase of the COVID-19 pandemic in Germany. *JMIR Hum Factors* 2018 Apr 17;5(2):e16 [FREE Full text] [doi: [10.31235/osf.io/m95df](https://doi.org/10.31235/osf.io/m95df)]
106. Senecal C, Widmer RJ, Larrabee BR, de Andrade M, Lerman LO, Lerman A, et al. A digital health weight loss program in 250,000 individuals. *J Obes* 2020 Mar 26;2020(5):1-8 [FREE Full text] [doi: [10.1155/2020/9497164](https://doi.org/10.1155/2020/9497164)]

107. Andone I, Błaszkiwicz K, Eibes M, Trendafilov B, Montag C, Markowetz A. How age and gender affect smartphone usage. *Proc Natl Acad Sci U S A* 2003 Sep 30;100(20):11696-11701 [[FREE Full text](#)] [doi: [10.1145/2968219.2971451](https://doi.org/10.1145/2968219.2971451)]
108. Yang Y, Koenigstorfer J. Determinants of fitness app usage and moderating impacts of education-, motivation-, and gamification-related app features on physical activity intentions: cross-sectional survey study. *J Med Internet Res* 2021 Jul 13;23(7):e26063 [[FREE Full text](#)] [doi: [10.2196/26063](https://doi.org/10.2196/26063)] [Medline: [34255656](https://pubmed.ncbi.nlm.nih.gov/34255656/)]
109. Smail-Crevier R, Powers G, Noel C, Wang J. Health-related internet usage and design feature preference for e-mental health programs among men and women. *J Med Internet Res* 2019 Mar 18;21(3):e11224-e11503 [[FREE Full text](#)] [doi: [10.2196/11224](https://doi.org/10.2196/11224)] [Medline: [30882361](https://pubmed.ncbi.nlm.nih.gov/30882361/)]
110. Escoffery C. Gender similarities and differences for e-health behaviors among U.S. adults. *Telemed J E Health* 2018 May 18;24(5):335-343. [doi: [10.1089/tmj.2017.0136](https://doi.org/10.1089/tmj.2017.0136)] [Medline: [28813630](https://pubmed.ncbi.nlm.nih.gov/28813630/)]
111. Bol N, Helberger N, Weert J. Differences in mobile health app use: a source of new digital inequalities? *The Information Society* 2018 Apr 26;34(3):183-193 [[FREE Full text](#)] [doi: [10.1080/01972243.2018.1438550](https://doi.org/10.1080/01972243.2018.1438550)]
112. Qan'ir Y, Khalifeh AH, Eid M, Hammad B, Al-Batran M. Mobile health apps use among Jordanian outpatients: a descriptive study. *Health Informatics J* 2021 May;27(2):14604582211017940-14604582vi, xxiii [[FREE Full text](#)] [doi: [10.1177/14604582211017940](https://doi.org/10.1177/14604582211017940)] [Medline: [34030504](https://pubmed.ncbi.nlm.nih.gov/34030504/)]
113. Carroll JK, Moorhead A, Bond R, LeBlanc WG, Petrella RJ, Fiscella K. Who uses mobile phone health apps and does use matter? A secondary data analytics approach. *J Med Internet Res* 2017 Apr 19;19(4):e125 [[FREE Full text](#)] [doi: [10.2196/jmir.5604](https://doi.org/10.2196/jmir.5604)] [Medline: [28428170](https://pubmed.ncbi.nlm.nih.gov/28428170/)]
114. Milne-Ives M, Lam C, De Cock C, Van Velthoven MH, Meinert E. Mobile apps for health behavior change in physical activity, diet, drug and alcohol use, and mental health: systematic review. *JMIR Mhealth Uhealth* 2020 Mar 18;8(3):e17046-e17511 [[FREE Full text](#)] [doi: [10.2196/17046](https://doi.org/10.2196/17046)] [Medline: [32186518](https://pubmed.ncbi.nlm.nih.gov/32186518/)]
115. Dunn CG, Turner-McGrievy GM, Wilcox S, Hutto B. Dietary self-monitoring through calorie tracking but not through a digital photography app is associated with significant weight loss: the 2SMART pilot study-a 6-month randomized trial. *J Acad Nutr Diet* 2019 Sep;19(9):1525-1532. [doi: [10.1016/j.jand.2019.03.013](https://doi.org/10.1016/j.jand.2019.03.013)] [Medline: [31155474](https://pubmed.ncbi.nlm.nih.gov/31155474/)]
116. Kliemann N, Croker H, Johnson F, Beeken RJ. Development of the top tips habit-based weight loss app and preliminary indications of its usage, effectiveness, and acceptability: mixed-methods pilot study. *JMIR Mhealth Uhealth* 2019 May 10;7(5):e12326 [[FREE Full text](#)] [doi: [10.2196/12326](https://doi.org/10.2196/12326)] [Medline: [31094352](https://pubmed.ncbi.nlm.nih.gov/31094352/)]
117. López D, Torres M, Vélez J, Grullon J, Negrón E, Pérez CM, et al. Development and evaluation of a nutritional smartphone application for making smart and healthy choices in grocery shopping. *Healthc Inform Res* 2017 Jan;23(1):16-24 [[FREE Full text](#)] [doi: [10.4258/hir.2017.23.1.16](https://doi.org/10.4258/hir.2017.23.1.16)] [Medline: [28261527](https://pubmed.ncbi.nlm.nih.gov/28261527/)]
118. Acharya SD, Elci OU, Sereika SM, Music E, Styn MA, Turk MW, et al. Adherence to a behavioral weight loss treatment program enhances weight loss and improvements in biomarkers. *Patient Prefer Adherence* 2009 Nov 03;3(3):151-160 [[FREE Full text](#)] [doi: [10.2147/ppa.s5802](https://doi.org/10.2147/ppa.s5802)] [Medline: [19936157](https://pubmed.ncbi.nlm.nih.gov/19936157/)]
119. Alnasser A, Kyle J, Alkhalifah A, Marais D. Relationship between evidence requirements, user expectations, and actual experiences: usability evaluation of the Twazon Arabic weight loss app. *JMIR Hum Factors* 2018 Apr 17;5(2):e16-e18 [[FREE Full text](#)] [doi: [10.2196/humanfactors.9765](https://doi.org/10.2196/humanfactors.9765)] [Medline: [29666042](https://pubmed.ncbi.nlm.nih.gov/29666042/)]
120. Ferrara G, Kim J, Lin S, Hua J, Seto E. A focused review of smartphone diet-tracking apps: usability, functionality, coherence with behavior change theory, and comparative validity of nutrient intake and energy estimates. *JMIR Mhealth Uhealth* 2019 May 17;7(5):e9232 [[FREE Full text](#)] [doi: [10.2196/mhealth.9232](https://doi.org/10.2196/mhealth.9232)] [Medline: [31102369](https://pubmed.ncbi.nlm.nih.gov/31102369/)]

## Abbreviations

- EE:** emotional eating
- FA:** food addiction
- IPQ-R:** revised illness perception questionnaire
- mHealth:** mobile health
- WLP:** weight loss program

*Edited by A Mavragani; submitted 05.04.22; peer-reviewed by H Meggy, H Ranjani; comments to author 11.07.22; revised version received 11.08.22; accepted 29.08.22; published 27.10.22.*

### *Please cite as:*

Pape M, Färber T, Seiferth C, Roth T, Schroeder S, Wolstein J, Herpertz S, Steins-Loeber S  
A Tailored Gender-Sensitive mHealth Weight Loss Intervention (I-GENDO): Development and Process Evaluation  
*JMIR Form Res* 2022;6(10):e38480  
URL: <https://formative.jmir.org/2022/10/e38480>  
doi: [10.2196/38480](https://doi.org/10.2196/38480)  
PMID: [36301614](https://pubmed.ncbi.nlm.nih.gov/36301614/)

©Magdalena Pape, Tanja Färber, Caroline Seiferth, Tanja Roth, Stefanie Schroeder, Joerg Wolstein, Stephan Herpertz, Sabine Steins-Loeber. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 27.10.2022. 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

# From Real-world Individuals' Data to National Health Indicators: Multiphase Pilot Study in Gabon

Aimé Patrice Koumamba<sup>1</sup>, PhD; Edgard Brice Ngougou<sup>2</sup>, PhD; Jean Engohang-Ndong<sup>3</sup>, PhD; Euloge Ibinga<sup>2</sup>, PhD; Raymond Ondzigue Mbenga<sup>4</sup>, MPH; Gayo Diallo<sup>1</sup>, PhD

<sup>1</sup>Centre de recherche sur la santé des populations de Bordeaux, Institut national de la santé et de la recherche médicale 1219, Université de Bordeaux, Bordeaux, France

<sup>2</sup>Research Unit in Epidemiology of Chronic Diseases and Environmental Health, University of Health Sciences, Libreville, Gabon

<sup>3</sup>Department of Biological Sciences, Kent State University at Tuscarawas, Kent, OH, United States

<sup>4</sup>Laboratoire d'Informatique fondamentale appliquée de Tours, University of Tours, Tours, France

**Corresponding Author:**

Aimé Patrice Koumamba, PhD

Centre de recherche sur la santé des populations de Bordeaux

Institut national de la santé et de la recherche médicale 1219

Université de Bordeaux

146 Rue Léo Saignat 11

Bordeaux, 33076

France

Phone: 33 77921571

Fax: 33 77921571

Email: [aime-patrice.koumamba@u-bordeaux.fr](mailto:aime-patrice.koumamba@u-bordeaux.fr)

## Abstract

**Background:** Achieving health goals requires informed decision-making supported by transparent, reliable, and relevant health information. This helps decision makers, such as health managers, to better understand the functioning of their health system and improve their ability to respond quickly to health demands. To achieve this, the health system needs to be supported by a digitized decision-making information system. In Sub-Saharan African countries, inadequate digital infrastructure, including limited internet connectivity and insufficient access to appropriate computer software, makes it difficult to collect, process, and analyze data for health statistics. The processing of data is done manually in this case; however, this situation affects the quality of the health statistics produced and compromises the quality of health intervention choices in these countries.

**Objective:** This study aimed to describe the conceptual approach of a data production and dissemination platform model proposed and implemented in Gabon. More precisely, it aimed to present the approach applied for the multidimensional analysis of the data production and dissemination process in the existing information system and present the results of an evaluation of the proposed model implemented in a real context.

**Methods:** The research was carried out in 3 phases. First, a platform was designed and developed based on the examination of the various data production and indicator generation procedures. Then, the platform was implemented in chosen health facilities in Gabon. Finally, a platform evaluation was carried out with actual end users.

**Results:** A total of 14 users with 12 years of average experience in health data management were interviewed. The results show that the use of the proposed model significantly improved the completeness, timeliness, and accuracy of data compared with the traditional system (93% vs 12%,  $P<.001$ ; 96% vs 18%,  $P<.001$ ; and 100% vs 18%,  $P<.001$ ; respectively).

**Conclusions:** The proposed model contributes significantly to the improvement of health data quality in Gabon.

(JMIR Form Res 2022;6(10):e35176) doi:[10.2196/35176](https://doi.org/10.2196/35176)

**KEYWORDS**

health data quality; decision support information system; low- and middle-income countries; LMICs; mobile phone

## Introduction

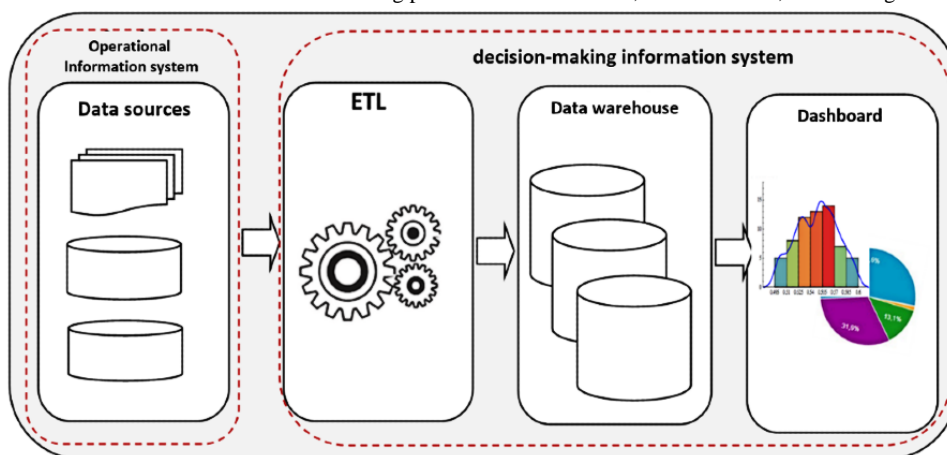
### Background

In the 21st century, a reliable health information system (HIS) must rely on a digitalized decision-making information system (DMIS), which is “a bundle of subject-oriented, integrated, time-evolving, and nonvolatile data designed to aid managers in their decision-making” [1]. The performance of such a DMIS depends on its ability to manage data from heterogeneous sources, and it uses a data warehousing approach to help the

HIS perform well. Thus, it is important to consider transactional information systems in the development of a DMIS (Figure 1).

In any organization, the quality of the choice of interventions to be undertaken is linked to the visibility that decision makers have of the performance of their organizations. For any health system, the decision-making level needs to rely on a DMIS, which, by means of an extraction, transformation, and loading (ETL) approach, can process data from the transactional systems [2]. This then allows the understanding that the performance of a DMIS in terms of producing quality data depends on its best interaction with transactional systems.

**Figure 1.** Simplified workflow of a data-enabled decision-making process. ETL: extraction, transformation, and loading.



### Challenges

Owing to restricted digital infrastructures, including limited internet connectivity or insufficient access to proper computer software, it is difficult for many health systems to obtain high-quality data. According to the Global Health Observatory, 81 nations (42%) gather data of extremely low quality. This data quality issue is a significant hindrance to the management of health services, the creation of health statistics, and even the monitoring of health. In Sub-Saharan African (SSA) countries, the problem is even more significant and undermines the quality of choices for health interventions [3]. In these countries, the organization of HIS seems to have some shortcomings, which could explain the data quality issue [4].

### HIS Implementation in SSA Countries

The implementation of HISs in SSA has been largely supported by the recent version of District Health Information Software 2 (DHIS2; Health Information Systems Programme) [5]. The latter was developed under the leadership of the University of Oslo in Norway [6,7]. The management of health data through the DHIS2 platform is now effective in many SSA countries [8]. However, despite this, the performance of information systems remains problematic in many SSA countries, including in Gabon. Indeed, in most health facilities, aggregated data are collected manually by data managers [9], who are in charge of integrating these aggregated data into the DHIS2 platform, contextualized according to the country and health system. Nevertheless, this approach to manual upstream data processing often results in a large number of human errors, including incomplete data. Therefore, the manual data collection

techniques are onerous and time-consuming and increase the workload of data managers who must collect data from the many health programs that generate field data [10]. Although DHIS2 is widely implemented in several SSA health systems, this platform does not appear to account for data production at the fundamental level, that is, at the health facilities where patients are seen by a health care provider (apart from the recent e-Tracker service in the context of the COVID-19 pandemic). However, because of the huge volume of data produced, their exploitation would be facilitated by the use of data ETL technologies to permit multidimensional analysis at the smallest feasible resolution if they were digitized. The poor use of digitalized hospital information systems is not only a source of demotivation for data managers but also a significant barrier to the development of trustworthy indicators for directing decision-making in health services. To rectify this predicament, some health system officers have decided to use the e-Tracker DHIS2 technology to digitize all data management tasks [7]. However, despite such initiatives to help improve the quality of data for decision support, the difficulties in getting integrated information systems that take into account individual data are still present. In most resource-limited countries, specifically in SSA, the implementation of decision support tools is largely supported by the DHIS2 system and its data warehouse (DW) platform. Unfortunately, this platform is not focused on data processing at the operational level, particularly at the level of primary data sources (consisting of transactional systems and where data are produced). However, the DHIS2 platform facilitates the collection of aggregated data at the level of the health department, health region, and central programs that are vertically oriented (HIV, tuberculosis, malaria, etc). The

operational level such as hospitals, clinics, and health centers, which beyond the operational (medical care), which ensures the day-to-day management of the organization, is rarely (or even not) taken into account. The process of data processing, also known as ETL of data into the DW, is done manually in most cases [11].

Similar to many other SSA countries, Gabon has a data quality issue [12]. To make the issue even more complex, Gabon has a fragmented information system without any formal coordination. In a context in which a health system generates voluminous, very diverse, and heterogeneous data, the information system should be based on well-structured transactional information systems and be continually fed. In Gabon, unfortunately, upstream of the HIS, data processing at the operating system level is carried out manually. As a result, the data quality is poor [4]. Overall, the quality of the data produced may be affected by the lack of data completeness (ie, the data produced are incomplete), data accuracy (ie, data recorded are different from those actually produced), and data timeliness (ie, data are made available to decision makers after a long delay).

According to the above observations, the following question arises: what information system should be adopted to improve the performance of the system and provide quality data? In order to answer this question, a pilot study was undertaken in Gabon. The main purpose was to implement a platform integrating both health-related operational and decision-making data.

## Objectives

The main objective of this study was to describe the design of the conceptual approach of the platform and its implementation

in Gabon. It enabled the multidimensional analysis of health information for management support to analyze the data production and dissemination processes in the existing information system and evaluate the multidimensional model and framework implemented in a real context setting.

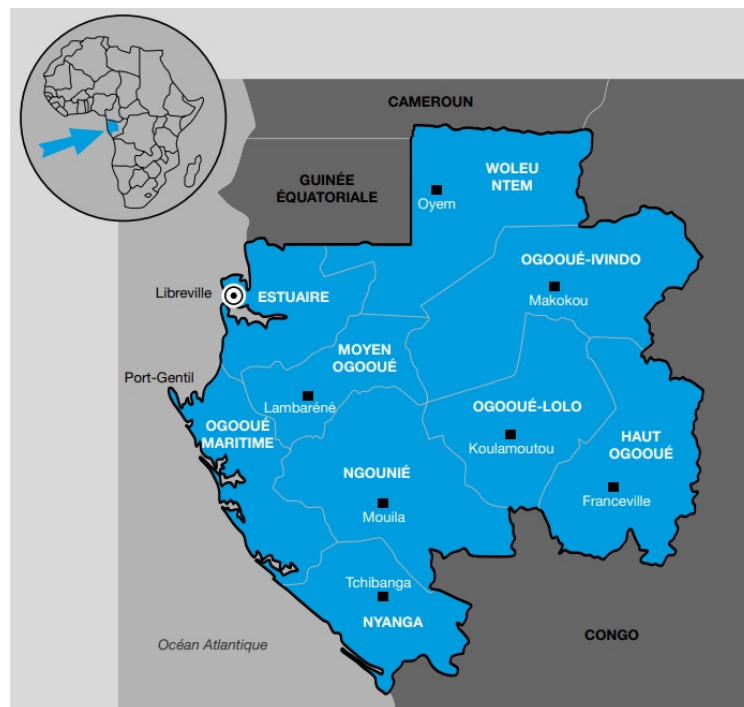
## Gabon and Its Health System Context

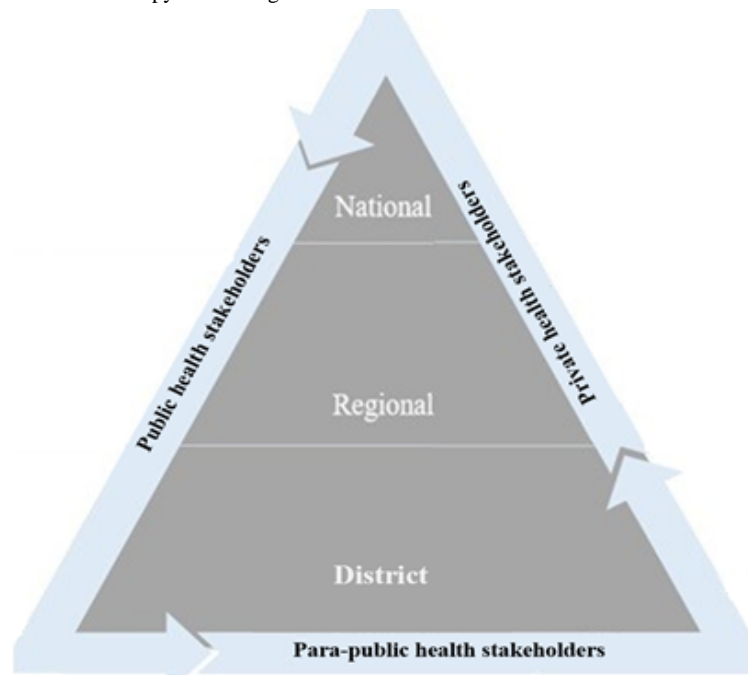
Located in Central Africa, Gabon is a country with a surface area of 267,667 km<sup>2</sup> and an estimated population of 1,811,079 as of 2021. It is bordered by Equatorial Guinea, Cameroon, Congo Brazzaville, and the Atlantic Ocean, which extends for about 800 km along its coast (Figure 2) [13].

The health system in Gabon comprises 3 types of stakeholders: public, private, and para-public. The entire system is structured as a pyramidal organization, with health districts at the bottom, regions in the middle, and the national level at the top, which coordinates and centralizes all health care activities and data in the country (Figure 3).

When the digital infrastructure is adequate, health care institutions are typically digitalized with a reliable internet connection, computers, software, and health care management software. Moreover, all patient care operations, including data-generating activities, are frequently combined in a digitized patient file. This is the primary data source used by ETL systems to automatically populate clinical DWs. However, in Gabon, this is not always the case. Thus, health facilities are compelled to collect data using paper records, which makes it challenging to use these data for decision-making [12].

**Figure 2.** Geographical location of Gabon.



**Figure 3.** The health system's stakeholders and pyramidal organization in Gabon.

Currently, the processes used in Gabon do not permit decision makers to differentiate between aggregated data collected from the district and individual health information data collected from each health facility inside the district. As an illustration, in 2019, the Komo-Mondah Health District, which consists of 162 health facilities, reported malaria as the leading cause of medical visits in the district. It accounted for 49% of all recorded health consultations (2966 cases out of 6054 medical consultations). However, investigators (statisticians) were unable to determine the distribution of malaria cases by health facility within the health district. This lack of detail does not promote targeted and relevant health management decision-making.

Therefore, it was necessary to brainstorm and create a system or model that would be suited for the automatic gathering and storage of health data and could link these data with their production sources. A system that could function effectively even in the absence of a dependable internet connection or limited access to computers. Such a system could enable the end user to enter data into the system using a smartphone connected to the national mobile network. This option is widely available in Gabon, as more than 80% of the mobile phone users have access to a smartphone with a reliable internet connection that enables the transfer of data in almost real time, thus eliminating the need to record health data on paper-based forms and allowing health professionals to instantly access the tool using previously assigned credentials.

The purpose was to minimize manual and delayed data processing, both of which are potential sources of errors that could influence data quality. The model suggested in this study includes built-in variables that provide the entry of health data contextualized by health facility and by health service, thereby enabling data processing employees and investigators to analyze digitized data with relevant details.

The Gabonese health system is organized into 10 health regions and 51 health districts. It comprises 1043 health facilities, the

main ones being 3 university hospital centers, 9 regional hospital centers, 51 medical centers, and 2 large hospitals with a capacity comparable with that of a university hospital center [14].

## Methods

### Overview

The implementation of any HIS requires the consideration of all the processes that contribute to the production of data and the dissemination of information. This makes it possible to highlight the sequence of activities enabling the transformation of data into information, information into knowledge, and knowledge into action [14]. Thus, to define the different relationships between the data structures, in this study, data were first collected from semidirective interviews with the different health actors involved in these different processes. Following data collection, a model of DMIS based on 2 modeling methods was proposed. First, the process method aims to break down activities step by step to study their functioning and their interactions to improve the organization of the overall system. For this study, we used Bizagi, a tool for designing architecture models in information systems development [15]. This tool allowed the research team to describe and represent the different data collection processes in the Gabonese HIS. Second, the method of study and computer implementation for business systems is a popular method in French-speaking countries, whereby an information system is designed independently of the technical choices for its implementation [16]. This allowed us to design the conceptual data model while identifying the different actors and how they interacted with each other at different levels of the system.

### Analysis of the Different Interactions in the Data Production Process and Modeling of the DMIS

Owing to the fact that operational systems such as HIS feed DMIS with data, it was vital for the purposes of this study to

analyze the numerous activities that occur within the health care facilities, where the primary data are produced. This allows for a realistic depiction of the many interactions among the actors involved in data production. This description emphasizes the sequence of medical care activities that generate diverse sorts of data. The operational level, which represents the numerous information sources, was now digitalized, and the data integration requirements were easily met by designing an ETL procedure. The latter ensures that the data are processed and loaded into a specific DW [17]. In cases where data production processes were not digitalized at the initial production level, it was determined that for the purposes of this pilot study specifically, which was focused on enabling decision support, epidemiological surveillance, and monitoring and evaluation of the performance of the health system (as opposed to medical follow-up or care), the framework should address all processes from data production to information dissemination.

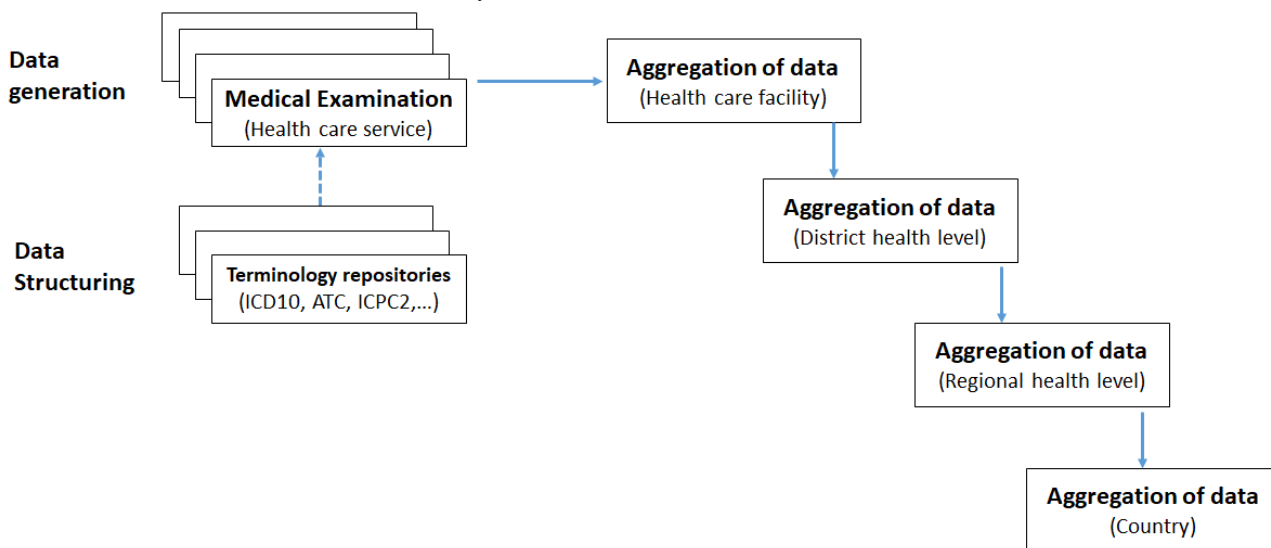
### Data Production Process

The production of data that were fed into the DMIS was based on a set of activities that were associated with medical diagnoses performed in a given medical service (operational system). During the medical visit phase, the patient was identified as the central participant around whom the medical staff performed all activities associated with the medical consultation, which produced a set of data. These data consisted mainly, but not exclusively, of 3 types of records:

- *Sociodemographic data*, which were information related to age, gender, nationality, place of residence, and date of consultation of the patient. In addition, the records indicated the name and role of the health provider.
- *Economic data*, which were information associated with the health insurance status, occupation, and copayment rate of the patient. With regard to health insurance, it was systematically indicated whether the patient received health coverage from the government.
- *Clinical data*, which included vital signs, medical diagnoses, drugs prescribed, and types of laboratory tests prescribed.

Traditionally, when the system is not digitalized, all raw data are recorded on paper-based forms. Data processing at all levels becomes problematic. Indeed, at all decision-making levels, data are processed and aggregated manually (Figure 4). In this case, their semantic structuring, at least to give them a semantic coherence, might be done in a very subjective way. Manual matching is sometimes carried out subjectively by data managers or statisticians who proceed to perform the semantic matching of local medical concepts with other standard concepts of terminological reference systems such as the International Classification of Diseases (ICD)-10 for diagnoses and the Anatomical Therapeutic Chemical for drugs. The construction of the database model, on which the DMIS to be set up will be based, has, therefore, taken all these elements (as described in Figure 4) into account.

**Figure 4.** Production process of the source and group requirements graph. ATC: Anatomical Therapeutic Chemical; ICD: International Classification of Diseases; ICPC: International Classification of Primary Care.



### Conceptual Modeling of the DMIS

The design of the DMIS considered two aspects: (1) the data production processes (bottom-up approach) and (2) the needs of the users (top-down approach). This led to the identification of 5 decision-making levels for the construction of the DMIS (Table 1).

Conceptually, this model is based on a relational database model (Figure 5). The objective is to propose a DMIS that integrates all the different data for the relational online analytical

processing (ROLAP) solution to be performed to allow a multidimensional analysis of the data.

A digital interface was integrated into this relational database model, into which data managers enter data from the various hospital registers. This interface is simply a digital replica of the different hospital registers, using the same items, except for the data on patient identity. The data entered are detailed at the finest possible level of granularity and stored in the various relational tables of the tool. Then, a Structured Query Language (SQL) query was integrated to execute and generate a data table

(a fact table) from which all multidimensional analyses can be performed.

The model summarized in Figure 5 integrates all the processes contributing to the performance of an HIS, from data collection to the dissemination of useful information for decision support. In this conceptual data model, each decision-making level was organized as a relational table. The tables were then linked together by logical relationships. These relationships made it possible to identify all the existing correlations among the data stored in different tables. This later facilitated their extraction and loading into a table generated using SQL queries.

The structuring of data before they are loaded into the DW has always been one of the major obstacles to the successful implementation of DMIS [18-21]. In organizations where the operational level is digitalized, this constraint is often solved by using an ETL tool that extracts and transforms the data before

they are loaded into the DW. This allows a significant gain in data production time, with efficiency and precision in terms of data quality.

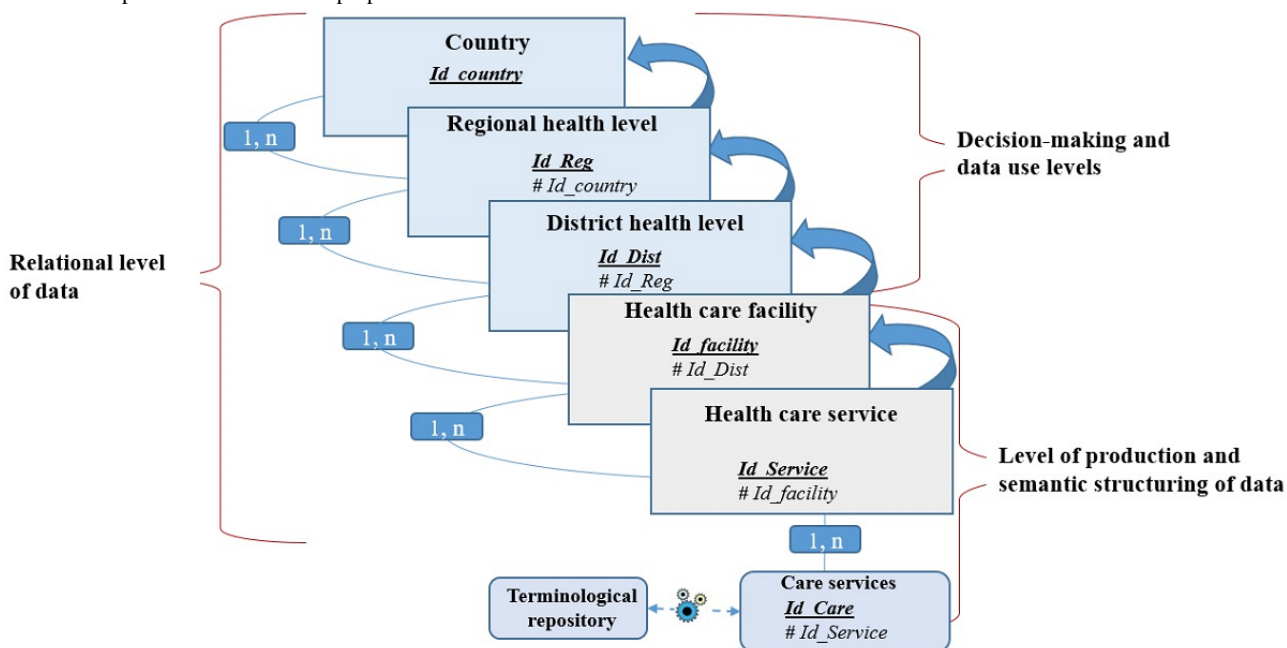
This study was set up in a context where most operational systems (care structures) did not have patient records digitalized, with terminological standards such as ICD-10 integrated to structure medical data and facilitate their automatic processing [22].

Often, to semantically standardize vocabulary terms, data managers or statisticians had no choice but to manually execute their correspondence with standardized terms of ICD-10, which not only made the task of collecting, processing, and disseminating data very laborious but also affected the quality of data in terms of completeness and delayed the production of statistics.

**Table 1.** Levels of data production and use of the produced data for decision-making.

Levels	Types of links between the different levels	Types of data production and use of the produced data for decision-making
Country	Composed of one or more health regions, but a health region belongs to only one country	Forth level of data aggregation. The country aggregates data from all health regions for national decision-making.
Regional health level	Composed of one or more health districts, but a health district belongs to only one health region	Third level of data aggregation. The health region aggregates data from all health districts for regional decision-making.
District health level	Composed of one or more health facilities, but a health facility belongs to only one health district	Second level of data aggregation. The health district aggregates data from all health care facilities for decision-making at the health district level.
Health care facility	Composed of one or more health care services, but a health care service belongs to only one health care facility	First level of data aggregation. The facility aggregates data from all health care services for decision-making at the facility level.
Health care service	Produces data during one or several medical consultations per day, but a medical consultation can occur at only one medical service	The health care service is the primary producer of data, as it provides various medical services including but not limited to drug prescriptions, medical examinations, and medical diagnoses on behalf of the medical facility to which it belongs. Therefore, the service is the producer of all health information in the health system.

**Figure 5.** Conceptual data model of the proposed data warehouse.



To address this issue, an algorithm was defined to help the semantic matching of data and was integrated into a tool that enables the structuring of data and facilitates their automatic processing.

The proposed tool provides a graphical user interface, allowing the user to map an unstructured term with a structured term of ICD-10 from other platforms of assistance to coding or semantic annotation based on medical ontologies. This was done to match medical terms with each other and strengthen their understanding by gaining access to more detailed knowledge elements, such as definitions and synonyms [17]. A match score between 0 and 1 was automatically generated to evaluate the quality of a given match. This score was calculated from the total number of matches established between a given local term and a given standard term. The closer its value was to 1, the better the match.

The ROLAP approach was used for the multidimensional representation of data. The ROLAP, using an SQL query-generated table containing the results (primary keys and structured data of the different tables) (Figure 6), allowed all analyses to be performed from several decision angles.

Figure 6 describes the architecture of the platform implemented in Gabon as well as the SQL query that allows the raw data to be stored at the finest possible level of granularity in a materialized view, which is assimilated to a fact table and integrates the data contained in the various tables of the database deployed in the platform. This materialized view avoids manual processing (ETL) of data before integration into the DW. It uses the primary keys of all other tables as foreign keys to retrieve data from all the relational tables (Figure 7). Thus, a simple multidimensional star model containing the following dimensions was designed as shown in Figure 7.

Figure 6. Architecture of the proposed platform model and SQL query generating the fact table. SQL: Structured Query Language.

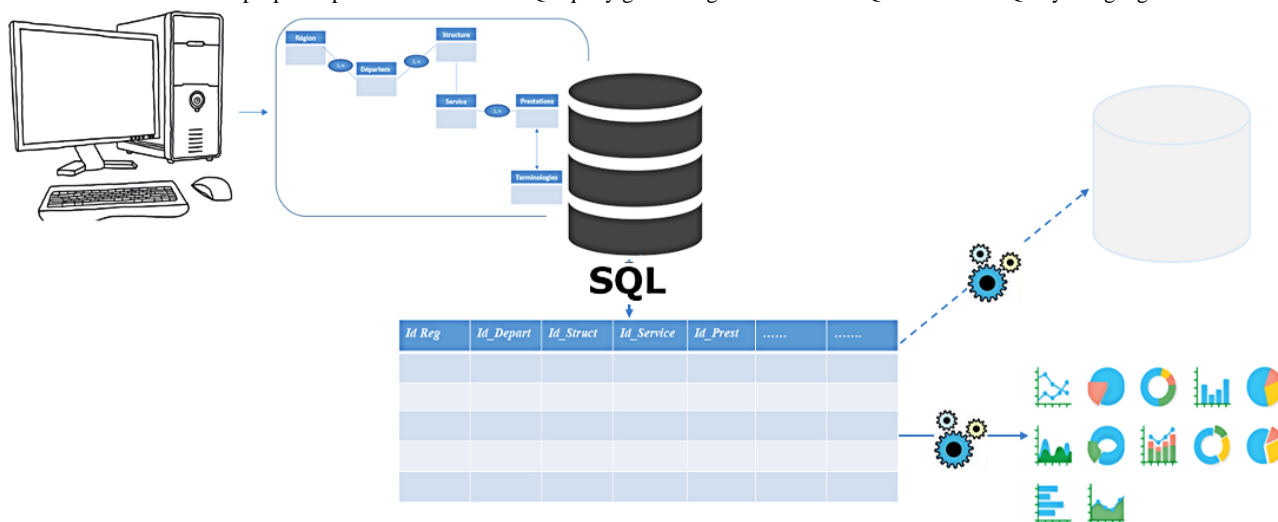
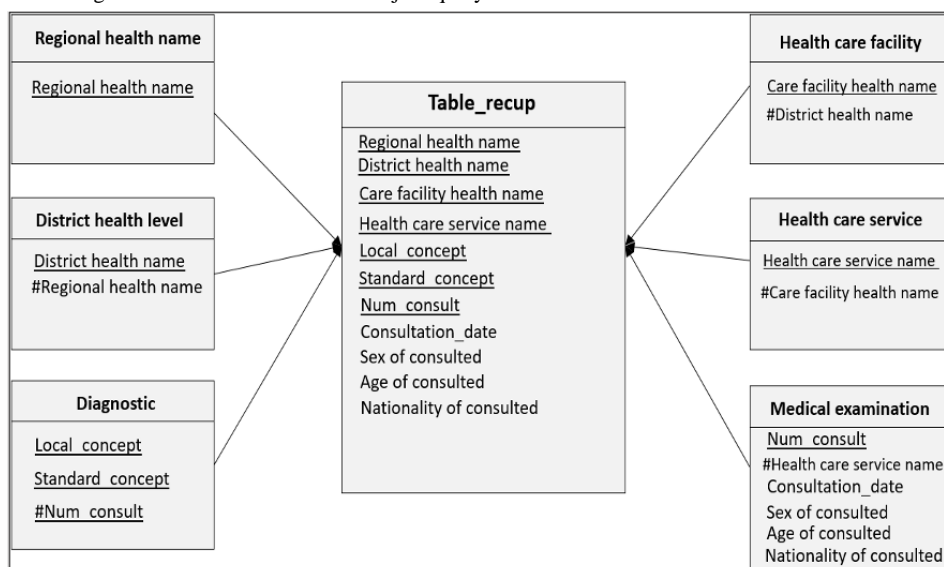


Figure 7. Example of a star diagram of the model centered on a join query.



Geographical location dimensions were as follows:

- Regional health level (regional health name)
- Health district level (health district name and regional health name)
- Health care facility (health care facility name, administrative category, and district health name)

- Health care service (health care service name and health care facility name)

Medico-administrative dimensions were as follows:

- Consultation (Num examination, examination date, patient sex, patient age, patient nationality, Consultation Residence, and Service ID)
- Diagnosis (CodeICD10 or ICPC2, local concept (diagnosis), standard concept (diagnosis), and Num examination)
- Drug (Anatomical Therapeutic Chemical code, local concept, standard concept, and Num examination)
- Laboratory (CodeLOINC, local concept, standard concept, and Num examination)

The table of facts contained the following:

- Retrieval table (Regional health name, health district name, health care facility name, health care service name, Num examination, CodeICD10 or ICPC2, CodeATC, CodeLOINC, Examination date, patient sex, patient age, patient nationality, patient habitation, local concept [diagnosis] standard concept [diagnosis], local concept [medicine], standard concept [medicine], local concept [biology], standard concept [biology], etc)

In contrast to the current HIS in Gabon, where data are aggregated manually, the model developed in this study allows data integration automation and the performance of relevant analyses and visualizations with more detailed levels of granularity. Using this model, there is a possibility to provide meaningful information from the data by considering the time such as day, week, month, and year, and the provenance of the data, including the health department, health region, and country [23].

**Design of the DMIS**

The DW was implemented using the MySQL relational database management system with a multidimensional ROLAP data model.

The implemented system, referred to as *Routine Info*, has a general interface allowing the user to connect using a log-in and password attached to one of the decision levels (health service, health facility, health district, regional health, and country), which constitute, at the same time, different connection profiles to the tool with functionalities distributed as shown in Table 2.

**Table 2.** Accessibility level of functionalities according to profiles.

Functionality profiles	Create a record	Record clinical data	Match local concepts to standard concepts in terminology repositories	Update data entered (delete, modify, or add)	Analyze and perform multidimensional cross-checking of data	Visualize data	Exporting service data
Health care service profile	✓	✓	✓	✓	✓	✓	✓
Health care facility profile				✓	✓	✓	✓
District health profile				✓	✓	✓	✓
Regional health profile				✓	✓	✓	✓
Country profile				✓	✓	✓	✓

**Implementation of the DMIS Pilot**

For the implementation and testing of the *Routine Info* platform, 2 health regions of Gabon were chosen. For 2 months, the platform was routinely used simultaneously in a health center, the regional hospital, and a university hospital to respond to the 3 levels of the health pyramid (department, region, and country).

The data managers of these facilities were asked to fill in the platform with the raw data available in the medical visit registers. The data filled in included clinical, demographic, and socioeconomic data. The objective was to compare the quality of the data reported using the *Routine Info* platform with that of the data produced by the traditional system used within the HIS in Gabon. During the 2 months of use, actual data from various medical consultations were integrated in near real time. For reporting purposes, Tableau 2020.1 software was connected to this database for data analysis and visualization.

The implementation of the platform in these 2 regions made it possible to obtain data associated with medical consultations recorded from January 20, 2018, to July 12, 2018. The results gathered made it possible to carry out multivariate analyses at

all levels of the health pyramid. These results have been shown to support the decisions of decision makers at all levels.

**Scalability and Evaluation of the Pilot**

To scale up the model, a pilot implementation phase was carried out over a period of 2 months in health facilities in 2 of the 10 health regions of Gabon. These 2 health regions serve 49% (8,87,428/1,811,079) of the general population, employ 34% (2125/6250) of health professionals, and have 31.95% (333/1042) of the health facilities in Gabon [13].

At the end of this pilot phase, the *Routine Info* tool was evaluated through a survey interview of all the users who participated in the pilot phase.

This evaluation was based on 3 determinants contributing to the achievement of data quality, namely the ability to have fully recorded data, the ability to facilitate the verification of data accuracy, and the ability to have all data available in a timely manner.

The evaluation method used was a comparative one, whereby the surveyed users filled in 2 different questionnaires with identical questions to evaluate and compare *Routine Info* with



the traditional system. The questionnaires consisted of the following questions:

- Does the proposed system have a database that stores all data recorded in the facility daily?
- Does the proposed system allow for automatic data recording?
- Does the system allow for real-time data entry of all the data contained in the register?
- Does the system allow for the timely production of statistics?
- In how many days, on average, can the user have the statistics with the proposed system?
- With the proposed system, did the user produce the report for the previous month in a timely manner, meaning by the 5th of the current month?
- Is it possible to disaggregate the data by health facility and by service in the system?
- With the proposed system, is it possible to disaggregate the data and check its conformity with the data contained in the register of a given service?
- Is it possible to check the accuracy of the data with the proposed system?
- What percentage of completeness does the user think is achievable with the system?
- Is it possible to automatically aggregate the data of a given health service with the help of the system?
- How do you rate the process of processing National Health Information (NHI) data with the proposed system?

The collected data made it possible to carry out a qualitative analysis with the calculation of the frequencies and scores associated with these assessment criteria. Using the McNemar test, the aggregate data scores obtained for the 2 systems were compared.

**Ethics Approval**

After approval was obtained from the ethics board of each hospital, all the persons participating in the evaluation were

fully informed about the evaluation being conducted, and they explicitly agreed through a consent form. Further, they participated in the evaluation free from any coercion and were informed about their rights to be free to withdraw their participation at any time without negatively impacting their involvement in their respective hospital activities.

We provided an anonymous web-based form to be filled out for the evaluation and for further analysis. No personal or identification information was collected during this study. The participants gave their free and informed consent to the publication of the results of this study.

**Results**

**Feedback on the Implementation of the Pilot and Evaluation**

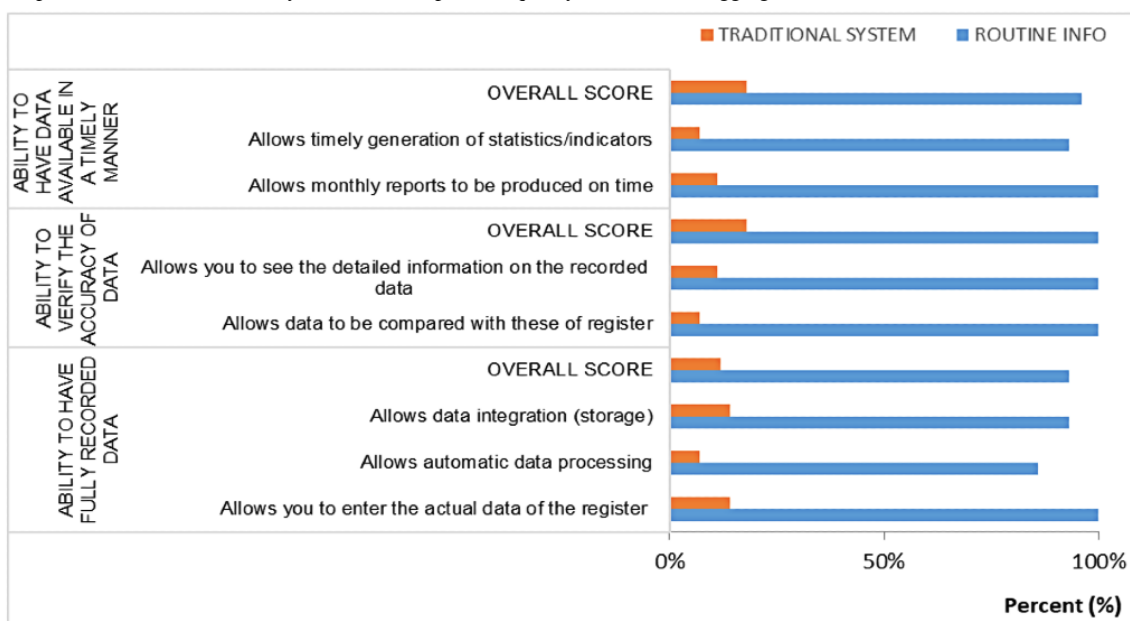
To test and evaluate the proposed system, a set of health professionals were first randomly selected from the participating health facilities in the 2 health regions of Gabon. Then, these professionals participated in various test sessions. Finally, we interviewed them to collect their opinions. In total, 14 health professionals (3 women and 11 men) were interviewed and participated in the various testing and evaluation sessions of the prototype. Overall, they had a mean age of 43 (SD 7) years and 12 (SD 1) years of experience in HIS data management. Specifically, there was no statistically significant difference between women and men in terms of age ( $P > .05$ ) or experience in the NHI system ( $P > .05$ ).

The 14 health professionals interviewed were distributed as follows: statistics specialist, 36% (5/14); data managers, 29% (4/14); activity coordinator, 14% (2/14); physician: 7% (1/14); epidemiologist: 7% (1/14); and computer scientist, 7% (1/14).

**Capabilities of the Tools to Produce Quality Data**

Figure 8 shows a short description of the results derived from the analysis of the questionnaires about the capacity of the tools to produce quality data.

**Figure 8.** Comparative results on the ability of the tools to produce quality data from the aggregate data score.



The ability to obtain all health data in a timely manner resulted in an overall response score of 96% (27/28) for all data quality criteria for the Routine Info tool compared with an overall response score of 18% (5/28) for all data quality criteria for the traditional system.

Overall, 2 criteria were used to assess the ability of the system to provide all data in a timely manner. These 2 criteria were the ability to produce monthly reports on time and the ability to generate statistics or indicators on time. These 2 criteria had scores of 93% (13/14) and 100% (14/14), respectively, for the Routine Info tool compared with 7% (1/14) and 28% (4/14), respectively, for the traditional system that met these criteria.

With regard to the ability to facilitate the verification of data accuracy, the Routine Info tool had an overall response score of 100% (28/28) for all data quality criteria compared with a score of 18% (5/28) for the traditional system.

The 2 criteria that were used to assess this dimension were the ability to compare the tool's data with the registry data and the ability to have the registered data detailed at the finest possible level of granularity. These 2 criteria had scores of 100% (14/14) each for the Routine Info tool and 28% (4/14) and 7% (1/14), respectively, for the traditional system.

Regarding the ability to have fully recorded data, the results showed that the Routine Info tool had an overall response score of 93% (39/42) for all data quality criteria, compared with a score of 12% (5/42) for the traditional system.

A total of 3 criteria were used to assess this dimension. Measured individually, these criteria had scores of 100%

(14/14), 86% (12/14), and 93% (13/14), respectively, for the Routine Info tool, compared with the scores of 14% (2/14), 7% (1/14), and 14% (2/14), respectively, for the traditional system.

The results presented in Table 3 show that according to the users surveyed, the quality of the data (completeness, accuracy, and timeliness) produced by the 2 systems was statistically different ( $P < .001$ ). Routine Info had a higher capacity to produce quality data than the traditional system.

The respondents felt that the use of Routine Info allowed them to have access to data in a shorter period, which allowed them to make decisions quickly. They estimated that with Routine Info, it is possible to obtain data within an average of 1 day, whereas it takes an average of 21 days to obtain data with the traditional system. This feedback shows that the respondents agree that Routine Info significantly ( $P < .001$ ) improves the time taken to provide data compared with the traditional system.

Following the test phase, 86% (12/14) of the respondents declared that Routine Info allows them to reach a completeness of more than 80%, contrary to the traditional system, where only 7% (1/14) of the respondents declared its capacity to allow the same completeness (Table 4).

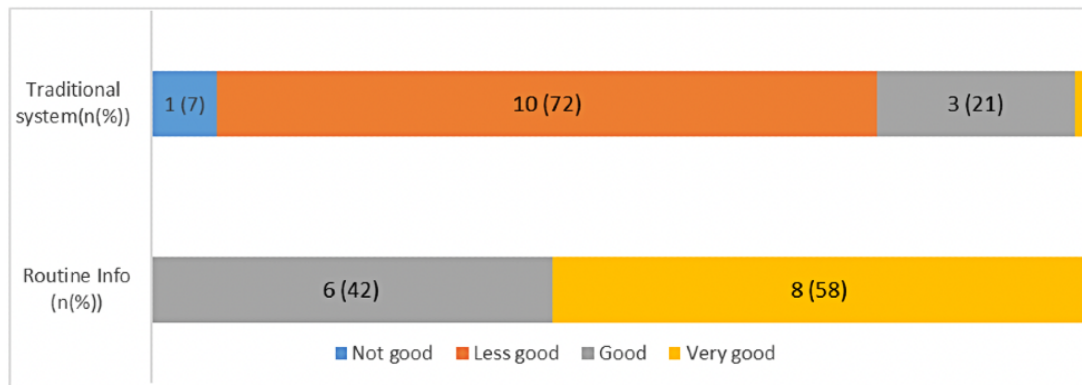
Overall, the users were satisfied with Routine Info, as shown in Figure 9. Indeed, when asked, "How do you rate the NHI data processing process with the proposed system?" 43% (6/14) and 57% (8/14) of the users thought that the process was good and very good, respectively, with Routine Info. However, with the former traditional system, only 21% (3/14) of the users thought that the process was good, and 72% (10/14) declared it to be less good.

**Table 3.** Comparison between the capacities of the tools to produce quality data.

Score criteria	Routine info	Traditional system	<i>P</i> value
Score in relation to the ability to have completely recorded data	93	12	<.001
Data readiness score	96	18	<.001
Score in relation to the verification of the accuracy of the data	100	18	<.001

**Table 4.** Comparison of the completeness of the data estimated for each tool by the 14 users surveyed.

Estimated completeness achievable by the tool (%)	Traditional system, n (%)	Routine info, n (%)
<30	1 (7)	0 (0)
30-50	11 (78)	0 (0)
50-80	1 (7)	2 (14)
≥80	1 (7)	12 (86)

**Figure 9.** Satisfaction of users with the 2 tools.

## Discussion

### Overview

The work reported in this paper illustrates a model and framework for improving the adequacy of information systems in low- and middle-income countries with limited digital infrastructure. A particular focus is placed on Gabon, where a pilot has been implemented and tested. The model proposed here constitutes the first initiative of this scale in the country. The inadequacy of the currently existing tools is often characterized by a low level of digitalization of the main data sources (operational systems). This leads to manual processing of health data. As a result, the quality is seriously affected to such a degree that the data are often incomplete, inconsistent, subject to long delays, and very often inaccurate. After testing and evaluating the proposed system in 2 regions representative of the 10 health regions in Gabon, feedback from the surveyed users showed that the proposed approach contributes to improving data quality and helps address the problems around data collection, data storage, and multidimensional data processing, processes which are often found to be not available in the health systems in developing countries [11]. Although various initiatives were conducted with international donors to improve the HISs and, in turn, the quality of data in these countries, these initiatives unfortunately lack sustainability after the donors withdrew. Indeed, the presence of a multitude of nonharmonized tools complicates the data collection process for local data managers. Thus, they have to collect all the data sent to them or available to them by manually filling in all the tools provided to them.

### Value of the Model Compared With Existing Models

The design and implementation of a system based on a relational database approach, which integrates all data production processes, is a real challenge, as it pertains to facilitating the collection and structuring of data available on paper registers. It helps to structure different data and make semantically coherent data recorded in the different tables of the system. Furthermore, it helps to meet another major challenge of improving data quality. This systematically excludes manual and subjective manipulation despite the absence of digitalized operational systems, as is the case in the context of an assertive digital transformation. Indeed, in a context of sufficient digital maturity, many platforms such as i2b2 [20] and Ehop [17] have

simply used ETL tools to extract and transform data to the platform model before loading [16].

As an ETL tool cannot be used because the data are not digitized at the source level, the problem of centralizing the data is therefore taken into account by an SQL query that carries out the selection of the relevant health information in the various tables and makes them persistent in an extraction table. The latter is a materialized view of the data on which several analyses can be carried out later on. As the common model of the Observational Medical Outcomes Partnership platform represents the common structure of the data in the Observational Medical Outcomes Partnership on which all possible processing of the data is performed, the materialized view (considered as a retrieval table) resulting from the SQL query of Routine Info is, therefore, the main table on which all processing around the data will be done. However, the platform has the distinct advantage of integrating a natural language processing program to automatically structure the free-text data of electronic patient records [18].

### Benefit in Data Storage and Processing From the Proposed Model

It is in fact from the retrieval materialized view of the data that all the relevant data for decision support are centralized, thus allowing users to dynamically generate various statistics. This model, which records the data from the department in which it is produced, now excludes manual processing to fill in the multiple tools that make the work of data managers repetitive and laborious. Furthermore, the verification of the accuracy of the data by comparing the observations of the reports and registers is therefore taken into account upstream. This makes it easier to assess the quality of data, for example, with the Performance of Routine Information System Management approach [3].

The implemented system makes it possible to disaggregate data from the country level down to the service level and assess whether the data contained therein really reflect those present in the production sources (eg, consultation registers). In a context of insufficient digital infrastructure at the level of the operating systems with, for example, nondigitalized HISs, as can be observed in several information systems [18], this model responds well to the question of automatic extraction and integration of data by avoiding manual processing (and preprocessing) of data. Indeed, because of the nonexistence of

digitized transactional databases, an initial processing consisting of counting the data by type of diagnosis and type of drug prescription is often carried out by staff members before filling in the aggregated totals in DWs. This approach, which our present framework avoids, did not allow for the most detailed analyses. This is because the data aggregated manually and entered into the DW are not related to their primary sources, which makes it very difficult to check the accuracy or completeness of the data, for example.

Therefore, this proposed model and framework has the advantage of acting as a transactional database for data integration into other platforms. This is specifically the case of HISs in many developing countries that use the DHIS2 platform [10]. Indeed, the DHIS2 platform is often populated at the district level from preprocessed, manually aggregated data, whereas for more detailed and better quality data [4], it is necessary to integrate the health structure level and, when existing, the health service level, which are both already taken into account in our model. In terms of the fragmentation of information systems, which multiplies the collecting tools for health structures, this study provides a solution in that a data manager will only have to query the database to fill in the various collecting tools.

## Limitations

A scale-up was carried out only in 2 pilot health regions, while reducing the size of the sample of interviewees; this could constitute a limitation in the exploitation of the results of this work. A countrywide or subregional implementation is envisaged to continue the testing and validation process of the proposed model.

Another limitation is the fact that only the general admission service with routine consultations and diagnostics is taken into account. It is worth implementing the framework in various units of hospitals, including those of intensive care.

## Conclusions

The proposed model and framework integrates all the processes of data collection, processing, and dissemination, thus providing complete, accurate, and near real-time data availability. Indeed, because data collection is mainly done at the level of the service in which these data are produced and the data are used at all levels for decision-making, our approach contributes significantly to ensuring the improvement of data quality for the management of health systems in the context of limited infrastructure. In addition, the approach also allows for multidimensional analysis and the provision of dashboards necessary for monitoring and evaluating health program indicators, both at the national and local levels.

## Acknowledgments

This work was financially supported by the e-Gabon National Health Information System project and the World Bank. The lead author, a PhD student in public health specializing in medical informatics, received funding from the project to support this work. However, the project played no role in the design of the study, data collection, analysis and interpretation of the data, or writing of the manuscript. The National Health Information System's e-Gabon project and the World Bank found no objections to the publication of the results of this study. The authors would like to thank the World Bank, the health information system e-Gabon project, the project team, the managers of the Libreville-Owendo and Ouest-Estuaire health regions, and their respective health staff.

## Authors' Contributions

APK and GD designed the research project, and APK wrote the first version of the manuscript. ROM contributed to modeling and computer development. GD and EBN validated the methodological approach and improved the language of the manuscript. JEN and EI participated in the scientific revision and linguistic improvement of the manuscript. All authors participated in the final revision of the manuscript and corrected and approved the manuscript for submission.

## Conflicts of Interest

None declared.

## References

1. Inmon WH. The data warehouse and data mining. *Commun ACM* 1996 Nov;39(11):49-50. [doi: [10.1145/240455.240470](https://doi.org/10.1145/240455.240470)]
2. Vassiliadis P, Simitsis A, Skiadopoulos S. Conceptual modeling for ETL processes. In: *Proceedings of the 5th ACM international workshop on Data Warehousing and OLAP*. 2002 Presented at: CIKM02: Eleventh ACM International Conference on Information and Knowledge Management; Nov 8, 2002; McLean Virginia USA. [doi: [10.1145/583890.583893](https://doi.org/10.1145/583890.583893)]
3. Glèlè Ahanhanzo Y, Ouendo E, Kpozèhouen A, Levêque A, Makoutodé M, Dramaix-Wilmet M. Data quality assessment in the routine health information system: an application of the lot quality assurance sampling in Benin. *Health Policy Plan* 2015 Sep 24;30(7):837-843. [doi: [10.1093/heapol/czu067](https://doi.org/10.1093/heapol/czu067)] [Medline: [25063699](https://pubmed.ncbi.nlm.nih.gov/25063699/)]
4. Karuri J, Waiganjo P, Orwa D, Many A. DHIS2: the tool to improve health data demand and use in Kenya. *J Health Inform Dev Countries* 2014 Mar 18;8(1).
5. Koumamba AP, Bisvigou UJ, Ngoungou EB, Diallo G. Health information systems in developing countries: case of African countries. *BMC* 2021;21(232). [doi: <https://doi.org/10.1186/s12911-021-01597-5>]

6. Msendema MB. Interfacing research and policy in informing data management for quality data in health information systems: case of DHIS2in Malawi. *J Health Informatics Africa* 2020;7(2). [doi: [10.12856/JHIA-2020-v7-i2-272](https://doi.org/10.12856/JHIA-2020-v7-i2-272)]
7. Asah FN, Kanjo C, Msendema MB, Addo H, Logo DD. The digitalization of routine data management at the point-of-care: the case of Ghana. *J Health Informatics Africa* 2020;7(2). [doi: [10.12856/JHIA-2020-v7-i2-285](https://doi.org/10.12856/JHIA-2020-v7-i2-285)]
8. Mettler T, Rohner P, Baacke L. Improving data quality of health information systems: a holistic design- oriented approach. In: *Proceedings of the 16th European Conference on Information Systems, ECIS 2008*. 2008 Presented at: 16th European Conference on Information Systems, ECIS 2008; 2008; Galway, Ireland URL: <https://aisel.aisnet.org/ecis2008/164>
9. Developing integrated national health information systems in malawi: challenges and south-south collaboration. In: *Exploring Success and Failure in Development Informatics: Innovation, Research and Practice*. Cape Town, South Africa: IDIA2011 Conference Proceedings; 2010.
10. Lippeveld T, Sauerborn R, Bodart C, World Health Organization. *Design and Implementation of Health Information Systems*. Geneva: World Health Organization; 2000.
11. Mutale W, Chintu N, Amoroso C, Awoonor-Williams K, Phillips J, Baynes C, et al. Improving health information systems for decision making across five sub-Saharan African countries: implementation strategies from the African Health Initiative. *BMC Health Serv Res* 2013 May 31;13(Suppl 2):S9. [doi: [10.1186/1472-6963-13-s2-s9](https://doi.org/10.1186/1472-6963-13-s2-s9)]
12. Koumamba AP, Lipenguet GM, Mbenga RO, Bisvigou UJ, Assoum-Mve FU, Effame YP, et al. State of play of Gabon's health information system. *Public Health* 2020;32(4):407-417. [doi: [10.3917/spub.204.0407](https://doi.org/10.3917/spub.204.0407)] [Medline: [33512108](https://pubmed.ncbi.nlm.nih.gov/33512108/)]
13. Makita-Ikouaya E, Mombo JB, Milleliri JM, Rudant JP. [Study of morbidity in Libreville (Gabon) in 2008]. *Med Sante Trop* 2013;23(3):324-327 [FREE Full text] [doi: [10.1684/mst.2013.0231](https://doi.org/10.1684/mst.2013.0231)] [Medline: [24126119](https://pubmed.ncbi.nlm.nih.gov/24126119/)]
14. Ministère de la santé. *Rapport de la cartographie des équipements usages et perceptions des TIC au Gabon des Gabon*. 2022:2016.
15. Gjoni O. Bizagi process management suite as an application of the model driven architecture approach for developing information systems. *Acad J Interdisciplinary Stud* 2014 Nov 1;3(6):475. [doi: [10.5901/ajis.2014.v3n6p475](https://doi.org/10.5901/ajis.2014.v3n6p475)]
16. Lescourret F, Pérochon L, Coulon J, Faye B, Landais E. Modelling an information system using the MERISE method for agricultural research: the example of a database for a study on performances in dairy cows. *Agricultural Syst* 1992 Jan;38(2):149-173. [doi: [10.1016/0308-521x\(92\)90038-p](https://doi.org/10.1016/0308-521x(92)90038-p)]
17. Bruley M. About the development of a decision-making information system. *Business Intelligence Information System*. 2010. URL: [https://www.decideo.fr/bruley/m/Propos-sur-le-developpement-d-un-systeme-d-information-decisionnel\\_a24.html](https://www.decideo.fr/bruley/m/Propos-sur-le-developpement-d-un-systeme-d-information-decisionnel_a24.html) [accessed 2022-09-27]
18. Delamarre D, Bouzille G, Dalleau K, Courtel D, Cuggia M. Semantic integration of medication data into the EHOP Clinical Data Warehouse. *Stud Health Technol Inform* 2015;210:702-706. [Medline: [25991243](https://pubmed.ncbi.nlm.nih.gov/25991243/)]
19. Dinh D, Tamine L. Vers un modèle d'indexation sémantique adapté aux dossiers médicaux de patients. *CiteSeerX*. URL: <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.184.192&rep=rep1&type=pdf> [accessed 2022-09-27]
20. Si Y, Weng C. An OMOP CDM-based relational database of clinical research eligibility criteria. *Stud Health Technol Inform* 2017;245:950-954 [FREE Full text] [Medline: [29295240](https://pubmed.ncbi.nlm.nih.gov/29295240/)]
21. Klann JG, Abend A, Raghavan VA, Mandl KD, Murphy SN. Data interchange using i2b2. *J Am Med Inform Assoc* 2016 Sep;23(5):909-915 [FREE Full text] [doi: [10.1093/jamia/ocv188](https://doi.org/10.1093/jamia/ocv188)] [Medline: [26911824](https://pubmed.ncbi.nlm.nih.gov/26911824/)]
22. Bagayoko C. *Implementation of a hospital information system in French-speaking Africa: cinz@n, study and validation of the model in Mali*. thesis.fr. 2010 Oct 4. URL: <http://www.theses.fr/2010AIX20680> [accessed 2022-09-27]
23. Zemri FA, Hamdadou D, Zeitouni K. *Towards a multicriteria and spatiotemporal decision support system for epidemiological surveillance. GAST-Spatial and Temporal Data Management and Analysis*. 2015. URL: [https://scholar.google.com/citations?view\\_op=view\\_citation&hl=fr&user=jR7-9BEAAA&citation\\_for\\_view=jR7-9BEAAA:d1gkVwhDp10C](https://scholar.google.com/citations?view_op=view_citation&hl=fr&user=jR7-9BEAAA&citation_for_view=jR7-9BEAAA:d1gkVwhDp10C) [accessed 2022-09-27]

## Abbreviations

- DHIS2:** District Health Information Software 2
- DMIS:** decision-making information system
- DW:** data warehouse
- ETL:** extraction, transformation, and loading
- HIS:** health information system
- ICD:** International Classification of Diseases
- NHI:** National Health Information
- ROLAP:** relational online analytical processing
- SQL:** Structured Query Language
- SSA:** Sub-Saharan African

*Edited by A Mavragani; submitted 24.11.21; peer-reviewed by J Olsen, J Bowen; comments to author 04.03.22; revised version received 14.05.22; accepted 24.08.22; published 07.10.22.*

*Please cite as:*

*Koumamba AP, Ngoungou EB, Engohang-Ndong J, Ibinga E, Ondzigue Mbenga R, Diallo G  
From Real-world Individuals' Data to National Health Indicators: Multiphase Pilot Study in Gabon  
JMIR Form Res 2022;6(10):e35176*

*URL: <https://formative.jmir.org/2022/10/e35176>*

*doi: [10.2196/35176](https://doi.org/10.2196/35176)*

*PMID: [36206045](https://pubmed.ncbi.nlm.nih.gov/36206045/)*

©Aimé Patrice Koumamba, Edgard Brice Ngoungou, Jean Engohang-Ndong, Euloge Ibinga, Raymond Ondzigue Mbenga, Gayo Diallo. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 07.10.2022. 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 Peer-Supported Self-Management Intervention Codesigned by People With Long COVID: Mixed Methods Proof-of-Concept Study

Hayley Wright<sup>1\*</sup>, BSc, MSc, PhD; Andrew Turner<sup>1\*</sup>, BSc, PhD; Stuart Ennis<sup>2,3,4\*</sup>, MSc; Carol Percy<sup>5</sup>, PhD; Garry Loftus<sup>6\*</sup>; Wendy Clyne<sup>7</sup>, PhD; Gabriela Matouskova<sup>6</sup>, MBA; Faith Martin<sup>1</sup>, PhD, DCLinPsy

<sup>1</sup>Centre for Intelligent Healthcare, Coventry University, Coventry, United Kingdom

<sup>2</sup>Department of Cardiac Rehabilitation, University Hospitals Coventry & Warwickshire NHS Trust, Coventry, United Kingdom

<sup>3</sup>Warwick Clinical Trials Unit, Warwick Medical School, University of Warwick, Coventry, United Kingdom

<sup>4</sup>Atrium Health, Centre for Exercise and Health, Coventry, United Kingdom

<sup>5</sup>School of Psychological, Social and Behavioural Sciences, Faculty of Health and Life Sciences, Coventry University, Coventry, United Kingdom

<sup>6</sup>Hope For The Community Community Interest Company, Enterprise Hub, Coventry, United Kingdom

<sup>7</sup>Peninsula Medical School, University of Plymouth, Plymouth, United Kingdom

\*these authors contributed equally

**Corresponding Author:**

Hayley Wright, BSc, MSc, PhD  
Centre for Intelligent Healthcare  
Coventry University  
Richard Crossman Building  
Jordan Well  
Coventry, CV1 5RW  
United Kingdom  
Phone: 44 02477659509  
Email: [ab7764@coventry.ac.uk](mailto:ab7764@coventry.ac.uk)

## Abstract

**Background:** There are around 1.3 million people in the United Kingdom with the devastating psychological, physical, and cognitive consequences of long COVID (LC). UK guidelines recommend that LC symptoms be managed pragmatically with holistic support for patients' biopsychosocial needs, including psychological, emotional, and physical health. Self-management strategies, such as pacing, prioritization, and goal setting, are vital for the self-management of many LC symptoms.

**Objective:** This paper describes the codevelopment and initial testing of a digital intervention combining peer support with positive psychology approaches for self-managing the physical, emotional, psychological, and cognitive challenges associated with LC. The objectives of this study were to (1) codesign an intervention with and for people with LC; (2) test the intervention and study methods; (3) measure changes in participant well-being, self-efficacy, fatigue, and loneliness; and (4) understand the types of self-management goals and strategies used by people with LC.

**Methods:** The study used a pre-post, mixed methods, pragmatic, uncontrolled design. Digital intervention content was codeveloped with a lived-experience group to meet the needs uncovered during the intervention development and logic mapping phase. The resulting 8-week digital intervention, Hope Programme for Long COVID, was attended by 47 participants, who completed pre- and postprogram measures of well-being, self-efficacy, fatigue, and loneliness. Goal-setting data were extracted from the digital platform at the end of the intervention.

**Results:** The recruitment rate (n=47, 83.9%) and follow-up rate (n=28, 59.6%) were encouraging. Positive mental well-being (mean difference 6.5,  $P<.001$ ) and self-efficacy (mean difference 1.1,  $P=.009$ ) improved from baseline to postcourse. All goals set by participants mapped onto the 5 goal-oriented domains in the taxonomy of everyday self-management strategies (TEDSS). The most frequent type of goals was related to activity strategies, followed by health behavior and internal strategies.

**Conclusions:** The bespoke self-management intervention, Hope Programme for Long COVID, was well attended, and follow-up was encouraging. The sample characteristics largely mirrored those of the wider UK population with LC. Although not powered

to detect statistically significant changes, the preliminary data show improvements in self-efficacy and positive mental well-being. Our next trial (ISRCTN: 11868601) will use a nonrandomized waitlist control design to further examine intervention efficacy.

(*JMIR Form Res* 2022;6(10):e41410) doi:[10.2196/41410](https://doi.org/10.2196/41410)

## KEYWORDS

long COVID; self-management; peer support; digital intervention; goal setting; psychological; physical; cognitive; intervention; United Kingdom; UK; efficacy; COVID-19

## Introduction

“Long COVID” (LC) is the term commonly used when symptoms continue or develop after acute COVID - 19. It includes ongoing symptomatic COVID - 19 (from 4 to 12 weeks) and post-COVID - 19 syndrome (12 weeks or more), where symptoms cannot be explained by an alternative diagnosis [1]. There are around 1.3 million people (2% of the population) with LC in the United Kingdom [1]. Of those, 40% report symptoms lasting at least 1 year and 64% report symptoms that adversely affect their daily activities [1]. The most common symptoms are fatigue (51%), loss of smell (37%), breathlessness (36%), and difficulty concentrating (28%)—also known as “brain fog” [1]. Mental well-being can also be impacted through the distress of living with a long-term health condition [2].

Symptoms vary within and between individuals and in severity. Patients often feel dismissed by medical professionals and seek validation elsewhere, such as peer support groups [3,4]. People with LC can also experience social stigma and invalidation in the wider community, further adding to the burden of coping [5-7]. LC has devastating psychological, physical, and cognitive consequences that disrupt lives and livelihoods. There are currently 90 specialist post-COVID-19 clinics across the United Kingdom, but there is wide local variation in referral rates, waiting times, and access across demographic groups [8]. Many services were suspended due to staff shortages following the rise in cases of the Omicron variant [9].

Since the first wave of the pandemic, people with LC have sought virtual peer support for managing their symptoms, driven by the lack of formal support in the early days of the pandemic [4]. In response to the pandemic generally, there has been a rapid and essential growth in the provision of digital health care for long-term conditions to allow remote care [10,11], and patients are more motivated, accepting, and familiar with digital technologies for health care and social connection [12]. The main digital resource for people with LC in the United Kingdom is “Your COVID Recovery” [13], a comprehensive National Health Service (NHS) website providing general symptom management advice and limited professional support. However, this broad resource lacks the social connection that patients desire [4], and evaluations of patient experiences of the intervention have yet to be published.

UK COVID-19 guidelines recommend that LC symptoms be managed pragmatically with holistic support for patients’ biopsychosocial needs, including psychological, emotional, and physical health [14]. Activities should be *paced* and only performed if they do not worsen symptoms, such as breathlessness or fatigue [15,16]. Self-management strategies,

such as pacing, prioritization, and goal setting, are vital for the self-management of many LC symptoms, such as fatigue [15]. However, there is still no NHS-endorsed, evidence-based, peer-supported self-management program for the multifaceted issues faced by people with LC.

NHS England recommends that patients with LC symptoms be treated in accordance with the national post-COVID-19 syndrome pathway, which encourages self-care through engaging patients in techniques such as specific, measurable, achievable, relevant, and timely (SMART) *goal setting* and *peer support* [17]. Goal setting is a core component of self-management of long-term conditions and can support patients to change behaviors. Setting goals empowers patients to maintain or improve management of their own physical, psychological, and social health. SMART goal setting helps to break down a desired achievement (ie, goal) into manageable chunks (ie, action plan), and SMARTER goal setting incorporates “enjoyable” and “reward” elements [18]. However, peer support and goal setting have not yet been evaluated in LC self-management programs. Improvement is nonlinear, and recovery to pre-COVID-19 levels of health and fitness is not guaranteed, so evaluation of goal setting is imperative.

This paper describes the codevelopment and testing of a digital intervention combining peer support with evidence-based behavior change techniques for self-managing the physical, emotional, psychological, and cognitive challenges associated with LC. The Hope Programme for Long COVID provides a proactive, timely solution in the most urgent public health crisis in a generation. We repurposed an existing digital self-management intervention—the digital Hope Programme—which has shown postcourse improvements in depression, anxiety, and mental well-being in multiple participant groups, such as cancer survivors [19,20], people with multiple sclerosis [21], and parents of children with autism [22]. The Hope Programme for Long COVID combines positive psychology and cognitive behavioral approaches, providing essential psychological support for people to live with the physical symptoms of LC and is delivered by trained facilitators.

This paper describes the codesign process and pre- and postprogram changes in mental well-being, self-efficacy, fatigue, and loneliness. In this study, we aimed to codesign and test a bespoke self-management intervention with and for people with LC. Specifically, the objectives of the study were to:

- Codesign bespoke intervention content specifically for LC.
- Test the intervention and study methods to inform future study design (calculate recruitment and follow-up rates and summarize demographics and symptoms of the sample).



- Measure changes in key indicators of well-being, self-efficacy, fatigue, and loneliness to signal efficacy of the intervention (calculate the difference in pre- and postprogram outcomes).
- Examine goal-setting data to better understand the types of self-management goals and strategies used by participants (perform a thematic analysis of the types of goals set by participants).

## Methods

### Study Design

The wider project was led by Hope For The Community (H4C) Community Interest Company, with the aim to empower people across Coventry, Warwickshire, and Rugby (CWR) to self-manage their health and well-being and to develop social connections and peer support opportunities. The study reported here adopted a pre-post, mixed methods, pragmatic, uncontrolled design to codesign and test a digital self-management intervention for people with LC.

### Participants

Adults with self-reported LC were eligible to enroll into the Hope Programme for Long COVID. Participants were informed that they must be over the age of 18 years and have access to an internet-enabled device to access the intervention content. No further exclusion criteria were applied. The intervention delivery partner (H4C) recruited people to enroll into the intervention using a range of social media, posters, flyers, and mailshots in the CWR area. Subsequently, those who had enrolled were contacted by email and asked whether they wanted to also participate in the research. Nonparticipation in the research had no impact on their receipt of the intervention. The email invitation to the study contained the necessary participant information sheet, online consent forms, and a link to complete baseline research questionnaires (via Qualtrics survey software). The intervention was hosted on the H4C digital platform.

### Intervention

The digital Hope Programme for Long COVID was cocreated to enable a rapid response to the ongoing public health crisis in the wake of the COVID-19 pandemic. The Hope Programme for Long COVID shares the same underlying theoretical framework as the digital Hope Programme, which has been described in detail elsewhere [19,20]. In summary, the Hope

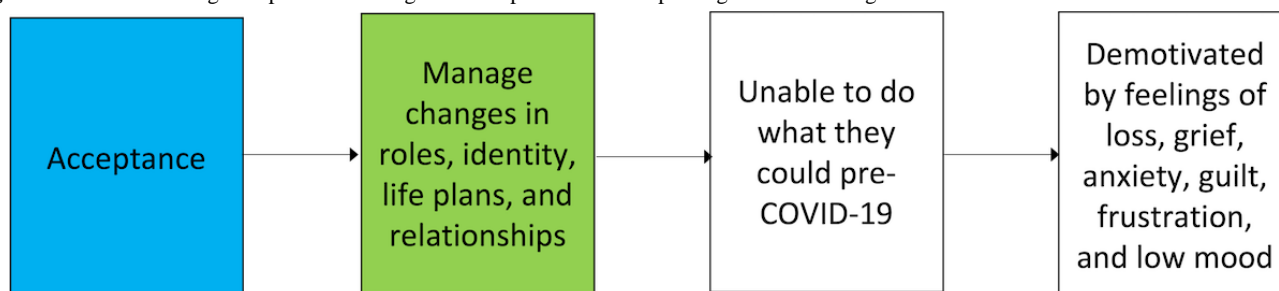
Programme for Long COVID focuses on strengths rather than deficits and uses group curative factors, including instilling hope, universality (realizing you are not alone), and altruism [23]. By sharing successful coping strategies, the peer support of the group achieves something greater than the sum of its parts. Through positive psychology, the Hope Programme for Long COVID cultivates an upward spiral of positive emotions [24] to improve well-being and coping. Hope and gratitude—2 important concepts in positive psychology—are core themes embedded within the Hope Programme for Long COVID. They are addressed through the experience of group curative factors [23] and through specific activities, such as goal setting and feedback, identification of strengths, and gratitude diaries [25-27]. The Hope Programme for Long COVID content was codesigned by people with LC and health care professionals who support them.

### Codesign Methods

We conducted 3 workshops: 2 workshops with a total of 8 people with LC and 1 workshop with 6 professionals involved in LC health care. In addition, 1-to-1 sessions were completed with 2 people with LC who could not join the group sessions. A total of 16 participants took part in the codesign activities.

A conversational and informal approach was used to facilitate discussion and open expression of views. Workshops were conducted online via Zoom. Participants were asked to reflect upon what “antecedents” caused the problem statement “People living with long COVID face many challenges in doing self-management” and then were asked to talk about what causes the antecedents, working back each time something was mentioned by asking “What causes this?” This process is based on elements of the antecedent target measurement process described by Renger and Hurley [28]. Responses were captured by the researcher in a logic map (see Figure 1, for example), noting each chain of antecedents that lead to the problem. Reading from right (white boxes) to left (green box), the map shows that antecedents (root causes) of challenges in engaging in self-management stem from negative feelings caused by a shift in abilities. The green box shows the need for support to manage change, and the blue box represents the target behavior of acceptance, which is addressed by the intervention. Participants were invited to add further antecedents or details throughout the discussion and on checking the logic map with the researcher.

**Figure 1.** Extract from logic map created during co-development of the Hope Programme for Long COVID.



A rapid literature search of the research databases PsycInfo, Cumulative Index to Nursing and Allied Health Literature (CINAHL), and Medline identified literature relating to support

or self-management needs for people with LC. LC is a new condition, and there is a relative paucity of research, so we also searched for literature on chronic fatigue syndrome (CFS) and

myalgic encephalomyelitis (ME), which share many of the key symptoms of LC (eg, fatigue, breathlessness, brain fog). Titles and abstracts of papers were assessed for relevance, prioritizing reports of recent systematic reviews of needs and interventions in LC. Research papers were then examined to ensure the needs identified by our stakeholder groups were supported by the research evidence. Reports of existing interventions for conditions with similar symptoms helped identify components that may be effective in supporting people with LC. This information was used to ensure an underlying evidence base for the foci and components of the Hope Programme for Long COVID.

The digital Hope Programme content [19-21,29] was repurposed to reflect the needs and individual lived experience of people with LC. New content was developed to meet the needs uncovered during the intervention development phase and logic mapping. A cognitive psychologist (author HW) led the codesign of the module addressing the management and understanding of brain fog, and a clinical psychologist (author FM) codeveloped new material on acceptance, self-compassion, and managing fatigue. Evidence-based breathing and physical exercises were provided by an exercise physiologist (author SE). All the intervention content was developed specifically to address the needs of people with LC, with goal setting embedded within each module. All modules were reviewed by the multidisciplinary team and patients with lived experience of LC, prior to testing the full program.

Key features of the Hope Programme for Long COVID, as identified by stakeholders and supported by the literature, include:

- Peer support has been integral to patients' recovery and understanding of COVID-19 throughout the pandemic [4], helping patients feel less lonely and isolated. Our workshop participants asked for this essential source of support to be included in the Hope Programme for Long COVID. Participants can interact with peers in an optional weekly Hope Café via Zoom or through online discussion forums and email. They are encouraged to support each other's goals and gratitude posts by posting likes and comments, where feedback strengthens group cohesiveness and a sense of belonging [23]. The Hope Programme for Long COVID is asynchronous, so it can be accessed as and when required, and text content is presented in concise sections measured against readability indices, such as the Gunning Fog Index

[30]. These design features support participants in pacing their learning and progress through the course at their own speed.

- Facilitated delivery: Two exercise specialists from Atrium Health were trained in health coaching and motivational interviewing to become Hope Programme facilitators. They were available to answer participants' questions and stimulate discussions between participants throughout the Hope Programme for Long COVID.
- Evidence-based guidelines relating to recovery and rehabilitation for COVID-19 are still emerging and can be inconsistent at best [31]. Evidence-based information about how to self-manage symptoms of LC and signposting to further information are provided by the NHS Your COVID Recovery web-based resource for UK patients [13]. However, during codesign, our stakeholders reported that patients experiencing fatigue may be overwhelmed by the amount of information to read. To mitigate these limitations, the Hope Programme for Long COVID provides bite-sized, evidence-based information, embedded in interactive videos, diagrams, and discussion forums, which is consolidated in the processes of social networking and peer support.
- Across our workshops, stakeholders were unanimous in their support for pacing [32] for people with LC. Many had experienced the "boom-and-bust cycle," where overexertion in times of high energy (boom) leads to periods of extreme fatigue (bust). Pacing materials and activities were codesigned with a clinical psychologist (FM) with experience of supporting patients in an NHS LC clinic. Participants described how fatigue can be caused by physical, mental, or emotional exertion. In collaboration with exercise physiologists (SE and team), we developed bespoke, expert-delivered physical activity content for the Hope Programme for Long COVID, with a clear focus on pacing and participants staying within their "energy envelope." Prioritization and goal setting [29,33] are central to self-management approaches and are core elements of pacing. These are all vital techniques for fatigue management in LC [15].

Access to the Hope Programme for Long COVID is unrestricted, sits outside of the NHS and so does not require participants to have a clinical diagnosis or referral, and is free of charge at the point of access to users. Weekly program content and activities are summarized in [Table 1](#).

**Table 1.** Weekly topics and activities.

Session	Examples of content, in addition to weekly goal setting	Examples of exercises and activities
Week 1: Introduction: Instilling Hope	<ul style="list-style-type: none"> <li>Welcome and introductions</li> <li>The benefits of positive emotions</li> <li>The power of gratitude</li> <li>Self-compassion</li> <li>Personalized goal setting</li> <li>Video: How to Set Achievable Goals</li> <li>Dates for live sessions</li> </ul>	Self-management tools: <ul style="list-style-type: none"> <li>Interactive gratitude diary</li> <li>SMARTER<sup>a</sup> goal setting</li> <li>Rectangle breathing exercise for LC<sup>b</sup></li> <li>Self-test: How are you feeling?</li> </ul>
Week 2: Long COVID Symptoms	<ul style="list-style-type: none"> <li>Fatigue management, including the boom-and-bust cycle, prioritizing, planning, and pacing</li> <li>What is brain fog, and what can we do to help ourselves?</li> <li>Forum topics: sharing experiences of managing fatigue and brain fog</li> <li>Further resources and links to LC information and support</li> </ul>	Self-management tools: <ul style="list-style-type: none"> <li>Interactive gratitude diary</li> <li>SMARTER goal setting and goal feedback</li> <li>Activity and fatigue diary</li> <li>Pacing planner</li> </ul>
Week 3: Managing Stress	<ul style="list-style-type: none"> <li>Coping with unhelpful thinking patterns</li> <li>Understanding and managing stress</li> <li>Self-compassion and acceptance</li> <li>Mindfulness for stress management and meditation</li> <li>Video: How to Be Kind to Yourself</li> <li>Further resources and links (eg, videos, podcasts, and websites) to self-compassion, mindfulness, and stress management</li> </ul>	Self-management tools: <ul style="list-style-type: none"> <li>Interactive gratitude diary</li> <li>SMARTER goal setting and goal feedback</li> <li>Guided relaxation and meditation</li> <li>Self-care checklist (worksheet)</li> </ul>
Week 4: Communication	<ul style="list-style-type: none"> <li>Communication skills and tips for talking with health professionals, your employer, and your family</li> <li>Preparing for difficult conversations</li> <li>Asking for and accepting help</li> <li>Compassion for worries</li> <li>Further resources and links (eg, videos, podcasts, and websites) to LC support groups</li> </ul>	Self-management tools: <ul style="list-style-type: none"> <li>Interactive gratitude diary</li> <li>SMARTER goal setting and goal feedback</li> <li>“Ask the Expert” live session</li> <li>Self-test: How are you feeling?</li> </ul>
Week 5: Sleep and Mindfulness	<ul style="list-style-type: none"> <li>How does COVID-19 affect our sleep?</li> <li>Tips for sleeping better</li> <li>Tips to aid relaxation</li> <li>Introduction to mindfulness, meditation, and relaxation</li> </ul>	Self-management tools: <ul style="list-style-type: none"> <li>Interactive gratitude diary</li> <li>SMARTER goal setting and goal feedback</li> <li>Raisin meditation</li> <li>Breathing exercises</li> <li>Live mindfulness meditation session via Zoom</li> </ul>
Week 6: Move Better, Feel Better	<ul style="list-style-type: none"> <li>Keeping active with LC</li> <li>Postexertional malaise</li> <li>Tips for getting active</li> <li>Eating well for physical and mental health</li> <li>Managing changes to taste and smell after COVID-19</li> </ul>	Self-management tools: <ul style="list-style-type: none"> <li>Interactive gratitude diary</li> <li>SMARTER goal setting and goal feedback</li> <li>Quiz: What contributes to happiness?</li> <li>Activity: gentle stretches and exercises</li> </ul>
Week 7: Happiness and Strengths	<ul style="list-style-type: none"> <li>LC and effects on mood</li> <li>Happiness and hope</li> <li>Identifying your character strengths</li> <li>Understanding how using your strengths can lead to a more fulfilling life</li> <li>Video: The Science of Character Strengths</li> <li>Managing setbacks</li> <li>Tips for authentic happiness, managing setbacks, and keeping going</li> </ul>	Self-management tools: <ul style="list-style-type: none"> <li>Interactive gratitude diary</li> <li>SMARTER goal setting and goal feedback</li> <li>Video: guided meditation activity</li> </ul>
Week 8: Moving Forward With Hope	<ul style="list-style-type: none"> <li>Hopes and dreams</li> <li>Doing something for yourself</li> <li>Planning pleasant activities</li> <li>Keeping in touch with peers</li> <li>Review of the program</li> <li>Feedback</li> <li>Staying hopeful</li> </ul>	Self-management tools: <ul style="list-style-type: none"> <li>Interactive gratitude diary</li> <li>SMARTER goal setting and goal feedback</li> <li>Self-test: How are you feeling?</li> </ul>

<sup>a</sup>SMARTER: specific, measurable, achievable, relevant, time-based, enjoyable, rewarded.

<sup>b</sup>LC: long COVID.

## Outcome Measures

### Sociodemographics

Participants were asked their name, email address, gender, date of birth, postcode, occupation, highest level of education attained, and details of formal treatment received for COVID-19-related symptoms.

### Symptom Logging

The COVID-19 Yorkshire Rehabilitation Screening Tool (C19-YRS [34]) screens for the most common complications after COVID-19 (eg, breathlessness, fatigue, cognitive problems) and rates the severity or significance of each symptom on a scale of 0-10 (0=no impact to 10=significant impact).

The logic map from the cocreation activities set out what issues were to be addressed by the intervention, and this study measured the impact of the intervention on these issues. Validated, standardized questionnaires were used to measure changes in key factors commonly affected by LC. The intervention components were designed to:

- Increase positive mental well-being
- Increase confidence to self-manage (self-efficacy)
- Reduce fatigue
- Reduce feelings of loneliness

### Positive Mental Well-being

The Warwick-Edinburgh Mental Wellbeing Scale (WEMWBS [35]) has previously been used in research on people with long-term conditions [19,20]. The items (questions) are worded positively and refer to function and feelings. For each item, there is a choice of 5 responses ranging from 1 for “none of the time” to 5 for “all the time.” The responses are summed to obtain a total score ranging from 14 to 70, where a higher score indicates a higher level of mental well-being. A score of 40 or less is indicative of probable depression, and a change of 3 or more in the total score represents a minimally important level of change. Two versions of WEMWBS were used in this study: The full measure was used at baseline and end of the intervention, and the shorter 7-item version of the WEMWBS (score range 7-35) was embedded within the Hope Programme for Long COVID in sessions 1, 4, and 7, allowing participants to monitor their own mental well-being throughout the course.

### Self-efficacy

The 6-item Self-Efficacy for Managing Chronic Disease scale (SEMCD6 [36]) provides a robust measure of participants' confidence to self-manage their symptoms of LC. The questions relate to participants' confidence that they can keep issues relating to their condition from interfering with daily life. Responses to each of the 6 questions are on a 10-point scale, ranging from 1 for “not at all confident” to 10 for “totally confident.” The total ranges from 6 to 60, and the self-efficacy score is calculated as the mean score across all items (eg, total/6), with a range of 1-10. There are no recommended cut-off scores for this scale, but data from a population of 489

individuals with chronic disease show an average self-efficacy score of 5.2 [37].

### Fatigue

The Fatigue Severity Scale (FSS [38]) measures fatigue and its effect on activities and lifestyle in patients with a variety of conditions. It comprises 9 questions answered on a 7-point scale, ranging from 1 for “strongly disagree” to 7 for “strongly agree.” The questions relate to how fatigue interferes with certain activities, and severity is indicated by the 7-point scale. Summed answers give a total score ranging from 9 to 63, and mean scores range from 1 to 7. Higher scores indicate higher fatigue severity. Research suggests that a mean score of 4 or more is indicative of fatigue [39].

### Loneliness

The University of California, Los Angeles (UCLA) Loneliness Scale, Version 3 [40], is a 3-item scale to assess loneliness and is recommended by the Office of National Statistics [41]. The 3 questions are as follows:

- How often do you feel that you lack companionship?
- How often do you feel left out?
- How often do you feel isolated from others?

The response scale is scored as 1 for “hardly ever,” 2 for “some of the time,” and 3 for “often.” The total score is the sum of all items (ranging from 3 to 9), where higher scores indicate greater feelings of loneliness.

### Sample Size

A target sample size of N=40 was deemed adequate for this study, informed by evidence that sample sizes of 24-50 are sufficient to estimate the key parameters of an efficacy trial [42]. Attendance on the Hope Programme for Long COVID was capped at 60 people to allow manageable facilitation and communication between participants.

### Analysis

Descriptive statistics were used to summarize recruitment and retention rates. All statistical data analyses were conducted using IBM SPSS Statistics version 26. The sample was not fully powered to detect significance in the outcome measures, but we presented changes in scores and preliminary *P* values to demonstrate the potential of the intervention (ie, to signal efficacy) and provide data on which to base power calculations for a larger study of efficacy. The level of statistical significance was set at  $P < .05$ . Outcomes were analyzed using paired-sample *t* tests. We reported the means with SDs and effect size estimates.

Goal-setting data were analyzed using deductive content analysis, with coding derived from the taxonomy of everyday self-management strategies (TEDSS) [43]. TEDSS focusses on self-management strategies for persistent symptoms, such as those that have become synonymous with LC (eg, fatigue), that lead to reduced physical and cognitive function. The taxonomy categorizes self-management strategies for conditions that are typically difficult to mitigate and have a significant impact on

everyday life [43], which again are typical of the fluctuating symptoms of LC. The types of goals set by participants were coded according to the 5 goal-oriented domains:

- Activity strategies
- Internal strategies
- Social interaction strategies
- Health behavior strategies
- Disease-controlling strategies

All goals were coded by 1 researcher, and a random selection of 25% of the data was second-coded by another researcher. An interrater reliability score (Cohen  $\kappa$ ) of  $<.70$  was used as a cut-off to indicate all data should be second-coded.

### Ethical Considerations

All participants received a Participant Information Sheet describing the purpose and anticipated outcomes of the research, including participants' right to decline to take part or withdraw from the research at any time, and data processing in accordance with the Data Protection Act of 2018. Participants were informed that their data would be treated confidentially and would be

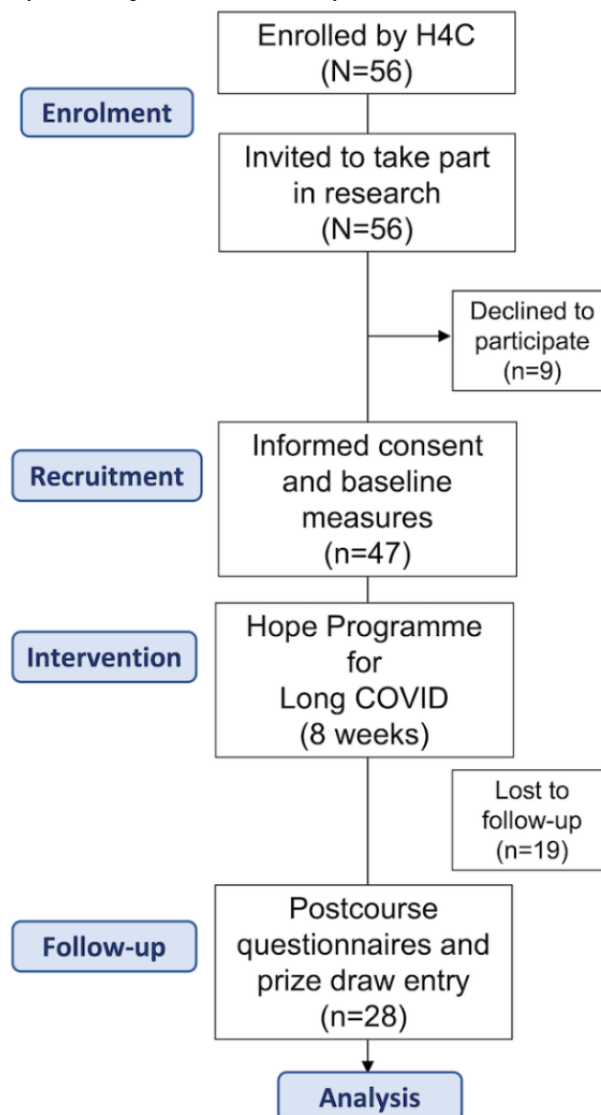
anonymized by the use of a randomly generated participant identification number in Qualtrics. The digital consent form required participants to confirm that they had read and understood the Participant Information Sheet and agreed to their anonymized data being used in research publications and dissemination. All statements had to be agreed upon before participants could progress into the study. Upon completion of pre- and postprogram questionnaires, participants were entered into a prize draw for a chance to win a £50 (US \$52.87) voucher. The research was reviewed and approved by the Black Country Research Ethics Committee (IRAS ID: 283172) and the Coventry University Research Ethics Committee (P106036).

## Results

### Recruitment and Follow-up

A total of 56 people enrolled into the Hope Programme for Long COVID, of which 47 (recruitment rate 83.9%) consented to take part in the study and provided baseline data. A total of 28 participants completed the postcourse questionnaires, giving a follow-up rate of 59.6% (see Figure 2 for participant flow).

**Figure 2.** Participant flow through the study. H4C: Hope For The Community.



## Sociodemographics and Symptoms

Sociodemographic data describing the characteristics of the sample and the severity of the impact of LC symptoms are shown in [Tables 2](#) and [3](#), respectively. Baseline data suggested that the characteristics of the sample were similar to the general UK population with LC [1]. The mean age of this group was 48.8 years (SD 11.0), 39 (83%) were female, 42 (91.3%) described themselves as White British, 36 (78.2%) were employed, and 36 (76.6%) had reduced their working hours due to their COVID-19-related illness.

All participants (n=47, 100%) reported fatigue at baseline, and the majority reported cognitive issues, including problems with concentration (n=46, 98%), memory (n=44, 94%), and communication (n=34, 72%), such as word finding and following a conversation. On average, COVID-19-related symptoms had persisted for longer than a year across the sample. The severity or impact of symptoms was rated highest for fatigue, usual daily activities (eg, shopping, household tasks), and cognitive communication (all rated 7/10, on average), followed by breathlessness on walking up stairs and pain (both rated 6/10, on average). A summary of severity scores is shown in [Table 3](#).

**Table 2.** Sample characteristics at baseline.

Characteristics	Baseline (N=47)
Age <sup>a</sup> (years), mean (SD)	48.8 (11.0)
Female, n (%)	39 (83.0)
<b>Ethnicity<sup>a</sup>, n (%)</b>	
White British	42 (91.3)
Asian British	3 (6.5)
Prefer not to say	1 (2.2)
Post-16 education, n (%)	37 (78.7)
<b>Employment status<sup>a</sup>, n (%)</b>	
Full time, paid	22 (47.8)
Part time, paid	14 (30.4)
Unemployed/retired/stay-at-home parent	10 (21.7)
Cut work hours due to LC <sup>b</sup> , n (%)	36 (76.6)
At least 1 long-term condition prior to LC <sup>c</sup> , n (%)	30 (73.2)
LC symptom duration in days <sup>a</sup> , mean (SD)	377.3 (171.8)
Positive COVID-19 test or diagnosis, n (%)	34 (72.3)
Hospitalized with acute COVID-19, n (%)	2 (4.3)
Accessed NHS <sup>d</sup> LC clinic, n (%)	19 (40.4)

<sup>a</sup>N=46 due to missing data.

<sup>b</sup>LC: long COVID.

<sup>c</sup>N=41 due to missing data.

<sup>d</sup>NHS: National Health Service.

**Table 3.** LC<sup>a</sup> symptoms and mean severity scores across the sample at baseline.

LC symptoms/impact (N=47)	Severity score (range 0-10), mean (SD)
Fatigue	7.0 (2.1)
Cognitive/communication	6.6 (2.4)
Impacting daily activities	6.9 (2.4)
<b>Breathlessness</b>	
Walking up stairs	5.7 (2.8)
On dressing	3.3 (2.7)
At rest	2.4 (2.2)
Personal care	3.0 (3.1)
Pain/discomfort	5.5 (2.4)
Throat sensitivity	5.2 (2.0)
Voice changes	5.1 (1.7)
Swallowing difficulties	5.0 (1.7)
Reduced appetite	3.5 (2.6)
Reduced mobility	5.0 (2.7)
General health	5.1 (1.9)
Anxiety	5.1 (2.8)
Depression	4.1 (3.4)

<sup>a</sup>LC: long COVID.

## Health and Well-being

Outcome data from participants completing both baseline and postcourse questionnaires are summarized in Table 4. Paired *t* tests measured statistically significant changes in pre-post well-being scores to signal efficacy of the intervention. Positive mental well-being (WEMWBS) scores increased by 6.5 points, on average, from baseline to postcourse ( $P<.001$ ). In addition, the Short-Form Warwick-Edinburgh Mental Wellbeing Scale (SWEMWBS) was embedded within the Hope Programme for Long COVID in weeks 1, 4, and 7 (see Table 5).

Repeated-measures ANOVA showed a significant difference in SWEMWBS scores across the 3 time points ( $F_{1,4,22.1}=20.2$ ,  $P<.001$ ). Bonferroni-corrected comparisons showed a significant increase in scores between week 1 and week 4 ( $P<.001$ ) and between week 1 and week 7 ( $P=.001$ ) but no difference in scores between week 4 and week 7 ( $P=.25$ ). Self-efficacy also improved from baseline to postcourse (mean difference 1.1, SD 1.9,  $P=.01$ ). Fatigue severity scores were unchanged from baseline to postcourse (mean difference 0.4, SD 1.6,  $P=.25$ ). Similarly, loneliness scores remained unchanged from pre- to postcourse (mean difference 0.2, SD 1.8,  $P=.60$ ).

**Table 4.** Summary of baseline and postcourse well-being measures.

Measure	Baseline score (N=28), mean (SD)	Postcourse score (N=28), mean (SD)	Difference, mean (SD)	<i>t</i> (df)	<i>P</i> value	Effect size (Cohen <i>d</i> )
Positive mental well-being	40.1 (8.7)	47.1 (6.6)	6.5 (8.5)	-4.1 (27)	<.001	0.8
Self-efficacy <sup>a</sup>	4.8 (2.0)	5.9 (1.9)	1.1 (1.9)	-2.8 (26)	.01	0.5
Fatigue	5.9 (1.3)	5.5 (1.2)	0.4 (1.6)	1.2 (27)	.25	0.3
Loneliness	5.9 (2.0)	5.7 (1.9)	0.2 (1.8)	0.5 (27)	.60	0.1

<sup>a</sup>N=27 due to missing data.

**Table 5.** SWEMWBS<sup>a</sup> scores across the program.

Measure	Week 1 score (N=19), mean (SD)	Week 4 score (N=16), mean (SD)	Week 7 score (N=16), mean (SD)	$F_{1,4,22.1}$ value	<i>P</i> value
SWEMWBS	19.5 (4.5)	22.6 (4.9)	24.1 (4.6)	20.2	<.001

<sup>a</sup>SWEMWBS: Short-Form Warwick-Edinburgh Mental Wellbeing Scale.

## Goal Setting

Of the 47 participants, 26 (55.3%) used the goal-setting tool at least once, and a total of 97 goals were set across the 8-week course. The mean and median number of goals set by participants using the tool was 4 (range 1-9). All 97 goals were coded according to the 5 goal-oriented domains in TEDSS (see Table 6) [43]. Cohen  $\kappa$  was used to determine whether there was agreement between the 2 researchers' judgement on whether participants' goals were categorized as activity, internal, social interaction, health behavior, or disease-controlling strategies.

There was good agreement between the 2 researchers ( $\kappa=.77$ , 95% CI 0.557-0.973,  $P<.001$ ).

The goal-setting tool within the intervention was based on SMARTER goal setting. In personal communication, 1 participant emphasized the importance of setting *realistic* goals (eg, lowering expectations in line with what is possible now, not what was achievable pre-COVID-19). Further, the concept of *time* may be different depending on individual circumstances, symptoms, physical ability, lifestyle, etc, and goals are *measurable* in terms of improvement rather than being time oriented.

**Table 6.** Frequency of goal type in each domain of TEDSS<sup>a</sup>, with examples extracted from participant entries in the goal-setting tool embedded within the Hope Programme for Long COVID.

Domain	Goal type frequency (N=97), n (%)	Examples ("My goal is...")
Activity strategies: finding ways to participate in everyday activities despite experiencing problems, such as fatigue, pain, or memory loss	47 (48.5)	<ul style="list-style-type: none"> <li>• "Plan my routine more each day...knowing what I am doing saves energy and reduces stress" [JH]</li> <li>• "Going back to work for 2 hours a day...to feel useful again" [LG]</li> <li>• "To get my house ready for Christmas...because I loved Christmas" [TC]</li> <li>• "Tonight I am going to paint my toenails...because it makes me feel nice when I care for myself" [G]</li> <li>• "I am going to begin singing lessons one day each week at a music studio...I want to surprise my daughter by singing at her wedding" [KW]</li> </ul>
Health behavior strategies: maintaining a healthy lifestyle to enhance health and limit the risk of lifestyle-related illness	18 (18.6)	<ul style="list-style-type: none"> <li>• "To walk more than I have already done...to build up my strength and stamina" [AS]</li> <li>• "Eat healthier snacks each day at work...what I eat can make a big difference to how I feel" [JH]</li> <li>• "To continue eating healthy and losing weight...because I know my health and recovery will benefit" [VL]</li> <li>• "To work on my breathing exercises more...it will help improve my general condition and health" [Ru]</li> </ul>
Internal strategies: preventing and managing stress, negative emotions, and internal distress and creating inner calm	16 (16.5)	<ul style="list-style-type: none"> <li>• "To be positive and have a purpose in life...to conquer fear, it helps my panic attacks and nervous tension" [DS]</li> <li>• "To increase mindfulness activities at home...switch my brain off and have a proper rest" [LT]</li> <li>• "To have a relaxing day today...because everyone needs to unwind" [TC]</li> </ul>
Social interaction strategies: managing social interactions and relationships to be able to participate without exposure to negative reactions	10 (10.3)	<ul style="list-style-type: none"> <li>• "To reach out to two people I haven't spoken to for a while on Whatsapp...because I value their friendship" [AM]</li> <li>• "To go out for lunch with my daughter on a Thursday lunchtime in a café...because I enjoy her company" [TC]</li> <li>• "I want to make sure I keep in touch with my friends and family, by phone or in person...because these relationships are very important to me" [IA]</li> </ul>
Disease-controlling strategies: preventing, controlling, and limiting symptoms, complications, or disease progression	6 (6.2)	<ul style="list-style-type: none"> <li>• "Starting a physical rehabilitation course...I so desperately want to get better" [SC]</li> <li>• "To feel less dizzy by doing my exercises...it will make me feel better overall with long covid symptoms" [KB]</li> </ul>

<sup>a</sup>TEDSS: taxonomy of everyday self-management strategies.



## Discussion

### Principal Findings

In this study, we aimed to codevelop and test a digital self-management intervention for people with LC and estimate the parameters for a full trial. Sixteen people with lived experience of LC (both patients and professionals) led the codevelopment phase through workshops and interviews to discuss unmet needs and areas in which they felt support was lacking. This allowed us to coproduce bespoke and relevant content aligned to the needs of many people with long-term effects of COVID-19.

Our target sample size of 40 was exceeded (N=47), suggesting the advertising campaign for enrollment and subsequent email invitation to participate in the study is a viable strategy. The postcourse questionnaire completion rate was suboptimal (60%), when compared to the >70% completion target in our previous Hope Programme for Long COVID study protocol [44] and 80% completion rates in other Hope Programme for Long COVID studies [20,22]. This will be addressed in future studies by enhanced incentives for completion (see later). Nevertheless, the postprogram outcome data indicated significant improvements in mental well-being and self-efficacy scores at the end of the program relative to the start. A change in the WEMWBS score of 3 or more points is a minimally important level of change, so an average group increase of 7 points from pre- to postprogram is extremely encouraging. However, this measure was taken immediately at the end of the intervention rather than at a later follow-up, so longer-term changes are yet to be determined. Furthermore, the finding that the most significant improvement in positive well-being is made within the first half of the course is essential for our planning and design of future interventions. For example, future courses may consider presenting modules in the first half of the intervention that participants rank as being the most useful or address the most common issues, such as fatigue. An alternative explanation could of course be that participants found the second half of the course to be less useful. Our next trial will incorporate follow-up interviews with participants and facilitators to gain a better understanding of their experience of the course and suggestions regarding the relevance of all modules.

The significant improvement in self-efficacy is promising, as the Hope Programme for Long COVID is a self-management course providing tools to support people to manage their own health and well-being. These early data indicate that participants feel more confident to manage their COVID-19-related symptoms after taking part in the course. However, we must exercise caution when interpreting these changes in participant scores, as the study was not powered to detect statistically significant changes. Full inferential analysis and statistical significance will be determined by a fully powered trial.

All goals set by participants mapped onto the 5 goal-oriented domains in TEDSS [43]. The most frequent type of goals related to activity strategies, such as using the 3 Ps (ie, pace, plan, prioritize), organizing time and routines, and engaging in meaningful activities. These relate to common self-management techniques and are embedded throughout the Hope Programme

for Long COVID. It is encouraging that participants appear to be using these strategies in setting and planning their own short-term goals, giving further assurance of the usefulness of the intervention for self-managing LC symptoms. Interestingly, goals relating to disease-controlling strategies were the least common types of goals set by participants in this study. This is perhaps because of the varied nature of LC symptoms and a lack of clear medication management that would be inherent in other complex conditions. As medical and scientific research on treatments for LC continues, disease management strategies may become more applicable within everyday self-management of LC symptoms.

The Hope Programme for Long COVID was delivered digitally to allow participants flexibility in when and how they accessed the intervention. The wider literature reports this as the preferred format of participants seeking peer support [3,4]. Although we did not offer an in-person option, digital face-to-face meetings were well attended in the form of Hope Cafés. It is not possible to know whether an in-person delivery would have different outcomes, but we believe that digital delivery removes barriers introduced by physical ill-health (eg, fatigue) and time constraints (eg, employment, care duties) in the population with LC.

### Strengths and Limitations

A major strength of this study is the rapid development and deployment of the Hope Programme for Long COVID, due to the repurposing of an existing digital intervention for long-term conditions. The course was developed and delivered by a multidisciplinary team and social enterprise, exploiting existing networks to draw upon professional and clinical expertise and maximize the impact of the Hope Programme for Long COVID. A further strength of this study is the depth of understanding gained about how the intervention is used by people with LC. That is, the data showed that the most significant change in mental well-being is likely to occur in the first half of the intervention, indicating that the topics covered in the first 4 weeks of the intervention are the most useful for participants or possibly that 4 weeks of peer support are sufficient to enable a positive change in mental well-being. It may be that participants gain a validation of their condition and a feeling of universality after only a few weeks on the program. This could be especially valuable, given the experience some patients have had of being dismissed by health professionals and stigmatized or invalidated socially [5-7]. Goals were most frequently related to activities and health behaviors, indicating that the motivations, priorities, and needs of participants center around staying physically and mentally healthy through planning meaningful activities. It must be noted that this was a self-selecting sample, where participants who saw the ads were motivated to join the intervention and the research. Improvements in mental well-being and self-efficacy may be inflated by scores from people who were simply more open to, or actively searching for, self-management support.

This study yielded encouraging data regarding improved mental well-being and self-efficacy scores after the course. This pre-post study was not powered, a priori, to detect a statistically significant difference in key outcome measures, nor to account

for potential covariates, such as age or education, which would ideally be included in a multivariate model to test the effect of the intervention. Nevertheless, we presented preliminary inferential statistics to signal efficacy. For a study of this nature to be sufficiently powered, based on the effect sizes noted within this study, and retaining an  $\alpha$  error rate of .05, a  $(1 - \beta)$  error of .95, and a 2-tailed hypothesis, we would need to recruit 32 participants based on the WEMWBS [35], 54 participants based on the Self-Efficacy in Managing Chronic Disease scale [36], 147 participants based on the FSS [38], and up to 1302 participants based on the UCLA Loneliness Scale [40]. (Note that these sample sizes are presented assuming each variable to be to primary outcome.) In a future study, the inclusion of a control group would allow comparison of postcourse scores between those who took part in the Hope Programme for Long COVID and those who had not received the intervention. A fully powered study using the WEMWBS as the primary outcome measure would require 87 participants per group (ie, intervention and control groups) to minimally detect moderate effect sizes (ie, Cohen  $f \geq 0.25$ , Cohen  $d \geq 0.5$ ) or 42 participants per group to minimally detect large effect sizes (ie, Cohen  $f \geq 0.4$ , Cohen  $d \geq 0.8$ ), both with a power of 0.95, and an a priori  $\alpha$  of .05 [45]. The results of the next study—a waitlist control trial—will allow us to specify with greater confidence that improvements in well-being and self-efficacy are related to the intervention, rather than natural, spontaneous recovery over 8 weeks. Collecting longer-term follow-up data in the next study

will show whether any improvements continue beyond the end of the intervention and whether participants are making use of the techniques and self-management strategies months later.

Participant interviews will be incorporated into the research design of the next trial to capture important data regarding acceptability, nonusage, or reasons for dropout. These rich, qualitative data would help us understand the key factors behind participant “improvements” (eg, whether it was peer support, particular activities, or goal setting that people found most useful). Such data may also indicate why greater changes in well-being are observed in the first half of the course. Better incentives (ie, a guaranteed £10 (US \$10.57) voucher for completing pre and postprogram surveys) will also be included to encourage participants to complete follow-up questionnaires regardless of whether they used or enjoyed the program.

## Conclusion

We codesigned a bespoke, digital self-management intervention with and for people with LC. Our recruitment rate was extremely encouraging, and the sample characteristics largely mirrored those of the wider UK population with LC. We enlisted an enhanced incentive strategy to improve completion rates of follow-up questionnaires in future research. We have already started our next trial (ISRCTN: 11868601) using a nonrandomized waitlist control design to further examine intervention efficacy.

## Acknowledgments

The Hope Programme for Long COVID is a place-based offering, delivered as part of the community “Living Well With Hope” project, funded by National Health Service (NHS) Charities Together, in conjunction with University Hospital Coventry and Warwickshire Charity.

Development of the physiological components of the intervention was led by Atrium Health, a Coventry-based social enterprise providing rehabilitation services.

## Conflicts of Interest

GM is CEO of Hope For The Community and AT is a non-executive director of the same company.

## References

1. Office for National Statistics. Prevalence of Ongoing Symptoms Following Coronavirus (COVID-19) Infection in the UK: 6 January 2022. 2020. URL: <https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/bulletins/prevalenceofongoingsymptomsfollowingcoronaviruscovid19infectionintheuk/6january2022> [accessed 2022-09-28]
2. Higgins V, Sohaei D, Diamandis EP, Prassas I. COVID-19: from an acute to chronic disease? Potential long-term health consequences. *Crit Rev Clin Lab Sci* 2021 Aug 21;58(5):297-310. [doi: [10.1080/10408363.2020.1860895](https://doi.org/10.1080/10408363.2020.1860895)] [Medline: [33347790](https://pubmed.ncbi.nlm.nih.gov/33347790/)]
3. Ladds E, Rushforth A, Wieringa S, Taylor S, Rayner C, Husain L, et al. Persistent symptoms after covid-19: qualitative study of 114 “long covid” patients and draft quality principles for services. *BMC Health Serv Res* 2020 Dec 20;20(1):1144 [FREE Full text] [doi: [10.1186/s12913-020-06001-y](https://doi.org/10.1186/s12913-020-06001-y)] [Medline: [33342437](https://pubmed.ncbi.nlm.nih.gov/33342437/)]
4. 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](https://pubmed.ncbi.nlm.nih.gov/32665317/)]
5. Van de Vyver J, Leite AC, Alwan NA. Navigating the social identity of long covid. *BMJ* 2021 Nov 26;375:n2933. [doi: [10.1136/bmj.n2933](https://doi.org/10.1136/bmj.n2933)] [Medline: [34836878](https://pubmed.ncbi.nlm.nih.gov/34836878/)]
6. Byrne EA. Understanding long covid: nosology, social attitudes and stigma. *Brain Behav Immun* 2022 Jan;99:17-24. [doi: [10.1016/j.bbi.2021.09.012](https://doi.org/10.1016/j.bbi.2021.09.012)] [Medline: [34563621](https://pubmed.ncbi.nlm.nih.gov/34563621/)]

7. Ballering A, Olde Hartman T, Rosmalen J. Long COVID-19, persistent somatic symptoms and social stigmatisation. *J Epidemiol Community Health* 2021 Feb 23;75(6):603-604 [FREE Full text] [doi: [10.1136/jech-2021-216643](https://doi.org/10.1136/jech-2021-216643)] [Medline: [33622802](https://pubmed.ncbi.nlm.nih.gov/33622802/)]
8. NHS England. 2022/23 Priorities and Operational Planning Guidance. 2022. URL: <https://www.england.nhs.uk/wp-content/uploads/2021/12/B1160-2022-23-priorities-and-operational-planning-guidance-v2.pdf> [accessed 2022-09-28]
9. Gallagher P, Saunders T. Doctors Told to stop Sending Patients to Long Covid Clinics at Large NHS Trusts. URL: <https://www.inews.co.uk/news/health/doctors-told-stop-sending-patients-long-covid-clinics-large-nhs-trusts-1416130> [accessed 2022-09-28]
10. Lonergan PE, Washington Iii SL, Branagan L, Gleason N, Pruthi RS, Carroll PR, et al. Rapid utilization of telehealth in a comprehensive cancer center as a response to covid-19: cross-sectional analysis. *J Med Internet Res* 2020 Jul 06;22(7):e19322 [FREE Full text] [doi: [10.2196/19322](https://doi.org/10.2196/19322)] [Medline: [32568721](https://pubmed.ncbi.nlm.nih.gov/32568721/)]
11. Salako O, Okunade K, Allsop M, Habeebu M, Toye M, Oluyede G, et al. Upheaval in cancer care during the COVID-19 outbreak. *Ecancermedicalscience* 2020;14:ed97 [FREE Full text] [doi: [10.3332/ecancer.2020.ed97](https://doi.org/10.3332/ecancer.2020.ed97)] [Medline: [32269597](https://pubmed.ncbi.nlm.nih.gov/32269597/)]
12. Morrow-Howell N, Galucia N, Swinford E. Recovering from the covid-19 pandemic: a focus on older adults. *J Aging Soc Policy* 2020 Apr 26;32(4-5):526-535. [doi: [10.1080/08959420.2020.1759758](https://doi.org/10.1080/08959420.2020.1759758)] [Medline: [32336225](https://pubmed.ncbi.nlm.nih.gov/32336225/)]
13. NHS England. Your COVID Recovery. 2021. URL: <https://www.yourcovidrecovery.nhs.uk> [accessed 2022-09-28]
14. National Institute for Health and Care Excellence (NICE), Scottish Intercollegiate Guidelines Network (SIGN), Royal College of General Practitioners (RCGP). COVID-19 Rapid Guideline: Managing the Long-Term Effects of COVID-19. 2022. URL: <https://www.nice.org.uk/guidance/ng188/resources/covid19-rapid-guideline-managing-the-longterm-effects-of-covid19-pdf-51035515742> [accessed 2022-09-28]
15. Greenhalgh T, Knight M, A'Court C, Buxton M, Husain L. Management of post-acute covid-19 in primary care. *BMJ* 2020 Aug 11;370:m3026. [doi: [10.1136/bmj.m3026](https://doi.org/10.1136/bmj.m3026)] [Medline: [32784198](https://pubmed.ncbi.nlm.nih.gov/32784198/)]
16. Torjesen I. NICE advises against using graded exercise therapy for patients recovering from covid-19. *BMJ* 2020 Jul 21;370:m2912. [doi: [10.1136/bmj.m2912](https://doi.org/10.1136/bmj.m2912)] [Medline: [32694164](https://pubmed.ncbi.nlm.nih.gov/32694164/)]
17. NHS England. National Guidance for Post-COVID Syndrome Assessment Clinics. 2021. URL: <https://www.england.nhs.uk/coronavirus/wp-content/uploads/sites/52/2020/11/C1248-national-guidance-post-covid-syndrome-assessment-clinics-v2.pdf> [accessed 2022-09-28]
18. Wade DT. Goal setting in rehabilitation: an overview of what, why and how. *Clin Rehabil* 2009 Apr 17;23(4):291-295. [doi: [10.1177/0269215509103551](https://doi.org/10.1177/0269215509103551)] [Medline: [19293289](https://pubmed.ncbi.nlm.nih.gov/19293289/)]
19. Martin F, Wright H, Moody L, Whiteman B, McGillion M, Clyne W, et al. Help to overcome problems effectively for cancer survivors: development and evaluation of a digital self-management program. *J Med Internet Res* 2020 May 19;22(5):e17824 [FREE Full text] [doi: [10.2196/17824](https://doi.org/10.2196/17824)] [Medline: [32209529](https://pubmed.ncbi.nlm.nih.gov/32209529/)]
20. Wright H, Martin F, Clyne W, Clark CCT, Matouskova G, McGillion M, et al. A digital self-management program (Help to Overcome Problems Effectively) for people living with cancer: feasibility randomized controlled trial. *J Med Internet Res* 2021 Nov 05;23(11):e28322 [FREE Full text] [doi: [10.2196/28322](https://doi.org/10.2196/28322)] [Medline: [34738912](https://pubmed.ncbi.nlm.nih.gov/34738912/)]
21. Anderson JK, Turner A, Clyne W. Development and feasibility of the Help to Overcome Problems Effectively (HOPE) self-management intervention for people living with multiple sclerosis. *Disabil Rehabil* 2017 Jun 09;39(11):1114-1121. [doi: [10.1080/09638288.2016.1181211](https://doi.org/10.1080/09638288.2016.1181211)] [Medline: [27278670](https://pubmed.ncbi.nlm.nih.gov/27278670/)]
22. Martin F, Clyne W, Pearce G, Turner A. Self-management support intervention for parents of children with developmental disorders: the role of gratitude and hope. *J Child Fam Stud* 2019 Feb 13;28(4):980-992. [doi: [10.1007/s10826-018-01308-1](https://doi.org/10.1007/s10826-018-01308-1)]
23. Yalom I, Leszcz M. *The Theory and Practice of Group Psychotherapy*. 5th ed. New York, NY: Basic Books; 2005.
24. Tugade MM, Fredrickson BL, Barrett LF. Psychological resilience and positive emotional granularity: examining the benefits of positive emotions on coping and health. *J Pers* 2004 Dec;72(6):1161-1190 [FREE Full text] [doi: [10.1111/j.1467-6494.2004.00294.x](https://doi.org/10.1111/j.1467-6494.2004.00294.x)] [Medline: [15509280](https://pubmed.ncbi.nlm.nih.gov/15509280/)]
25. Snyder CR, Lopez SJ, editors. *Oxford Handbook of Positive Psychology*. Oxford; New York: Oxford University Press; 2009.
26. Wood AM, Maltby J, Gillett R, Linley PA, Joseph S. The role of gratitude in the development of social support, stress, and depression: two longitudinal studies. *J Res Pers* 2008 Aug;42(4):854-871. [doi: [10.1016/j.jrp.2007.11.003](https://doi.org/10.1016/j.jrp.2007.11.003)]
27. Larsen DJ, King RL, Stege R, Egeli NA. Hope in a strengths-based group activity for individuals with chronic pain. *Couns Psychol Q* 2015 Mar 13;28(2):175-199. [doi: [10.1080/09515070.2015.1007444](https://doi.org/10.1080/09515070.2015.1007444)]
28. Renger R, Hurley C. From theory to practice: Lessons learned in the application of the ATM approach to developing logic models. *Eval Program Plann* 2006 May;29(2):106-119. [doi: [10.1016/j.evalprogplan.2006.01.004](https://doi.org/10.1016/j.evalprogplan.2006.01.004)]
29. Turner A, Martin F. Spreading HOPE: the development of a hope-based self-management intervention. In: Martz E, editor. *Promoting Self-Management of Chronic Health Conditions*. Oxford, UK: Oxford University Press; 2017.
30. Gunning R. *The Technique of Clear Writing*. New York, NY: McGraw Hill Higher Education; 1968.
31. Editorial. Understanding long COVID: a modern medical challenge. *Lancet* 2021 Aug;398(10302):725. [doi: [10.1016/S0140-6736\(21\)01900-0](https://doi.org/10.1016/S0140-6736(21)01900-0)]

32. Garner P. Long Covid and Self-Help Pacing Groups—Getting by with a Little Help from Our Friends. 2020. URL: <https://blogs.bmj.com/bmj/2020/09/29/long-covid-and-self-help-pacing-groups-getting-by-with-a-little-help-from-our-friends/> [accessed 2022-09-28]
33. 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/)]
34. O'Connor RJ, Preston N, Parkin A, Makower S, Ross D, Gee J, et al. The COVID-19 Yorkshire Rehabilitation Scale (C19-YRS): application and psychometric analysis in a post-COVID-19 syndrome cohort. *J Med Virol* 2022 Mar;94(3):1027-1034 [FREE Full text] [doi: [10.1002/jmv.27415](https://doi.org/10.1002/jmv.27415)] [Medline: [34676578](https://pubmed.ncbi.nlm.nih.gov/34676578/)]
35. Tennant R, Hiller L, Fishwick R, Platt S, Joseph S, Weich S, et al. The Warwick-Edinburgh Mental Well-being Scale (WEMWBS): development and UK validation. *Health Qual Life Outcomes* 2007 Nov 27;5:63 [FREE Full text] [doi: [10.1186/1477-7525-5-63](https://doi.org/10.1186/1477-7525-5-63)] [Medline: [18042300](https://pubmed.ncbi.nlm.nih.gov/18042300/)]
36. Ritter PL, Lorig K. The English and Spanish Self-Efficacy to Manage Chronic Disease Scale measures were validated using multiple studies. *J Clin Epidemiol* 2014 Nov;67(11):1265-1273. [doi: [10.1016/j.jclinepi.2014.06.009](https://doi.org/10.1016/j.jclinepi.2014.06.009)] [Medline: [25091546](https://pubmed.ncbi.nlm.nih.gov/25091546/)]
37. Lorig KR, Sobel DS, Ritter PL, Laurent D, Hobbs M. Effect of a self-management program on patients with chronic disease. *Eff Clin Pract* 2001;4(6):256-262. [Medline: [11769298](https://pubmed.ncbi.nlm.nih.gov/11769298/)]
38. Krupp LB, Alvarez LA, LaRocca NG, Scheinberg LC. Fatigue in multiple sclerosis. *Arch Neurol* 1988 Apr 01;45(4):435-437. [doi: [10.1001/archneur.1988.00520280085020](https://doi.org/10.1001/archneur.1988.00520280085020)] [Medline: [3355400](https://pubmed.ncbi.nlm.nih.gov/3355400/)]
39. Valko P, Bassetti C, Bloch K, Held U, Baumann C. Validation of the Fatigue Severity Scale in a Swiss cohort. *Sleep* 2008 Nov;31(11):1601-1607 [FREE Full text] [doi: [10.1093/sleep/31.11.1601](https://doi.org/10.1093/sleep/31.11.1601)] [Medline: [19014080](https://pubmed.ncbi.nlm.nih.gov/19014080/)]
40. 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/)]
41. Snape D, Martin G. Measuring Loneliness: Guidance for Use of the National Indicators on Surveys. 2018. URL: <https://tinyurl.com/3w8ah8e6> [accessed 2022-09-28]
42. NIHR, RDS London. Justifying Sample Size for a Feasibility Study. 2019. URL: <https://www.rds-london.nihr.ac.uk/resources/justify-sample-size-for-a-feasibility-study/> [accessed 2022-09-28]
43. Audulv Å, Ghahari S, Kephart G, Warner G, Packer TL. The Taxonomy of Everyday Self-management Strategies (TEDSS): a framework derived from the literature and refined using empirical data. *Patient Educ Couns* 2019 Feb;102(2):367-375 [FREE Full text] [doi: [10.1016/j.pec.2018.08.034](https://doi.org/10.1016/j.pec.2018.08.034)] [Medline: [30197252](https://pubmed.ncbi.nlm.nih.gov/30197252/)]
44. Wright H, Martin F, Clyne W, Clark CCT, McGillion M, Matouskova G, et al. A digital program (Hope) for people living with cancer during the covid-19 pandemic: protocol for a feasibility randomized controlled trial. *JMIR Res Protoc* 2020 Dec 4;9(12):e24264. [doi: [10.2196/24264](https://doi.org/10.2196/24264)]
45. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. 2nd ed. Hillsdale, NJ: L. Erlbaum Associates; 1988.

## Abbreviations

**CWR:** Coventry, Warwickshire, and Rugby

**FSS:** Fatigue Severity Scale

**H4C:** Hope For The Community

**LC:** long COVID

**NHS:** National Health Service

**SMART:** specific, measurable, achievable, relevant, and timely

**SMARTER:** specific, measurable, achievable, relevant, time-based, enjoyable, rewarded

**SWEMWBS:** Short-Form Warwick-Edinburgh Mental Wellbeing Scale

**TEDSS:** taxonomy of everyday self-management strategies

**UCLA:** University of California, Los Angeles

**WEMWBS:** Warwick-Edinburgh Mental Wellbeing Scale

*Edited by A Mavragani; submitted 25.07.22; peer-reviewed by A Finny, Z Li; comments to author 17.08.22; revised version received 07.09.22; accepted 07.09.22; published 14.10.22.*

*Please cite as:*

*Wright H, Turner A, Ennis S, Percy C, Loftus G, Clyne W, Matouskova G, Martin F*

*Digital Peer-Supported Self-Management Intervention Codesigned by People With Long COVID: Mixed Methods Proof-of-Concept Study*

*JMIR Form Res* 2022;6(10):e41410

URL: <https://formative.jmir.org/2022/10/e41410>

doi:[10.2196/41410](https://doi.org/10.2196/41410)

PMID:[36166651](https://pubmed.ncbi.nlm.nih.gov/36166651/)

©Hayley Wright, Andrew Turner, Stuart Ennis, Carol Percy, Garry Loftus, Wendy Clyne, Gabriela Matouskova, Faith Martin. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 14.10.2022. 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 First Asynchronous Online Evidence-Based Medicine Course for Syrian Health Workforce: Effectiveness and Feasibility Pilot Study

Yazan Kenjrawi<sup>1,2</sup>, DDS, MedEd; Mayssoon Dashash<sup>1,3</sup>, DDS, MedEd, MSc, PhD, MFDS RCS

<sup>1</sup>Medical Education Program, Syrian Virtual University, Damascus, Syrian Arab Republic

<sup>2</sup>Faculty of Dentistry, Tishreen University, Latakia, Syrian Arab Republic

<sup>3</sup>Pediatric Dentistry Department, Faculty of Dentistry, Damascus University, Damascus, Syrian Arab Republic

**Corresponding Author:**

Yazan Kenjrawi, DDS, MedEd

Medical Education Program

Syrian Virtual University

Omar Bin Abdulaziz Street

Damascus

Syrian Arab Republic

Phone: 963 948379097

Email: [yazankenjrawi@gmail.com](mailto:yazankenjrawi@gmail.com)

## Abstract

**Background:** Evidence-based medicine (EBM) is critical for providing the best scientifically proven patient health care, and it is implemented worldwide in order to improve the quality of the delivered health care. However, not all Syrian health professionals are knowledgeable about the importance, methodology, and implementation of EBM. Providing web-based learning courses on EBM might be effective in improving the EBM knowledge of health care professionals.

**Objective:** This study was performed to test the effectiveness and the feasibility of an asynchronous web-based course on EBM in improving the competencies of Syrian health care professionals in terms of EBM.

**Methods:** A web-based course on EBM was developed in Arabic and uploaded onto the Syrian Virtual University platform. An electronic registration form was designed and distributed to medical groups on social media for registration to this web-based course. Both the pretest and posttest had the same 3 sections to measure the impact of this web-based EBM program on the knowledge, skills, and attitudes of the Syrian health care professionals. The posttest had an additional section for measuring the efficacy and ease of use of this program. Student paired 1-tailed *t* test was used to analyze the differences in the different assessment sections among the participants.

**Results:** Nineteen participants filled the electronic registration form, but 8 participants did not meet the inclusion criteria. Therefore, the pretest was sent to the remaining 11 participants (7 men and 4 women) who graduated from Syrian universities. Ten of them completed the pretest, while 7 of them completed the posttest. The web-based course was found to be effective in improving the participants' EBM knowledge, skills, and attitudes at  $P > .05$ . Further, the web-based EBM course was feasible and easy-to-use.

**Conclusions:** In order for EBM to be implemented in Syria, continuous medical education training programs should be designed for clinical practitioners. Our study shows that asynchronous web-based medical education is an effective and a feasible means for introducing the concept of EBM, improving practitioners' skills, and promoting the positive attitudes of Syrian clinical practitioners toward EBM.

(*JMIR Form Res* 2022;6(10):e36782) doi:[10.2196/36782](https://doi.org/10.2196/36782)

**KEYWORDS**

evidence-based medicine; continuous medical education; web-based learning; Syria; medical graduates; medical education; web-based course

## Introduction

Evidence-based medicine (EBM) is the integration of well-designed research with clinical experience and patients' beliefs that aim at developing decision-making through the utilization of the best available scientific evidence [1]. The practice of EBM requires formulating a question based on the clinical problem faced, searching the available literature to find evidences, and critically appraising the evidence in order to find the best application on the patient [2]. This process is followed by the evaluation of each of the last 4 steps of EBM (explained in the subsequent section) to make sure that the goal of EBM in improving patient care has been achieved [2].

EBM promotes a patient-centered approach by making patients' preferences and values an important component of the treatment process [3]. EBM is also a cost-effective approach since it promotes what has already been proven to be effective [4]. This will be reflected not only on the health care provider but also on the patient, which makes its implementation of special importance in limited-resource countries. EBM provides practitioners with an effective continuous medical education strategy by offering the required tools needed to keep up with the most recent scientific evidence [5] and benefits patients as it will provide them with the best scientifically approved health care.

EBM was first introduced in the medical programs in 1990 at McMaster University [6]. Since then, it has been a part of the medical programs offered in universities worldwide [7]. In limited-income countries, however, teaching EBM is faced with major barriers such as the lack of educational resources and the language barrier [5]. These barriers are more obvious in Syria, as EBM was still not part of the undergraduate medical program in public universities until 2017 [8]. Moreover, medical education in Syria is delivered in Arabic, with the vast majority of students having the desire to improve their English skills to be able to raise their scientific standards, but the English language is still a significant barrier to learning [9]. This barrier has resulted in producing doctors who are not familiar with EBM and who lack the skills needed to learn about this concept on their own, since educational resources on EBM are rare in Arabic. Another barrier is that clinical practitioners usually lack the needed time to master EBM skills [6].

e-Learning is as an approach that utilizes the internet and digital technologies to deliver mediated, well-designed, learner-centered, and interactive learning to anyone at any place and anytime according to instructional design principles [10]. Several studies have shown that e-learning is as effective as the traditional methods of learning EBM [11-13]. In fact, asynchronous web-based learning is superior to face-to-face learning, because it offers the opportunity for learners to learn based on their availability without worrying about a specific time or place [14], which makes it the perfect option for busy clinical practitioners. Additionally, social distancing procedures due to the COVID-19 pandemic make web-based learning the safest option, specifically for clinical practitioners [15]. Web-based education facilitates learning from different geographical places [14], which is particularly useful in war-torn

Syria, where moving across the country is not an option for many Syrians [16]. In addition, web-based learning in medical education has already been proven to be successful in Syria during the ongoing crisis and is suggested to be further implemented to cover the other health topics [17]. Regarding EBM, peer-taught web-based courses should be implemented in limited-income countries as this is found to be an effective approach [18].

In light of the above reasons, we developed an asynchronous web-based EBM course in Arabic that would be delivered by a peer of the targeted study population in order to investigate its effects on Syrian medical graduates' competencies and to study the course's efficacy and feasibility.

## Methods

### Electronic Registration Form and Participant Selection Criteria

An electronic registration form was designed using Google forms. This form included an introduction about the web-based course and the reason for choosing the internet to deliver this course. The form also included the learning objectives of the web-based course in addition to how those objectives are connected to the learning contents, the specified duration of each video in the asynchronous web-based session, and a detailed description of the expected time for the whole web-based course. This form worked as an advertising tool in addition to its use as a registration form. The e-form was distributed to medical groups on both Telegram and WhatsApp. To be accepted in this web-based course, the participant had to be a clinical practitioner who had graduated from a Syrian medical faculty.

### Web-Based EBM Course Design

The design of the web-based course followed the objective-based approach. The learning objectives were set to meet the need of medical practitioners who were currently working in clinical settings and who did not have the opportunity to learn about EBM in their undergraduate medical program. The web-based EBM course had 3 main learning objectives. The first objective was to define and introduce the concept of EBM and to explain its 5 steps. The second objective focused on improving the knowledge and skills about the methodology and how to apply EBM while approaching clinical problems in clinical settings. The last learning objective was related to the clinical practice guidelines and its importance in clinical settings.

After setting those objectives, YK, who was a peer of the targeted study population, developed a screening process for the medical literature to design the material of the web-based course. Considering the time problem and the language barrier that deprive clinical practitioners from learning this concept, 11 presentations were designed in Arabic using the PowerPoint 2016 software. Those were turned into 11 video lessons by YK who explained the presentation screens in Arabic by using 2 free video recording software, namely, Loom and iFun Screen Recorder. For editing purposes, OpenShot Video Editor, which is an open-source software, was used to remove the noise,

silence, and errors clips from the videos. On average, a lesson lasted for 6 minutes.

The lessons were categorized under 4 learning units. The first learning unit was based on the first learning objective. At the start of the second unit, participants were introduced to a clinical problem, and they were asked to turn this problem into the PICO (Patient, Intervention, Comparison, and Outcome) format question and to search for an answer in one of the medical search engines (eg, Cochrane, PubMed). This was a voluntary task and participants were able to find the best solution for that PICO task problem in the last lesson of the second unit. In the same unit, there was also a practical session where a clinical problem was presented, turned into a PICO question, and then a medical search engine (PubMed) was searched for an answer to that

PICO question. The third learning unit was also related to the second learning objective. It focused on the last 3 steps of the EBM implementation. It started by introducing the scientific evidence pyramid and then moving onto defining the concept of validity and reliability. This unit also dealt with the application of the evidence and the evaluation process of each step of EBM. The fourth learning unit covered the last learning objective, wherein the clinical practice guidelines were defined, and the components, databases, and the benefits of following the recommendations of clinical practice guidelines were listed. Participants needed approximately 70 minutes to finish all the materials of the web-based EBM course. [Table 1](#) shows the web-based course map ([Multimedia Appendices 1-Multimedia Appendices 11](#)).

**Table 1.** The web-based course map of evidence-based medicine in clinical settings ([Multimedia Appendices 1-Multimedia Appendices 11](#)).

Objective, learning unit and objective, lesson title	Duration	PowerPoint Presentation
<b>Learning objective 1<sup>a</sup></b>		
<b>First learning unit: Introduction to evidence-based medicine</b>		
Defining evidence-based medicine	5 min	<a href="#">Multimedia Appendix 1</a>
The importance of evidence-based medicine	8 min	<a href="#">Multimedia Appendix 2</a>
Explanation of the 5 steps of evidence-based medicine	5 min	<a href="#">Multimedia Appendix 3</a>
<b>Learning objective 2<sup>b</sup></b>		
<b>Second learning unit: Clinical problem solving</b>		
Patient, Intervention, Comparison, and Outcome format	8 min	<a href="#">Multimedia Appendix 4</a>
How to search in medical search engines	7 min	<a href="#">Multimedia Appendix 5</a>
Practical training	5 min	<a href="#">Multimedia Appendix 6</a>
<b>Third learning unit: Appraise, apply, and evaluate our approach to solving a clinical problem<sup>c</sup></b>		
Critical appraising 1	8 min	<a href="#">Multimedia Appendix 7</a>
Critical appraising 2	7 min	<a href="#">Multimedia Appendix 8</a>
Application guide and process calendar	4 min	<a href="#">Multimedia Appendix 9</a>
<b>Learning objective 3<sup>d</sup></b>		
<b>Fourth learning unit: Clinical guidelines</b>		
Clinical practice guidelines	6 min	<a href="#">Multimedia Appendix 10</a>
The importance of clinical practice guideline	4 min	<a href="#">Multimedia Appendix 11</a>

<sup>a</sup>Define the concept of evidence-based medicine, summarizing its benefits, and explaining its 5 steps.

<sup>b</sup>Apply the 5 steps of evidence-based medicine to solve clinical problems encountered during clinical work. Step 1: convert your clinical problem into an answerable question. Step 2: Searching medical search engines for evidence using the Patient, Intervention, Comparison, and Outcome format.

<sup>c</sup>Apply the 5 steps of evidence-based medicine to solve clinical problems encountered during clinical work. Step 3: critically evaluate studies. Step 4: application of the practical guide. Step 5: defining the process of evaluating the steps of evidence-based medicine.

<sup>d</sup>Define the clinical practice guideline and explain its components and importance for health care.

## Knowledge, Skills, and Attitudes of Participants Regarding EBM

The EBM knowledge, skills, and attitudes were defined in the context of this study as follows.

### *EBM Knowledge*

This involves the understanding of the EBM triad and pointing out its importance, in addition to being able to explain the 5

steps of EBM and the evidence pyramid, and defining the clinical practice guideline, its components, and importance.

### *EBM Skills*

This involves effectively being able to turn a clinical problem into a PICO format and to search for an answer in a web-based medical database and applying the critical appraisal skills and implementing the evidence and the evaluation of the entire process.



### ***Attitude of Participants Toward EBM***

This involves the willingness to learn and improve knowledge and skill with regard to EBM and support the implementation of EBM on a large scale in the Syrian community.

### **Measurements**

The pretests and the posttests were drafted in Arabic and designed using Google forms. The first section of the pretest focused on the demographic information of the participants. The other 3 sections were designed to test the participants' knowledge, skills, and attitudes regarding EBM in the context of this study. All 3 sections were designed based on the learning objectives of the web-based EBM course. The knowledge section included 20 questions: 12 multiple choice questions and 8 true or false questions. For the skills section, a clinical problem was introduced followed by 8 multiple choice questions about its PICO elements in addition to the best study type to answer that type of PICO question. The last section consisted of 20 statements on the Likert scale to test the attitudes of the participants regarding EBM. The posttest differed from the pretest in that it had an extra section that measured the usefulness and ease of use of the web-based course. This section included 14 statements evaluated on the Likert scale [19]. The assessment was checked for content and face validity by the supervisor MD, an expert in the field of EBM and a member of the Cochrane Collaboration. However, no criterion or construct validity of the questions was attempted.

### **Study Design and Procedures**

The contents of the course were validated by the supervisor MD according to the essential information presented by Cochrane Collaboration and all related international publications. After that, the course videos were uploaded to the learning management system of the Syrian Virtual University. The web-based EBM course was designed as a 1-group pretest and posttest study. Replies to the electronic registration form expired after 72 hours. Then, participants were asked to answer the pretest. After that, they were provided with the login data to access the web-based course on the learning management system. On the first day, all asynchronous sessions were available for the participants so that they could attend the web-based course according to their own pace and schedule. On the second day, reading materials were provided to the participants, which included the presentations and the scientific resources used to develop these presentations. The reading materials provided more in-depth information, and participants were informed as such. On the third day, an encouraging message and a reminder of the deadline (ie, the fourth day) was sent to participants. On the fourth and last day, the posttest

assessment was shared with the participants along with an encouraging message to fill it even if they did not attend all the sessions. Participants were kept motivated during these 4 days by daily messaging, explaining the significance of this web-based course and its benefits to their practice and the health care system. Further, they were constantly encouraged to share their opinions and suggestions about the web-based EBM course. The degree of commitment of the participants and interest in the EBM course were monitored by the researchers (YK and MD) by observing and counting the number of downloads, videos watched, and duration of access to the course through a specific tracking system related to the Syrian Virtual University platform in collaboration with the technical support staff in the university without violation of private data and information of participants.

### **Statistical Analysis**

All data obtained from the pretests and posttests were retrieved and put into Excel 2016. Paired 1-tailed *t* test was used to analyze the differences between the findings of the participants in the pretest and posttest.  $P < .05$  was used to determine the statistical significance threshold.

### **Ethics Approval**

This study was reviewed and approved by the ethics approval committee of the Syrian Virtual University on November 15, 2021 (ID 2083/0). Informed consent was obtained from all potential participants through the electronic registration form, wherein they were informed that the data shared for the purpose of this research would be kept confidential and would solely be used for this research.

## **Results**

### **Characteristics of the Participants**

Nineteen potential participants filled out the electronic registration form. Eight potential participants did not meet the inclusion criteria, and the pretest was sent to the remaining 11 participants (7 men and 4 women). Eight of them were dentists, 2 of them were physicians, and 1 was a pharmacist. The majority of the participants (9/11, 82%) graduated from Tishreen University, 1 graduated from Aleppo University, and 1 graduated from the International Private University for Science and Technology in Damascus. All of them were studying for their master's degree in their fields. Table 2 shows the demographic characteristics of the participants in detail. Ten participants filled out the pretest, while 7 completed the final test.

**Table 2.** Demographic characteristics of the participants.

Characteristic	Participants (n)
<b>Gender</b>	
Female	4
Male	7
<b>Age</b>	
25 years	1
26 years	6
27 years	2
31 years	1
40 years	1
<b>Graduation year</b>	
2021	1
2020	1
2019	2
2018	4
2017	1
2014	1
2004	1
<b>Medical specialty</b>	
Dentistry	8
Medicine	2
Pharmacy	1
<b>Master's study</b>	
Medical education	4
Quality assurance	1
Cosmetic dentistry	1
Prosthodontics	1
Oral maxillofacial surgery	1
Otolaryngology	1
General surgery	1
Microbiology, Hematology, and Immunology	1
<b>Participants' self-perceived expertise in evidence-based medicine (scale of 1-10; 1: novice, 10: expert)</b>	
1	1
3	2
4	1
5	2
7	2
8	3

### Efficiency and Ease of Use of the Web-Based Course

Seven participants filled out the posttest. The majority of the participants agreed on the positive statements regarding the effectiveness and ease of use of the web-based course. [Table 3](#)

shows the statements and the number of participants who agreed to them.

The overall satisfaction with the web-based course was high, as 4 participants rated their satisfaction as 9 out of 10, while the other 3 participants rated their satisfaction as 10. Most of

the participants attended all the asynchronous web-based sessions, while only 2 could not attend the last one. Most of them tried to solve the PICO task, and only 1 participant did not attempt to do so. Further, all participants agreed on receiving future communications in case another web-based course was developed in the future. The practical training session and the PICO format session in the second learning unit were the most interesting and useful for 4 participants. Two of them found that the entire training program was useful and interesting, while 1 participant liked the topics of the third and fourth educational

units. As for the least interesting and useful aspect, 3 participants answered “nothing” on the web-based course being not interesting or useful. Two of them commented that having many definitions and terms affected their enjoyment and clarity of the learning process. Only 1 participant commented that the first lesson of the third unit was the least interesting and useful. Further, the web-based course was feasible, as the design of its materials and presenting them in asynchronous web-based sessions were free of charge.

**Table 3.** Participants who agreed with the statements used to test the efficacy and ease of use of the web-based course.

Statement	Agree/totally agree (n=7)
The objective of the web-based course was clear.	7
The content of the web-based course was understandable.	7
The web-based course met my expectations.	7
The web-based course offered me sufficient information for an introductory course.	7
The content of the web-based course matched my educational level.	7
I will be able to use what I learned in this web-based course.	7
The content of the web-based course fits to my work setting.	5
This web-based course will help me to better practice evidence-based medicine.	7
This web-based course will help me to improve the quality of my work.	7
The difficulty level of the web-based course was appropriate.	6
The web-based course was presented in a clear logical manner.	7
The web-based format was a good way for me to learn evidence-based medicine.	7
I enjoyed taking the web-based course.	7
I would like to take more of this kind of web-based courses.	7

### Knowledge, Skills, and Attitudes of the Participants Regarding EBM

Only the results of the participants who completed both the pretest and posttest assessment were included. The data were exported from the Google form in Excel format. Paired 1-tailed Student *t* test was performed 3 times to each group of the data representing each of the 3 sections of the assessment. The results showed that the web-based course is effective in terms of improving EBM competencies. The average score of the participants in the EBM knowledge section was clearly improved from 8.4 to 15 after attending the web-based

course—an improvement that was statically significant ( $P=.002$ ). Although the increase in the average score of the participants for EBM skills between the pretest and the posttest was quite small (from 3.1 to 4.4), it was still a statistically significant improvement ( $P=.03$ ). The average score of the participants' attitudes toward EBM was improved from 66.5 to 79.5 among medical graduates who did not receive EBM training during their undergraduate years, which was also statistically significant ( $P=.001$ ). The standard deviation, however, was always greater in the posttest scores when compared with that in the pretest scores, with the knowledge section having the greatest difference between the pretest and posttest (Table 4).

**Table 4.** Participants' scores in the pretest and posttest.

Section	X <sup>a</sup> <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>6</sub>	X <sub>7</sub>	Mean (SD)	P value
<b>Knowledge section (x<sup>b</sup>/20)</b>									.002
Pretest	7.00	6.00	8.00	10.00	9.00	9.00	10.00	8.4 (1.5)	
Posttest	18.00	14.00	14.00	18.00	15.00	7.00	19.00	15 (4)	
<b>Skills section (x/8)</b>									.03
Pretest	3.00	1.00	3.00	4.00	3.00	3.00	5.00	3.1 (1.2)	
Posttest	6.00	3.00	3.00	7.00	4.00	2.00	6.00	4.4 (1.9)	
<b>Attitudes section (x/100)</b>									<.001
Pretest	66	68	65	63	74	60	70	66.5 (4.6)	
Posttest	83	80	79	75	77	71	92	79.5 (6.6)	

<sup>a</sup>X indicates participant number.

<sup>b</sup>x indicates participant's score.

## Discussion

### Main Findings

Web-based medical education can provide many learning and training opportunities for medical students. Several web-based learning modalities have shown to improve clinical training, such as simulation technology, synchronous learning delivery, and web-based or videoconferencing for standardized patient-based training [20]. However, in Syria, despite the growing demand for web-based medical learning resources, e-learning is still in its early stage [21]. Moreover, since English is a language barrier in Syria, Syrian clinical practitioners are unlikely to benefit from the large number of already existing English web-based learning resources [9]. Furthermore, the current unstable political situation in Syria makes it hard to organize a traditional course [17].

In this study, our web-based EBM course is the first one in Syria, to the best of our knowledge, to rely entirely on asynchronous web-based sessions and to be designed and delivered in Arabic by a peer of the targeted study population. Additionally, this is the first web-based EBM learning course in Syria that targeted clinical practitioners who did not have the opportunity to be introduced to EBM during their undergraduate studies, and our study reveals the opportunities that e-learning offers in a low-resource and an unstable environment. The goal of our study was to test the feasibility of this web-based course and to develop an effective method to promote EBM practice in Syria. This goal could be achieved because a significant improvement in EBM learning was observed after the clinicians attended our web-based course that was designed using open-sourced software.

Kirkpatrick's model for the evaluation of training programs represent one of the most comprehensive strategies for evaluating organizational training [22]. Our web-based EBM course was found to be successful when evaluated on the basis of the adaptation of Kirkpatrick's model to the e-learning environment [22]. Course assessment showed a clear satisfaction, enjoyment, improvement in participants' learning, and willingness to apply what they learned after attending this

course. Our web-based EBM course was not only feasible but also extremely useful during the COVID-19 pandemic, wherein social distancing regulations had to be followed. Further, asynchronous web-based learning was found to be advantageous for busy clinical practitioners with crowded schedules, as they were able to attend approximately all sessions in just 4 days.

This web-based course focused on practitioners acquiring the needed skills to practice EBM in clinical settings. Given the fact that there is preappraised evidence available online, this web-based course focused on the first 2 sets of EBM skills in formatting a clinical question and searching the literature online. For that purpose, the PICO task was designed and presented to be solved voluntarily by the participants followed by general feedback on that task. In addition to designing a practical session, the process of searching a web-based search engine was presented. Moreover, a population of clinical practitioners was targeted to try to make this web-based EBM course clinically integrated, as it was found that EBM skills can be best fostered during ward rotation [23]. These were found to be effective techniques, as EBM skills were significantly improved in this 70-minute web-based course, which was provided over a period of 4 days. Further, the electronic registration form with its content worked as an effective advertising tool for the web-based course, as more than half of the potential participants rated their self-perceived expertise regarding EBM as 5 or more on the scale of 1-10.

### Limitations and Suggestions

This was a pilot study conducted with only a few medical graduates. Thus, the effectiveness of this web-based learning method could not be determined among older clinical practitioners who often have technophobia and will therefore find web-based learning to be a challenging task [24]. Further, it is assumed that only participants who were interested in learning the concept of EBM had joined, which could raise bias in the results. However, future studies will be performed on a more diverse sample, and critical appraisal tasks will be designed with general feedback to be delivered in the same technique as done for the PICO task to reveal more benefits of the e-learning in terms of EBM.

In the future, it is possible to adopt a peer-review strategy to communicate feedback provided by participants to their peers' work before introducing the general feedback and studying the effectiveness of such an approach and its reflection on the participants' EBM skills. Specifically, the peer's role was found to be effective in improving EBM skills, which can overcome the barrier of the limited number of EBM experts in Syria [18]. The effect of including a test after each asynchronous session is also worth studying, as it enhances the retention rates to make the participants feel more engaged [25]. Since this web-based course was found to be effective for the introduction of EBM, it can be disseminated further as part of a larger continuous medical education program to make all clinical practitioners familiar with the concept of EBM. This web-based EBM course can also be presented at the undergraduate level as an introductory session before proceeding to the newly developed official curriculum.

### Comparison With Prior Work

The results obtained in our study are consistent with the previously obtained findings [11,12,26] that e-medical education is effective and feasible for developing EBM knowledge and improving practitioners' attitudes toward EBM. Previous studies [11,12,26] have focused on comparisons between e-learning and lecture-based methods, but they have not shown the effects of e-learning on EBM skills. A study showed that e-learning was found to be effective for developing EBM skills; however, in the 8-week-long e-module in that study, a tutor with experience in teaching EBM was responsible for providing feedback on the multiple exercises and assignments incorporated

in that e-module [13]. Another study found that e-learning can improve EBM skills. However, that e-module contained summative, formative, and individual assessments [27]. In this study, EBM skills were improved by providing general feedback on a given task. Another study showed that a 90-minute introductory EBM e-module, which consists of set of screens without explanation for those screens, failed to improve EBM skills but succeeded in improving EBM knowledge and practitioners' attitudes [19]. Our web-based course improved EBM skills significantly in 70-minute video recordings of explanations for a set of presentations. In Syria, 2 other studies [8,18] were found to be associated with EBM teaching: the first one was based on traditional teaching methods by experienced EBM instructors for the undergraduate population [8], and the second one relied on both asynchronous and synchronous sessions of web-based learning and the web-based course was delivered over 6 weeks [18].

### Conclusions

Transforming health care in Syria from relying on expert opinion to relying on the latest scientific evidence starts by teaching the concept of EBM at the undergraduate level and continuing the medical education trainings throughout the study program. Our study showed that asynchronous web-based medical education introduces the concept of EBM effectively, provides adequate EBM skill training, and promotes the positive attitudes of Syrian clinical practitioners toward EBM. Further, this web-based course was designed with simple and accessible means, and it was highly feasible and well-received by the participants.

---

### Acknowledgments

The authors would like to thank all the participants in this study. This study was supported by the Syrian Virtual University.

---

### Conflicts of Interest

None declared.

---

#### Multimedia Appendix 1

First lesson in the first learning unit.

[[PPTX File , 2207 KB - formative\\_v6i10e36782\\_app1.pptx](#) ]

---

#### Multimedia Appendix 2

Second lesson in the first learning unit.

[[PPTX File , 768 KB - formative\\_v6i10e36782\\_app2.pptx](#) ]

---

#### Multimedia Appendix 3

Third lesson in the first learning unit.

[[PPTX File , 609 KB - formative\\_v6i10e36782\\_app3.pptx](#) ]

---

#### Multimedia Appendix 4

First lesson in the second learning unit.

[[PPTX File , 483 KB - formative\\_v6i10e36782\\_app4.pptx](#) ]

---

#### Multimedia Appendix 5

Second lesson in the second learning unit.

[[PPTX File , 2450 KB - formative\\_v6i10e36782\\_app5.pptx](#) ]

## Multimedia Appendix 6

Third lesson in the second learning unit.

[[PPTX File , 105 KB - formative\\_v6i10e36782\\_app6.pptx](#) ]

---

## Multimedia Appendix 7

First lesson in the third learning unit.

[[PPTX File , 952 KB - formative\\_v6i10e36782\\_app7.pptx](#) ]

---

## Multimedia Appendix 8

Second lesson in the third learning unit.

[[PPTX File , 675 KB - formative\\_v6i10e36782\\_app8.pptx](#) ]

---

## Multimedia Appendix 9

Third lesson in the third learning unit.

[[PPTX File , 1149 KB - formative\\_v6i10e36782\\_app9.pptx](#) ]

---

## Multimedia Appendix 10

First lesson in the fourth learning unit.

[[PPTX File , 96 KB - formative\\_v6i10e36782\\_app10.pptx](#) ]

---

## Multimedia Appendix 11

Second lesson in the fourth learning unit.

[[PPTX File , 613 KB - formative\\_v6i10e36782\\_app11.pptx](#) ]

---

## References

1. Reddy KR. Evidence Based Medicine: A Paradigm for Clinical Practice. *J Gandaki Med Coll Nepal* 2018 Dec 31;11(02):74-81 [[FREE Full text](#)] [doi: [10.3126/jgmcn.v11i02.22989](https://doi.org/10.3126/jgmcn.v11i02.22989)]
2. Akobeng AK. Principles of evidence based medicine. *Arch Dis Child* 2005 Aug;90(8):837-840 [[FREE Full text](#)] [doi: [10.1136/adc.2005.071761](https://doi.org/10.1136/adc.2005.071761)] [Medline: [16040884](https://pubmed.ncbi.nlm.nih.gov/16040884/)]
3. Karagiannis T. The Importance of Applying Evidence-Based Medicine in Clinical Practice. *Management of Hypertension* 2019;3-17. [doi: [10.1007/978-3-319-92946-0\\_1](https://doi.org/10.1007/978-3-319-92946-0_1)]
4. Gray GE, Pinson LA. Evidence-based medicine and psychiatric practice. *Psychiatr Q* 2003;74(4):387-399. [doi: [10.1023/a:1026091611425](https://doi.org/10.1023/a:1026091611425)] [Medline: [14686461](https://pubmed.ncbi.nlm.nih.gov/14686461/)]
5. Prasad K. Teaching evidence-based medicine in resource-limited countries. *JAMA* 2012 Dec 05;308(21):2248-2249. [doi: [10.1001/jama.2012.74124](https://doi.org/10.1001/jama.2012.74124)] [Medline: [23212503](https://pubmed.ncbi.nlm.nih.gov/23212503/)]
6. Kang H. How to understand and conduct evidence-based medicine. *Korean J Anesthesiol* 2016 Oct;69(5):435-445 [[FREE Full text](#)] [doi: [10.4097/kjae.2016.69.5.435](https://doi.org/10.4097/kjae.2016.69.5.435)] [Medline: [27703623](https://pubmed.ncbi.nlm.nih.gov/27703623/)]
7. Maggio LA, Tannery NH, Chen HC, ten Cate O, O'Brien B. Evidence-based medicine training in undergraduate medical education: a review and critique of the literature published 2006-2011. *Acad Med* 2013 Jul;88(7):1022-1028. [doi: [10.1097/ACM.0b013e3182951959](https://doi.org/10.1097/ACM.0b013e3182951959)] [Medline: [23702528](https://pubmed.ncbi.nlm.nih.gov/23702528/)]
8. Alahdab F, Firwana B, Hasan R, Sonbol MB, Fares M, Alnahhas I, et al. Undergraduate medical students' perceptions, attitudes, and competencies in evidence-based medicine (EBM), and their understanding of EBM reality in Syria. *BMC Res Notes* 2012 Aug 12;5:431 [[FREE Full text](#)] [doi: [10.1186/1756-0500-5-431](https://doi.org/10.1186/1756-0500-5-431)] [Medline: [22882872](https://pubmed.ncbi.nlm.nih.gov/22882872/)]
9. Dashash M, Youssef M. Does the level of knowledge delivered in English in dentistry affect the quality of future dental education. *ICERI 2009 Proceedings*. URL: [https://mdashash.com/images/pup\\_pdf/Nimagehc61a354e2ba3752.35961246.pdf](https://mdashash.com/images/pup_pdf/Nimagehc61a354e2ba3752.35961246.pdf) [accessed 2022-10-17]
10. Khan B. *Managing e-learning: Design, delivery, implementation, and evaluation*. In: Khan B, editor. IGI Global. Pennsylvania: Information Science Publishing; 2005.
11. Davis J, Chryssafidou E, Zamora J, Davies D, Khan K, Coomarasamy A. Computer-based teaching is as good as face to face lecture-based teaching of evidence based medicine: a randomised controlled trial. *BMC Med Educ* 2007 Jul 20;7:23 [[FREE Full text](#)] [doi: [10.1186/1472-6920-7-23](https://doi.org/10.1186/1472-6920-7-23)] [Medline: [17659076](https://pubmed.ncbi.nlm.nih.gov/17659076/)]
12. Hadley J, Kulier R, Zamora J, Coppus SFPJ, Weinbrenner S, Meyerrose B, et al. Effectiveness of an e-learning course in evidence-based medicine for foundation (internship) training. *J R Soc Med* 2010 Jul;103(7):288-294 [[FREE Full text](#)] [doi: [10.1258/jrsm.2010.100036](https://doi.org/10.1258/jrsm.2010.100036)] [Medline: [20522698](https://pubmed.ncbi.nlm.nih.gov/20522698/)]
13. Kulier R, Gülmezoglu AM, Zamora J, Plana MN, Carroli G, Cecatti JG, et al. Effectiveness of a clinically integrated e-learning course in evidence-based medicine for reproductive health training: a randomized trial. *JAMA* 2012 Dec 05;308(21):2218-2225. [doi: [10.1001/jama.2012.33640](https://doi.org/10.1001/jama.2012.33640)] [Medline: [23212499](https://pubmed.ncbi.nlm.nih.gov/23212499/)]

14. Jingyu L. Pros and cons: Web-based education. 2014 Presented at: Proceedings of the 2014 International Conference on Education, Management and Computing Technology; June; Tianjin, China. [doi: [10.2991/icemct-14.2014.111](https://doi.org/10.2991/icemct-14.2014.111)]
15. Dinh L, Nguyen T. Pandemic, social distancing, and social work education: students' satisfaction with online education in Vietnam. *Social Work Education* 2020 Sep 16;39(8):1074-1083 [FREE Full text] [doi: [10.1080/02615479.2020.1823365](https://doi.org/10.1080/02615479.2020.1823365)]
16. Stone-Brown K. Syria: a healthcare system on the brink of collapse. *BMJ* 2013 Dec 10;347:f7375. [doi: [10.1136/bmj.f7375](https://doi.org/10.1136/bmj.f7375)] [Medline: [24327182](https://pubmed.ncbi.nlm.nih.gov/24327182/)]
17. Dashash M. CRISIS Criteria for Effective Continuous Education in Traumatic Dental Injuries During Syrian Crisis. *American Journal of Health Research. Special Issue: Medical Education in Emergency* 2016:1-6. [doi: [10.11648/j.ajhr.s.2016040601.11](https://doi.org/10.11648/j.ajhr.s.2016040601.11)]
18. Sabouni A, Bdaiwi Y, Janoudi SL, Namous LO, Turk T, Alkhatib M, et al. Multiple strategy peer-taught evidence-based medicine course in a poor resource setting. *BMC Med Educ* 2017 May 04;17(1):82 [FREE Full text] [doi: [10.1186/s12909-017-0924-1](https://doi.org/10.1186/s12909-017-0924-1)] [Medline: [28472986](https://pubmed.ncbi.nlm.nih.gov/28472986/)]
19. Hugenholtz NI, Sluiter JK, van Dijk FJ, Nieuwenhuijsen K. EBM E-learning: Feasible and Effective for Occupational Physicians in Different Countries. *Saf Health Work* 2012 Sep;3(3):199-208 [FREE Full text] [doi: [10.5491/SHAW.2012.3.3.199](https://doi.org/10.5491/SHAW.2012.3.3.199)] [Medline: [23019532](https://pubmed.ncbi.nlm.nih.gov/23019532/)]
20. Ayoub F, Moussa M, Papatsoris A, Chakra M, Chahine N, Fares Y. The online learning in medical education: A novel challenge in the era of COVID-19 pandemic. *Hellenic Urology* 2020;32(2):89. [doi: [10.4103/huaj.huaj\\_9\\_20](https://doi.org/10.4103/huaj.huaj_9_20)]
21. Kouba L, Amin B, Azzam A. Online education opportunity for Syria's future doctors. *Lancet* 2019 Nov 16;394(10211):1805-1806. [doi: [10.1016/S0140-6736\(19\)31899-9](https://doi.org/10.1016/S0140-6736(19)31899-9)] [Medline: [31679944](https://pubmed.ncbi.nlm.nih.gov/31679944/)]
22. Hamtini TM. Evaluating E-learning Programs: An Adaptation of Kirkpatrick's Model to Accommodate E-learning Environments. *J Comp Sci* 2008 Aug 1;4(8):693-698. [doi: [10.3844/jcssp.2008.693.698](https://doi.org/10.3844/jcssp.2008.693.698)]
23. Coomarasamy A, Khan KS. What is the evidence that postgraduate teaching in evidence based medicine changes anything? A systematic review. *BMJ* 2004 Oct 30;329(7473):1017 [FREE Full text] [doi: [10.1136/bmj.329.7473.1017](https://doi.org/10.1136/bmj.329.7473.1017)] [Medline: [15514348](https://pubmed.ncbi.nlm.nih.gov/15514348/)]
24. Rajab MH, Gazal AM, Alkattan K. Challenges to Online Medical Education During the COVID-19 Pandemic. *Cureus* 2020 Jul 02;12(7):e8966 [FREE Full text] [doi: [10.7759/cureus.8966](https://doi.org/10.7759/cureus.8966)] [Medline: [32766008](https://pubmed.ncbi.nlm.nih.gov/32766008/)]
25. Deepika V, Soundariya K, Karthikeyan K, Kalaiselvan G. 'Learning from home': role of e-learning methodologies and tools during novel coronavirus pandemic outbreak. *Postgrad Med J* 2021 Sep;97(1151):590-597. [doi: [10.1136/postgradmedj-2020-137989](https://doi.org/10.1136/postgradmedj-2020-137989)] [Medline: [33154099](https://pubmed.ncbi.nlm.nih.gov/33154099/)]
26. Kulier R, Coppus SFPJ, Zamora J, Hadley J, Malick S, Das K, et al. The effectiveness of a clinically integrated e-learning course in evidence-based medicine: a cluster randomised controlled trial. *BMC Med Educ* 2009 May 12;9:21 [FREE Full text] [doi: [10.1186/1472-6920-9-21](https://doi.org/10.1186/1472-6920-9-21)] [Medline: [19435520](https://pubmed.ncbi.nlm.nih.gov/19435520/)]
27. Rohwer A, Young T, van Schalkwyk S. Effective or just practical? An evaluation of an online postgraduate module on evidence-based medicine (EBM). *BMC Med Educ* 2013 May 27;13:77 [FREE Full text] [doi: [10.1186/1472-6920-13-77](https://doi.org/10.1186/1472-6920-13-77)] [Medline: [23710548](https://pubmed.ncbi.nlm.nih.gov/23710548/)]

## Abbreviations

**EBM:** evidence-based medicine

**PICO:** Patient, Intervention, Comparison, and Outcome

*Edited by A Mavragani; submitted 25.01.22; peer-reviewed by B Foroutan, O Las Vergnas, E Toki, S Ganesh; comments to author 18.07.22; revised version received 13.08.22; accepted 22.08.22; published 25.10.22.*

*Please cite as:*

Kenjrawi Y, Dashash M

*The First Asynchronous Online Evidence-Based Medicine Course for Syrian Health Workforce: Effectiveness and Feasibility Pilot Study*

*JMIR Form Res* 2022;6(10):e36782

URL: <https://formative.jmir.org/2022/10/e36782>

doi: [10.2196/36782](https://doi.org/10.2196/36782)

PMID: [36282556](https://pubmed.ncbi.nlm.nih.gov/36282556/)

©Yazan Kenjrawi, Mayssoon Dashash. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 25.10.2022. 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

# Mobile App to Enhance Patient Activation and Patient-Provider Communication in Major Depressive Disorder Management: Collaborative, Randomized Controlled Pilot Study

Maggie McCue<sup>1</sup>, RD, MS; Christopher Blair<sup>1</sup>, MS; Ben Fehnert<sup>2,3</sup>, MA; James King<sup>2,3</sup>, BSc; Francesca Cormack<sup>2,4</sup>, PhD; Sara Sarkey<sup>1</sup>, PhD; Anna Eramo<sup>5</sup>, MD; Christopher Kabir<sup>6</sup>, MS; Rasha Khatib<sup>6</sup>, PhD; David Kemp<sup>7</sup>, MS, MD

<sup>1</sup>Takeda Pharmaceuticals USA, Inc, Lexington, MA, United States

<sup>2</sup>Cognition Kit, Cambridge, United Kingdom

<sup>3</sup>Ctrl Group, London, United Kingdom

<sup>4</sup>Cambridge Cognition, Cambridge, United Kingdom

<sup>5</sup>Lundbeck LLC, Deerfield, IL, United States

<sup>6</sup>Advocate Research Institute, Advocate Health Care, Downers Grove, IL, United States

<sup>7</sup>Advocate Aurora Health, Downers Grove, IL, United States

**Corresponding Author:**

Maggie McCue, RD, MS

Takeda Pharmaceuticals USA, Inc

95 Hayden Ave

Lexington, MA, 02421

United States

Phone: 1 224 554 2820

Email: [Maggie.McCue@takeda.com](mailto:Maggie.McCue@takeda.com)

## Abstract

**Background:** Enhanced patient-provider engagement can improve patient health outcomes in chronic conditions, including major depressive disorder (MDD).

**Objective:** We evaluated the impact of a digitally enabled care mobile app, Pathway, designed to improve MDD patient-provider engagement. Patients used a mobile interface to assess treatment progress and share this information with primary care providers (PCPs).

**Methods:** In this 52-week, real-world effectiveness and feasibility study conducted in primary care clinics, 40 patients with MDD who were recently prescribed antidepressant monotherapy were randomized to use a mobile app with usual care (20/40, 50%) or usual care alone (20/40, 50%). Patients in the app arm engaged with the app daily for 18 weeks; a report was generated at 6-week intervals and shared with the PCPs to facilitate shared treatment decision-making discussions. The patients discontinued the app at week 18 and were followed through year 1. Coprimary outcome measures, assessed via research visits, included change from baseline in the 13-item Patient Activation Measure (PAM-13) and 7-item Patient-Provider Engagement Scale scores at week 18. Additional outcome measures included depression severity (9-item Patient Health Questionnaire [PHQ-9]) and cognitive symptoms (5-item Perceived Deficits Questionnaire–Depression).

**Results:** All 37 patients (app arm: n=18, 49%; usual care arm: n=19, 51%) who completed the 18-week follow-up period (n=31, 84% female, mean age 36, SD 11.3 years) had moderate to moderately severe depression. Improvements in PAM-13 and PHQ-9 scores were observed in both arms. Increases in PAM-13 scores from baseline to 18 weeks were numerically greater in the app arm than in the usual care arm (mean 10.5, SD 13.2 vs mean 8.8, SD 9.4;  $P=.65$ ). At 52 weeks, differences in PAM-13 scores from baseline demonstrated significantly greater improvements in the app arm than in the usual care arm (mean 20.2, SD 17.7 vs mean 1.6, SD 14.2;  $P=.04$ ). Compared with baseline, PHQ-9 scores decreased in both the app arm and the usual care arm at 18 weeks (mean 7.8, SD 7.2 vs mean 7.0, SD 6.5;  $P=.73$ ) and 52 weeks (mean 9.5, SD 4.0 vs mean 4.7, SD 6.0;  $P=.07$ ). Improvements in 7-item Patient-Provider Engagement Scale and WHO-5 scores were observed in both arms at 18 weeks and were sustained through 52 weeks in the app arm. Improvements in WHO-5 scores at 52 weeks were significantly greater in the app arm than in the usual care arm (41.5 vs 20.0;  $P=.02$ ).



**Conclusions:** Patients with MDD will engage with a mobile app designed to track treatment and disease progression. PCPs will use the data generated as part of their assessment to inform clinical care. The study results suggest that an app-enabled clinical care pathway may enhance patient activation and benefit MDD management.

**Trial Registration:** ClinicalTrials.gov NCT03242213; <https://clinicaltrials.gov/ct2/show/NCT03242213>

(*JMIR Form Res* 2022;6(10):e34923) doi:[10.2196/34923](https://doi.org/10.2196/34923)

## KEYWORDS

depression; major depressive disorder; depression management; patient engagement; patient satisfaction; mobile app; patient-reported outcomes; mobile phone

## Introduction

### Background

Major depressive disorder (MDD) is among the most common mental health disorders in the United States, with a lifetime prevalence of up to 20% [1]. Although MDD is a chronic and recurrent disorder that frequently requires long-term treatment with antidepressants [2], the rates of nonadherence and premature discontinuation of antidepressant therapy are high and associated with worse clinical outcomes [3-5]. In addition, the impact of nonadherence on clinical outcomes subsequently translates to increased medical and total health care use [3]. In the United States and across the world, most patients with depression are treated by primary care providers (PCPs) [6,7]. Consequently, nearly 10% of all primary care visits are related to depression [8]. Given the recent decline in the number of psychiatrists practicing in the United States, the demand for MDD care in primary care likely will continue to surge [9].

Effective management of MDD in primary care requires a systematic approach to diagnosis, patient education, treatment, close follow-up, and a commitment to adjusting care or consulting with specialists when needed [8,10]. One systematic approach to improving MDD outcomes is measurement-based care, which involves the use of rating scales to monitor symptoms, adherence, and side effects combined with guideline-dependent resources to inform treatment [11]. However, increasing time constraints and the infrequency of visits in primary care settings may make it difficult for clinicians to practice measurement-based care and fully engage with patients with MDD [7,11]. Consequently, patients with MDD and other chronic conditions may benefit from additional support and increased engagement with their PCPs [8,12].

One proposed strategy to increase patient engagement, improve patient-provider communication to facilitate MDD management, and improve patient outcomes in primary care is the use of mobile health apps and digital platforms [13,14]. To improve patient-provider engagement, apps need to be easy to use and easily integrated into the workflow of traditional clinical care and care teams [15]. Currently available apps supporting MDD management are heterogeneous in features and quality, possibly because of the absence of standards governing their development, evaluation, and use [16]. Moreover, the number of apps developed for depression exceeds the number of studies that have demonstrated their efficacy and feasibility. Although few apps have the ability to transmit data directly to PCPs, apps that feature the active involvement of mental health

professionals may also increase patient engagement more than presentation enhancements of the technology platform [17]. Digital tools that can share data with PCPs may enable PCPs to more easily and efficiently embrace measurement-based care.

The development of any mobile health and information technology tool, including those supporting MDD management, may benefit from collaboration between industry, app developers, and the health care team representing large health care systems [18]. Research has demonstrated that involving target users and stakeholders in the development of such tools yields a higher acceptance of apps by clinicians. To help meet the needs of PCPs and patients with MDD and improve patient-provider engagement, Takeda, Lundbeck, and Advocate Aurora Health (AAH) worked together with software developers (Ctrl Group, Fora Health, and Cognition Kit) to develop the patient app and care team view of the patient data. The process followed user-centered design principles, in which the designs and development were iterated based on user input and feedback. The app enables patients to track their symptoms, monitor their treatment progress, and share data collected by the app with their care team.

### Objectives

The primary objective of this study was to determine whether the addition of the Pathway mobile app to usual clinical care improves patient-provider engagement in the management of MDD over an 18-week period. The secondary objectives of the study were to evaluate the impact of the Pathway mobile app on changes in certain patient-reported outcomes, including self-reported clinical depression severity (via the 9-item Patient Health Questionnaire [PHQ-9]), cognitive dysfunction (via the 5-item Perceived Deficits Questionnaire–Depression [PDQ-D5]), quality of life (via the 5-item World Health Organization Well-Being Index [WHO-5]), and patient satisfaction, as well as changes in medication and medication adherence. Additional retrospective assessments were also planned to evaluate the impact of the app on measures of health care use 1 year after enrollment.

## Methods

### Study Design

In this randomized controlled pilot study (ClinicalTrials.gov NCT03242213), we enrolled patients diagnosed with MDD who were receiving primary care services at Advocate Health Care, which is now part of AAH. The study took place between July 2017 and January 2019 and involved 4 study sites in

suburban and urban settings within metropolitan Chicago, Illinois, including Advocate Medical Group–Huntley, Advocate Medical Group–Hometown Family Medicine, Family Medicine Center in Ravenswood, and Adult Medicine Center in Oak Lawn. On identifying a patient with MDD who met the criteria for study participation, the physician providing care for that patient introduced the study; a designated research study coordinator then explained the study and obtained informed consent. Patients were then randomized to receive either the Pathway mobile app along with usual care or usual care alone. Participating patients were randomized based on the results from a randomized study list created using serially generated random numbers obtained by the study staff using a random number generator. Patients in the mobile app arm were encouraged to engage with the mobile app daily for 18 weeks (Figure 1). At week 18, the use of the mobile app was discontinued. Patients in the usual care arm did not receive study-related interventions.

The mobile app was specifically designed to enhance patient-provider engagement, promote shared decision-making, and support measurement-based care in the management of clinical depression. It also provided patients with a way to track changes in clinical depression severity, cognitive symptoms, quality of life, emotional well-being, and adherence to medications; set up medication reminders; and record side-effect experiences. App functions included PHQ-9 and PDQ-D5 assessments conducted every 2 weeks, daily assessments of depression using 2 questions from the PHQ-9 and 1 question

from the PDQ-D5, daily assessments of emotional well-being using a visual analog measurement of global well-being on a scale of 0 to 100, and daily cognitive symptoms assessed using the Cognition Kit 2-back test, in which patients indicate whether the symbol (usually an abstract shape) matches the symbol 2 items back [19]. Patients reporting any change in suicidal ideation during PHQ-9 assessments were instructed to contact their health care provider or emergency services immediately because data from the app were not monitored; these instructions were reviewed with patients during the consent process. In addition, patients were able to review a graphical summary of their data (Figure 2), allowing them to review their symptom progression and the treatment’s effectiveness and side effects. A graphical summary of the data was also shared with the care team every 6 weeks to reinforce measurement-based care. No further instructions were provided to the care team.

The single-phase study included 2 follow-up periods: the primary follow-up period and the long-term follow-up period. The primary follow-up period began with randomization and continued through week 18. The long-term follow-up period began after the final visit at week 18 and continued through the 1-year follow-up phone interview (34 weeks after the use of the mobile app was discontinued). At year 1, the results from the follow-up phone interview and extraction of data from patients’ electronic medical records (EMRs) were analyzed to evaluate the residual impact of the app on patient-reported outcomes and health care use.

**Figure 1.** Study design. Long-term follow-up period indicates that no visits, calls, or use of the mobile app occurred during this phase in either treatment arm. The app arm also included usual care. ED: emergency department; LOS: length of stay; PAM-13: 13-item Patient Activation Measure; PDQ-D5: 5-item Perceived Deficits Questionnaire–Depression; PHQ-9: 9-item Patient Health Questionnaire; PPES-7: 7-item Patient-Provider Engagement Scale; WHO-5: 5-item World Health Organization Well-Being Index.

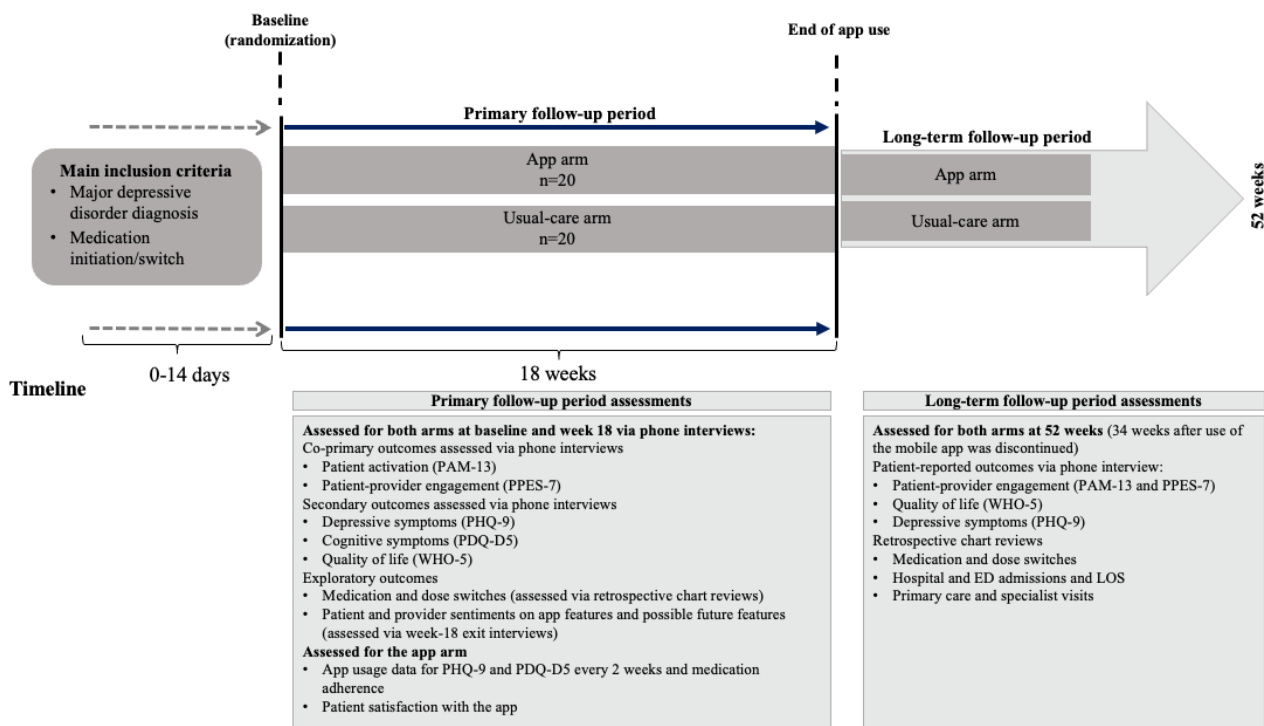
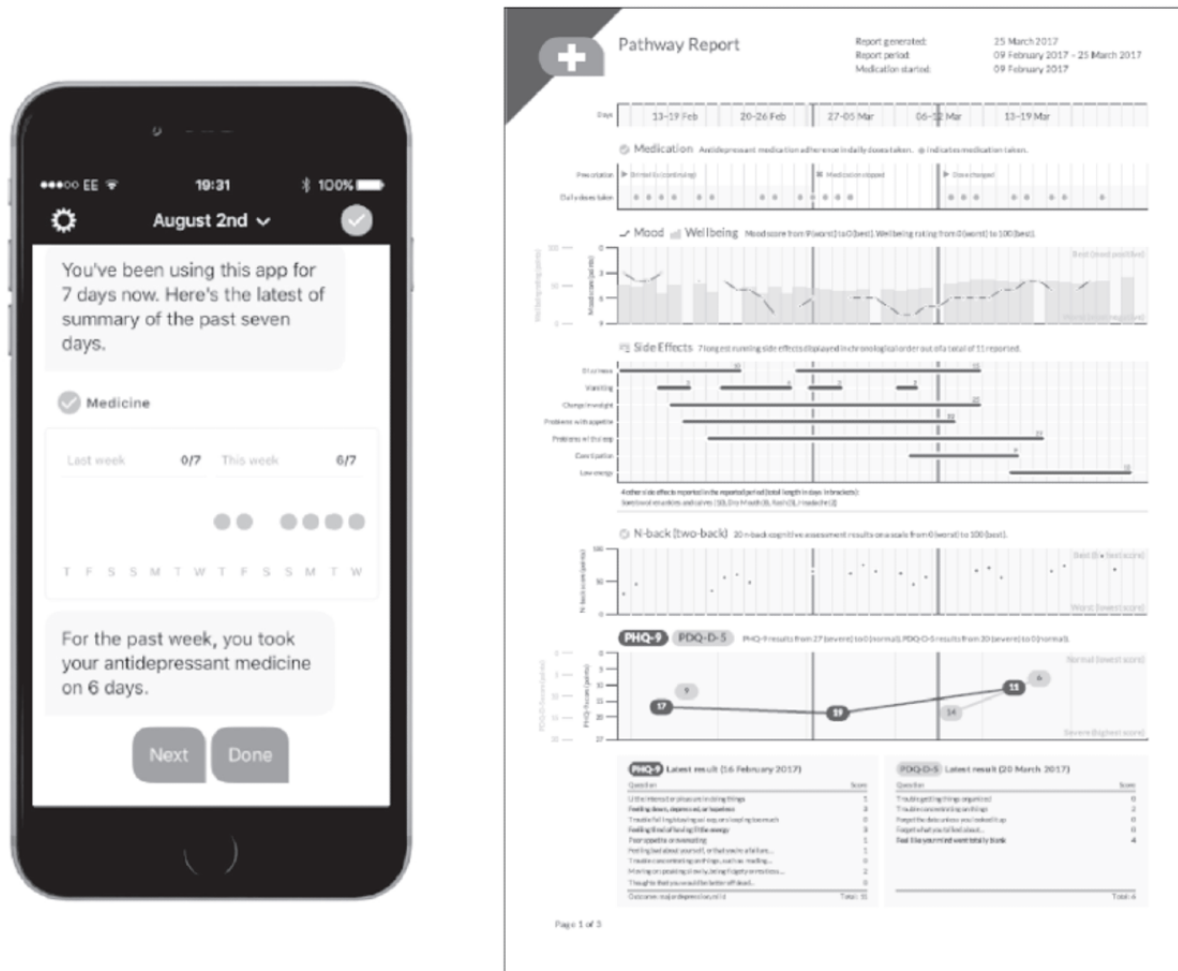


Figure 2. Screenshots of the app and a sample report.



**Participants**

Patients were eligible to participate in the study if they were aged 18 to 70 years, were receiving primary care for MDD at AAH, used an iPhone version 5 or later or a smartphone with an Android operating system, and had an active cellular data plan or regular Wi-Fi access. Patients were also required to have a PHQ-9 score of >5 at baseline, indicating at least mild depression severity [20], and to have initiated monotherapy with a new antidepressant (either a new prescription or a switch from another antidepressant medication) in the previous 0 to 14 days.

Patients were excluded from the study if they had a diagnosis of a major psychiatric disorder other than MDD, were considered to be at imminent risk of hospitalization due to MDD, had been hospitalized due to MDD within 3 months, had a significant risk of suicide or had a previous suicide attempt within 6 months, or had a history of responding only to combination or augmentation therapy in their current depressive episode. Patients for whom the use of antidepressants was contraindicated were not eligible for the study.

**Study Procedures**

Eligible patients were randomized to usual care (20/40, 50%) or usual care plus the mobile app (20/40, 50%) for 18 weeks (unblinded for both participants and researchers). An in-person

introduction to the app and instructional handouts were provided to the patients at the time of enrollment. Although the use of the mobile app was encouraged, it was not required. Patients assigned to usual care received regular care as needed from their PCP; no specific interventions were mandated. At the end of the long-term follow-up period (1 year), patient-reported outcome measures were collected via follow-up phone calls. Data on health care resource use and medication changes during the long-term follow-up period were collected from medical chart reviews. A quality control committee reviewed the data for adequate completion and integrity.

**Study End Points**

Study coprimary, secondary, and exploratory outcomes were assessed at 18 weeks for each arm via in-person research visits and phone interviews. Coprimary outcomes included changes in the 13-item Patient Activation Measure (PAM-13) scale and 7-item Patient Provider Engagement Scale (PPES-7) between baseline and 18 weeks. The PAM-13 scale was developed and validated to assess patient engagement and confidence in self-management of the disease [21,22]. The PAM-13 scores ranged from 0 to 100, with higher values reflecting greater activation. The PPES-7 is an assessment that was developed for this study and has not yet been validated. In the PPES-7, scores range from 7 to 28, with higher values reflecting more engagement.

The secondary outcomes included changes (between baseline and 18 weeks) in depression severity evaluated using the PHQ-9 measure [20], cognitive symptoms measured using the PDQ-D5 scale [23], and quality of life measured using the WHO-5 assessment, a 5-item questionnaire that measures the subjective quality of life. In the WHO-5, scores range from 0 to 100, with higher scores reflecting the best imaginable quality of life [24].

Exploratory outcomes included medication and dose switches, defined as medication switches, dosage changes, medication add-ons, or discontinuations (assessed via retrospective chart review), and patient and provider satisfaction with the care and use of the app. Additional assessments were collected throughout the study in the app arm via app functions such as the PHQ-9, PDQ-D5, emotional well-being, and cognitive symptom assessments. In addition, information on patient satisfaction with the app and app use data were collected from the patients in the app arm.

After the 18-week end-of-study visit, patients and providers were invited to participate in a remote, qualitative, semistructured interview using a digital tool to discuss sentiments on app features and future features. This qualitative tool allowed the interviewer to observe and talk to the patients as they looked at the app's features on their own devices, record the interview, and capture time-stamped notes.

At the end of the long-term follow-up phase of the study (at year 1), a phone interview was conducted (34 weeks after the use of the mobile app was discontinued) to assess patient-reported outcomes, including patient and provider engagement (using PAM-13 and PPES-7), quality of life (using WHO-5), and depressive symptoms (using PHQ-9). At this time, a retrospective analysis that compared health care use between the app arm and the usual care arm was also conducted. Data were collected on inpatient visits, including depression-related hospitalizations; emergency department (ED) visits; outpatient visits, including visits to PCPs, psychiatrists, behavioral therapy specialists, and other health care providers; and medication and dose switches. Spontaneously reported serious adverse events were also recorded during the study period.

### Statistical Analyses

This was a pilot study, and thus no sample size estimation was conducted. A sample size of 20 patients per group was expected to be sufficient to provide initial information about the potential

effects and benefits of the app and the feasibility of its real-world use to inform future larger-scale studies. Patients were included in the analysis based on treatment allocation, and an intent-to-treat analysis was used. As 18-week follow-up data were not available for 3 randomized patients, these patients were dropped from the analysis, and an intent-to-treat analysis (with the exclusion of missing data) was conducted on the remaining population. For the primary and long-term follow-up periods, between-group differences in changes in continuous variables were evaluated using the 2-tailed Student *t* test or Mann-Whitney *U* test. Changes in categorical variables were compared using Fisher exact test or Pearson chi-square test. Two-tailed tests using a significance threshold of  $P < .05$  were performed.

A retrospective comparison of health care use was conducted between patients in the usual care arm and those in the mobile app arm using medical record extraction. This comparison included data from the time of consent to 1 year after enrollment for each patient. Health care use was compared between groups for long-term differences at 1 year after each patient's study enrollment, overall, by cause (depression-related or not) and by category (inpatient via ED, outpatient, or specialty). Sensitivity analysis was used to assess attrition bias among patients lost to follow-up in the long-term follow-up period; these outcomes were used to assess the generalizability of the 52-week results across the original study group. Statistical analysis was performed using SAS software (version 9.4; SAS Institute).

### Ethics Approval

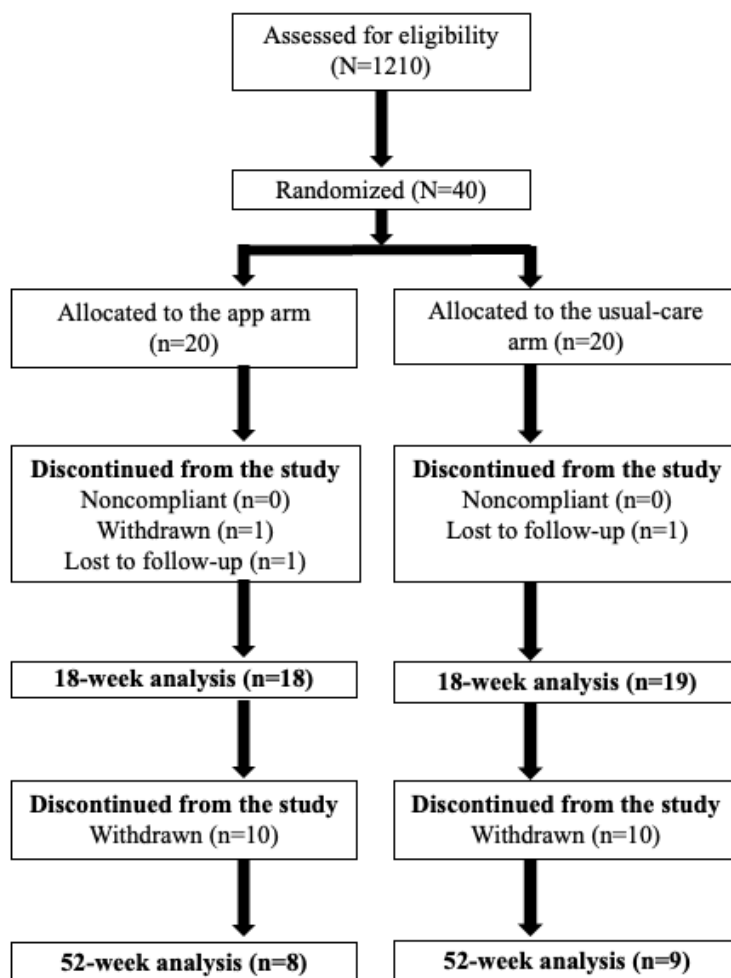
This study (#Y5000249) was approved by the Advocate Health Care Institutional Review Board.

## Results

### Patient Disposition

A total of 40 patients were enrolled, and of them, 37 (93%) completed the 18-week primary follow-up period (Figure 3) and were included in the main analysis based on treatment allocation. In the app arm, 18 (90%) patients completed the primary follow-up period, 1 (5%) withdrew, and 1 (5%) was lost to follow-up. In the usual care arm, 19 (95%) patients completed the primary follow-up period and 1 (5%) was lost to follow-up. At year 1 (the long-term follow-up phase), data were available for 43% (17/40) of patients, including 8 (47%) patients in the app arm and 9 (53%) in the usual care arm.

Figure 3. CONSORT (Consolidated Standards of Reporting Trials) diagram.



### Baseline Characteristics

The baseline characteristics of patients who completed 18 weeks of the study are shown in Table 1. Demographic categories were well represented and balanced between each treatment arm, including race (Hispanic, non-Hispanic Black, and non-Hispanic White), geographic area of residence (rural, urban, and suburban), income range, work type, marital status (single, married, or living as a couple; widowed; divorced; or separated), and type of health insurance. The mean age was 33.8 years in the app arm and 38.9 (SD 11.0) years in the usual care arm; 84% (31/37) were women, and 16% (6/37) were men. The mean PHQ-9 score at baseline was 15.3 in the app arm and 14.1 in

the usual care arm, indicating moderate to moderately severe depression. Although certain socioeconomic characteristics (eg, education level and annual income) differed between the groups, the groups were similar in terms of MDD severity and treatment history. The baseline demographic characteristics of the 17 patients who completed the long-term follow-up period (the 52-week completers) were also similar between the groups (Table S1 in Multimedia Appendix 1). When the 52-week completers were compared with noncompleters with regard to baseline demographics or patient-reported outcomes (PHQ-9, PAM-13, PPES-7, and WHO-5), no statistically significant differences were observed.

**Table 1.** Demographics and baseline characteristics for patients completing 18 weeks.

	App (n=18)	Usual care (n=19)	Total (N=37)
Sex (female), n (%)	14 (78)	17 (90)	31 (84)
Age (years), mean (SD)	33.8 (11.4)	38.9 (11.0)	36.4 (11.3)
Aged ≥45 years, n (%)	2 (11)	5 (26)	7 (19)
<b>Race or ethnicity, n (%)</b>			
Hispanic	8 (44)	6 (32)	14 (38)
Non-Hispanic Black	2 (11)	5 (26)	7 (19)
Non-Hispanic White	7 (39)	8 (42)	15 (41)
Non-Hispanic multiracial	1 (6)	0 (0)	1 (3)
<b>Employment status, n (%)</b>			
Employed full-time	6 (33)	10 (53)	16 (43)
Employed part-time	5 (28)	3 (16)	8 (22)
Self-employed	1 (6)	2 (11)	3 (8)
Not employed	3 (17)	2 (11)	5 (14)
Student	3 (17)	2 (11)	5 (14)
Nonworking spouse, retired, or other	2 (11)	2 (11)	4 (11)
Annual income <US \$40,000, n (%)	13 (72)	8 (42)	21 (57)
Education (associate's degree or higher), n (%)	3 (17)	12 (63)	15 (41)
<b>Number of non-MDD<sup>a</sup>-related medications currently taken, n (%)</b>			
0	6 (33)	7 (37)	13 (35)
1-3	8 (44)	9 (47)	17 (46)
≥4	4 (22)	3 (16)	7 (19)
PHQ-9 <sup>b</sup> , mean (SD)	15.3 (5.1)	14.1 (5.0)	14.7 (5.0)
<b>Antidepressant use at baseline, n (%)</b>			
SSRIs <sup>c</sup>	13 (72)	15 (79)	28 (76)
Bupropion	2 (11)	3 (16)	5 (14)
SNRIs <sup>d</sup>	2 (11)	1 (5)	3 (8)
TCAs <sup>e</sup> , MAOIs <sup>f</sup> , SMSs <sup>g</sup> , or other	1 (6)	0 (0)	1 (3)
<b>Years on antidepressants, mean (SD)</b>			
None	10 (56)	11 (58)	21 (57)
<1	2 (11)	2 (11)	4 (11)
≥1	6 (33)	4 (21)	10 (27)
Unknown	0 (0)	2 (11)	2 (5)

<sup>a</sup>MDD: major depressive disorder.

<sup>b</sup>PHQ-9: 9-item Patient Health Questionnaire.

<sup>c</sup>SSRI: selective serotonin reuptake inhibitor.

<sup>d</sup>SNRI: serotonin and norepinephrine reuptake inhibitor.

<sup>e</sup>TCA: tricyclic antidepressant.

<sup>f</sup>MAOI: monoamine oxidase inhibitor.

<sup>g</sup>SMS: serotonin modulator and stimulator.

## Primary Follow-up Results (18-Week Analysis)

### Coprimary Outcomes

At week 18, both arms exhibited an increase in patient activation based on the PAM-13 scores (Table S2 in [Multimedia Appendix 1](#)), with greater improvement in the app arm than in the usual care arm although this difference was not statistically significant (mean change from baseline 10.5, SD 13.2 vs 8.8, SD 9.4;  $P=.65$ ). Patient-provider engagement improved in both arms based on changes in PPES-7 (Table S2 in [Multimedia Appendix 1](#)), with greater improvement in the usual care arm than in the app arm, although again, this difference was not statistically significant (mean change from baseline 1.7, SD 2.7 vs 0.6, SD 3.1;  $P=.27$ ).

### Secondary Outcomes

Depression severity (as measured by the PHQ-9 score) decreased in both arms (Table S2 in [Multimedia Appendix 1](#)). Although the decrease was greater in the app arm (mean change from baseline  $-7.8$ , SD  $7.2$  vs  $-7.0$ , SD  $6.5$ ), no statistically significant differences were observed between the groups ( $P=.73$ ). In addition, no significant differences in the rates of depression response or remission between groups were reported. Response, defined as a  $\geq 50\%$  decrease in the PHQ-9 score from baseline, was achieved in 56% (10/18) of the patients in the app arm and 58% (11/19) in the usual care arm. Remission, defined as a PHQ-9 score of  $\leq 5$ , was achieved by 39% (7/18) in the app arm and 44% (8/18) in the usual care arm. Cognitive symptoms

(PDQ-D5) and quality of life (WHO-5) improved in both the app arm and the usual care arm (mean change from baseline for PDQ-D5 was  $-2.6$ , SD  $5.6$  vs  $-5.5$ , SD  $4.3$ , respectively;  $P=.08$ ; mean change from baseline for WHO-5 was  $31.8$ , SD  $19.7$  vs  $30.7$ , SD  $23.4$ , respectively;  $P=.88$ ).

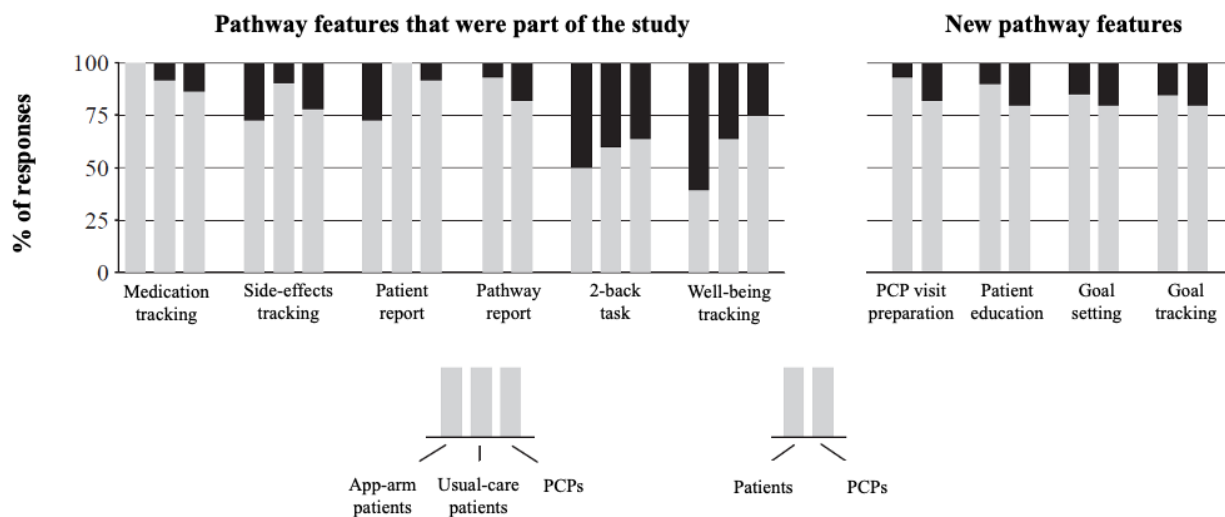
### Exploratory Outcomes

A total of 11% (2/18) of patients in the app arm and 0% (0/18) of patients in the usual care arm switched medications during the study. One serious adverse event (inpatient hospitalization related to depression) was reported in the app arm.

All patients randomized to the app arm (20/20, 100%) completed at least one app assessment during the study period. A majority of patients (12/20, 60%) completed the PHQ-9 and PDQ-D5 assessments biweekly for at least 12 weeks. A total of 70% (14/20) of the app users completed the self-report of medication assessment daily for  $>100$  days.

Patient satisfaction in the app arm was high, as shown in the results of the patient satisfaction survey administered at 18 weeks (Table S3 in [Multimedia Appendix 1](#)). In remote interviews,  $>70\%$  of the patients and PCPs provided positive feedback on most of the app's features and potential new features, including its ability to track medication use and side effects and provide reports (Figure 4). The 2-back task and well-being tracking features scored the lowest, highlighting the need to update these features during the next iteration.

**Figure 4.** Positive and negative patient and PCP sentiments on Pathway app features (at 18 weeks). Responses were classified as follows: Gray=positive responses: app features were well received; black=negative responses: app features need further work. Data from remote interviews with providers (n=15) and patients (n=20). PCP: primary care provider.



## Long-term Follow-up Results (1-Year Analysis)

### Patient-Reported Outcomes

At year 1, a significant increase in patient activation (PAM-13) was observed for patients in the app arm (Table S2 in [Multimedia Appendix 1](#)), with a greater improvement in the app arm than in the usual care arm (mean change from baseline 20.2, SD 17.7 vs 1.6, SD 14.2;  $P=.04$ ). The quality of life (WHO-5) improved in both arms (mean change from baseline for app vs usual care 41.5, SD 12.3 vs 20.0, SD 19.5), with a

significantly greater improvement observed in the app arm ( $P=.02$ ; Table S2 [Multimedia Appendix 1](#)). Patients in the app arm experienced improvements in patient-provider engagement as assessed by the PPES-7 although the improvement was not significantly different from that observed in the usual care arm (mean change from baseline for app vs usual care 1.5, SD 2.6 vs 0.1, SD 3.1;  $P=.33$ ). Depression severity (as measured by the PHQ-9 score) decreased in both arms at year 1 (mean change from baseline for app arm vs usual care arm was  $-9.5$ , SD  $4.0$  vs  $-4.7$ , SD  $6.0$ ;  $P=.07$ ). The longitudinal patient-reported

outcomes among patients who completed the 52-week trial are shown in Table S4 in [Multimedia Appendix 1](#).

### Outcomes From Medical Chart Review

Among the 17 patients assessed during the 1-year follow-up period, there were no inpatient hospitalizations. One patient in the usual care arm visited the ED twice; however, neither visit was considered related to depression. Patients in the usual care arm ( $n=9$ ) had more outpatient clinic visits to any provider than those in the app arm ( $n=8$ ; 88 visits vs 49 visits), including visits to PCPs (59 visits vs 38 visits). At 1-year follow-up, 3 patients in each group had a medication change: 1 patient in the mobile app arm and 2 in the usual care arm switched medications; 1 patient in each group had a medication dose change; and 1 patient in the mobile app arm added a new medication to their regimen.

## Discussion

### Principal Findings

The results of this pilot study suggest that the Pathway mobile app may facilitate the systematic use of measurement-based care in MDD management, which can enhance shared decision-making and patient-provider communication, with improved medication adherence and treatment outcomes [10]. The small sample size of this study prevents interpretation based on individual characteristics; however, a larger implementation study is underway (NCT04891224) [25]. Patients in the mobile app arm exhibited a greater change in patient activation (PAM-13) from baseline, with a 10.5-point increase over 18 weeks; this was numerically better than that observed for the usual care arm, although these differences were not statistically significant. The PAM-13 is scored on a scale from 0 to 100, and PAM-13 results are categorized into 4 levels of patient activation: level 1 (0-47.0), which suggests that patients may not yet understand that their role is important; level 2 (47.1-55.1), which indicates that patients lack the confidence or knowledge to take action; level 3 (55.2-72.4), which suggests that patients are beginning to engage in recommended health behaviors; and level 4 (72.5-100), which indicates that patients are proactive about health and engage in many recommended health behaviors [26].

In this study, the mean PAM-13 scores suggested that patients in the app arm, on average, moved from PAM-13 level 3 (58.2) to level 4 (74.2) at the end of the primary follow-up period, whereas patients in the usual care group, on average, started and remained at level 3. A cross-sectional study of patients visiting a primary care clinic reported that every 10-point increase in the PAM-13 score was associated with a 1% reduction in the predicted probability of having an ED visit or being obese and a 1% increase in the predicted probability of having clinical indicators (eg, hemoglobin A<sub>1c</sub>) in the normal range [22]. Taken together, these results may suggest that patients who used the app have a slightly greater likelihood of engaging in proactive health behaviors. In addition, the increase in patient-provider engagement (as measured by the PPES-7 score) noted for both arms may further increase this likelihood by helping patients and clinicians make better care decisions,

thus improving the ability of patients to effectively manage their own care [27].

The severity of depression, quality of life, and subjective cognitive symptoms improved in both arms, with no statistically significant differences between groups from baseline to week 18. Although not statistically significant, a trend toward improvement in depressive symptoms (PHQ-9) was observed among the patients in the app arm. Patient-provider engagement also showed small improvements in both arms. These results suggest that a larger study is warranted to determine whether the use of the app is associated with a clinically meaningful improvement in the symptoms of depression or patient-provider engagement.

At year 1, greater improvement in patient activation (PAM-13;  $P=.04$ ) and quality of life (WHO-5;  $P=.02$ ) was observed for patients in the mobile app arm than in the usual care arm, indicating that the app was associated with a long-term impact on patient activation that was sustained for at least 34 weeks after the app was discontinued. Small improvements were observed in both treatment arms with regard to patient-provider engagement (PPES-7) from baseline to week 52. It is possible that the reported improvement at year 1 among patients who completed the 1-year trial may be influenced by attrition bias. However, patients who were lost to follow-up were similar in most baseline characteristics to those who completed the study although they did have higher PHQ-9 scores at baseline and were more likely to have received education beyond high school. The similarity between the populations that completed the study and those that were lost to follow-up suggests that the year 1 results may be generalizable to the original study population.

Moreover, although the overall number of medication changes was similar in both groups at 52 weeks, 2 switches occurred in the app group before the week 18 assessment, with no observed switches in the usual care group in that time frame. In addition, the number of outpatient visits (overall and PCP visits) was greater for patients in the usual care arm than for those in the app arm. These examples may suggest that the app, through improved patient-provider communication, allowed for a more rapid response to changes in patient status while reducing the burden of in-person office visits.

Several systematic reviews and meta-analyses have shown that the use of digital mental health interventions such as apps can aid in the reduction of depressive symptoms, with larger effects seen in patients with more severe symptoms [28,29]. In addition, a randomized clinical trial of a mobile intervention app-based platform, IntelliCare, in primary care patients positive for depression or anxiety demonstrated a greater reduction in symptoms, with sustained changes over a 2-month follow-up period compared with participants in the control arm [30]. Furthermore, a pilot study in 23 women with postpartum depression demonstrated that the enhancement of clinical care with ecological momentary assessment using a wearable device to track daily symptoms, depression, anxiety, and maternal functioning was found to be clinically useful by both study participants and the study clinician [31]. In this study, the digitally enabled care pathway showed sustained effects for up to 1 year (34 weeks after the mobile app intervention was



discontinued) on patient-provider engagement, clinical symptoms, quality of life, and resource use. These results confirm the findings of previous studies that demonstrated that app use in patients with MDD and other chronic conditions can have a positive effect on patient adherence, symptoms of depression and anxiety, and patient engagement with therapeutic interventions [13,28-33]. In summary, digital mental health interventions to support clinical decisions and empower shared decision-making may confer solutions to existing barriers and high discontinuation rates observed in psychotherapy.

Although the app users used the app for only 18 weeks, the 34-week follow-up period enabled us to determine whether the benefits of the app were sustainable. The length of the follow-up period is unique in the field of mobile health research on MDD. In fact, a recent review of the effectiveness of apps targeting patients with MDD identified 18 studies evaluating their impact on depression [16], and none of the studies was for >4 months in duration.

Two important strengths of our pilot study were its randomized controlled study design and the long-term 1-year follow-up period. Another strength was that the app was developed and piloted in collaboration with end users in the health care team and cocreated with patient end users. Research has demonstrated that the effectiveness of digital technology resources can depend on the extent to which end users are included as active participants in their design [18]. Moreover, approximately 60% of the study population was Hispanic, non-Hispanic Black, or multiracial, suggesting that the patients included in the study were largely representative of the racial and ethnic diversity observed in the US population.

### Limitations and Future Directions

Potential limitations of the study included its small sample size, which limited our ability to identify statistically significant differences between groups, and the relatively short duration of app use although we were able to maintain follow-up with nearly half of the study's participants after they discontinued its use. In addition, although patients and providers generally expressed high satisfaction with the app and interest in its features, both groups received limited education on how to use the app or its associated reports. The provision of additional education about the functionality and reporting features of the app might help increase patient-provider engagement and lead to improvements in the overall management of MDD.

Embedding the reporting feature into the EMR rather than it being a stand-alone report might also improve the ability of PCPs to make real-time decisions about treatment. Additional work on the Pathway platform informed by the results of this study will help integrate the digitally enabled MDD care pathway into the current AAH system by assessing process and workflow improvements, clinician-patient experiences, collaborative care model enhancements, EMR integration, and efficiencies with other platforms.

Our study data and qualitative insights informed the design of a real-world, prospective, interventional study of the app currently underway at AAH, designed to test the scaling and integration of the Pathway platform, along with educational interventions, at multiple primary care sites (ClinicalTrials.gov NCT04891224). The goal of this study was to determine whether the use of the app can improve adherence to measurement-based care practices in primary care to help improve outcomes for patients with MDD.

### Conclusions

This pilot study demonstrated that patients with MDD will engage with a mobile app designed to track treatment and disease progression and that health care providers will use the data generated as part of their assessment to inform care. The study results demonstrate that it is feasible to conduct an innovative app intervention in this diverse patient population with moderate to moderately severe depression. Introducing a customized, cocreated patient app into the care pathway can provide both patients and clinicians with greater details and trend data related to the disease state outside the traditional in-person visit. Enhanced use of patient-reported data within real-world health care settings can help support measurement-based care practices by making patient self-reported data and summaries of the data easy to interpret and easy to access within existing EMR instances.

Although the sample size was small for the long-term follow-up phase of the study, the results of this feasibility study suggest that this digitally enabled MDD clinical care pathway approach may support shared decision-making and help provide sustainable benefits over at least 1 year. The impact of the app on patient activation and MDD management will be further explored in a larger prospective study of its real-world use in patients with MDD.

### Acknowledgments

The authors thank all patients for their participation in this study. The authors also gratefully acknowledge the contributions of Ellen Rhodes, Lambros Chrones, and Anit Roy, who were involved in the project management and medical review aspects of the study. The authors also thank the study site subinvestigators (Ramon A Gonzalez, Kevin T Loughry, Jacqueline Ivey-Brown, Roxanne Smith, Robert Zitter, and Gina Schueneman) for their contributions to the study and manuscript. This study was funded by Takeda Pharmaceuticals USA, Inc and Lundbeck LLC. Medical writing assistance was provided by Nicole Cooper on behalf of Syneos Health and supported by Takeda Pharmaceuticals USA, Inc and Lundbeck LLC.

## Data Availability

The data supporting the findings of this study are included in the published paper (and its supplementary information files). The authors may be contacted for further data sharing. The CONSORT - eHEALTH (Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth) checklist is available as [Multimedia Appendix 2](#).

## Authors' Contributions

MM contributed to the study design, data interpretation, editing, critical review, and approval of the manuscript. CB contributed to the study design, data interpretation, critical review, and approval of the manuscript. BF contributed to the study design, data interpretation, critical review, and approval of the manuscript. JK contributed to the study design, data interpretation, critical review, and approval of the manuscript. FC contributed to the study design, data interpretation, critical review, and approval of the manuscript. SS contributed to the study design, data interpretation, critical review, and approval of the manuscript. AE contributed to the study design, data interpretation, critical review, and manuscript approval. CK contributed to the study design; data collection, analysis, and interpretation; and drafting, editing, critical review, and approval of the manuscript. RK contributed to the study design, data interpretation, critical review, and approval of the manuscript. DK contributed to the study design, data interpretation, critical review, and approval of the manuscript.

## Conflicts of Interest

MM and SS are employees of Takeda Pharmaceuticals USA, Inc. JK, BF, and FC are employees of Cognition Kit. CB was an employee of Advocate Aurora Health at the time of the study and is now an employee of Takeda Pharmaceuticals USA, Inc. CK, RK, and DK are employees of Advocate Aurora Health. DK has received remuneration from Takeda for activities unrelated to the conduct of the study. AE was an employee of Lundbeck LLC at the time of the study and is now an employee of LB Pharmaceuticals.

### Multimedia Appendix 1

Key study demographics, patient-reported outcomes, and longitudinal results.

[\[PDF File \(Adobe PDF File\), 194 KB - formative\\_v6i10e34923\\_app1.pdf\]](#)

### Multimedia Appendix 2

CONSORT-eHEALTH (Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth) checklist.

[\[PDF File \(Adobe PDF File\), 284 KB - formative\\_v6i10e34923\\_app2.pdf\]](#)

## References

1. Hasin DS, Sarvet AL, Meyers JL, Saha TD, Ruan WJ, Stohl M, et al. Epidemiology of adult DSM-5 major depressive disorder and its specifiers in the United States. *JAMA Psychiatry* 2018 Apr 01;75(4):336-346 [[FREE Full text](#)] [doi: [10.1001/jamapsychiatry.2017.4602](https://doi.org/10.1001/jamapsychiatry.2017.4602)] [Medline: [29450462](https://pubmed.ncbi.nlm.nih.gov/29450462/)]
2. Gelenberg AJ, Freeman MP, Markowitz JC, Rosenbaum JF, Thase ME, Trivedi MH, et al. Practice guideline for the treatment of patients with major depressive disorder. 3rd edition. American Psychiatric Association. 2010 Oct. URL: [https://psychiatryonline.org/pb/assets/raw/sitewide/practice\\_guidelines/guidelines/mdd.pdf](https://psychiatryonline.org/pb/assets/raw/sitewide/practice_guidelines/guidelines/mdd.pdf) [accessed 2021-08-20]
3. Ho SC, Chong HY, Chaiyakunapruk N, Tangiisuran B, Jacob SA. Clinical and economic impact of non-adherence to antidepressants in major depressive disorder: a systematic review. *J Affect Disord* 2016 Mar 15;193:1-10. [doi: [10.1016/j.jad.2015.12.029](https://doi.org/10.1016/j.jad.2015.12.029)] [Medline: [26748881](https://pubmed.ncbi.nlm.nih.gov/26748881/)]
4. Semahegn A, Torpey K, Manu A, Assefa N, Tesfaye G, Ankomah A. Psychotropic medication non-adherence and its associated factors among patients with major psychiatric disorders: a systematic review and meta-analysis. *Syst Rev* 2020 Jan 16;9(1):17 [[FREE Full text](#)] [doi: [10.1186/s13643-020-1274-3](https://doi.org/10.1186/s13643-020-1274-3)] [Medline: [31948489](https://pubmed.ncbi.nlm.nih.gov/31948489/)]
5. Sansone RA, Sansone LA. Antidepressant adherence: are patients taking their medications? *Innov Clin Neurosci* 2012 May;9(5-6):41-46 [[FREE Full text](#)] [Medline: [22808448](https://pubmed.ncbi.nlm.nih.gov/22808448/)]
6. Barkil-Oteo A. Collaborative care for depression in primary care: how psychiatry could "troubleshoot" current treatments and practices. *Yale J Biol Med* 2013 Jun;86(2):139-146 [[FREE Full text](#)] [Medline: [23766735](https://pubmed.ncbi.nlm.nih.gov/23766735/)]
7. Alson AR, Robinson DM, Ivanova D, Azer J, Moreno M, Turk ML, et al. Depression in primary care: strategies for a psychiatry-scarce environment. *Int J Psychiatry Med* 2016;51(2):182-200. [doi: [10.1177/0091217416636580](https://doi.org/10.1177/0091217416636580)] [Medline: [27079777](https://pubmed.ncbi.nlm.nih.gov/27079777/)]
8. Unützer J, Park M. Strategies to improve the management of depression in primary care. *Prim Care* 2012 Jun;39(2):415-431 [[FREE Full text](#)] [doi: [10.1016/j.pop.2012.03.010](https://doi.org/10.1016/j.pop.2012.03.010)] [Medline: [22608874](https://pubmed.ncbi.nlm.nih.gov/22608874/)]
9. Bishop TF, Seirup JK, Pincus HA, Ross JS. Population of US practicing psychiatrists declined, 2003-13, which may help explain poor access to mental health care. *Health Aff (Millwood)* 2016 Jul 01;35(7):1271-1277. [doi: [10.1377/hlthaff.2015.1643](https://doi.org/10.1377/hlthaff.2015.1643)] [Medline: [27385244](https://pubmed.ncbi.nlm.nih.gov/27385244/)]

10. Trivedi MH. How can measurement-based care help improve treatment outcomes for major depressive disorder in primary care? *J Clin Psychiatry* 2020 Mar 17;81(2):UT17042BR2C [[FREE Full text](#)] [doi: [10.4088/JCP.UT17042BR2C](https://doi.org/10.4088/JCP.UT17042BR2C)] [Medline: [32220156](https://pubmed.ncbi.nlm.nih.gov/32220156/)]
11. Kroenke K, Unutzer J. Closing the false divide: sustainable approaches to integrating mental health services into primary care. *J Gen Intern Med* 2017 Apr;32(4):404-410 [[FREE Full text](#)] [doi: [10.1007/s11606-016-3967-9](https://doi.org/10.1007/s11606-016-3967-9)] [Medline: [28243873](https://pubmed.ncbi.nlm.nih.gov/28243873/)]
12. Hibbard JH, Greene J. What the evidence shows about patient activation: better health outcomes and care experiences; fewer data on costs. *Health Aff (Millwood)* 2013 Feb;32(2):207-214. [doi: [10.1377/hlthaff.2012.1061](https://doi.org/10.1377/hlthaff.2012.1061)] [Medline: [23381511](https://pubmed.ncbi.nlm.nih.gov/23381511/)]
13. Bakker D, Rickard N. Engagement in mobile phone app for self-monitoring of emotional wellbeing predicts changes in mental health: MoodPrism. *J Affect Disord* 2018 Feb;227:432-442. [doi: [10.1016/j.jad.2017.11.016](https://doi.org/10.1016/j.jad.2017.11.016)] [Medline: [29154165](https://pubmed.ncbi.nlm.nih.gov/29154165/)]
14. Batra S, Baker RA, Wang T, Forma F, DiBiasi F, Peters-Strickland T. Digital health technology for use in patients with serious mental illness: a systematic review of the literature. *Med Devices (Auckl)* 2017 Oct 4;10:237-251 [[FREE Full text](#)] [doi: [10.2147/MDER.S144158](https://doi.org/10.2147/MDER.S144158)] [Medline: [29042823](https://pubmed.ncbi.nlm.nih.gov/29042823/)]
15. Lee JA, Choi M, Lee SA, Jiang N. Effective behavioral intervention strategies using mobile health applications for chronic disease management: a systematic review. *BMC Med Inform Decis Mak* 2018 Feb 20;18(1):12 [[FREE Full text](#)] [doi: [10.1186/s12911-018-0591-0](https://doi.org/10.1186/s12911-018-0591-0)] [Medline: [29458358](https://pubmed.ncbi.nlm.nih.gov/29458358/)]
16. Porras-Segovia A, Díaz-Oliván I, Gutiérrez-Rojas L, Dunne H, Moreno M, Baca-García E. Apps for depression: are they ready to work? *Curr Psychiatry Rep* 2020 Feb 05;22(3):11. [doi: [10.1007/s11920-020-1134-9](https://doi.org/10.1007/s11920-020-1134-9)] [Medline: [32025826](https://pubmed.ncbi.nlm.nih.gov/32025826/)]
17. Torous J, Nicholas J, Larsen ME, Firth J, Christensen H. Clinical review of user engagement with mental health smartphone apps: evidence, theory and improvements. *Evid Based Ment Health* 2018 Aug;21(3):116-119. [doi: [10.1136/eb-2018-102891](https://doi.org/10.1136/eb-2018-102891)] [Medline: [29871870](https://pubmed.ncbi.nlm.nih.gov/29871870/)]
18. Kushniruk A, Nøhr C. Participatory design, user involvement and health IT evaluation. *Stud Health Technol Inform* 2016;222:139-151. [Medline: [27198099](https://pubmed.ncbi.nlm.nih.gov/27198099/)]
19. Snyder HR. Major depressive disorder is associated with broad impairments on neuropsychological measures of executive function: a meta-analysis and review. *Psychol Bull* 2013 Jan;139(1):81-132 [[FREE Full text](#)] [doi: [10.1037/a0028727](https://doi.org/10.1037/a0028727)] [Medline: [22642228](https://pubmed.ncbi.nlm.nih.gov/22642228/)]
20. 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/)]
21. Hibbard JH, Mahoney ER, Stockard J, Tusler M. Development and testing of a short form of the patient activation measure. *Health Serv Res* 2005 Dec;40(6 Pt 1):1918-1930 [[FREE Full text](#)] [doi: [10.1111/j.1475-6773.2005.00438.x](https://doi.org/10.1111/j.1475-6773.2005.00438.x)] [Medline: [16336556](https://pubmed.ncbi.nlm.nih.gov/16336556/)]
22. Greene J, Hibbard JH. Why does patient activation matter? An examination of the relationships between patient activation and health-related outcomes. *J Gen Intern Med* 2012 May;27(5):520-526 [[FREE Full text](#)] [doi: [10.1007/s11606-011-1931-2](https://doi.org/10.1007/s11606-011-1931-2)] [Medline: [22127797](https://pubmed.ncbi.nlm.nih.gov/22127797/)]
23. Lam RW, Lamy FX, Danchenko N, Yaras A, White MK, Rive B, et al. Psychometric validation of the Perceived Deficits Questionnaire-Depression (PDQ-D) instrument in US and UK respondents with major depressive disorder. *Neuropsychiatr Dis Treat* 2018 Oct 29;14:2861-2877 [[FREE Full text](#)] [doi: [10.2147/NDT.S175188](https://doi.org/10.2147/NDT.S175188)] [Medline: [30464471](https://pubmed.ncbi.nlm.nih.gov/30464471/)]
24. 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/)]
25. Kemp D. Use of a Digitally Enabled App With Clinical Team Interface in the Management of Depression. *Clinical Trials*. 2021 May 18. URL: <https://clinicaltrials.gov/ct2/show/NCT04891224> [accessed 2022-03-30]
26. Greene J, Hibbard JH, Sacks R, Overton V, Parrotta CD. When patient activation levels change, health outcomes and costs change, too. *Health Aff (Millwood)* 2015 Mar;34(3):431-437. [doi: [10.1377/hlthaff.2014.0452](https://doi.org/10.1377/hlthaff.2014.0452)] [Medline: [25732493](https://pubmed.ncbi.nlm.nih.gov/25732493/)]
27. Bombard Y, Baker GR, Orlando E, Fancott C, Bhatia P, Casalino S, et al. Engaging patients to improve quality of care: a systematic review. *Implement Sci* 2018 Jul 26;13(1):98 [[FREE Full text](#)] [doi: [10.1186/s13012-018-0784-z](https://doi.org/10.1186/s13012-018-0784-z)] [Medline: [30045735](https://pubmed.ncbi.nlm.nih.gov/30045735/)]
28. Serrano-Ripoll MJ, Zamanillo-Campos R, Fiol-DeRoque MA, Castro A, Ricci-Cabello I. Impact of smartphone app-based psychological interventions for reducing depressive symptoms in people with depression: systematic literature review and meta-analysis of randomized controlled trials. *JMIR Mhealth Uhealth* 2022 Jan 27;10(1):e29621 [[FREE Full text](#)] [doi: [10.2196/29621](https://doi.org/10.2196/29621)] [Medline: [35084346](https://pubmed.ncbi.nlm.nih.gov/35084346/)]
29. Wu A, Scult MA, Barnes ED, Betancourt JA, Falk A, Gunning FM. Smartphone apps for depression and anxiety: a systematic review and meta-analysis of techniques to increase engagement. *NPJ Digit Med* 2021 Feb 11;4(1):20 [[FREE Full text](#)] [doi: [10.1038/s41746-021-00386-8](https://doi.org/10.1038/s41746-021-00386-8)] [Medline: [33574573](https://pubmed.ncbi.nlm.nih.gov/33574573/)]
30. Graham AK, Greene CJ, Kwasny MJ, Kaiser SM, Lieponis P, Powell T, et al. Coached mobile app platform for the treatment of depression and anxiety among primary care patients: a randomized clinical trial. *JAMA Psychiatry* 2020 Sep 01;77(9):906-914 [[FREE Full text](#)] [doi: [10.1001/jamapsychiatry.2020.1011](https://doi.org/10.1001/jamapsychiatry.2020.1011)] [Medline: [32432695](https://pubmed.ncbi.nlm.nih.gov/32432695/)]
31. Krohn H, Guintivano J, Frische R, Steed J, Rackers H, Meltzer-Brody S. App-based ecological momentary assessment to enhance clinical care for postpartum depression: pilot acceptability study. *JMIR Form Res* 2022 Mar 23;6(3):e28081 [[FREE Full text](#)] [doi: [10.2196/28081](https://doi.org/10.2196/28081)] [Medline: [35319483](https://pubmed.ncbi.nlm.nih.gov/35319483/)]

32. Adam A, Jain A, Pletnikova A, Bagga R, Vita A, N Richey L, et al. Use of a mobile app to augment psychotherapy in a community psychiatric clinic: feasibility and fidelity trial. *JMIR Form Res* 2020 Jul 03;4(7):e17722 [FREE Full text] [doi: [10.2196/17722](https://doi.org/10.2196/17722)] [Medline: [32618572](https://pubmed.ncbi.nlm.nih.gov/32618572/)]
33. Armitage LC, Kassavou A, Sutton S. Do mobile device apps designed to support medication adherence demonstrate efficacy? A systematic review of randomised controlled trials, with meta-analysis. *BMJ Open* 2020 Jan 30;10(1):e032045 [FREE Full text] [doi: [10.1136/bmjopen-2019-032045](https://doi.org/10.1136/bmjopen-2019-032045)] [Medline: [32005778](https://pubmed.ncbi.nlm.nih.gov/32005778/)]

## Abbreviations

**AAH:** Advocate Aurora Health

**CONSORT-eHEALTH:** Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth

**ED:** emergency department

**EMR:** electronic medical record

**MDD:** major depressive disorder

**PAM-13:** 13-item Patient Activation Measure

**PCP:** primary care provider

**PDQ-D5:** 5-item Perceived Deficits Questionnaire–Depression

**PHQ-9:** 9-item Patient Health Questionnaire

**PPES-7:** 7-item Patient-Provider Engagement Scale

**WHO-5:** 5-item World Health Organization Well-Being Index

*Edited by A Mavragani; submitted 12.11.21; peer-reviewed by H Ayatollahi, D Chao, R Marshall; comments to author 14.02.22; revised version received 21.06.22; accepted 27.06.22; published 27.10.22.*

*Please cite as:*

*McCue M, Blair C, Fehnert B, King J, Cormack F, Sarkey S, Eramo A, Kabir C, Khatib R, Kemp D*

*Mobile App to Enhance Patient Activation and Patient-Provider Communication in Major Depressive Disorder Management: Collaborative, Randomized Controlled Pilot Study*

*JMIR Form Res* 2022;6(10):e34923

URL: <https://formative.jmir.org/2022/10/e34923>

doi: [10.2196/34923](https://doi.org/10.2196/34923)

PMID: [36301599](https://pubmed.ncbi.nlm.nih.gov/36301599/)

©Maggie McCue, Christopher Blair, Ben Fehnert, James King, Francesca Cormack, Sara Sarkey, Anna Eramo, Christopher Kabir, Rasha Khatib, David Kemp. Originally published in *JMIR Formative Research* (<https://formative.jmir.org>), 27.10.2022. 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

# Telehealth Movement-to-Music to Increase Physical Activity Participation Among Adolescents With Cerebral Palsy: Pilot Randomized Controlled Trial

Byron Lai<sup>1</sup>, PhD; Laura Vogtle<sup>2</sup>, PhD, OTR/L; Raven Young<sup>1</sup>, BS; Mary Craig<sup>1</sup>, MD; Yumi Kim<sup>3</sup>, PhD; Marissa Gowey<sup>4</sup>, PhD; Erin Swanson-Kimani<sup>1</sup>, MD; Drew Davis<sup>1</sup>, MD; James H Rimmer<sup>5</sup>, PhD

<sup>1</sup>Division of Pediatric Rehabilitation Medicine, Department of Pediatrics, University of Alabama at Birmingham, Birmingham, AL, United States

<sup>2</sup>Department of Occupational Therapy, University of Alabama at Birmingham, Birmingham, AL, United States

<sup>3</sup>Department of Physical Medicine and Rehabilitation, University of Alabama at Birmingham, Birmingham, AL, United States

<sup>4</sup>Division of Pediatric Gastroenterology, Hepatology and Nutrition, Department of Pediatrics, University of Alabama at Birmingham, Birmingham, AL, United States

<sup>5</sup>Dean's Office, University of Alabama at Birmingham, Birmingham, AL, United States

**Corresponding Author:**

Byron Lai, PhD

Division of Pediatric Rehabilitation Medicine

Department of Pediatrics

University of Alabama at Birmingham

1600 7th Avenue South

5 Dearth, McWane 5601

Birmingham, AL, 35209

United States

Phone: 1 205 638 9790 ext 8

Email: [blai@uabmc.edu](mailto:blai@uabmc.edu)

## Abstract

**Background:** Adolescents with cerebral palsy (CP) who have mobility limitations have almost no access to inexpensive and enjoyable home-based programs that can be disseminated on a large scale to help them independently manage their health through participation in leisure-time physical activity (LTPA).

**Objective:** The primary aim of this study was to determine the preliminary efficacy of the early adoption phase of an adult Movement-to-Music (M2M) program with behavioral telecoaching for increasing LTPA and activity participation compared with a waitlist control group in adolescents with CP. The secondary aim was to explore the effects of the program on perceived levels of pain and fatigue. The tertiary aim was to qualitatively evaluate the factors that influenced adherence and develop a theory that would inform the development of a more targeted M2M telehealth program for this group.

**Methods:** This randomized controlled trial piloted a 4-week M2M program with weekly behavioral telecoaching among 58 adolescents with CP who walked or used wheelchairs. The participants were randomized into one of 2 groups: M2M or control, which maintained their daily activities. M2M included videos that participants were asked to complete 3 times each week at home (asynchronous training). Adherence to video minutes was objectively measured using cloud-based analytics. Changes in activity and LTPA participation were measured before and after the intervention using the Children's Assessment of Participation and Enjoyment total domain scores and active physical recreation domain scores, respectively. Perceived pain and fatigue were measured using the National Institutes of Health Neuro-QoL short forms. The changes in scores were compared between the groups using analysis of covariance. A grounded theory approach was used to analyze one-on-one interviews, coaching notes, and feedback surveys.

**Results:** A total of 58 people were enrolled, of which 49 (84%) completed the primary outcome follow-up assessment. The mean adherence to the prescribed exercise video minutes across all 4 weeks was 68%, starting from 90% in week 1 and gradually declining to 43% in week 4. Mean adherence to coaching calls was 91%. Analysis of covariance revealed a statistically significant difference between the pre- to postchange scores for Children's Assessment of Participation and Enjoyment Active Physical Recreation-Intensity domain scores in favor of the intervention group ( $F_{1,47}=8.76$ ;  $P=.005$ ; effect size=0.17, also known as volume

of LTPA). The qualitative findings highlighted 5 critical factors that influenced participants' adherence to the program: caregiver support, video elements, suitable exercises, music, and behavioral coaching.

**Conclusions:** This project determined that adolescents with CP responded well to an M2M telehealth program that could enhance their LTPA levels. This paper describes a theory in which adherence to a telehealth LTPA program can be optimized through functional and age-specific modifications for adolescents with CP.

**Trial Registration:** ClinicalTrials.gov NCT04264390; <https://clinicaltrials.gov/ct2/show/NCT04264390>

(*JMIR Form Res* 2022;6(10):e36049) doi:[10.2196/36049](https://doi.org/10.2196/36049)

## KEYWORDS

exercise; developmental disability; cerebral palsy; telemedicine; telerehabilitation; therapy; mobile phone

## Introduction

### Background

Cerebral palsy (CP) is a common motor disorder in childhood [1] and “describes a group of permanent disorders of the development of movement and posture, causing activity limitations that are attributed to nonprogressive disturbances that occurred in the developing fetal or infant brain” [2,3]. Adolescence is a critical period for health professionals to engage youth with CP in regular leisure-time physical activities (LTPAs). LTPA is defined as energy expenditure from sports, conditioning, exercise, household tasks, and other daily activities [4]. Reports have indicated that young people with CP have alarmingly low rates of LTPA, far exceeding those observed among typically developing young people [5,6]. The coronavirus pandemic has further reduced participation in LTPA among youth with CP [7]. However, physically active youth with CP are more likely to live healthy active lifestyles as adults [6], which has important consequences for healthy aging with CP.

Previous studies have suggested that LTPA interventions can improve gross motor function [8], reduce morbidity risk [9], and potentially assist in the management of secondary conditions, such as pain and fatigue [10,11]. Participation in LTPA is further supported by over 3 decades of published clinical studies that reported benefits to physical fitness and well-being in children with CP [8,12-15]. Collectively, these findings demonstrate a gap between successful clinically tested programs and those that impact larger populations within the community. Thus, there is a need to identify programs that can easily be translated beyond clinical and research-supported settings to reach larger populations of people with CP.

Clinical LTPA studies targeting youth with disabilities are often not designed to be translated into the home or community and therefore *fall short* of reaching larger populations [16]. Owing to the inability to reach and include large sample sizes within randomized controlled trials (RCTs), there is little high-quality evidence (meta-analyses) that confirms the effects of LTPA interventions in CP [8]. Moreover, interventions typically prescribe programs that are supervised and delivered on-site at a fitness or clinical facility and often include only ambulatory individuals (ie, studies often exclude people who use wheelchairs) [8,17-19]. Achieving LTPA guidelines is far more difficult for people with CP who may use wheelchairs or have difficulty walking for long periods, regardless of whether general guidelines [20] or CP-specific guidelines [15] are used. In

addition, there are often geographical, environmental (eg, lack of transportation), or economic (eg, cannot afford a fitness membership or one-on-one supervision by a therapist) reasons for isolation from participating in on-site LTPA programs at local facilities or clinics [21,22]. Therefore, there is an urgent need to translate successful clinical exercise studies into more widely available exercise programs that target underserved and physically inactive adolescents with CP who have mobility limitations.

Home-based exercise programs that incorporate “virtual” behavioral coaching through videoconferencing (telecoaching) are a desirable approach for people with disabilities who may not have convenient access to other means of exercise. The advantages of a telehealth approach over usual care include increased social support, quality of care, cost-effectiveness, and access to services [23]. These advantages have made telehealth an instrumental part of health services since the onset of the COVID-19 pandemic [24]. According to the supportive accountability theory [25], programs that use telehealth technology can foster strong adherence by overcoming barriers to on-site health services and promoting strong relationships with health professionals. Telehealth with behavioral coaching explains the success of 2 of the largest exercise trials for youth with disabilities [26,27], which reached sample sizes of 101 and 92 by incorporating telehealth technology to reduce the burden on both the participant and the LTPA trainer. The average sample size for the LTPA trials among youth with disabilities is 27 people [16].

To develop and implement large-scale trials for adolescents with CP, lessons can be learned from the adult literature, which has been more successful in reaching larger study samples. One ongoing telehealth trial is testing the effectiveness of a novel Movement-to-Music (M2M) video program that includes more than 400 adults with physical disabilities (NCT03024320). M2M is a 12-week program that incorporates 83 enjoyable movement routines accompanied by music to improve cardiovascular capacity, muscular strength and endurance, and range of motion. M2M combines elements of music with health-enhancing exercises, which may improve attention, communication, brain plasticity [28,29], and functional outcomes [29]. Evidence supports the efficacy of M2M in enhancing functional mobility (the Timed Up and Go and 6-Minute Walk Test) among adults with multiple sclerosis [30]. Although M2M is primarily implemented through telehealth channels, M2M was originally modified and scaled up from an on-site program that was

delivered at a state-of-the-art exercise facility for people with disabilities (Lakeshore Foundation).

To date, there is no evidence-based program that can be easily disseminated by health professionals across multiple sites and settings to promote LTPA among adolescents with CP who have mobility limitations. A low-cost M2M program has the potential to address this need, but it has not been tested among children with disabilities.

## Objectives

The purpose of this study was to examine the preliminary efficacy of a youth-based adapted M2M intervention for increasing both activity and LTPA participation among adolescents with CP, compared with a 4-week waitlist control (WC) group. A secondary purpose was to explore the potential effects of the program on the levels of perceived pain and fatigue, which are 2 important secondary conditions that can likely be improved through LTPA participation. A tertiary aim was to qualitatively evaluate the critical factors that influenced adherence and to develop a theory that would inform the development of a more targeted M2M telehealth program for CP. The original M2M program included movement exercises for a larger, more general population of adults with physical disabilities.

To determine whether M2M requires specific modifications for adolescents with CP, this study aimed to address the following research question: Can the early adoption phase of an adult-based program increase physical activity, LTPA, pain, and fatigue levels among adolescents with CP? Given that the intervention included exercise videos and behavioral coaching aimed at increasing community activity, we hypothesized that participants in M2M would increase their levels of LTPA and activity after the intervention compared with WC participants.

## Methods

### Study Design and RCT Considerations

This study was a parallel-arm pilot RCT (immediate M2M intervention group vs WC group). The person entering the data was blinded to group allocation. Randomization was performed by the project statistician, who was not involved in recruitment or intervention. This trial was registered a priori as a clinical trial (NCT04264390). The trial was conducted from February 2020 to September 2021 during the COVID-19 pandemic. Owing to COVID-19 university-related delays, the length of the intervention was shortened to 4 months from the originally proposed 8 months before conducting the trial (approved by the study sponsor). In accordance with the National Institutes of Health Stages of Intervention Development, this study was a stage 1 trial. The trial focused on testing a new behavioral intervention in a pilot test and testing implementation issues that assist in the development of training materials for a larger trial. This study also included a preliminary examination of outcomes that will provide estimates of efficacy for a larger trial.

## Recruitment

Participants were recruited via telephone through medical and billing records and physician referrals from a Children's Hospital. Recruitment was conducted between September 2020 and September 2021. The eligibility criteria included (1) a diagnosis of CP as determined by the International Classification of Disease codes from electronic medical records, (2) the ability to exercise with arms because most of the movements include arm exercises, (3) aged between 10 and 19 years (adolescent age range, as defined by the World Health Organization) [31], (4) access to a Wi-Fi internet connection at home, and (5) ability to use a device capable of viewing internet video content (television, computer tablet, laptop computer, or desktop computer). The exclusion criteria were as follows: (1) physically active (defined as >60 minutes per day of moderate to vigorous intensity exercise in a typical week) [32] and (2) complete blindness or deafness.

## Assessment Procedures

Screening, recruitment, and data collection were remotely performed. This was performed to bypass the need for transportation. Transportation to a facility deters a substantial proportion of people from participating (even for data collection) [33,34]. The assessment procedures were as follows. First, prospective participants were prescreened for age and disability type based on medical and billing records. Second, the child, adult, or caregiver was contacted to confirm their interest and eligibility. Third, prospective participants completed the assessments. They had 2 options to complete the assessments: (1) physically mailed documents that were in a large envelope containing the informed consent document and all outcome measures or questionnaires or (2) signature of a digital consent document through a secure web database REDCap (Research Electronic Data Capture; Vanderbilt University) and completion of the questionnaires through phone calls with the research staff. Children were instructed to complete the questionnaires with the assistance of a caregiver. Fourth, once consent documents were signed, participants were randomized to either the intervention or control group. After the 4-week intervention or wait period, participants in both the intervention and control groups completed a second set of questionnaires using the same option as they had done at baseline.

In addition, at the end of the 4-week intervention, all participants provided written feedback on an open-response survey asking questions related to likes, dislikes, and recommendations to improve the program. They were also provided with the option of participating in a one-on-one phone call or Zoom (Zoom Communications) semistructured interview with the principal investigator (PI). The interviews included questions related to general perceptions of the program, likes and dislikes, preferences for music and exercises, factors that affected attendance, and recommendations to improve the program. The interviews lasted no longer than 45 minutes. The participants were given US \$20 to complete the interviews. The interviews were audio recorded and transcribed for analysis.

## Randomization

After baseline assessments were completed and returned, the recruiting staff contacted the project statistician to determine the participants' allocation to a group. Participants were randomized into 1 of 2 groups (M2M or WC) with a 1:1 allocation ratio using a permuted block randomization approach. Participants were stratified based on their functional level into either a standing (Gross Motor Function Classification System [GMFCS] levels I-III) or a seated M2M program (GMFCS levels IV-V). The randomization sequence was generated a priori by the project statistician using a computer-generated random schedule in permuted blocks (SAS, version 9.4; SAS Institute). Only the statistician knew the randomization sequence (the recruitment staff, investigators, and interventionists were blinded to the randomization order), and the statistician was not involved in any other part of the intervention or recruitment process. Given the nature of the intervention and control, it was not possible to blind the participants or telecoach during the intervention.

## Intervention Procedures

### Overview

If randomized to the M2M group, the participants were mailed the intervention equipment and instructions. Participants randomized to the WC group received the intervention after a 4-week wait period, during which they continued their normal daily activities or therapies. The participants were given US \$45 for completing the questionnaires.

### Home-Based M2M Intervention

M2M incorporated a variety of music-oriented movement routines, which were accompanied by music, to improve cardiovascular capacity, muscular strength and endurance, range of motion, and general physical function. This M2M intervention included the early adoption phase (first 4 weeks) of a larger 12-week M2M program for adults (5R01HD085186-02) [35]. M2M includes three types of video levels that are given to participants based on their GMFCS level and functional ability: (1) videos with arm and leg exercises (level 1, designated for GMFCS levels I-II), (2) videos with only arm exercises (level 2, designated for GMFCS levels III-V or wheelchair users), and (3) videos with arm and leg exercises for one side of the body (level 3, GMFCS levels I-III with weakness or movement limitation on one side of the body, ie, hemiparesis). GMFCS levels were determined during the screening call using the expanded and revised version of the GMFCS (GMFCS-E&R) [36]. Specifically, parents were verbally guided through either the GMFCS-E&R for children aged 6 to 12 years or the GMFCS-E&R for children aged 12 to 19 years [36].

The present intervention included compilations of M2M videos organized into weekly playlists. Participants were instructed to complete all videos within their playlist 3 times per week on nonconsecutive days. In week 1, the prescription generally included a total of 48 minutes of video time, which included a

mixed range of motion exercise routines with guided instructions. In weeks 2 and 3, the patients were prescribed 70 minutes of video time. Week 2 included 2 mixed range of motion and muscle strengthening routines. Week 3 included 2 range of motion routines and 1 strengthening routine. In week 4, participants were prescribed 100 minutes of video time, which included 2 range of motion exercise routines, one for strength, and one for cardiovascular exercise. A new movement routine was introduced each week in a video taught by an adult instructor. The following week, the instructional video was removed and replaced with another video, which was demonstrated by an adult actor with a disability, who guided the participants through the routine. The videos have been described in detail elsewhere [35]. The music incorporated into the videos was noncopyrighted. Songs with relaxing tones and slower tempos were selected for the range of motion. Songs with faster tempos were chosen for muscle strengthening and cardiovascular exercises. If participants had difficulty using the videos prescribed, the research staff replaced them with videos that were more suitable for their functional mobility level. For example, if a participant had difficulty with the arm or leg movements from video level 1, videos in their program could be added from videos of level 2 (arm exercise only) or level 3 (exercise with one side of the body), as appropriate. The intervention doses per program level are displayed in terms of frequency, intensity, time, and type (Table 1). The participants were mailed adjustable wrist weights for the strength exercise videos. Participants were instructed to call the lead investigator immediately if they had any concerns regarding safety or had experienced any condition or injury resulting from the intervention.

This intervention used consumer-available equipment and software that could be readily replicated by other interventionists. All M2M videos were stored on a free, secure, publicly available, cloud-based video sharing service (YouTube). The YouTube account was set to "private": only users to whom the PI gave access could watch the videos. A telecoach monitored participant adherence using the YouTube Analytics web application. YouTube Analytics automatically records watch time, type of device used to stream content, audience retention, location, date or time frame, devices, likes and dislikes, and user comments.

To access the videos on YouTube, participants were allowed to use laptop computer, desktop computer, tablet, mobile phone, or television. To increase the consistency of how the videos were watched, all participants were provided with a Google Chromecast device so that they could stream the videos to their home television, and this was the method that participants were encouraged to use. Chromecast is a low-cost device that allows users to stream content from another computer device (eg, smartphone) onto the television, and it can be purchased from most major retailers. No changes were made to the equipment or the web-based protocol throughout the intervention period.



**Table 1.** Intervention dose.

	Week 1	Week 2	Week 3	Week 4
<b>Level 1 (standing)</b>				
Frequency	3	3	3	3
Intensity	Low	Low	Low to moderate	Moderate
Time (mm:ss)	16:20	25:05	21:08	33:17
Type (video, mm:ss)	ROM <sup>a</sup> I (instruct)	ROM I (part, 04:55); ROM II (Instr <sup>b</sup> , 20:10)	ROM I (part, 04:55); ROM II (part, 05:05); STR <sup>c</sup> (Instr, 11:17)	ROM I (part, 04:55); ROM II (Instr, 05:05); STR (part, 03:40); CV <sup>d</sup> (Instr, 19:37)
<b>Level 2 (seated)</b>				
Frequency	3	3	3	3
Intensity	Low	Low	Low to moderate	Moderate
Time (mm:ss)	16:20	23:43	21:59	33:08
Type (video, mm:ss)	ROM I (instruct)	ROM I (part, 04:55); ROM II (Instr, 18:48)	ROM I (part, 04:55); ROM II (part, 05:05); STR (Instr, 11:58)	ROM I (part, 04:55); ROM II (part, 05:05); STR (part, 03:35); CV (Instr, 19:33)
<b>Level 3 (hemiparesis)</b>				
Frequency	3	3	3	3
Intensity	Low	Low	Low to moderate	Moderate
Time (mm:ss)	15:50	21:26	21:28	30:13
Type (video, mm:ss)	ROM I (instruct)	ROM I (part, 04:55); ROM II (Instr, 16:33)	ROM I (part, 04:55); ROM II (part, 05:05); STR (Instr, 11:27)	ROM I (part, 04:55); ROM II (part, 05:05); STR (part, 03:40); CV (Instr, 16:32)

<sup>a</sup>ROM: range of motion.

<sup>b</sup>Instr: instructor-guided movements.

<sup>c</sup>STR: strength.

<sup>d</sup>CV: cardiovascular.

In addition to the M2M program, the intervention included behavioral coaching calls with a telecoach using the Zoom videoconference software. The participants received a total of 4 coaching calls (1 per week). Each coaching call (phone or videoconference) lasted approximately 20 minutes. Caregivers were instructed to support their children during the coaching call, considering that parental knowledge plays a significant role in determining the extent to which adolescents participate in LTPA [22,37]. Parental awareness of the benefits of LTPA, a positive attitude, and perseverance in exploring LTPA options have been identified as facilitators of LTPA behavior [37]. Parental barriers include beliefs that LTPA and sport are unimportant, concerns about the adolescent fitting in with typically developing children, and difficulties with watching the child struggle in competitive sports [37]. A caregiver was encouraged to attend coaching calls with the adolescent participant. The participant was permitted to attend the call without a caregiver in the room, but a caregiver had to be present in the home throughout the duration of the call.

The goal of the coaching calls was to enhance adherence to exercise videos and promote general LTPA and activities within the community. The calls included behavior change techniques framed within the social cognitive theory [38]. Specific strategies included confidence building, goal setting (specific, measurable, attainable, realistic, and timely goals), planned

steps toward achieving the goals, instructions on proper movement techniques to increase mastery, discussions on methods to overcome barriers to participation, resolving questions related to exercise, and discussing realistic benefits that can be obtained from exercise (review recent systematic reviews of exercise for CP). The coaching call procedures are described in [Multimedia Appendix 1](#). Coaching sessions were conducted by 2 research staff. The lead telecoach (BL) had a research background in adapted LTPA and behavioral coaching through tele-exercise for people with disabilities, in addition to 10 years of hands-on experience in exercise training for various disability groups. The assistant telecoach was a physical medicine and rehabilitation resident physician at the Children's, who was instructed by the lead telecoach. Together, the telecoaches reviewed the procedures biweekly to maintain coaching fidelity. The telecoaches took written notes on an Excel sheet, which was archived on the university's secure BOX server.

### **WC Participants**

Participants in the WC group were instructed to maintain their usual activities for 4 weeks. After completing this period of nonintervention, the WC participants received the M2M intervention.

## Summary of Qualitative Procedures

The summary of qualitative procedures is as follows:

1. The participants who completed the intervention were asked to participate in an interview to provide feedback on how to improve the program.
2. Interested participants were scheduled for a 45-minute phone or Zoom call (participant preference).
3. The interviews were conducted by the PI, who asked 10 questions related to general perceptions of the program, likes and dislikes, preferences for music and exercises, factors that affected attendance, and recommendations to improve the program.
4. The interview was audio recorded and transcribed by a third-party company.
5. Transcriptions were double-checked for accuracy.
6. Data were coded (see the Aim 3 Analysis section).
7. A theory was created based on coding results.
8. Participants were mailed an electronic debit card for completing the interview.

## Measures or Assessments

The study outcomes were assessed using questionnaires. If the participants were unable to complete the questionnaires independently, they were permitted to assist their children.

### *Aim 1: Primary Outcome Measures*

The primary outcome was pre- to postchanges in LTPA after the 4-week M2M intervention. At weeks 0 and 4, LTPA was measured using the Children's Assessment of Participation and Enjoyment (CAPE) [39]. CAPE is used to document changes in everyday activities. CAPE provides three levels of scoring: (1) overall participation scores, (2) domain scores that reflect participation in formal and informal activities, and (3) scores that reflect participation in 5 types of activities (active physical recreation and social, skill-based, and self-improvement activities). We aimed to compare pre- and postchanges in active physical recreation activity (a measure of LTPA) and the overall total CAPE score (a measure of activity) between the study groups. Systematic reviews have reported that CAPE has adequate evidence of validity (construct and content) and reliability (internal consistency, intrarater reliability, and test-retest reliability) to support its use among adolescents with CP [40,41].

### *Aim 2: Potential Effects on Pain and Fatigue*

The secondary aim of this study was to explore the potential effects of the program on perceived pain and fatigue. Perceived pain and fatigue were measured using the National Institutes of Health Neuro-QoL Pediatric Pain and Fatigue short forms [42]. The Neuro-QoL pediatric assessments are a set of health-related quality of life instruments designed to be used for children and adolescents (aged 8-17 years) with neurological conditions or disorders [43]. The psychometric properties of the Neuro-QoL measures have been tested in a variety of clinical populations but not in CP. The rationale for examining pain and fatigue was that treatments for managing these conditions have been identified as a top research priority by CP stakeholders [44]. The Neuro-QoL Pain short form includes 8 questions pertaining to perceived pain level in the past 7 days. Questions probe

perceived pain intensity, frequency, pain location, pain presence, and how pain affects lifestyle activities. The fatigue short form includes 8 questions pertaining to fatigue presence and how fatigue levels affected lifestyle activities and mood in the past 7 days. A higher score indicates a lower level of pain and fatigue.

### **Aim 3: Theory Generation to Inform Future Telehealth LTPA Programs**

The grounded theory framework by Charmaz [45] underpins the qualitative component of this study. This method posits that theories are constructed "through our past and present involvements and interactions with people, perspectives, and research practices." Thus, all study components were guided by the following philosophical assumptions: a critical realism ontological perspective (ie, there is a singular reality that can be understood through subjective perceptions or experiences of events that are linked with reality) [46] and an interpretivism epistemological perspective (ie, knowledge is socially constructed by both the participant and researcher) [47]. In other words, participants recalled reality when thinking about the questions in relation to their past experiences, and this reality was influenced by the interaction of the interviewer, caregiver, and participant. The theory was informed by 3 sets of data: the one-on-one postintervention interviews, the postprogram feedback survey, and notes taken by the 2 telecoaches after each behavioral coaching call.

## Ethical Considerations

The study protocol was approved by the University Institutional Review Board for Human Use at the University of Alabama at Birmingham (300004608). Approval was obtained from the university institutional review board before starting the study. Written informed consent was obtained from all participants or their caregivers before enrollment.

## Statistical Analysis

### *Overview*

Statistical significance was evaluated at a family-wise error rate of 0.05. Quality control included data double-checking as well as descriptive and graphical approaches to summarize the baseline characteristics of all key variables, followed by independent 2-tailed *t* tests and chi-square statistics to assess baseline group differences. All data were double-checked for their accuracy.

### *Power*

A power estimate calculation determined that a sample size of 54 was needed to detect a group difference, given the following variables: significance level of 0.05, Cohen *d* of 0.75, and power of 80%. The calculation was based on data obtained from an RCT of exercise for children and adolescents with CP that used the same measure for LTPA participation (CAPE) [48]. Although the duration of the intervention was longer than that in this study, the team decided that the intervention and measures were most relevant compared with other randomized trials [19].

### Primary Analysis

We examined the effects of the intervention versus control condition on pre- to postchanges in LTPA (CAPE total scores, domain scores, and CAPE Active Physical Recreation [CAPE-APR] domain scores) using analysis of covariance (ANCOVA) based on the univariate *F*-statistic (conditions: intervention and control, change score, and baseline variable data as the covariate). ANCOVA has been found to be valuable as an effective and nonbiased method of analyzing pre- to postchange scores [49], and it has the smallest variance, highest power, and nominal 95% CI coverage compared with ANOVA and linear mixed models [49,50]. This method was also used to compare group differences between pre- and postintervention changes in secondary outcomes. If ANCOVA assumptions were not met, 2-tailed *t* tests were conducted.

### Tertiary Qualitative Analysis

Audio transcriptions, participant feedback surveys, and telecoach notes were considered qualitative data that were analyzed using a grounded theory approach. The coding process included 3 phases: generation of (1) initial codes (ie, phrases that represent lines of text), (2) focused codes (ie, phrases that represent one or more initial codes), and (3) conceptual categories (ie, higher-order phrases that represent one or more focused codes) [45]. Conceptual categories and codes were arranged into a conceptual map (substantive theory) to address the study objectives. This process has been described in detail elsewhere [51]. The 2 analysts were the telecoaches. The lead analyst was a qualitative researcher with experience conducting more than 350 interviews for people with disabilities. The second analyst, a medical resident, was trained by the analyst.

### Data Exclusion

Data from participants who did not complete the postintervention assessments were excluded from the analyses.

## Results

### Quantitative Results

#### Overview

The participant characteristics are presented in Table 2. Detailed recruitment and enrollment information is displayed in a CONSORT (Consolidated Standards of Reporting Trials) diagram (Multimedia Appendix 2). Of the 521 people contacted, 82 (15.7%) were assessed for eligibility, 58 (11.1%) were enrolled, 51 (9.8%) completed the intervention, and 49 (9.4%) completed the intervention and assessments and were included in the analyses. A total of 4 participants did not want to complete the pain and fatigue measures because they felt overburdened by the assessments and the paperwork. At baseline, there were no statistically significant differences between the groups in terms of study outcomes or characteristics (all  $P > .05$ ).

The mean adherence to the prescribed exercise video minutes across all 4 weeks of the program was 68% (all 49 people, immediate start and control). The mean adherence to the videos was 90% (44/49 minutes) in week 1, 83% (56/68 minutes) in week 2, and 69% (45/65 minutes) and 43% (40/95 minutes) in weeks 3 and week 4, respectively. Adherence to coaching calls was 98% in week 1, 90% in week 2, and 90% and 86% in weeks 3 and 4, respectively. No adverse events (eg, accidents, injuries, or conditions related to the intervention) were reported by the participants.

**Table 2.** Participant characteristics (n=49).

Characteristics	Values
Age (years), mean (SD)	14 (3)
<b>Sex, n (%)</b>	
Male	25 (51)
Female	24 (49)
<b>Ethnicity, n (%)</b>	
White	30 (61)
Black	17 (35)
Asian	2 (4)
Hispanic or Latino	1 (2)
<b>Gross Motor Function Classification System level I-V, n (%)</b>	
I	7 (14)
II	16 (33)
III	6 (12)
VI	11 (22)
V	9 (18)

### Aim 1 and 2 Outcomes

The pre- to postchanges in outcomes between the groups are displayed in Table 3 for outcome variables. ANCOVA revealed statistically significant between-group differences in pre- to postchange scores for CAPE-Intensity ( $F_{1,47}=5.63$ ;  $P=.02$ ; effect size=0.11; also known as volume of activity) and CAPE-APR-Intensity ( $F_{1,47}=8.76$ ;  $P=.005$ ; effect size=0.17; also known as volume of LTPA). The estimated marginal means

revealed that these were small changes in favor of the intervention group. Interpretation of CIs revealed that changes in activity intensity could be classified as minimal to no changes because 0 was included within the CI [52]. There were no statistically significant group differences in overall CAPE score or CAPE-APR-Diversity, -With Whom, -Where, or -Enjoyment scores (all  $P>.05$ ). There were no statistically significant group differences in the pain or fatigue raw scores (all  $P>.05$ ).

**Table 3.** Pre- to postchanges in outcomes by intervention group and time.

Variable	M2M <sup>a</sup> change in score, mean (SD; CI)	M2M (n=23), n (%)	WC <sup>b</sup> change in score, mean (SD; CI)	WC (n=26), n (%)	F value (df)	P value
CAPE <sup>c</sup> -Diversity	-0.43 (3.59; -2.0 to 1.47)	23 (47)	0.19 (4.72; -1.59 to 1.67)	26	0.065 (1,47)	.80
CAPE-Intensity	0.10 (0.31 <sup>d</sup> ; -0.06 to 0.26)	23 (47)	-0.16 (0.43 <sup>d</sup> ; -0.31 to -0.01)	26	5.75 (1,47)	.02 <sup>d</sup>
CAPE-With Whom	0.02 (0.41; -0.18 to 0.23)	23 (47)	-0.05 (0.66; -0.24 to 0.14)	26	0.29 (1,47)	.59
CAPE-Where	-0.03 (0.52; -0.20 to 0.19)	23 (47)	0.11 (0.42; -0.09 to 0.27)	26	0.52 (1,47)	.47
CAPE-Enjoyment	-0.24 (0.99; -0.54 to 0.06)	23 (47)	-0.05 (0.44; -0.33 to 0.23)	26	0.93 (1,47)	.34
CAPE-APR <sup>e</sup> -Diversity	0.35 (1.56; -0.32 to 1.02)	23 (47)	-0.15 (1.05; -0.57 to 0.27)	26	1.79 (1,47)	.18
CAPE-APR-Intensity	0.35 (0.62 <sup>d</sup> ; 0.11 to 0.58)	23 (47)	-0.18 (0.54 <sup>d</sup> ; -0.4 to 0.041)	26	5.76 (1,47)	.02 <sup>d</sup>
CAPE-APR-With Whom	0.03 (1.15; -0.69 to 0.6)	23 (47)	0.48 (1.74; -0.11 to 1.07)	26	1.45 (1,47)	.23
CAPE-APR-Where	-0.03 (1.23; -0.66 to 0.64)	23 (47)	0.54 (1.75; -0.1 to 1.13)	26	1.37 (1,47)	.24
CAPE-APR-Enjoyment	-0.10 (1.43; -0.66 to 0.44)	23 (47)	0.02 (1.18; -0.49 to 0.53)	26	0.12 (1,47)	.72
Pain	2.48 (5.86; 0.3 to 6.43)	21 (46)	-0.44 (7.5; -0.64 to 5.32)	25 (54)	1.75 (1,47)	.19
Fatigue	-0.09 (7.9; -2.55 to 2.79)	21 (46)	0.03 (4.68; -0.48 to 0.25)	25 (54)	0.22 (1,47)	.88

<sup>a</sup>M2M: Movement-to-Music.

<sup>b</sup>WC: waitlist control.

<sup>c</sup>CAPE: Children's Assessment of Participation and Enjoyment.

<sup>d</sup>Significant difference at  $\alpha$  level .05.

<sup>e</sup>CAPE-APR: Children's Assessment of Participation and Enjoyment Active Physical Recreation.

## Qualitative Results

### Overview

A total of 28 people who completed the intervention also completed the interviews, and we chose to halt the interviews at this sample size because the conceptual categories appeared saturated (ie, the data analysts felt confident that the conceptual categories were adequately represented and explored) [45,53]. The results are organized into 2 sections. The first section briefly

describes the conceptual categories and codes that represent participants' preferences for an ideal home-based telehealth LTPA program. There were 6 resultant conceptual categories that are listed in Table 4 and described in further sections. The second section elaborates on the relationships between the conceptual categories and describes a theoretical framework that can be used by interventionists and health professionals aiming to develop or implement a telehealth LTPA program for adolescents with CP.

**Table 4.** Qualitative themes.

Conceptual categories	Focused code
Enjoyment is influenced by music, suitability of the exercises, and caregiver support	<ul style="list-style-type: none"> <li>• Enjoyment increases attentional focus on the videos</li> <li>• Music was the core influencer of video enjoyment and the decision to join the program</li> <li>• Caregivers participating with their child led to noticeable improvements in a child's mood and attention to the videos</li> <li>• Physical assistance from a caregiver created feelings of happiness and enhanced caregiver-child bonds</li> </ul>
Caregiver support increases the usability and suitability of the exercises	<ul style="list-style-type: none"> <li>• Caregivers are instrumental in addressing physical, cognitive, and motivational demands and can modify exercises</li> <li>• Caregivers are the primary method for exercise adaptations in a home-based environment</li> <li>• Caregiver scheduling was a direct influencer of attendance</li> </ul>
Music is instrumental for joining and sustaining an exercise regime	<ul style="list-style-type: none"> <li>• Participants desired music that matched the theme of the movements (tempo and movements that are performed rhythmically to the music)</li> <li>• Personalized music can increase enjoyment and adherence. Music need not be genre specific but should be upbeat</li> <li>• Music was identified as critical component of the children's lifestyles and was incorporated in most enjoyable activities they performed.</li> </ul>
Video elements influence usability and suitability of the exercises	<ul style="list-style-type: none"> <li>• Most participants appreciated and preferred short and asynchronous (video-based) home exercise program because of the convenience of fitting it into their busy schedules</li> <li>• Exercises should be adapted to include slow verbal instructions, different body positions (eg, bed or chair), and repetitive movements. These adaptations will help accommodate cognitive needs in the presence of intellectual disability (which influences the suitability of the exercises) or physical needs in the presence of impaired physical function. Doing so provides a sense of accomplishment and happiness while avoiding frustration and loss of attention</li> <li>• Videos should include child-appropriate themes (eg, superhero, sci-fi, western, pop dance, cartoons, and sports) and relatable actors (eg, children of similar ability and mobility as the participant or cartoon characters)</li> <li>• Exercises should be perceived as beneficial for range of motion or mobility</li> </ul>
Behavioral coaching influences caregiver support, usability, and adherence	<ul style="list-style-type: none"> <li>• Caregivers strongly noted that meeting with a coach created a sense of accountability that was a key factor that affected adherence to the program</li> <li>• Knowing that the program was offered by a trusted provider and created by a recognized community fitness facility for people with disabilities enhanced attendance and was a key factor in their decision to join the program</li> <li>• Developing a social bond with the participant and caregiver increased accountability</li> <li>• The coach resolved technical issues throughout the first 2 weeks of the program that had negatively impacted adherence</li> <li>• Participants and caregivers relied on the knowledge of the coach for guidance on the prescription (instructions at baseline were not sufficient)</li> </ul>

### ***Enjoyment Is Influenced by Music, Suitable Exercises, and Caregiver Support***

Adherence to the program was highly dependent on the children's enjoyment of the program. Enjoyment with the program increased the child's attentional focus on the movements and instructions within each video. High levels of enjoyment led to laughter and smiles when performing the movements, whereas low levels of enjoyment led to low levels of video minutes performed and noninterest in participation or even dropout (2 cases). The factors that contributed to enjoyment included elements of music, suitable exercises, and caregiver support.

### ***Caregiver Support Increases Adherence and Suitability of the Exercises***

Caregivers were a key factor in adherence and could perform supplemental adaptations that were needed for the child (in addition to those shown in the videos). Among participants with

GMFCS levels IV to V, caregivers were instructed to modify the program for their child by physically assisting them through the movements. In these cases, caregivers reported that they noticed that the children became happy when the caregiver physically put their hands on them to perform fun and enjoyable movements, as opposed to care activities. Moreover, caregivers often managed the child's exercise schedule and coaching calls, which were difficult because participants often had busy daily routines with school and care activities. In some cases, caregivers performed the videos with the child and used verbal cues and positive reinforcement to increase the child's adherence.

### ***Music Is Instrumental for Joining and Sustaining an Exercise Regime***

Music was identified as a core component of both the child's and caregiver's interest in joining and participating in the program. Participants reported a love for music that was incorporated into their most enjoyable activities they did at

home. Some participants loved dancing to music at home or listening to music to ease their anxiety and improve their mood. Participants expressed a variety of musical interests and noted that music aligned with their interests could enhance adherence. Some enjoyed the noncopyrighted music within the intervention, while others preferred music that aligned with their interests. However, for exercise, they reported that the music does not need to be limited to a specific genre, instead the music should coincide with the rhythm of the exercise movements (eg, an “upbeat” tempo for cardiovascular exercise). Participants strongly noted that an exercise program without a substantial musical component would not be appealing to them to participate in or perform for a prolonged period.

**Video Elements Influence Usability and Suitability of the Exercises**

Participants and caregivers reported several video elements that affected the child’s adherence. First, participants preferred short and asynchronous videos to real-time instructions (either in-person or videoconference instruction). Several participants suggested a duration <20 minutes. This was because caregivers and children had busy schedules with school, care activities, and frequent medical or therapeutic appointments or concerns. The ability to pause the videos and work at the child’s own pace was a strong advantage of asynchronous training. Second, videos were recommended to contain repetitive exercises and visual-verbal guidance to accommodate various learning and physical needs. Specifically, the fourth week of the cardiovascular exercise videos had too many transitions between exercise movements. In addition, the movements were too fast. These 2 factors were the reasons for the low video adherence at week 4. Participants also desired videos of shorter duration to accommodate short attention spans, which was the rationale for declining adherence in weeks 3 and 4. Notably, caregivers desired more program adaptations for people with GMFCS levels IV to V. In these cases, caregivers and participants desired a caregiver-child–assisted M2M program. Such a program could display a caregiver physically assisting the child throughout the movements and more appropriate exercise positions to avoid the need for transfer (eg, exercises performed in a bed, power wheelchair, or floor). For children with GMFCS levels I to III, exercises often appeared suitable and when performed

successfully, provided participants with a sense of happiness and accomplishment that increased enjoyment. “Honestly, he reacted better to the videos than I anticipated...He uses a wheelchair and enjoyed them. Doing something he knows he can do. He looks forward to it” [caregiver of participant 22]. Third, participants and caregivers reported that age-appropriate themes and relatable actors could increase their interest. Actors could have a similar functional ability or be superheroes or animated characters from a show or movie. Participants desired videos based on age-appropriate themes (eg, superhero themes, pop dance themes, or sports). Fourth, exercises that were perceived as beneficial to physical function appealed to both the participants and caregivers. As stated by participant 47, “I really liked doing the exercises because I had arm surgery, and I feel like the exercises helped my arms and made me feel better.”

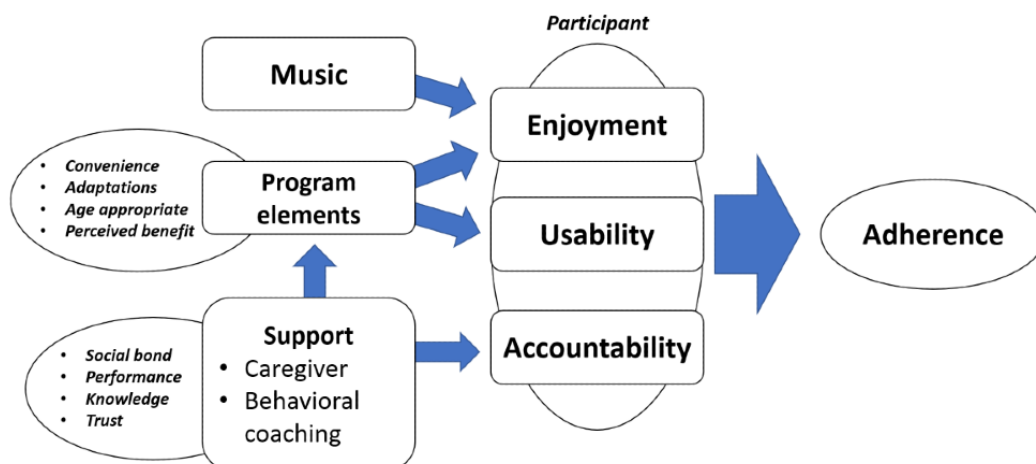
**Behavioral Coaching Increases Adherence Through Accountability and Support**

Both caregivers and participants reported that the program would not have adhered well without behavioral coaching support. Coaching helped in identifying movement adaptations, understanding the importance of exercise for CP, and reviewing the program structure or instructions. Most importantly, the coaching provided participants and caregivers with a sense of accountability. They felt that they had made a commitment and built a social bond with a trusted health professional, which made them want to try their best to attend the sessions. They further appreciated that a health professional was aiming to improve their child’s health outside the medical and school setting, which they felt was underserved. In addition, the coach was relied on to resolve streaming issues and provide guidance for completing the exercise prescription.

**Theory for Increasing Adherence to Telehealth LTPA Programs**

This section describes a theory that can be used by interventionists and health professionals (Figure 1) for the development and implementation of a telehealth program for adolescents with CP. The theory displays a framework that depicts linkages between conceptual categories that may lead to increased adherence to a telehealth LTPA program.

Figure 1. Theory for maximizing adherence in telehealth physical activity among adolescents with cerebral palsy.



The theory posits that 3 elements influence the adherence of adolescents with CP to a telehealth LTPA program: support, video elements, and music. A successful telehealth program aimed at promoting self- or family-regulated home exercise behavior should address all 3 elements. Adherence is influenced by enjoyment, usability (ie, successful video performance), and accountability. Music is a strong factor of enjoyment and should be considered mandatory in any home-based exercise program that aims to promote sustainable exercise behavior. Enjoyment is further influenced by program elements (eg, children are happy to perform movements that match their functional ability and include age-appropriate themes and actors) and support from both the caregiver and behavioral coach.

The program elements influence adherence and usability. If an exercise is not adapted for the individual, they will not be able to perform the movement successfully and as a result, rapidly lose interest. Moreover, despite suitable exercises, if the video instructions were not presented appropriately, the exercise would not be suitable for the participants' needs. Thus, video elements are indirectly linked to usability.

Caregiver influence was critical to participant adherence and often outweighed other factors. For example, a highly involved caregiver could increase adherence to less enjoyable music and unsuitable exercises or video elements. Behavioral coaching was critical for maintaining the caregivers and participants involved by providing a sense of accountability through performance monitoring, knowledge of benefits, and a bond with a trusted health professional or organization. This theory was underpinned by the supportive accountability theory [24]. In addition, coaching helped resolve programmatic and technological issues and helped caregivers engage in program adaptations. Thus, behavioral coaching influences adherence through accountability and indirectly through caregiver support, along with an indirect influence on adherence through program elements.

Conceptually, all 3 constructs of the theory, namely music, program elements, and support, should be addressed to enhance the likelihood that adolescents with CP adhere strongly to a telehealth LTPA trial.

## Discussion

### Principal Findings

Although there are efficacious on-site LTPA interventions for adolescents with CP, the effectiveness and dissemination of these interventions are low because of accessibility and expense. Recent systematic reviews have demonstrated that LTPA trials are only effective at increasing LTPA participation among ambulatory people with CP and typically include small sample sizes [16,19]. This project piloted a low-cost behavioral coaching M2M program among adolescents with CP to determine whether the telehealth program could potentially improve activity, LTPA, pain, and fatigue and whether modifications were necessary.

Quantitative findings demonstrated that the M2M program resulted in a small increase (effect size=0.17) in the volume of LTPA performed by the individual in the home setting compared

with WC. This was demonstrated by a statistically significant improvement in CAPE-APR–Intensity but no improvement or difference in CAPE-APR–Diversity (types of activities), –With Who (exercising with others), –Where (exercise setting), and –Enjoyment scores. A statistically significant increase in nonsupervised LTPA behavior is meaningful, given that very few RCTs have successfully promoted the adoption of independent exercise behavior within the community [16,18,19], particularly among people with CP who have higher mobility disability. A meta-analysis by Reedman et al [18] demonstrated that behavioral interventions are efficacious in improving “habitual physical activity,” measured via steps per day, but interventions have not been effective in improving LTPA. In other words, interventions with behavioral coaching can improve step count among people who are ambulatory. However, further research is warranted into methods of promoting LTPA within the community among people who are nonambulatory. Although there was a statistically significant improvement in general activity (CAPE-Intensity), the difference was not strong enough to be termed an effect.

People with CP who use wheelchairs have largely been excluded from past LTPA interventions. A scoping review of LTPA studies reported that of the 48 RCTs including 1513 people with CP, only 16 people were reportedly included with a GMFCS level V, and 64 people with a GMFCS level IV were included [19]. In a recent report by Gross et al [44] and the Cerebral Palsy Research Network, the development and evaluation of effective methods for increasing LTPA among people who are less ambulatory have been ranked as a top research priority by a large cohort of CP stakeholders.

A strength of this study was that there was a strong representation of GMFCS levels IV and V. People with these levels were able to be included by adapting the exercise movements. Telecoaches guided caregivers to physically assist the participants with their exercises. This was done by the caregiver moving the participant's arm while following the video. This often required several pauses by the caregiver to allow adequate time to learn a movement or adjust the participant's position. A valuable and unexpected outcome was that the caregiver physically assisting the child created sensations of happiness and joy, which was caused by the child associating physical touch with something fun and playful, as opposed to daily care. Nevertheless, caregivers and children expressed a strong desire for caregiver-assisted videos and general exercise knowledge for their children, which was directly applicable to their scenarios, and this should be a research priority.

Regarding program adherence, we deemed an overall adherence rate of 69% to be acceptable. This determination was based on evidence that adherence rates between 50% and 69% to an on-site 12-week M2M program resulted in improved functional mobility among adults with physical disabilities [30,54]. Nevertheless, adherence was strong during the first 2 weeks of the program and declined gradually throughout weeks 3 and 4. Qualitative study findings demonstrated that adherence declines were due to the increasing duration of the videos as well as the too fast and complex cardiovascular routine provided in week 4. On the basis of the qualitative study findings, we recommend

that asynchronous video programs should be brief in duration and include repetitive cardiovascular exercise routines with age-appropriate themes.

Although we were slightly shy of our target recruitment goal, another strength of this study was that a respectable sample size was obtained within a 1-year time frame. The achieved sample size was nearly twice the mean sample size of LTPA trials for children with disabilities [16]. Nevertheless, the findings demonstrated that the intervention had no effect on the level of perceived pain or fatigue. There are several possibilities for this. First, the dose was likely insufficient to improve the pain. Only a handful of studies have suggested potentially beneficial effects of exercise on perceived pain and fatigue among people with CP [11,19,55]. The 2 studies that reported benefits to pain and fatigue included resistance training programs that were higher in intensity than in the first 4 weeks of M2M. A second explanation was that the baseline values for pain were low, leaving little room for improvement. The mean baseline scores were 1.6 and 2.4 for pain and fatigue, respectively. These scores indicated “never/almost never” feeling pain and “a little bit/some of the time” feeling fatigue. Nevertheless, further efforts are needed to examine the effects of exercise on pain and fatigue in people with CP [19].

Several novel features were embedded in the project. First, the telehealth intervention enhanced participant adherence by providing remote support from a disability exercise professional. Second, the telehealth project overcame common logistical barriers (eg, transportation, time, and program costs) that prevented adolescents with CP from participating. Third, the intervention procedures were specifically designed to be implemented with minimal resources (eg, low-cost prerecorded videos, no transportation required, and minimal equipment) and burden on research staff to facilitate replication of the project in a larger future trial.

### Recommendations for Future Telehealth LTPA Trials

Qualitative findings identified critical factors that contributed to participants' adherence, and these findings were packaged into a theory that can maximize adherence in future telehealth LTPA trials. The 5 constructs of the theory, namely music, caregiver support, video elements, suitable exercises, and coaching, provide a general guide for replicating the study procedures in a future telehealth trial that can foster enjoyment, usability, and adherence. Specific intervention content (eg, choice of music and specific exercises) should be tailored to the needs and preferences of the specific target population. For example, the participants in this study were located within the Southeast United States, and their preferences will certainly not be generalizable to all adolescents with CP. This information can be obtained from a quick survey or usability study. In further sections, we elaborate on how these factors affect adherence to our intervention and list recommendations for future LTPA telehealth trials.

Attendance at the program was strong during the first 2 weeks of M2M. Participants had passion for music and felt that the program came from a trusted source (the hospital and adapted fitness facility). These factors also heavily influenced their decisions to participate in the program. However, adherence to

week 3 declined as the movements became more numerous and complex, and the videos became longer. A large drop in attendance was observed during week 4, when the first cardiovascular exercise video was introduced, which had many movements that transitioned too quickly for many participants to follow. Caregiver support was often sufficient to counterbalance these inadequacies. Several lessons can be learned to further increase participant adherence in future trials:

1. Exercises and their instructions should be repetitive with slow transitions between different movements.
2. Exercises should always be visually guided, in addition to verbal instruction.
3. Music should be integrated into all components of the program (even the instructions).
4. Music preferences vary, and thus, the choice of music should emphasize the upbeat tempos that are specific to the target group.
5. Videos should include actors with disabilities that match the mobility and functional abilities of the participants.
6. Telecoaches should have behavior change techniques readily available for both the caregiver and adolescent.
7. Videos should be presented with adolescent-appropriate themes.

### Limitations

This pilot study has several limitations. First, the study did not include any objective measures of health or physical activity, which should be the focus of a future trial testing a CP-specific M2M program. Second, the study was implemented during the COVID-19 pandemic, which could have influenced the response rates and changes in outcomes. For example, participation in community activities or physical activities may have been resistant to change, considering that the state was in lockdown, and many community activities were closed or had occupancy limits. By contrast, this could have inflated our response rates, as participants may have been more interested in a home exercise program during this period. Third, our sample was short of the target by 5 people. Fourth, this project initially aimed to test an 8-week M2M program. Owing to the COVID-19 restrictions, we experienced a 6-month delay in university-related operations, which required us to shorten the intervention to complete the study within a 1.5-year period. There is a need to test for longer intervention and follow-up periods to assess sustainability. Fifth, with only 2 study groups (intervention and control), we were unable to discern how physical activity levels would differ with and without behavioral coaching (ie, how behavioral coaching affects adherence), which should be tested in a future trial with 3 study groups (intervention with behavioral coaching, intervention alone, and control). Finally, the intervention required participants to have access to wireless internet at home, which may affect the generalizability of the study findings. No participants were excluded because they did not have access to wireless internet at home.

### Conclusions

This project established the preliminary efficacy of a scalable telehealth M2M program for increasing LTPA behavior among adolescents with CP. Overall, participants responded well to



the program, which improved their LTPA levels at home. Adherence to the program was strong during the first few weeks but declined as the exercises became less suited to the participants' needs. These findings demonstrated that M2M may require additional CP-specific modifications before

implementation in a larger multisite trial. Ideally, the final product will result in an exercise program that can be implemented across a variety of settings to reach and include an underserved population of adolescents with mobility limitations.

---

## Acknowledgments

This project was funded by the American Academy for Cerebral Palsy and Developmental Medicine Junior Investigator Award (grant AACPD MJIA012001). The interviews were funded by a pilot grant from the Center for Adherence in Disability Health and Rehabilitation Sciences.

---

## Authors' Contributions

All authors contributed to the study design, reviewed the paper in full, and made revisions that were incorporated into the final draft of the manuscript. BL, YK, and MG contributed to initial draft of the manuscript. JHR provided all prerecorded M2M video content. RY helped manage and implement the program. MC was the behavioral coach. LV, DD, and ES-K provided recommendations, from which BL tailored the exercise and technology content for adolescents with cerebral palsy.

---

## Conflicts of Interest

None declared.

---

### Multimedia Appendix 1

Coaching operation procedures.

[[DOCX File, 19 KB - formative\\_v6i10e36049\\_app1.docx](#)]

---

### Multimedia Appendix 2

CONSORT (Consolidated Standards of Reporting Trials) diagram.

[[PDF File \(Adobe PDF File\), 82 KB - formative\\_v6i10e36049\\_app2.pdf](#)]

---

### Multimedia Appendix 3

CONSORT-eHEALTH checklist (V 1.6.1).

[[PDF File \(Adobe PDF File\), 1231 KB - formative\\_v6i10e36049\\_app3.pdf](#)]

---

## References

1. Accardo PJ. Capute & Accardo's Neurodevelopmental Disabilities in Infancy and Childhood, Third Edition: Volume I: Neurodevelopmental Diagnosis and Treatment. Baltimore, MD, USA: Brookes Publishing; 2008.
2. Bax M, Goldstein M, Rosenbaum P, Leviton A, Paneth N, Dan B, Executive Committee for the Definition of Cerebral Palsy. Proposed definition and classification of cerebral palsy, April 2005. *Dev Med Child Neurol* 2005 Aug;47(8):571-576 [[FREE Full text](#)] [doi: [10.1017/s001216220500112x](https://doi.org/10.1017/s001216220500112x)] [Medline: [16108461](https://pubmed.ncbi.nlm.nih.gov/16108461/)]
3. Rosenbaum P, Paneth N, Leviton A, Goldstein M, Bax M, Damiano D, et al. A report: the definition and classification of cerebral palsy April 2006. *Dev Med Child Neurol Suppl* 2007 Feb;109:8-14. [Medline: [17370477](https://pubmed.ncbi.nlm.nih.gov/17370477/)]
4. Caspersen CJ, Powell KE, Christenson GM. Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. *Public Health Rep* 1985;100(2):126-131 [[FREE Full text](#)] [Medline: [3920711](https://pubmed.ncbi.nlm.nih.gov/3920711/)]
5. Carlon SL, Taylor NF, Dodd KJ, Shields N. Differences in habitual physical activity levels of young people with cerebral palsy and their typically developing peers: a systematic review. *Disabil Rehabil* 2013 Apr;35(8):647-655. [doi: [10.3109/09638288.2012.715721](https://doi.org/10.3109/09638288.2012.715721)] [Medline: [23072296](https://pubmed.ncbi.nlm.nih.gov/23072296/)]
6. Waltersson L, Rodby-Bousquet E. Physical activity in adolescents and young adults with cerebral palsy. *Biomed Res Int* 2017;2017:8080473 [[FREE Full text](#)] [doi: [10.1155/2017/8080473](https://doi.org/10.1155/2017/8080473)] [Medline: [29423412](https://pubmed.ncbi.nlm.nih.gov/29423412/)]
7. Lai B, Wen H, Sinha T, Davis D, Swanson-Kimani E, Wozow C, et al. The impact of COVID-19 on the lifestyles of adolescents with cerebral palsy in the Southeast United States. *Disabil Health J* 2022 Apr;15(2):101263 [[FREE Full text](#)] [doi: [10.1016/j.dhjo.2021.101263](https://doi.org/10.1016/j.dhjo.2021.101263)] [Medline: [35000878](https://pubmed.ncbi.nlm.nih.gov/35000878/)]
8. Ryan JM, Cassidy EE, Noorduyt SG, O'Connell NE. Exercise interventions for cerebral palsy. *Cochrane Database Syst Rev* 2017 Jun 11;6(6):CD011660 [[FREE Full text](#)] [doi: [10.1002/14651858.CD011660.pub2](https://doi.org/10.1002/14651858.CD011660.pub2)] [Medline: [28602046](https://pubmed.ncbi.nlm.nih.gov/28602046/)]
9. Hallal PC, Victora CG, Azevedo MR, Wells JC. Adolescent physical activity and health: a systematic review. *Sports Med* 2006;36(12):1019-1030. [doi: [10.2165/00007256-200636120-00003](https://doi.org/10.2165/00007256-200636120-00003)] [Medline: [17123326](https://pubmed.ncbi.nlm.nih.gov/17123326/)]

10. Slaman J, Roebroek M, van der Slot W, Twisk J, Wensink A, Stam H, LEARN 2 MOVE Research Group. Can a lifestyle intervention improve physical fitness in adolescents and young adults with spastic cerebral palsy? A randomized controlled trial. *Arch Phys Med Rehabil* 2014 Sep;95(9):1646-1655. [doi: [10.1016/j.apmr.2014.05.011](https://doi.org/10.1016/j.apmr.2014.05.011)] [Medline: [25067790](https://pubmed.ncbi.nlm.nih.gov/25067790/)]
11. Vogtle LK, Malone LA, Azuero A. Outcomes of an exercise program for pain and fatigue management in adults with cerebral palsy. *Disabil Rehabil* 2014;36(10):818-825. [doi: [10.3109/09638288.2013.821181](https://doi.org/10.3109/09638288.2013.821181)] [Medline: [23924251](https://pubmed.ncbi.nlm.nih.gov/23924251/)]
12. Gillett JG, Lichtwark GA, Boyd RN, Barber LA. Functional anaerobic and strength training in young adults with cerebral palsy. *Med Sci Sports Exerc* 2018 Aug;50(8):1549-1557. [doi: [10.1249/MSS.0000000000001614](https://doi.org/10.1249/MSS.0000000000001614)] [Medline: [29557839](https://pubmed.ncbi.nlm.nih.gov/29557839/)]
13. Johnson CC. The benefits of physical activity for youth with developmental disabilities: a systematic review. *Am J Health Promot* 2009;23(3):157-167. [doi: [10.4278/ajhp.070930103](https://doi.org/10.4278/ajhp.070930103)] [Medline: [19149420](https://pubmed.ncbi.nlm.nih.gov/19149420/)]
14. Park EY, Kim WH. Meta-analysis of the effect of strengthening interventions in individuals with cerebral palsy. *Res Dev Disabil* 2014 Feb;35(2):239-249. [doi: [10.1016/j.ridd.2013.10.021](https://doi.org/10.1016/j.ridd.2013.10.021)] [Medline: [24291625](https://pubmed.ncbi.nlm.nih.gov/24291625/)]
15. Verschuren O, Peterson MD, Balemans AC, Hurvitz EA. Exercise and physical activity recommendations for people with cerebral palsy. *Dev Med Child Neurol* 2016 Aug;58(8):798-808 [FREE Full text] [doi: [10.1111/dmcn.13053](https://doi.org/10.1111/dmcn.13053)] [Medline: [26853808](https://pubmed.ncbi.nlm.nih.gov/26853808/)]
16. Lai B, Lee E, Wagatsuma M, Frey G, Stanish H, Jung T, et al. Research trends and recommendations for physical activity interventions among children and youth with disabilities: a review of reviews. *Adapt Phys Activ Q* 2020 Apr 01;37(2):211-234. [doi: [10.1123/apaq.2019-0081](https://doi.org/10.1123/apaq.2019-0081)] [Medline: [32131052](https://pubmed.ncbi.nlm.nih.gov/32131052/)]
17. Pauwels C, Roren A, Gautier A, Linières J, Rannou F, Poiraudou S, et al. Home-based cycling program tailored to older people with lumbar spinal stenosis: barriers and facilitators. *Ann Phys Rehabil Med* 2018 May;61(3):144-150 [FREE Full text] [doi: [10.1016/j.rehab.2018.02.005](https://doi.org/10.1016/j.rehab.2018.02.005)] [Medline: [29499383](https://pubmed.ncbi.nlm.nih.gov/29499383/)]
18. Reedman S, Boyd RN, Sakzewski L. The efficacy of interventions to increase physical activity participation of children with cerebral palsy: a systematic review and meta-analysis. *Dev Med Child Neurol* 2017 Oct;59(10):1011-1018 [FREE Full text] [doi: [10.1111/dmcn.13413](https://doi.org/10.1111/dmcn.13413)] [Medline: [28318009](https://pubmed.ncbi.nlm.nih.gov/28318009/)]
19. Lai B, Lee E, Kim Y, Matthews C, Swanson-Kimani E, Davis D, et al. Leisure-time physical activity interventions for children and adults with cerebral palsy: a scoping review. *Dev Med Child Neurol* 2021 Feb;63(2):162-171 [FREE Full text] [doi: [10.1111/dmcn.14751](https://doi.org/10.1111/dmcn.14751)] [Medline: [33241561](https://pubmed.ncbi.nlm.nih.gov/33241561/)]
20. 2008 Physical Activity Guidelines for Americans: Be Active, Healthy, and Happy!. U.S. Department of Health and Human Services. 2008. URL: <https://health.gov/sites/default/files/2019-09/paguide.pdf> [accessed 2022-09-27]
21. Martin Ginis KA, Ma JK, Latimer-Cheung AE, Rimmer JH. A systematic review of review articles addressing factors related to physical activity participation among children and adults with physical disabilities. *Health Psychol Rev* 2016 Dec;10(4):478-494. [doi: [10.1080/17437199.2016.1198240](https://doi.org/10.1080/17437199.2016.1198240)] [Medline: [27265062](https://pubmed.ncbi.nlm.nih.gov/27265062/)]
22. Bloemen MA, Backx FJ, Takken T, Wittink H, Benner J, Mollema J, et al. Factors associated with physical activity in children and adolescents with a physical disability: a systematic review. *Dev Med Child Neurol* 2015 Feb;57(2):137-148 [FREE Full text] [doi: [10.1111/dmcn.12624](https://doi.org/10.1111/dmcn.12624)] [Medline: [25403649](https://pubmed.ncbi.nlm.nih.gov/25403649/)]
23. Jennett PA, Affleck Hall L, Hailey D, Ohinmaa A, Anderson C, Thomas R, et al. The socio-economic impact of telehealth: a systematic review. *J Telemed Telecare* 2003;9(6):311-320. [doi: [10.1258/135763303771005207](https://doi.org/10.1258/135763303771005207)] [Medline: [14680514](https://pubmed.ncbi.nlm.nih.gov/14680514/)]
24. Friedman AB, Gervasi S, Song H, Bond AM, Chen AT, Bergman A, et al. Telemedicine catches on: changes in the utilization of telemedicine services during the COVID-19 pandemic. *Am J Manag Care* 2022 Jan 01;28(1):e1-e6 [FREE Full text] [doi: [10.37765/ajmc.2022.88771](https://doi.org/10.37765/ajmc.2022.88771)] [Medline: [35049260](https://pubmed.ncbi.nlm.nih.gov/35049260/)]
25. 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/)]
26. Mitchell LE, Ziviani J, Boyd RN. A randomized controlled trial of web-based training to increase activity in children with cerebral palsy. *Dev Med Child Neurol* 2016 Jul;58(7):767-773 [FREE Full text] [doi: [10.1111/dmcn.13065](https://doi.org/10.1111/dmcn.13065)] [Medline: [26877078](https://pubmed.ncbi.nlm.nih.gov/26877078/)]
27. Lin HC, Wuang YP. Strength and agility training in adolescents with Down syndrome: a randomized controlled trial. *Res Dev Disabil* 2012;33(6):2236-2244. [doi: [10.1016/j.ridd.2012.06.017](https://doi.org/10.1016/j.ridd.2012.06.017)] [Medline: [22820064](https://pubmed.ncbi.nlm.nih.gov/22820064/)]
28. Bringas ML, Zaldivar M, Rojas PA, Martinez-Montes K, Chongo DM, Ortega MA, et al. Effectiveness of music therapy as an aid to neurorestoration of children with severe neurological disorders. *Front Neurosci* 2015 Nov 4;9:427 [FREE Full text] [doi: [10.3389/fnins.2015.00427](https://doi.org/10.3389/fnins.2015.00427)] [Medline: [26582974](https://pubmed.ncbi.nlm.nih.gov/26582974/)]
29. Wang TH, Peng YC, Chen YL, Lu TW, Liao HF, Tang PF, et al. A home-based program using patterned sensory enhancement improves resistance exercise effects for children with cerebral palsy: a randomized controlled trial. *Neurorehabil Neural Repair* 2013 Oct;27(8):684-694. [doi: [10.1177/1545968313491001](https://doi.org/10.1177/1545968313491001)] [Medline: [23757295](https://pubmed.ncbi.nlm.nih.gov/23757295/)]
30. Young HJ, Mehta TS, Herman C, Wang F, Rimmer JH. The effects of M2M and adapted yoga on physical and psychosocial outcomes in people with multiple sclerosis. *Arch Phys Med Rehabil* 2019 Mar;100(3):391-400 [FREE Full text] [doi: [10.1016/j.apmr.2018.06.032](https://doi.org/10.1016/j.apmr.2018.06.032)] [Medline: [30092206](https://pubmed.ncbi.nlm.nih.gov/30092206/)]
31. Recognizing Adolescence. World Health Organization. URL: <http://apps.who.int/adolescent/second-decade/section2/page1/recognizing-adolescence.html> [accessed 2019-09-02]

32. Global recommendations on physical activity for health. World Health Organization. 2010 Jan 1. URL: <https://www.who.int/publications/i/item/9789241599979> [accessed 2022-09-27]
33. Lai B, Cederberg K, Vanderbom KA, Bickel CS, Rimmer JH, Motl RW. Characteristics of adults with neurologic disability recruited for exercise trials: a secondary analysis. *Adapt Phys Activ Q* 2018 Oct 01;35(4):476-497. [doi: [10.1123/apaq.2017-0109](https://doi.org/10.1123/apaq.2017-0109)] [Medline: [30382753](https://pubmed.ncbi.nlm.nih.gov/30382753/)]
34. Bezyak JL, Sabella S, Hammel J, McDonald K, Jones RA, Barton D. Community participation and public transportation barriers experienced by people with disabilities. *Disabil Rehabil* 2020 Nov;42(23):3275-3283. [doi: [10.1080/09638288.2019.1590469](https://doi.org/10.1080/09638288.2019.1590469)] [Medline: [30991852](https://pubmed.ncbi.nlm.nih.gov/30991852/)]
35. Rimmer JH, Mehta T, Wilroy J, Lai B, Young HJ, Kim Y, et al. Rationale and design of a scale-up project evaluating responsiveness to home exercise and lifestyle tele-health (SUPER-HEALTH) in people with physical/mobility disabilities: a type 1 hybrid design effectiveness trial. *BMJ Open* 2019 Mar 30;9(3):e023538 [FREE Full text] [doi: [10.1136/bmjopen-2018-023538](https://doi.org/10.1136/bmjopen-2018-023538)] [Medline: [30928927](https://pubmed.ncbi.nlm.nih.gov/30928927/)]
36. Palisano RJ, Rosenbaum P, Bartlett D, Livingston MH. Content validity of the expanded and revised Gross Motor Function Classification System. *Dev Med Child Neurol* 2008 Oct;50(10):744-750 [FREE Full text] [doi: [10.1111/j.1469-8749.2008.03089.x](https://doi.org/10.1111/j.1469-8749.2008.03089.x)] [Medline: [18834387](https://pubmed.ncbi.nlm.nih.gov/18834387/)]
37. Verschuren O, Wiart L, Hermans D, Ketelaar M. Identification of facilitators and barriers to physical activity in children and adolescents with cerebral palsy. *J Pediatr* 2012 Sep;161(3):488-494. [doi: [10.1016/j.jpeds.2012.02.042](https://doi.org/10.1016/j.jpeds.2012.02.042)] [Medline: [22494875](https://pubmed.ncbi.nlm.nih.gov/22494875/)]
38. Bandura A. Health promotion by social cognitive means. *Health Educ Behav* 2004 Apr;31(2):143-164. [doi: [10.1177/1090198104263660](https://doi.org/10.1177/1090198104263660)] [Medline: [15090118](https://pubmed.ncbi.nlm.nih.gov/15090118/)]
39. King GA. CAPE/PAC Manual: Children's Assessment of Participation and Enjoyment & Preferences for Activities of Children. San Antonio, TX, USA: PsychCorp; 2004.
40. Capio CM, Sit CH, Abernethy B, Rotor ER. Physical activity measurement instruments for children with cerebral palsy: a systematic review. *Dev Med Child Neurol* 2010 Oct;52(10):908-916 [FREE Full text] [doi: [10.1111/j.1469-8749.2010.03737.x](https://doi.org/10.1111/j.1469-8749.2010.03737.x)] [Medline: [20646029](https://pubmed.ncbi.nlm.nih.gov/20646029/)]
41. Clanchy KM, Tweedy SM, Boyd R. Measurement of habitual physical activity performance in adolescents with cerebral palsy: a systematic review. *Dev Med Child Neurol* 2011 Jun;53(6):499-505 [FREE Full text] [doi: [10.1111/j.1469-8749.2010.03910.x](https://doi.org/10.1111/j.1469-8749.2010.03910.x)] [Medline: [21418195](https://pubmed.ncbi.nlm.nih.gov/21418195/)]
42. HealthMeasures. Neuro-QoL: List of Pediatric Measures. National Institutes of Health. 2019. URL: <https://www.healthmeasures.net/explore-measurement-systems/neuro-qol/obtain-and-administer-measures> [accessed 2022-09-27]
43. HealthMeasures. Neuro-QoL: Validation. National Institutes of Health. 2019. URL: <https://www.healthmeasures.net/explore-measurement-systems/neuro-qol/measure-development-research/validation> [accessed 2022-09-27]
44. Gross PH, Bailes AF, Horn SD, Hurvitz EA, Kean J, Shusterman M, cerebral palsy research network. Setting a patient-centered research agenda for cerebral palsy: a participatory action research initiative. *Dev Med Child Neurol* 2018 Dec;60(12):1278-1284 [FREE Full text] [doi: [10.1111/dmcn.13984](https://doi.org/10.1111/dmcn.13984)] [Medline: [30132826](https://pubmed.ncbi.nlm.nih.gov/30132826/)]
45. Charmaz K. *Constructing Grounded Theory*. 2nd edition. London, UK: Sage Publications; 2014.
46. Fletcher AJ. Applying critical realism in qualitative research: methodology meets method. *Int J Soc Res Methodol* 2016 Feb 29;20(2):181-194. [doi: [10.1080/13645579.2016.1144401](https://doi.org/10.1080/13645579.2016.1144401)]
47. Smith B, Sparkles AC. *Routledge Handbook of Qualitative Research in Sport and Exercise*. Milton Park, UK: Taylor & Francis; 2016.
48. Verschuren O, Ketelaar M, Gorter JW, Helder PJ, Uiterwaal CS, Takken T. Exercise training program in children and adolescents with cerebral palsy: a randomized controlled trial. *Arch Pediatr Adolesc Med* 2007 Nov;161(11):1075-1081. [doi: [10.1001/archpedi.161.11.1075](https://doi.org/10.1001/archpedi.161.11.1075)] [Medline: [17984410](https://pubmed.ncbi.nlm.nih.gov/17984410/)]
49. O'Connell NS, Dai L, Jiang Y, Speiser JL, Ward R, Wei W, et al. Methods for analysis of pre-post data in clinical research: a comparison of five common methods. *J Biom Biostat* 2017 Feb 24;8(1):1-8 [FREE Full text] [doi: [10.4172/2155-6180.1000334](https://doi.org/10.4172/2155-6180.1000334)] [Medline: [30555734](https://pubmed.ncbi.nlm.nih.gov/30555734/)]
50. Vickers AJ, Altman DG. Statistics notes: analysing controlled trials with baseline and follow up measurements. *BMJ* 2001 Nov 10;323(7321):1123-1124 [FREE Full text] [doi: [10.1136/bmj.323.7321.1123](https://doi.org/10.1136/bmj.323.7321.1123)] [Medline: [11701584](https://pubmed.ncbi.nlm.nih.gov/11701584/)]
51. Lai BW, Rimmer JH, Yates A, Jeter A, Young HJ, Thirumalai M, et al. Critical factors influencing the decision to enroll in a physical activity intervention among a predominant group of adults with spinal cord injury: a grounded theory study. *Spinal Cord* 2021 Jan;59(1):17-25 [FREE Full text] [doi: [10.1038/s41393-020-0530-7](https://doi.org/10.1038/s41393-020-0530-7)] [Medline: [32747672](https://pubmed.ncbi.nlm.nih.gov/32747672/)]
52. Lee EC, Whitehead AL, Jacques RM, Julious SA. The statistical interpretation of pilot trials: should significance thresholds be reconsidered? *BMC Med Res Methodol* 2014 Mar 20;14:41 [FREE Full text] [doi: [10.1186/1471-2288-14-41](https://doi.org/10.1186/1471-2288-14-41)] [Medline: [24650044](https://pubmed.ncbi.nlm.nih.gov/24650044/)]
53. Thomson SB. Sample size and grounded theory. Thomson, SB (2010). Grounded theory-sample size. *J Admin Gov* 2010 Sep 14;5(1):45-52. [doi: [10.4135/9781412959193.n112](https://doi.org/10.4135/9781412959193.n112)]
54. Young HJ, Mehta T, Herman C, Baidwan NK, Lai B, Rimmer JH. The effects of a movement-to-music (M2M) intervention on physical and psychosocial outcomes in people poststroke: a randomized controlled trial. *Arch Rehabil Res Clin Transl* 2021 Dec;3(4):100160 [FREE Full text] [doi: [10.1016/j.arrct.2021.100160](https://doi.org/10.1016/j.arrct.2021.100160)] [Medline: [34977542](https://pubmed.ncbi.nlm.nih.gov/34977542/)]

55. Colquitt G, Kiely K, Caciula M, Li L, Vogel RL, Moreau NG. Community-based upper extremity power training for youth with cerebral palsy: a pilot study. *Phys Occup Ther Pediatr* 2020;40(1):31-46. [doi: [10.1080/01942638.2019.1636924](https://doi.org/10.1080/01942638.2019.1636924)] [Medline: [31282292](https://pubmed.ncbi.nlm.nih.gov/31282292/)]

## Abbreviations

**ANCOVA:** analysis of covariance

**CAPE:** Children's Assessment of Participation and Enjoyment

**CAPE-APR:** Children's Assessment of Participation and Enjoyment Active Physical Recreation

**CONSORT:** Consolidated Standards of Reporting Trials

**CP:** cerebral palsy

**GMFCS:** Gross Motor Function Classification System

**GMFCS-E&R:** expanded and revised version of the Gross Motor Function Classification System

**LTPA:** leisure-time physical activity

**M2M:** Movement-to-Music

**PI:** principal investigator

**RCT:** randomized controlled trial

**REDCap:** Research Electronic Data Capture

**WC:** waitlist control

*Edited by A Mavragani; submitted 29.12.21; peer-reviewed by F Babatunde, J Dubu; comments to author 09.07.22; revised version received 24.08.22; accepted 12.09.22; published 28.10.22.*

*Please cite as:*

*Lai B, Vogtle L, Young R, Craig M, Kim Y, Gowe M, Swanson-Kimani E, Davis D, Rimmer JH*

*Telehealth Movement-to-Music to Increase Physical Activity Participation Among Adolescents With Cerebral Palsy: Pilot Randomized Controlled Trial*

*JMIR Form Res* 2022;6(10):e36049

URL: <https://formative.jmir.org/2022/10/e36049>

doi: [10.2196/36049](https://doi.org/10.2196/36049)

PMID: [36306154](https://pubmed.ncbi.nlm.nih.gov/36306154/)

©Byron Lai, Laura Vogtle, Raven Young, Mary Craig, Yumi Kim, Marissa Gowe, Erin Swanson-Kimani, Drew Davis, James H Rimmer. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 28.10.2022. 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

# Social Determinants of Health and Diabetes-Related Distress in Patients With Insulin-Dependent Type 2 Diabetes: Cross-sectional, Mixed Methods Approach

Natalie K Levy<sup>1</sup>, MD; Agnes Park<sup>1</sup>, BSc; Daniela Solis<sup>1</sup>, MPH; Lu Hu<sup>2</sup>, PhD; Aisha T Langford<sup>2</sup>, PhD; Binhuan Wang<sup>2</sup>, PhD; Erin S Rogers<sup>2</sup>, DPhil

<sup>1</sup>Department of Medicine, New York University Grossman School of Medicine, New York, NY, United States

<sup>2</sup>Department of Population Health, New York University Grossman School of Medicine, New York, NY, United States

**Corresponding Author:**

Natalie K Levy, MD

Department of Medicine

New York University Grossman School of Medicine

462 First Ave

New York, NY, 10016

United States

Phone: 1 212 263 8924

Email: [natalie.levy@nyulangone.org](mailto:natalie.levy@nyulangone.org)

## Abstract

**Background:** Social determinants of health (SDOH) refer to the social, economic, and psychosocial conditions that influence health. Lower levels of SDOH factors including income, education, and employment are associated with a higher prevalence of diabetes, poorer glycemic control, and increased diabetes-related mortality. Few studies have conducted a comprehensive evaluation of multiple SDOH factors in a population with type 2 diabetes mellitus (T2DM).

**Objective:** This study aimed to identify the range of SDOH challenges—including diabetes-related distress—that impact patients with insulin-dependent diabetes at an urban safety-net clinic using the 5-domain SDOH framework developed by the Healthy People 2020 initiative.

**Methods:** The pilot study used a cross-sectional, mixed methods approach. Participants were recruited from 3 programs within a general internal medicine clinic that provides ambulatory care for patients with uncontrolled T2DM. We administered an investigator-developed SDOH survey based on the Healthy People 2020 framework and the validated Diabetes Distress Scale (DDS), which assesses 4 domains of diabetes-related distress. One-on-one interviews were conducted to gain in-depth information about challenges.

**Results:** In total, 57 participants had an average hemoglobin A<sub>1c</sub> level of 11.0% (SD 2.6%). Overall, 92% (52/57) of participants had a barrier in at least one SDOH domain. SDOH challenges were most commonly reported in the domain of Health and Health Care (84%, 48/57), followed by Economic Stability (54%, n=31), Neighborhood and Built Environment (53%, n=30), Education and Health Literacy (47%, n=27), and Social and Community context (37%, n=21). The mean overall DDS score was 2.09 (SD 0.84), where scores of  $\geq 2$  indicate distress. Further, 79% (45/57) of participants had at least moderate diabetes-related distress in one of the 4 DDS domains. General themes that emerged from participant interviews included job interference with healthy behaviors, concerns about burdening others, challenges communicating with providers, and difficulty getting appointments in a timely manner.

**Conclusions:** We found high levels of SDOH barriers across all 5 domains of the Center for Disease Control and Prevention's Healthy People 2020 framework, including significant levels of diabetes-related distress. Future programs to address SDOH barriers in patients with uncontrolled insulin-dependent diabetes should consider screening for and focusing on a wide range of challenges.

(JMIR Form Res 2022;6(10):e40164) doi:[10.2196/40164](https://doi.org/10.2196/40164)

**KEYWORDS**

social determinants of health; income; socioeconomic; cross sectional; insulin; diabetic; HbA<sub>1c</sub>; barrier; diabetes-related distress; type 2 diabetes; ambulatory care; healthcare; health care; distress; epidemiology; T2DM; diabetes

## Introduction

Social determinants of health (SDOH) are broadly defined as the circumstances in which people are born, live, and work [1,2]; they include the social, economic, and psychosocial conditions that influence health [1-4]. In 2010, the US Department of Health and Human Services published the goals for its Healthy People 2020 (HP2020) initiative, which included a new section on SDOH; this framework organizes key issues into one of 5 domains: Health and Health Care, Economic Stability, Neighborhood and Built Environment, Education, and Social and Community Context [1].

Over 30 million Americans are estimated to have diabetes, and another 84 million are estimated to have prediabetes [5,6]. Diabetes increases the risk of heart disease, stroke, kidney failure, blindness, and lower limb amputation [5,6]. In 2017, the estimated cost of diagnosed diabetes in the United States was US \$327 billion [5,6]. Studies examining the connections between SDOH and diabetes have shown that lower levels of SDOH factors including income, education, and employment are associated with a higher prevalence of diabetes [7], increased diabetes-related mortality [8], and poorer glycemic control [9-11].

To date, most diabetes-related SDOH literature has generally assessed a narrow subset of SDOH challenges at one time [12-15]. A 2014 systematic review evaluating the impact of SDOH on outcomes for type 2 diabetes mellitus (T2DM) included articles whose focus was mostly clustered around a single SDOH domain: 28 focused on Health and Health Care, 17 on Social and Community Context, 11 on Economic Stability, 3 on Neighborhood and Built Environment, and 1 on Education [12]. Similarly, studies included in a 2021 review by the American Diabetes Association mainly assessed single SDOH domains [15]. It is less often that a single study evaluates multiple SDOH factors in the same population [9,14].

While not an individual category in the HP2020 SDOH framework, psychosocial distress is also considered an SDOH factor [2,3,16]. Diabetes-related psychosocial distress (hereinafter referred to as “diabetes-related distress” [DRD]) is associated with poor glycemic control and self-care [2,17-19]. Research estimates the prevalence of DRD in the United States to be 18%-48% [17,19-22]. However, the average hemoglobin A<sub>1c</sub> (HbA<sub>1c</sub>) levels in the populations studied ranged from 6.7% to 9.9% [17,19-22], where a level of <8% is generally considered to indicate good control and that of ≥11% is considered poor for individuals with diabetes. The prevalence of DRD in a population of patients with T2DM with poorer glycemic control is less studied.

To fill SDOH-related knowledge gaps for patient populations with T2DM, this pilot study sought to distinguish itself in several ways. First, this study evaluated SDOH barriers across a broad range of SDOH domains. Second, instead of looking at larger

upstream SDOH factors (eg, income, education, and employment), this study sought to evaluate how SDOH challenges affect patients and their diabetes management on a day-to-day basis. Finally, we evaluated SDOH barriers and DRD in a population likely to have poorer glycemic control than previously reported.

## Methods

### Ethical Considerations

Study approval was granted by the institutional review board of the New York University Grossman School of Medicine (NYUGSoM; s17-01553). All participants provided written informed consent in English or Spanish in the presence of a bilingual study team member.

### Study Setting

This study was conducted at Bellevue Hospital, New York City, New York, which is affiliated with NYUGSoM [23]. Bellevue is the oldest public hospital in the United States and is part of NYC Health + Hospitals, the largest municipal health care system in the United States. Bellevue provides safety-net care to a diverse population of over 30,000 patients—approximately one-third of whom are uninsured—through its Adult Primary Care Center.

### Participants and Recruitment

From January 2018 to November 2018, patients with insulin-dependent T2DM and an HbA<sub>1c</sub> level of ≥7.0% and who speak English or Spanish were recruited from 3 settings within Bellevue Hospital’s Adult Primary Care Center: (1) the High A1C Clinic, a referral program imbedded within primary care that focuses on the management of patients with poorly controlled diabetes; (2) the Diabetes Group Medical Visit, a 4-week, multidisciplinary program that provides comprehensive self-management education; and (3) the Mobile Insulin Titration Intervention (MITI) Program, a telehealth service that uses basic SMS text messages and phone calls to adjust basal insulin remotely [24-26]. To address selection bias, all physicians of the High A1C Clinic and Diabetes Group Medical Visit as well as the MITI enrollment coordinator were informed about the study and eligibility criteria (ie, insulin-dependent T2DM, an HbA<sub>1c</sub> level of ≥7.0%, and speaking English or Spanish). Physicians and the MITI enrollment coordinator were asked to provide information about the study to every patient who met the eligibility criteria. Frequent reminders were provided each week during physician huddles to encourage recruitment efforts.

Upon receiving informed consent, bilingual research assistants administered the study questionnaires and conducted the study interview. All study procedures were completed in one sitting, and participants were provided US \$30 in cash upon completion of study activities.

## Study Instruments

### *The SDOH Questionnaire*

#### Overview

We developed a 15-item questionnaire that assessed for SDOH factors across the 5 HP2020 SDOH domains that could affect the health and well-being of patients with T2DM. The majority of questions were adapted from the Protocol for Responding to and Assessing Patients' Assets, Risks, and Experiences (PRAPARE) survey, a standardized tool to screen for SDOH [27]. The SDOH questionnaire also examined participants' levels of physical activity. Questions were categorized into 6 sections.

#### Economic Stability

One question was adapted from PRAPARE to assess if participants had been unable to pay for an essential item when they really needed it in the past year. In total, 12 items were listed in the choices, including food, medication, and medical visits.

#### Education and Health Literacy

The following questions were adapted from a health literacy scale developed at the University of California, Los Angeles [28]: "How often do you have trouble: 1) explaining health concerns to a doctor or nurse; 2) understanding what a doctor or nurse says; 3) understanding written instructions on medication labels; 4) completing medical forms?" Participants had the option to respond with "often," "sometimes," and "never."

#### Social and Community Context

Two questions were adapted from the PRAPARE survey: "How often do you see or talk to people that you care about and feel close to?" and "Who are the people you speak to when you are feeling stressed?"

#### Health and Health Care

One question was adapted from the PRAPARE survey to assess whether lack of transportation prevented the patient from consulting a doctor within the past year. We assessed for factors beyond the PRAPARE survey, such as "lack of insurance," "can't get an appointment," and "hard to miss work."

#### Neighborhood and Built Environment

Two questions were adapted from the PRAPARE survey to assess for housing situation and housing security. Additional questions were added by the research team to assess for the availability and quality of fruits and vegetables in the neighborhood and the availability of safe places to exercise.

#### Physical Activity

We adapted questions from the US National Health Interview Survey (NHIS) [29] to ask if participants engaged in physical activity outside of any paid job and the frequency, duration, and vigorousness of the physical activity.

### *The DRD Scale*

DRD was measured using the validated 17-question Diabetes Distress Scale (DDS) [30,31]. The DDS measures distress in 4

areas: Emotional Burden, Regimen Distress, Interpersonal Distress, and Physician Distress. Responses are scored on a Likert scale. Total scores of 2-2.9 indicate moderate distress and those of  $\geq 3$  indicate severe distress. For this study, a total score of  $\geq 2$  was considered the threshold for describing patients as having distress, as advised by the DDS and prior studies [18,21,32,33].

### *Semistructured Interview Guide*

The first 31 participants who completed the 2 study questionnaires were also asked to complete a semistructured interview. This sample size was determined by the number of participants who were interviewed until thematic saturation was reached. The research team developed an interview guide to gather in-depth information about the SDOH challenges patients reported on the SDOH questionnaire and DDS. The questions were written to allow participants to expand upon the SDOH challenges they identified and how those challenges impacted their diabetes.

### **Electronic Medical Record Data Abstraction**

Sociodemographic data and HbA<sub>1c</sub> values were abstracted from electronic medical records. HbA<sub>1c</sub> values were the most recent values taken within 3 months from the date of enrollment.

### **Statistical Analyses**

The SDOH quantitative data are reported as the number and percentage of participants who had challenges within each HP2020 domain, its subcategories, and the number and percentage of participants who had at least one barrier across a certain number of HP2020 domains. DDS data are reported such that for each individual DDS domain, the average score as well as the number and percentage of patients having scores in at least the moderate ( $\geq 2$ ) and the severe ( $\geq 3$ ) range are reported. These data are also reported by age, language, and gender. The number and percentage of patients who have at least moderate distress across a certain number of domains is also reported. Spearman and Pearson correlations were used to measure the level of association between HbA<sub>1c</sub> values and SDOH data (number of challenge domains, presence or lack of challenges in individual domains), HbA<sub>1c</sub> and DDS data (overall and domain DDS scores, number of domains with distress, and number of domains with severe distress), and overall DDS scores and number of SDOH challenge domains.

For quantitative data, descriptive statistics were used to summarize demographic characteristics and other factors. Continuous variables (eg, age and HbA<sub>1c</sub>) are described as means and SDs, and categorical variables (such as gender) are summarized as frequencies and percentages. For qualitative data, analyses were conducted on Atlas.ti (version 8.1; ATLAS.ti Scientific Software Development GmbH). Interviews were transcribed verbatim, and Spanish-language transcripts were translated into English. In total, 3 investigators used both deductive and inductive (grounded theory) approaches to code the interviews. They created an initial codebook that included 5 main SDOH domains from the interview guide, domain and code definitions, and inclusion and exclusion criteria. They then independently coded 4 transcripts, followed by discussions of

coding agreement and disagreement and an updating of the codebook to include open coding of SDOH subcodes. When the codebook was complete, they independently coded the remaining transcripts, with co-coding of every fifth transcript to ensure intercoder reliability. Once coding was complete, they met to identify themes.

## Results

### Participant Characteristics

In total, 61 patients were screened for eligibility. Of those, 60 were eligible to participate. In total, 3 declined enrollment owing to scheduling difficulties. In total, 57 patients participated in the study. The mean age of participants was 54.8 (SD 12.1) years, 56% (n=32) were Hispanic, 28% (n=16) were uninsured, and the average HbA<sub>1c</sub> level was 11.0% (SD 2.6%). Additional characteristics are provided in [Table 1](#).

**Table 1.** Participant demographics (N=57).

Demographics	Values
Age (years), mean (SD)	54.8 (12.1)
<b>Gender, n (%)</b>	
Male	35 (61)
Female	22 (39)
<b>Language</b>	
English	33 (58)
Spanish	24 (42)
<b>Race and ethnicity, n (%)</b>	
White	6 (11)
Black	14 (25)
Hispanic	32 (56)
Asian	5 (9)
<b>Insurance status, n (%)</b>	
Medicaid	26 (46)
Medicare	6 (11)
Medicaid + Medicare	9 (16)
Uninsured	16 (28)
Hemoglobin A <sub>1c</sub> levels (%), mean (SD)	11.0 (2.6)

### SDOH Questionnaire Data

Data were obtained on the percentage of patients who had challenges within each of the 5 major HP2020 domains in the past year. All participants provided responses to every question on the SDOH questionnaire. Overall, 84% (48/57) of participants had a barrier in Health and Health Care in the past year, including 49% (28/57) who were unable to get an appointment. In total, 56% (32/57) of participants had a barrier in Economic Stability and reported  $\geq 1$  instance when they were unable to pay for an essential item when it was truly needed. This included 23% (13/57) of participants who were unable to pay rent and 23% (13/57) who were unable to pay for healthy food. Overall, 53% (30/57) of participants had a barrier in their Neighborhood and Built Environment including 32% (18/57) of participants with no or only a very small number of stores that sell produce in their neighborhood and 16% (9/57) having no safe place to exercise. Overall, 47% (27/57) of participants had an Education and Health Literacy challenge, including 42% (24/57) of those

who had difficulty explaining their health concerns to a doctor and 18% (10/57) of those who had trouble reading a medication label. Approximately one-third of participants had challenges in the domain of Social and Community Context, including 37% (21/57) of participants who shared that they speak to or see someone who they care about  $\geq 2$  times per week, and 30% (17/57) of them speak to no one when stressed. The presence or lack of challenges in the domain of Health and Health Care was significantly correlated with HbA<sub>1c</sub> values ( $\rho=0.32$ ,  $P=.02$ ). Correlation coefficients for all other domains were  $<0.5$  and nonsignificant ( $P>.05$ ).

Across the 5 HP2020 domains, 92% (52/57) of patients had a barrier in at least one domain, 73% (42/57) in at least 2 domains, 62% (35/57) in at least 3 domains, and 44% (25/57) in at least 4 domains. A further breakdown of challenges within each main domain are provided in [Table 2](#). There was no significant correlation between the number of SDOH challenge domains and HbA<sub>1c</sub> values ( $\rho=0.23$ ,  $P=.09$ ).



**Table 2.** Frequency of social determinants of health barriers reported by the study participants (N=57).

Domain	Participants, n (%)
<b>Health and Health Care</b>	
<b>Had <math>\geq 1</math> barrier to health care access in the past year</b>	84 (48)
Could not get appointment	49 (28)
Forgot the appointment	40 (23)
Difficult to miss work	21 (12)
Cost or lack of insurance	18 (10)
Transportation difficulties	14 (8)
Childcare difficulties	7 (4)
Other	12 (7)
<b>Economic Stability</b>	
<b>Unable to pay for <math>\geq 1</math> essential items in the past year</b>	56 (32)
Rent	23 (13)
Healthy food	23 (13)
Any food	18 (10)
Medical visit	16 (9)
Medical supplies	16 (9)
Utilities	16 (9)
Phone	14 (8)
Medication	14 (8)
Clothing	12 (7)
Mortgage	4 (2)
Childcare	0 (0)
<b>Neighborhood and Built Environment</b>	
<b>Has <math>\geq 1</math> issue with the built environment</b>	53 (30)
Worried about losing housing	19 (11)
No stores or only a small number of stores that sell produce in the neighborhood	30 (17)
Unsatisfied with the quality of produce available in the neighborhood	9 (5)
Do not have safe places to exercise in the neighborhood	16 (9)
<b>Education and Health Literacy</b>	
Has $\geq 1$ issue with health literacy	47 (27)
Has difficulty explaining health concerns to a doctor or nurse	42 (24)
Has difficulty understanding what a doctor or nurse is saying	19 (11)
Has difficulty filling out medical forms	19 (11)
Has difficulty understanding medication labels	18 (10)
<b>Social and Community Context</b>	
Receives social support $\leq 2$ times a week	37 (21)
Has nobody to speak to when stressed	30 (17)

### DDS Data

DDS scores were calculated for 56 participants and are listed in [Table 3](#). One participant did not complete the DDS owing to challenges with understanding the questions. The largest individual domain of DRD on the DDS was Emotional Burden

(mean 2.62, SD 1.37), with 64% of participants having at least moderate distress (ie, score  $\geq 2$ ) and 34% having severe distress (ie, score  $\geq 3$ ). The second most common distress domain was Regimen Distress (mean 2.41, SD 1.13), with 57% (32/57) of participants having at least moderate distress and 30% having severe distress. The areas of Interpersonal Distress and Physician

Distress were present but were generally less of a challenge for participants, with means of 1.68 (SD 0.91) and 1.33 (SD 0.55), respectively. Emotional Burden and Regimen Distress remained the largest domains of DRD after analysis by age, language, and gender. As with SDOH, participants had DRD across

multiple domains: 79% (45/57) of participants had at least moderate distress in  $\geq 1$  domain, 52% (30/57) in  $\geq 2$  domains, and 36% (21/57) in  $\geq 3$  domains. There were no significant correlations between HbA<sub>1c</sub> value and overall score ( $r=-0.006$ ,  $P=.96$ ) and individual domain scores.

**Table 3.** Mean participant Diabetes Distress Scale scores<sup>a</sup> and categorization by age, language, and gender (N=56<sup>b</sup>).

Variables	Emotional Burden	Physician Distress	Regimen Distress	Interpersonal Distress	Overall Score
Score, mean (SD)	2.62 (1.37)	1.33 (0.55)	2.41 (1.13)	1.68 (0.91)	2.09 (0.84)
Ages 29-40 years (n=8), mean	3.10	1.38	2.34	2.50	2.39
Ages 41-59 years (n=27), mean	2.71	1.50	2.49	1.52	2.15
Age $\geq 60$ years (n=21), mean	2.34	1.18	2.24	1.57	1.89
English speakers (n=32), mean	2.51	1.27	2.41	1.74	1.98
Spanish speakers (n=24), mean	2.62	1.42	2.42	1.60	2.04
Male (n=35), mean	2.47	1.29	2.13	1.69	1.95
Female (n=21), mean	2.87	1.40	2.85	1.64	2.32
Participants with at least moderate distress, n (%)	36 (64)	8 (14)	32 (57)	21 (38)	26 (46)

<sup>a</sup>2.0-2.9: moderate distress;  $\geq 3$ : high distress

<sup>b</sup>One participant was unable to complete the Diabetes Distress Scale.

### Participant Interview Data

Themes identified from participant interviews were grouped by SDOH domain and are presented alongside representative quotes in [Table 4](#).

Within the domain of Economic Stability, participants regularly experienced difficulty affording items that were necessary to manage their diabetes and care for their health. Consequences included delayed treatment and the foregoing of certain items, despite participants recognizing the impact these would have on their health. Participants also reported challenges stemming from their jobs. Many worked long hours (eg, late nights, 7 days a week) and had unpredictable schedules. As a result, they had little time or energy to cook and exercise. In addition, some participants reported difficulty taking time off from work, which impacted their ability to attend doctor's visits.

Within the Social and Community Context domain, more than half of participants reported feeling like they did not have a support system for their diabetes. In particular, these individuals felt that they lacked people in their personal lives who understood what diabetes was or who they could talk to about their experience living with diabetes. Some participants noted that their family or friends impeded healthy eating efforts by offering them unhealthy foods. Participants who had friends and family members with diabetes shared that they provided a source of knowledge and support for coping, regularly checked in with the participant, and helped participants with their

diabetes-related care. Many participants described experiencing a personal emotional toll from their diabetes. Finally, participants did not want to burden their loved ones with their diabetes and instead wanted to try to deal with their health on their own.

Within the domain of Neighborhood and Built Environment, several participants reported having limited access to healthy foods in their neighborhood, particularly fresh vegetables. Participants were frequently exposed to advertisements for fast foods, which made it difficult to resist unhealthy eating behaviors.

Within the Education and Health Literacy domain, some participants reported difficulty communicating with their providers. This included the following experiences with providers: the providers spoke too quickly, spoke too coldly, used medical terminology that the participants did not understand, or did not fully explain the participants' condition and how to take care of it.

Lastly, within the Health and Health Care domain, participants reported difficulty getting health care appointments. This impacted their diabetes because their physicians wanted them to return for follow-up visits every 2-3 months but it often took much longer to get an appointment. Participants also experienced significant delays and frustrations owing to perceived disorganization and a lack of communication among different departments within the health care system.

**Table 4.** Social determinants of health and themes in the Diabetes Distress Scale identified during semistructured interviews.

Social determinants of health domain and themes	Example quotes
<b>Economic Instability</b>	
Difficulty affording health-related needs	<ul style="list-style-type: none"> <li>“Sometimes the way you eat is based on survival because it’s like \$10 for a salad right? But it’s like \$5 for a steak sandwich. So it’s like when you making financial choices and you trying to eat based on your budget. It conflicts because you know that you can’t have a Philly cheesesteak because of your diabetes but then again, you gotta eat something...So it’s about survival then.” (M117)</li> <li>“My income is not enough. I have rent, food, everything, so I just cannot pay for the medical visit.” (M109)</li> </ul>
Job interferes with healthy behavior	<ul style="list-style-type: none"> <li>“I just don’t have the time [to exercise]. I work 11-12 hours daily, every day. I come home tired, I take a shower and go to sleep.” (M123)</li> <li>Interviewer: What are some challenges you have in being able to see your doctor for your diabetes? “Financially, permission from work, sometimes I don’t have money. It’s all together.” (M114)</li> </ul>
<b>Social and Community Context</b>	
Limited or no support for diabetes-related health	<ul style="list-style-type: none"> <li>“I had some family staying with me and because the type of food [they were cooking], I think it contributed to [my blood sugar getting out of control].” (M109)</li> <li>“Sometimes, people [don’t] understand diabetic people. Sometimes, I’m visit my friend, she’s having a birthday and gives me cake. Sometimes [she says], ‘You eat, you eat.’...Sometimes, people [don’t] understand.” (M105)</li> </ul>
Loved ones with diabetes are supportive of diabetes-related health	<ul style="list-style-type: none"> <li>“If I feel a concern, I would just always talk to my father, because he’s been with diabetes and he’ll give me an example [of what to do]... ‘cause he’s been through the same thing.” (M106)</li> <li>“I get a call from my cousin every day, so sometimes she’ll ask me if I did check my sugar.” (M107)</li> </ul>
Participant feels an emotional toll associated with having diabetes	<ul style="list-style-type: none"> <li>“You know you’re supposed to be doing better but it’s like, you know, how can you? It does put you in a—um, a clouded mental space. [...] Like ,you know you be beat. You don’t wanna beat yourself up but it’s like, how do you make better choices?” Interviewer: What is the emotion you feel the most often when you struggle with your diabetes? “Frustration and sadness.” (M104)</li> </ul>
Participant does not want to burden others	<ul style="list-style-type: none"> <li>“I have got plenty of [family], but, you know, sometimes everybody has their own thing to do. You don’t want to burden nobody with your things so try to get your thing over by yourself.” (M118)</li> </ul>
<b>Neighborhood and Built Environment</b>	
Environmental exposure to unhealthy food	<ul style="list-style-type: none"> <li>“[It’s harder] when watching TV and sometimes when commercials come on with food...psychologically it makes me hungry and I start eating the wrong things when I see commercials.” (M119)</li> </ul>
Limited access to healthy food in the neighborhood	<ul style="list-style-type: none"> <li>Interviewer: What are some reasons that you find eating healthy challenging? “It’s just finding a place to buy the vegetables.” (M108)</li> </ul>
<b>Education and Health Literacy</b>	
Challenges communicating with providers	<ul style="list-style-type: none"> <li>“There are some words I still do not know. It might [help if my doctor would] explain things more slowly or that is when he speak sometimes talking is very fast and very cold, indifferent.” (M116)</li> </ul>
Language discordance	<ul style="list-style-type: none"> <li>“The diabetes doctor for example speaks not an ounce of Spanish. But she also doesn’t use the translator phone.” (M135)</li> </ul>
<b>Health and Health Care</b>	
Cannot get appointments in a timely manner	<ul style="list-style-type: none"> <li>“I haven’t seen my primary in five months, so that’s half a year gone by there. You call and there’s no appointments available.” (M130)</li> </ul>
Disorganization of the health care system	<ul style="list-style-type: none"> <li>“They shouldn’t be giving me appointments that are too close together, because then I might miss one or literally run from one appointment to the next and get late. 15 minutes late, they don’t want to see you. I’ve seen people get frustrated...extremely to the point where they’re yelling and screaming.” (M130)</li> </ul>

## Discussion

### Principal Findings

This study used a mixed methods approach to identify challenges across the 5 domains of the Healthy People 2020 SDOH framework and the presence of DRD, measured by the validated DDS, in patients with insulin-dependent T2DM seeking care at a safety-net hospital. Participants had substantial SDOH barriers, which created regular and significant challenges for participants wanting to better manage their diabetes. Data demonstrated that the challenges were extensive, impacting nearly every aspect of participants' day-to-day lives—from difficulties affording health-related needs to encountering limited understanding of diabetes among loved ones—as well as their interactions with the health care system and health care teams. Given the association between glycemic control and DRD, it is noteworthy, but not surprising, that 46% of participants also had DRD. To our knowledge, our participant sample has the highest average HbA<sub>1c</sub> levels (mean 11.0%, SD 2.6%) out of any cohort for which DRD prevalence has been reported [17-22,32-37].

The lack of significant correlations between HbA<sub>1c</sub> levels and SDOH data (with the exception of the Health and Health Care SDOH domain) and DDS data may, in part, be due to the fact that all participants in our study were insulin-dependent. Generally, patients with T2DM are started on insulin when hyperglycemia is severe or when other therapies have not been successful at lowering blood sugar levels to a well-controlled range [38]. It is possible that correlations are more significant up until the point when diabetes is severe enough to warrant insulin (ie, insulin may subsequently lower blood sugar to more optimal levels but the conditions that contributed to needing insulin may still persist).

Prior studies have often focused on big-picture “upstream” SDOH factors (eg, education, poverty, and employment). For example, data from the 1990-2000 US NHIS combined with that from the NHIS Linked Mortality Files through 2002 showed that having less than a high school education and having a family income below the poverty line were each associated with a 2-fold higher mortality from diabetes, compared to adults with a college degree or higher or with those with the highest family income, respectively [8]. Data from the 2015 NHIS found that having less than a high school education, having a family income of less than US \$35,000 a year, and being “not employed but having worked previously” were each associated with about twice the risk of having diabetes compared to those who had graduated college, had a family income of US \$100,000 or greater, or who were employed full time [7]. Building upon these data, our study looked within the broad categories and drilled down to focus on the day-to-day effects of these SDOH challenges that are not typically reported. Our findings shed light on the potential mechanisms by which the broad SDOH factors studied previously may impact one's ability to manage their diabetes. For example, poverty can reduce one's ability to afford healthy food and live in safe neighborhoods with access to fresh produce, while increasing one's reliance on busy safety-net health systems with limited appointment availability.

By comprehensively assessing for SDOH challenges, we found that patients had barriers across several domains simultaneously: 73% (42/57) of patients having barriers across  $\geq 2$  domains, 62% (35/57) having barriers across  $\geq 3$  domains, and 44% (25/57) having barriers across  $\geq 4$  domains. These findings suggest that prior studies that assessed a single or small number of SDOH domains likely underestimated the prevalence of SDOH challenges in their participant samples. Clinical settings that rely on single-domain SDOH screeners (eg, food insecurity and risk for homelessness) also risk underestimating the social needs felt by a significant portion of their patients. Therefore, future SDOH research and SDOH-driven clinical program development should use comprehensive screening tools, such as the PRAPARE tool [27], the Accountable Health Communities Health-Related Social Needs Screening Tool [39], or the Health Leads Social Needs Screening Tool [40].

This work has clinical significance. As the movement to recognize the importance of SDOH grows stronger, not only will policy makers need to understand and focus on the big-picture SDOH (eg, poverty and education level) but local health systems gearing up to help patients with T2DM overcome such barriers need to know how these upstream SDOH challenges effects patients' daily lives. This work details such “day to day” challenges. In addition, the evaluation of barriers in multiple domains reinforces that patients with T2DM can have a wide range of concomitant challenges. Diabetes team members charged with helping at-risk patients will need to be prepared to address challenges that span across multiple SDOH domains. Future research is needed to help health systems identify best practices in addressing SDOH challenges identified in our study, which are not typically within the scope of health care, such as helping patients overcome work-related barriers to healthy eating or exercising (eg, working long hours), obtaining healthy food when living in neighborhoods with no fresh produce, or engaging loved ones in discussions about the need for support. In addition, this work can serve as an example to the entire diabetes team, that high levels of DRD may accompany such patients. Given that both SDOH and DRD are associated with glycemic control, both need to be recognized and addressed as part of care for the whole patient.

### Limitations

There are several limitations to this study. First, it was conducted at a single site—an ambulatory clinic at a safety-net hospital in New York City—and thus may not be generalizable to other settings. Second, in an effort to target a population with a high likelihood of SDOH challenges and DRD, the study sample was limited to patients with insulin-dependent T2DM. Our findings may not generalize to other patient groups, and we are not able to determine whether SDOH barriers are more severe in our study sample than in patients with T2DM who are not insulin-dependent. Third, owing to the study's cross-sectional design, we are unable to assess changes in participants' SDOH profiles over a period of time or establish relationships between SDOH and diabetes control. Lastly, while efforts were made to reduce bias in our recruitment methods and all participants responded to all questionnaire items, except for one who did not complete the DDS, our participants who opted to be part of our study may be different from those who did not participate.

## Conclusions

This pilot study found high levels of SDOH barriers across all 5 domains of the Center for Disease Control and Prevention's Health People 2020 SDOH framework, including significant

levels of DRD. Future programs to address SDOH barriers in patients with uncontrolled insulin-dependent diabetes in safety-net programs are needed and should be designed to screen for and address a wide range of challenges.

## Acknowledgments

We wish to thank the Health Homes, Social Work, and Care Management teams at Bellevue Hospital for their contributions to this project. We wish to thank the participants of our study, without whom this work would not have been possible. We would also like to acknowledge the staff of the Adult Primary Care Centers at Bellevue Hospital, as well as New York City Health and Hospitals, who provided tremendous support to the program. This work was supported by the Aetna Foundation, a national foundation based in Hartford, Connecticut, that supports projects to promote wellness, health and access to high-quality health care for everyone. The views presented here are those of the authors and not necessarily those of the Aetna Foundation, its directors, officers or staff.

## Conflicts of Interest

None declared.

## References

1. The Vision, Mission, and Goals of Healthy People 2020. Healthy People. URL: <https://www.healthypeople.gov/sites/default/files/HP2020Framework.pdf> [accessed 2022-08-15]
2. Asuzu CC, Walker RJ, Williams JS, Egede LE. Pathways for the relationship between diabetes distress, depression, fatalism and glycemic control in adults with type 2 diabetes. *J Diabetes Complications* 2017 Jan;31(1):169-174 [FREE Full text] [doi: [10.1016/j.jdiacomp.2016.09.013](https://doi.org/10.1016/j.jdiacomp.2016.09.013)] [Medline: [27746088](https://pubmed.ncbi.nlm.nih.gov/27746088/)]
3. Closing the gap in a generation: health equity through action on the social determinants of health - Final report of the commission on social determinants of health. World Health Organization. 2008 Aug 27. URL: <https://www.who.int/publications/i/item/WHO-IER-CSDH-08.1> [accessed 2022-08-15]
4. Fair Society, Healthy Lives: The Marmot Review. Strategic Review of Health Inequalities in England post-2010. URL: <https://www.instituteofhealthequity.org/resources-reports/fair-society-healthy-lives-the-marmot-review/fair-society-healthy-lives-full-report-pdf.pdf> [accessed 2022-08-15]
5. Statistics About Diabetes. American Diabetes Association. URL: <http://www.diabetes.org/diabetes-basics/statistics/> [accessed 2022-08-15]
6. Diabetes and Prediabetes. Centers for Disease Control and Prevention. URL: <https://www.cdc.gov/chronicdisease/resources/publications/factsheets/diabetes-prediabetes.htm> [accessed 2022-08-15]
7. Singh GK, Daus GP, Allender M, Ramey CT, Martin EK, Perry C, et al. Social Determinants of Health in the United States: Addressing Major Health Inequality Trends for the Nation, 1935-2016. *Int J MCH AIDS* 2017;6(2):139-164 [FREE Full text] [doi: [10.21106/ijma.236](https://doi.org/10.21106/ijma.236)] [Medline: [29367890](https://pubmed.ncbi.nlm.nih.gov/29367890/)]
8. Saydah S, Lochner K. Socioeconomic status and risk of diabetes-related mortality in the U.S. *Public Health Rep* 2010;125(3):377-388 [FREE Full text] [doi: [10.1177/003335491012500306](https://doi.org/10.1177/003335491012500306)] [Medline: [20433032](https://pubmed.ncbi.nlm.nih.gov/20433032/)]
9. Walker RJ, Gebregziabher M, Martin-Harris B, Egede LE. Independent effects of socioeconomic and psychological social determinants of health on self-care and outcomes in Type 2 diabetes. *Gen Hosp Psychiatry* 2014;36(6):662-668 [FREE Full text] [doi: [10.1016/j.genhosppsych.2014.06.011](https://doi.org/10.1016/j.genhosppsych.2014.06.011)] [Medline: [25103544](https://pubmed.ncbi.nlm.nih.gov/25103544/)]
10. Mayberry LS, Osborn CY. Family support, medication adherence, and glycemic control among adults with type 2 diabetes. *Diabetes Care* 2012 Jun;35(6):1239-1245 [FREE Full text] [doi: [10.2337/dc11-2103](https://doi.org/10.2337/dc11-2103)] [Medline: [22538012](https://pubmed.ncbi.nlm.nih.gov/22538012/)]
11. Reynolds DB, Walker RJ, Campbell JA, Egede LE. Differential effect of race, education, gender, and language discrimination on glycemic control in adults with type 2 diabetes. *Diabetes Technol Ther* 2015 Apr;17(4):243-247 [FREE Full text] [doi: [10.1089/dia.2014.0285](https://doi.org/10.1089/dia.2014.0285)] [Medline: [25549154](https://pubmed.ncbi.nlm.nih.gov/25549154/)]
12. Walker RJ, Smalls BL, Campbell JA, Strom Williams JL, Egede LE. Impact of social determinants of health on outcomes for type 2 diabetes: a systematic review. *Endocrine* 2014 Sep;47(1):29-48 [FREE Full text] [doi: [10.1007/s12020-014-0195-0](https://doi.org/10.1007/s12020-014-0195-0)] [Medline: [24532079](https://pubmed.ncbi.nlm.nih.gov/24532079/)]
13. Clark ML, Utz SW. Social determinants of type 2 diabetes and health in the United States. *World J Diabetes* 2014 Jun 15;5(3):296-304 [FREE Full text] [doi: [10.4239/wjd.v5.i3.296](https://doi.org/10.4239/wjd.v5.i3.296)] [Medline: [24936251](https://pubmed.ncbi.nlm.nih.gov/24936251/)]
14. Brady E, Bridges K, Murray M, Cheng H, Liu B, He J, et al. Relationship between a comprehensive social determinants of health screening and type 2 diabetes mellitus. *Prev Med Rep* 2021 Sep;23:101465 [FREE Full text] [doi: [10.1016/j.pmedr.2021.101465](https://doi.org/10.1016/j.pmedr.2021.101465)] [Medline: [34194965](https://pubmed.ncbi.nlm.nih.gov/34194965/)]

15. Hill-Briggs F, Adler NE, Berkowitz SA, Chin MH, Gary-Webb TL, Navas-Acien A, et al. Social Determinants of Health and Diabetes: A Scientific Review. *Diabetes Care* 2020 Nov 02 [FREE Full text] [doi: [10.2337/dci20-0053](https://doi.org/10.2337/dci20-0053)] [Medline: [33139407](https://pubmed.ncbi.nlm.nih.gov/33139407/)]
16. Marmot M. Social determinants of health inequalities. *Lancet* 2005;365(9464):1099-1104. [doi: [10.1016/S0140-6736\(05\)71146-6](https://doi.org/10.1016/S0140-6736(05)71146-6)] [Medline: [15781105](https://pubmed.ncbi.nlm.nih.gov/15781105/)]
17. Fisher L, Skaff MM, Mullan JT, Arean P, Glasgow R, Masharani U. A longitudinal study of affective and anxiety disorders, depressive affect and diabetes distress in adults with Type 2 diabetes. *Diabet Med* 2008 Sep;25(9):1096-1101 [FREE Full text] [doi: [10.1111/j.1464-5491.2008.02533.x](https://doi.org/10.1111/j.1464-5491.2008.02533.x)] [Medline: [19183314](https://pubmed.ncbi.nlm.nih.gov/19183314/)]
18. Aljuaid MO, Almutairi AM, Assiri MA, Almalki DM, Alswat K. Diabetes-Related Distress Assessment among Type 2 Diabetes Patients. *J Diabetes Res* 2018;2018:7328128 [FREE Full text] [doi: [10.1155/2018/7328128](https://doi.org/10.1155/2018/7328128)] [Medline: [29770340](https://pubmed.ncbi.nlm.nih.gov/29770340/)]
19. Aikens JE. Prospective associations between emotional distress and poor outcomes in type 2 diabetes. *Diabetes Care* 2012 Dec;35(12):2472-2478 [FREE Full text] [doi: [10.2337/dc12-0181](https://doi.org/10.2337/dc12-0181)] [Medline: [23033244](https://pubmed.ncbi.nlm.nih.gov/23033244/)]
20. Delahanty LM, Grant RW, Wittenberg E, Bosch JL, Wexler DJ, Cagliero E, et al. Association of diabetes-related emotional distress with diabetes treatment in primary care patients with Type 2 diabetes. *Diabet Med* 2007 Jan;24(1):48-54. [doi: [10.1111/j.1464-5491.2007.02028.x](https://doi.org/10.1111/j.1464-5491.2007.02028.x)] [Medline: [17227324](https://pubmed.ncbi.nlm.nih.gov/17227324/)]
21. Fisher L, Mullan JT, Arean P, Glasgow RE, Hessler D, Masharani U. Diabetes distress but not clinical depression or depressive symptoms is associated with glycemic control in both cross-sectional and longitudinal analyses. *Diabetes Care* 2010 Jan;33(1):23-28 [FREE Full text] [doi: [10.2337/dc09-1238](https://doi.org/10.2337/dc09-1238)] [Medline: [19837786](https://pubmed.ncbi.nlm.nih.gov/19837786/)]
22. Chew BH, Vos RC, Metzendorf M, Scholten RJ, Rutten GE. Psychological interventions for diabetes-related distress in adults with type 2 diabetes mellitus. *Cochrane Database Syst Rev* 2017 Sep 27;9:CD011469 [FREE Full text] [doi: [10.1002/14651858.CD011469.pub2](https://doi.org/10.1002/14651858.CD011469.pub2)] [Medline: [28954185](https://pubmed.ncbi.nlm.nih.gov/28954185/)]
23. NYC Health + Hospitals. Bellevue. The City of New York. URL: <https://www.nychealthandhospitals.org/bellevue/about-bellevue> [accessed 2022-08-15]
24. Levy N, Moynihan V, Nilo A, Singer K, Bernik LS, Etiebet M, et al. The Mobile Insulin Titration Intervention (MITI) for Insulin Glargine Titration in an Urban, Low-Income Population: Randomized Controlled Trial Protocol. *JMIR Res Protoc* 2015 Mar 13;4(1):e31 [FREE Full text] [doi: [10.2196/resprot.4206](https://doi.org/10.2196/resprot.4206)] [Medline: [25794243](https://pubmed.ncbi.nlm.nih.gov/25794243/)]
25. Levy N, Moynihan V, Nilo A, Singer K, Bernik LS, Etiebet M, et al. The Mobile Insulin Titration Intervention (MITI) for Insulin Adjustment in an Urban, Low-Income Population: Randomized Controlled Trial. *J Med Internet Res* 2015 Jul 17;17(7):e180 [FREE Full text] [doi: [10.2196/jmir.4716](https://doi.org/10.2196/jmir.4716)] [Medline: [26187303](https://pubmed.ncbi.nlm.nih.gov/26187303/)]
26. Levy NK, Orzeck-Byrnes NA, Aidasani SR, Moloney DN, Nguyen LH, Park A, et al. Transition of a Text-Based Insulin Titration Program From a Randomized Controlled Trial Into Real-World Settings: Implementation Study. *J Med Internet Res* 2018 Mar 19;20(3):e93 [FREE Full text] [doi: [10.2196/jmir.9515](https://doi.org/10.2196/jmir.9515)] [Medline: [29555621](https://pubmed.ncbi.nlm.nih.gov/29555621/)]
27. PRAPARE®: Protocol for Responding to and Assessing Patient Assets, Risks, and Experiences. URL: [http://www.nachc.org/wp-content/uploads/2018/05/PRAPARE\\_One\\_Pager\\_Sept\\_2016.pdf](http://www.nachc.org/wp-content/uploads/2018/05/PRAPARE_One_Pager_Sept_2016.pdf) [accessed 2022-08-15]
28. Bastani R. Informed Decision Making. In: *Dictionary of Marketing Communications*. Thousand Oaks, CA: Sage Publications; 2004:99.
29. National Health Interview Survey 1997-2018. National Center for Health Statistics. Centers for Disease Control and Prevention. URL: <https://www.cdc.gov/nchs/nhis/1997-2018.htm> [accessed 2022-08-15]
30. Tools to face the psychological demands of diabetes. Behavioral Diabetes Institute. URL: <https://behavioraldiabetes.org/scales-and-measures/#1448434304099-9078f27c-4106> [accessed 2022-08-15]
31. Polonsky WH, Fisher L, Earles J, Dudl RJ, Lees J, Mullan J, et al. Assessing psychosocial distress in diabetes: development of the diabetes distress scale. *Diabetes Care* 2005 Mar;28(3):626-631. [doi: [10.2337/diacare.28.3.626](https://doi.org/10.2337/diacare.28.3.626)] [Medline: [15735199](https://pubmed.ncbi.nlm.nih.gov/15735199/)]
32. Fisher L, Mullan JT, Skaff MM, Glasgow RE, Arean P, Hessler D. Predicting diabetes distress in patients with Type 2 diabetes: a longitudinal study. *Diabet Med* 2009 Jun;26(6):622-627 [FREE Full text] [doi: [10.1111/j.1464-5491.2009.02730.x](https://doi.org/10.1111/j.1464-5491.2009.02730.x)] [Medline: [19538238](https://pubmed.ncbi.nlm.nih.gov/19538238/)]
33. Chew B, Vos R, Mohd-Sidik S, Rutten GEHM. Diabetes-Related Distress, Depression and Distress-Depression among Adults with Type 2 Diabetes Mellitus in Malaysia. *PLoS One* 2016;11(3):e0152095 [FREE Full text] [doi: [10.1371/journal.pone.0152095](https://doi.org/10.1371/journal.pone.0152095)] [Medline: [27002728](https://pubmed.ncbi.nlm.nih.gov/27002728/)]
34. Stoop CH, Nefs G, Pop VJ, Wijnands-van Gent CJM, Tack CJ, Geelhoed-Duijvestijn PHLM, et al. Diabetes-specific emotional distress in people with Type 2 diabetes: a comparison between primary and secondary care. *Diabet Med* 2014 Oct;31(10):1252-1259. [doi: [10.1111/dme.12472](https://doi.org/10.1111/dme.12472)] [Medline: [24766062](https://pubmed.ncbi.nlm.nih.gov/24766062/)]
35. Karlsen B, Oftedal B, Bru E. The relationship between clinical indicators, coping styles, perceived support and diabetes-related distress among adults with type 2 diabetes. *J Adv Nurs* 2012 Feb;68(2):391-401. [doi: [10.1111/j.1365-2648.2011.05751.x](https://doi.org/10.1111/j.1365-2648.2011.05751.x)] [Medline: [21707728](https://pubmed.ncbi.nlm.nih.gov/21707728/)]
36. Browne JL, Scibilia R, Speight J. The needs, concerns, and characteristics of younger Australian adults with Type 2 diabetes. *Diabet Med* 2013 May;30(5):620-626. [doi: [10.1111/dme.12078](https://doi.org/10.1111/dme.12078)] [Medline: [23181664](https://pubmed.ncbi.nlm.nih.gov/23181664/)]
37. Nanayakkara N, Pease A, Ranasinha S, Wischer N, Andrikopoulos S, Speight J, et al. Depression and diabetes distress in adults with type 2 diabetes: results from the Australian National Diabetes Audit (ANDA) 2016. *Sci Rep* 2018 May 18;8(1):7846 [FREE Full text] [doi: [10.1038/s41598-018-26138-5](https://doi.org/10.1038/s41598-018-26138-5)] [Medline: [29777153](https://pubmed.ncbi.nlm.nih.gov/29777153/)]

38. Wexler DJ. Patient education: Type 2 diabetes: Insulin treatment (Beyond the Basics). UpToDate. 2022. URL: <https://www.uptodate.com/contents/type-2-diabetes-insulin-treatment-beyond-the-basics> [accessed 2022-08-15]
39. The Accountable Health Communities Health-Related Social Needs Screening Tool. Centers for Medicare & Medicaid Services. URL: <https://innovation.cms.gov/Files/worksheets/ahcm-screeningtool.pdf> [accessed 2022-08-15]
40. The Health Leads Screening Toolkit. Health Leads. 2018 Sep 17. URL: <https://healthleadsusa.org/resources/the-health-leads-screening-toolkit/> [accessed 2022-08-15]

## Abbreviations

**DDS:** Diabetes Distress Scale

**DRD:** diabetes-related distress

**HbA<sub>1c</sub>:** hemoglobin A<sub>1c</sub>

**HP2020:** Healthy People 2020

**MITI:** Mobile Insulin Titration Intervention

**NHIS:** National Health Interview Survey

**NYUGSoM:** New York University Grossman School of Medicine

**PRAPARE:** Protocol for Responding to and Assessing Patients' Assets, Risks, and Experiences

**SDOH:** social determinants of health

**T2DM:** type 2 diabetes mellitus

*Edited by A Mavragani; submitted 08.06.22; peer-reviewed by D Newman, ER Khalilian, SJC Soerensen ; comments to author 19.07.22; revised version received 02.08.22; accepted 09.08.22; published 12.10.22.*

*Please cite as:*

*Levy NK, Park A, Solis D, Hu L, Langford AT, Wang B, Rogers ES*

*Social Determinants of Health and Diabetes-Related Distress in Patients With Insulin-Dependent Type 2 Diabetes: Cross-sectional, Mixed Methods Approach*

*JMIR Form Res 2022;6(10):e40164*

*URL: <https://formative.jmir.org/2022/10/e40164>*

*doi: [10.2196/40164](https://doi.org/10.2196/40164)*

*PMID: [36222807](https://pubmed.ncbi.nlm.nih.gov/36222807/)*

©Natalie K Levy, Agnes Park, Daniela Solis, Lu Hu, Aisha T Langford, Binhuan Wang, Erin S Rogers. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 12.10.2022. 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

# Separating Features From Functionality in Vaccination Apps: Computational Analysis

George Shaw Jr<sup>1</sup>, BSc, MSc, PhD; Devaki Nadkarni<sup>2</sup>, PharmB; Eric Phann<sup>3,4</sup>; Rachel Sielaty<sup>5</sup>, BAsC; Madeleine Ledenyi<sup>2</sup>; Razaan Abnow<sup>6</sup>; Qian Xu<sup>7</sup>, PhD; Paul Arredondo<sup>8</sup>, BS, MS; Shi Chen<sup>1</sup>, PhD

<sup>1</sup>Public Health Sciences, School of Data Science, University of North Carolina, Charlotte, NC, United States

<sup>2</sup>Public Health Sciences, University of North Carolina at Charlotte, Charlotte, NC, United States

<sup>3</sup>Department of Computer Science, University of North Carolina at Charlotte, Charlotte, NC, United States

<sup>4</sup>Department of Language and Culture Studies, University of North Carolina at Charlotte, Charlotte, NC, United States

<sup>5</sup>Department of Biological Sciences, University of North Carolina at Charlotte, Charlotte, NC, United States

<sup>6</sup>Department of Global Studies, Belk College of Business, University of North Carolina at Charlotte, Charlotte, NC, United States

<sup>7</sup>School of Communications, Elon University, Elon, NC, United States

<sup>8</sup>School of Data Science, University of North Carolina, Charlotte, NC, United States

**Corresponding Author:**

George Shaw Jr, BSc, MSc, PhD

Public Health Sciences

School of Data Science

University of North Carolina

9201 University City Blvd.

Charlotte, NC, 28223-0001

United States

Phone: 1 704 562 0021

Email: [gshaw11@unc.edu](mailto:gshaw11@unc.edu)

## Abstract

**Background:** Some latest estimates show that approximately 95% of Americans own a smartphone with numerous functions such as SMS text messaging, the ability to take high-resolution pictures, and mobile software apps. Mobile health apps focusing on vaccination and immunization have proliferated in the digital health information technology market. Mobile health apps have the potential to positively affect vaccination coverage. However, their general functionality, user and disease coverage, and exchange of information have not been comprehensively studied or evaluated computationally.

**Objective:** The primary aim of this study is to develop a computational method to explore the descriptive, usability, information exchange, and privacy features of vaccination apps, which can inform vaccination app design. Furthermore, we sought to identify potential limitations and drawbacks in the apps' design, readability, and information exchange abilities.

**Methods:** A comprehensive codebook was developed to conduct a content analysis on vaccination apps' descriptive, usability, information exchange, and privacy features. The search and selection process for vaccination-related apps was conducted from March to May 2019. We identified a total of 211 apps across both platforms, with iOS and Android representing 62.1% (131/211) and 37.9% (80/211) of the apps, respectively. Of the 211 apps, 119 (56.4%) were included in the final study analysis, with 42 features evaluated according to the developed codebook. The apps selected were a mix of apps used in the United States and internationally. Principal component analysis was used to reduce the dimensionality of the data. Furthermore, cluster analysis was used with unsupervised machine learning to determine patterns within the data to group the apps based on preselected features.

**Results:** The results indicated that readability and information exchange were highly correlated features based on principal component analysis. Of the 119 apps, 53 (44.5%) were iOS apps, 55 (46.2%) were for the Android operating system, and 11 (9.2%) could be found on both platforms. Cluster 1 of the k-means analysis contained 22.7% (27/119) of the apps; these were shown to have the highest percentage of features represented among the selected features.

**Conclusions:** We conclude that our computational method was able to identify important features of vaccination apps correlating with end user experience and categorize those apps through cluster analysis. Collaborating with clinical health providers and public health officials during design and development can improve the overall functionality of the apps.



**KEYWORDS**

vaccines; mobile health; mHealth; principal component analysis; PCA; k-means clustering; information exchange; mobile phone

## Introduction

### Background

IT has revolutionized all aspects of the world, including our health care system. IT has enhanced the overall efficiency and accessibility of patient care [1]. Smartphones are a type of IT that has become important within health care [2]. Some of the latest estimates show that approximately 95% of Americans own a smartphone with numerous functions such as texting, the ability to take high-resolution pictures, and mobile software apps [3]. Owners of smartphones also use the available functions to manage various facets of their health [4]. Today, mobile health (mHealth) technology plays a crucial role in providing quality health care services by improving health outcomes and facilitating health care access. Istepanian et al [5] defined mHealth as mobile computing, medical sensor, and communication technologies designed for health care. The use of mHealth apps provides an efficient way for patients to share their medical information with providers, improves the collection of real-time health information, and supports vaccine uptake [6].

A significant concern that is often communicated by mHealth app users is data privacy. In the United States, the Health Insurance Portability and Accountability Act (HIPAA) ensures that health care entities provide adequate measures to protect patient data. Many consumer-based apps that track and monitor health data are not HIPAA compliant. Health care stakeholders in the United States recommend that mobile apps designed for health care be HIPAA compliant [2,7]. Another impediment faced by mHealth apps in the current dynamic immunization practice is a 1-sided vaccine delivery system or lack of bidirectional information exchange. There exists the opportunity to leverage mHealth apps as tools to support public navigation of complex health systems and promote bidirectional communication of information between the public and health providers. Factors such as lack of health care access, fragmented vaccine provider systems, and low vaccine literacy can lead to undervaccination in the community [8]. Moreover, vaccine hesitancy can further reduce vaccine uptake among populations and undermine previous gains in eradicating communicable diseases [6].

Vaccine hesitancy—the delay or refusal to be vaccinated despite available vaccination services—is a complex phenomenon that involves emotional, cultural, social, spiritual, and political factors [9,10]. When considering vaccine hesitancy and decision-making, parental vaccine hesitancy stems from a variety of reasons, and there is no one-size-fits-all *type* of parent who chooses to forgo vaccinating their child [8,11,12]. Suspected autism side effects, religious reasons, concerns over the “newness” of the vaccine, and inaccurate portrayal of vaccines in various media outlets are common factors that influence parental vaccine hesitancy [6,13]. The recent resurgence of outbreaks of whooping cough and measles in

children is a prime example of vaccination refusal associated with the resurgence of preventable communicable diseases in communities [13]. Although vaccination mHealth apps are attempting to address this issue, current results are mixed [14].

### Rationale and Aim

Results from a recent systematic review reported a lack of evidence supporting the use of vaccination apps geared toward children, as shown through vaccination uptake, knowledge, and decision-making [15]. Another systematic review reported that mHealth improved vaccination uptake among adults and children; however, there is inconclusive evidence that digital solutions will achieve optimal vaccination coverage [16]. Barriers such as technology hesitancy, complicated app navigation, and difficult app features can compromise vaccination app use. Security and storage compliance associated with HIPAA, along with transmission and protection of private health data collected through vaccination apps, is another pressing concern.

The primary aim of our study was to develop a computational method to explore the descriptive, usability, information exchange, and privacy features of various vaccination apps, which can inform vaccination app design. We also aimed to assess these apps using a content analysis approach and identify potential flaws in app functionality. This study analyzed these data according to their respective operating platforms and collectively. The content analysis approach used was adapted from previous studies [17,18].

## Methods

### Definition and Identification of Vaccination-Related Apps

For this study, vaccination-related apps were operationalized as apps that allowed tracking, scheduling, and general dissemination of vaccination information [6,19]. Apps were included if they were found on the Google Play Store and the Apple App Store. The query terms “vaccination,” “immunization,” “vaccine,” “immunization schedule,” and “vaccination schedule” were used in the search process to generate our sample of apps. English-language apps and English-language apps with a second language were both included. We did not search any cell phone manufacturer app stores (eg, Samsung Galaxy Store) as Google Play and the Apple App Store are prominent web-based marketplaces used by Android and iOS smartphone users to download apps. Apps characterized as sideload apps and homebrew apps were excluded. Sideload apps are apps that have not been certified to be included within an app store. We characterized homebrew apps as apps that can be downloaded using a computer terminal.

The search and selection process for the apps was conducted from March to May 2019. We identified a total of 211 apps across both platforms, with iOS and Android representing 62.1%

(131/211) and 37.9% (80/211) of the apps, respectively. As iOS apps represented more than half of the apps collected, we chose a random sample of apps from each platform to generate a sample of 132 apps (62.6% of the total sample). We oversampled the Android apps that were collected to provide a balance of Android apps that would be represented in the feature space. Moreover, oversampling is a common technique used when there is an underrepresentation of 1 class [20,21]. In addition, prior work has documented health apps that were discontinued within a 12-month time frame [22]. Our final list of apps included those that remained on the market for at least 12 months (as of May 2020). Using the criteria of meeting the 12-month time frame, content written in English, and the operationalized definition of vaccination-related apps as part of the 2 lead researchers' additional deliberation of the apps, the final number of apps included in the study was 119. To accurately classify the apps together and individually, they were categorized according to their status of iOS, Android, or both. Although we could have included apps designated as *both* in one platform, it would have been an inaccurate representation of the apps [23,24]. To evaluate and understand the nuances of vaccination apps' descriptive, usability, information exchange capability, and privacy features, we used a mixed methods approach to frame our work. First, we developed and evaluated a codebook and conducted a content analysis of the included

apps based on the codebook. This provided us with a general insight into the categories of the codebook. Second, we conducted dimensionality reduction on the features constructed from the codebook and content analysis results. Third, we wanted to identify those features that were more important for explaining variances in the data. Finally, through k-means clustering, we clustered the apps according to the feature dimension reduction results from step 3. The following sections provide additional details for each step.

### Codebook Development

To comprehensively characterize the features of the vaccination apps retrieved, we systematically developed an inclusive codebook with 4 categories (Table 1). These 4 main categories were developed during the app screening process. These broad categories have been used in similar vaccination app-related studies [6]. A total of 10 apps were randomly selected in June 2020 to evaluate the codebook, and the results were cross-validated to ensure a moderate level of agreement between the 2 coders using percentage agreement [25]. We achieved 90% agreement regarding the selected apps. Following the establishment of a stable version of the codebook, immunization apps were evaluated according to the major categories (Table 1). The 42 features across the 4 categories in Table 1 were used to represent the feature space for our computational analysis, as detailed in the following sections.

**Table 1.** Summary of the codebook features with descriptions.

Category	Features	Description
Descriptive	App name, developer; platform (iOS, Android, or both), category in the app store (medical, health and fitness, travel, or local), size in MB, ranking in its respective category if applicable, overall star rating if applicable, age rating if applicable, and cost (completely free, free to download with in-app purchase, or paid)	These descriptive characteristics gave an overview of the immunization app and such information could generally be found on the app store's description page without the need to download or install the app [26].
Users and diseases	Target users and target diseases; for target users, we analyzed whether the app provided information on a specific user group (eg, children, parents, women, physicians, and age group); target diseases pertained to the description of a specific disease or general information concerning vaccinations and scheduling	In this category, we evaluated the targeted users and diseases of the apps. Some apps could be used by multiple, potentially overlapping groups of adult users, such as travelers and women, for which we created a specific group with binary response only. The targeted users included the following: minors, parents, travelers, women, people of all ages, and health care providers and staff. For targeted diseases, 0 was associated with no user-defined diseases, and 1 was for specific diseases such as seasonal influenza and measles-mumps-rubella [27].
Information exchange features	Account requirement for full app functionality, information presented about specific types of vaccines, educational information about vaccination and immunization in general, immunization tracking, customization of schedule, identification of nearby vaccination clinics, reminders of upcoming vaccination events, and personalized vaccination recommendations	In this category, we further explored and quantified vaccination-related core features of the apps.
Privacy and readability	Health Insurance Portability and Accountability Act-compliance feature; presence of in-app privacy statement; presence of privacy statement in the app store; presence of multilingual (at least 2 languages) privacy statement; and the average length of the privacy content (in English) using the following 7 readability measures: Simple Measure of Gobbledygook, Flesch Reading Ease score, Gunning Fog Index, Flesch-Kincaid Grade, Coleman-Liau Index, Automated Readability Index, and Linsear Write Formula [28]	Here, we considered an important element in mobile health-related research and app development, which is privacy-related features to address privacy concerns around sensitive and private vaccine health information. These features would provide information on how user-generated data would be collected, stored, shared, and transmitted on the web and offline [29,30].

## Data Analysis

### Overview

We analyzed and evaluated the content of the apps using the aforementioned codebook through a combination of content analysis [26], descriptive statistics, and unsupervised machine learning. First, we used principal component analysis (PCA) to reduce the feature space from our original data set. Second, the apps were clustered using the k-means algorithm in R (R Foundation for Statistical Computing). The following sections will discuss in detail how PCA and k-means clustering were used in this study.

### PCA Process

PCA is an important preprocessing step. Prior studies have used PCA to show children's interactions with education apps [31] and reduce the context dimensions of data from smartphone apps [32]. After coding the 119 apps based on the 42 features, we conducted PCA to reduce the dimensionality space of our data. We used the *prcomp* function in R to explain the variance that was represented by the different principal components (PCs).

After identifying the proportion of variance, we determined the value of each feature contained within each PC. We used the loading values of each PC to determine this information. These values represent the correlations between the PC and the original used features. A correlation that is close to 1 or -1 indicates how important the feature is to the component. We extracted the top 5 features for each PC with the highest variance. Using these values, we reduced the number of features to represent the apps from 42 to 10. The key idea of PCA is to reduce the number of variables in the data set but preserve as much information or representation of that information in the new data set as possible [33]. Although there is no gold standard for determining the number of features to retain from this process [34], the retained features represented important components of many apps. The retained features were used to describe the data and conduct our k-means cluster analysis.

### K-Means Cluster Analysis

Cluster analysis is used to define classes within a set of data. Clustering can be conducted using supervised and unsupervised methods. We used the unsupervised k-means clustering method to group our apps. This clustering algorithm is well documented, with successfully separating data for analysis; moreover, it has performed similar to or better than other clustering approaches

[35,36]. This method uncovers latent patterns within the data and allows us to have a better understanding of which apps are associated with each other based on the selected features. To determine the number of clusters, we used the total within-cluster sum of squares (or elbow method) [37] and the silhouette method. The total within-cluster sum of squares measures how compact the clusters are. The silhouette method [38] seeks to measure the quality of the clustering. We examined how well the feature object lies within the clustering [39]. We analyzed both methods to determine the optimal number of topics to use for our k-means clustering analysis.

### Ethical Considerations

The data used in this study satisfied two research activities that did not require IRB approval, Quality Assurance and Improvement. IRB approval is not required if the study involves the practice of program evaluation, self-assessment of programs or business practices, and other quality improvement projects where methods rather than humans are the subject of the study. It also satisfies the conditions of a pilot study where the activities are intended to refine data collection procedures – time to participate, testing survey questions, etc. where any data collected are only used to plan and/or improve a future research study.

## Results

### Overview of Categories and Features

Of the 42 features, 12 (28%) were used for the descriptive app category. Of these 12 features, 9 (75%) were used for the (targeted) users and diseases category, and 8 (67%) were used for the information exchange category. Finally, 31% (13/42) of features represented the privacy and readability category. Of the 119 apps, 53 (44.5%) were iOS apps, 55 (46.2%) were for the Android operating system, and 11 (9.2%) could be found on both platforms. The Flesch-Kincaid Grade readability score (readability tests designed to indicate how difficult the content is to understand) had an average of 6.4 (SD 6.6) for both platforms combined (Table 2) [34]. Privacy statements on iOS had an average length of 850.38 (SD 1483.42) words, whereas the privacy statements of apps on Android had an average length of 790.42 (SD 1227.05) words. The user star rating was higher for the Android apps than for the iOS apps. There was a considerable difference in the sizes of apps, with iOS apps using more space (37.54 MB) than apps supported by Android (11.48 MB).

**Table 2.** Select app features characteristics (N=119).

App features	iOS (n=53)	Android (n=55)	Both (n=11)	Total
Number of ratings, mean (SD)	13.53 (62.34)	1772.8 (8136.84)	61.91 (79.07)	831.11 (5600.44)
Size in MB, mean (SD)	37.54 (41.2)	11.48 (17)	14.4 (19.47)	23.36 (32.66)
Star rating, mean (SD)	0.83 (1.59)	2.63 (2.08)	2.71 (2.07)	1.84 (2.06)
Age rating, mean (SD)	9.34 (5.4)	2.62 (4.99)	— <sup>a</sup>	5.37 (6.15)
Length of privacy policy (words) <sup>b</sup> , mean (SD)	850.38 (1483.42)	790.42(1227.05)	874.64 (1206.42)	824.91 (1329.78)
Flesch-Kincaid Grade, mean (SD)	6.13 (6.89)	6.01 (6.38)	9.63 (6.28)	6.4 (6.6)
<b>HIPAA<sup>c</sup> compliance, n (%)</b>				
Yes	7 (13)	2 (4)	0 (0)	9 (8)
No	46 (87)	53 (96)	13 (100)	110 (92)

<sup>a</sup>Not available.

<sup>b</sup>Several apps identified contained privacy policy statements written in a different language; *however*, some apps provided an English-translated version of the policy. All apps reviewed adopted, followed, or referenced a US-based vaccination schedule (ie, Centers for Disease Control and Prevention).

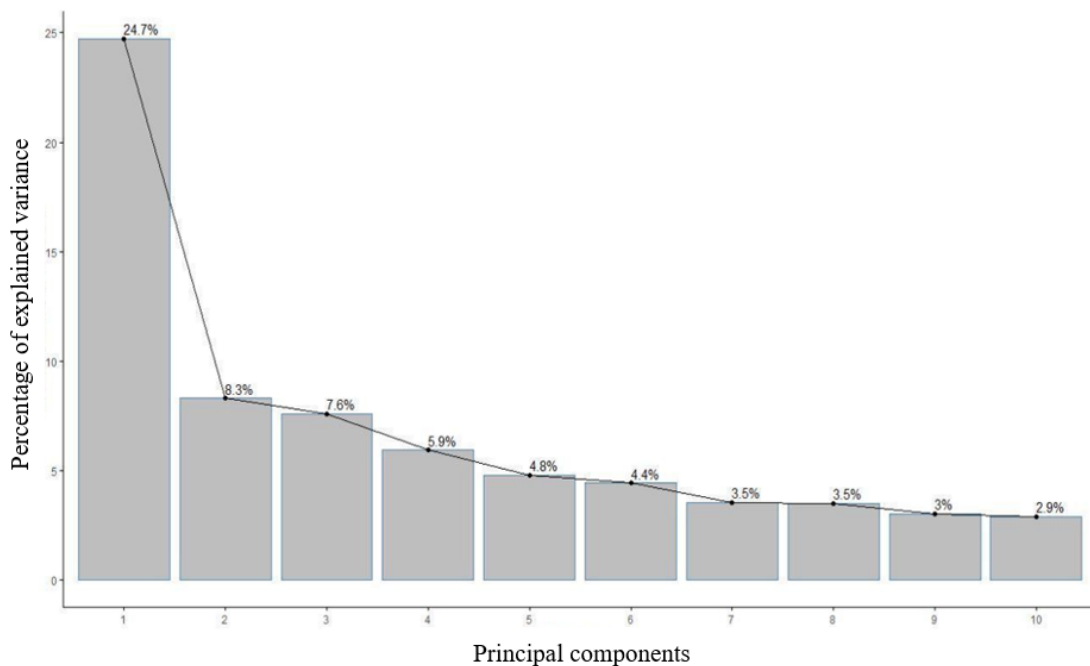
<sup>c</sup>HIPAA: Health Insurance Portability and Accountability Act.

## PCA Results

Results from the dimensionality reduction of the feature space showed that PC1 explained approximately 24.7% of the data and PC2 explained 8.3% of the data (Figure 1). The next step in our PCA involved reviewing the correlations between the PCs and the features [39]. Using the loading scores, we analyzed the values for PC1 and PC2. A review of the features for PC1 showed that the Automated Readability Index, Simple Measure of Gobbledygook, and Flesch-Kincaid Grade were the top 3

correlated features for PC1 (Textbox 1). Reminders of vaccinations, customized scheduling, and vaccination tracking were the most correlated features for PC2 based on the loading values (Textbox 1). Results from PC1 showed a high correlation between readability-related features, whereas results from PC2 showed a high correlation between customization-related features. PC2 highlights the importance of a consumer-focused approach to managing immunization schedules for children [40].

**Figure 1.** Principal component analysis showing that 24.7% of the data variance is explained by principal component 1 and 8.3% of the data variance is explained by principal component 2.



**Textbox 1.** Top 5 features correlated to their respective principal component (PC).

#### Top 5 features

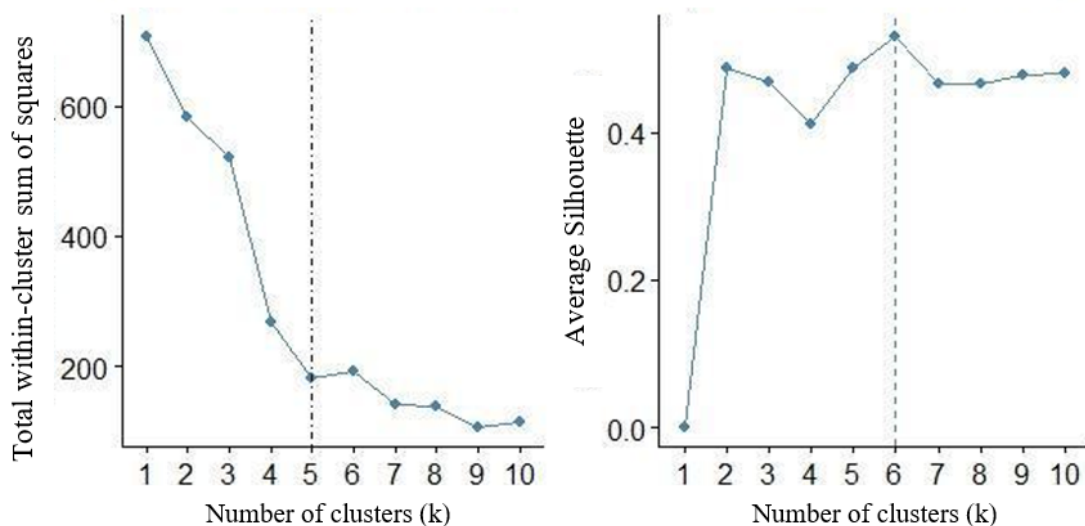
- PC1 (features related to readability)
  - Automated Readability Index
  - Simple Measure of Gobbledygook formula
  - Flesch-Kincaid Grade
  - Reading text page success
  - Linsear Write Formula
- PC2 (features related to user customization)
  - Reminder for vaccination
  - Customized schedule
  - Vaccination tracking
  - Personalized recommendations
  - Targeted at parents

### K-Means Cluster Analysis Results

The top 5 features from PC1 and PC2 were used to create a cluster graph that represented the optimal number of clusters for the new feature space ([Multimedia Appendix 1](#)). In [Figure 2](#), the dotted line represents the optimal number of clusters based on each measure. On the basis of the limited additional insight that would be derived from 6 clusters, 5 clusters were chosen as the optimal number of clusters to group the apps ([Figure 2](#)). [Table 3](#) displays the number of apps for each cluster in accordance with selected features from the new feature space

that includes the apps' target users (*targeted parents*), *customized schedule*, and *presence of privacy policy*. Cluster 1, with 22.7% (27/119) apps, had the highest percentage of apps with a user target focused on parents. Cluster 3, with 24.4% (29/119) apps, did not offer features of customizing a schedule or the presence of a privacy policy. Cluster 1 and cluster 2, (with 59/119, 49.6% apps in total), were the only clusters with the presence of a privacy policy. Cluster 5 did not include apps found on both platforms. The specific name of each app for each cluster can be found in [Multimedia Appendix 2](#).

**Figure 2.** Total within-cluster sum of squares and average silhouette width. The optimal number of clusters is 5 (left) for the total within-cluster sum of squares measure and 6 (right) for the average silhouette width measure.



**Table 3.** K-means clusters with selected new features represented (N=119).

Features	App cluster, n (%)				
	1 (n=27)	2 (n=32)	3 (n=29)	4 (n=19)	5 (n=12)
<b>Platform</b>					
Android	12 (44)	15 (47)	13 (45)	9 (47)	6 (50)
iOS	11 (41)	13 (41)	14 (48)	9 (47)	6 (50)
Both	4 (15)	4 (12)	2 (7)	1 (6)	0 (0)
<b>Targeted parents</b>					
Yes	16 (59)	4 (12)	8 (28)	6 (32)	5 (42)
No	11 (41)	28 (88)	21 (72)	13 (68)	7 (58)
<b>Customized schedule</b>					
Yes	25 (93)	4 (12)	0 (0)	18 (95)	10 (83)
No	2 (7)	28 (88)	29 (100)	1 (5)	2 (17)
<b>Reading text page (privacy policy)</b>					
Yes	27 (100)	32 (100)	0 (0)	0 (0)	0 (0)
No	0 (0)	0 (0)	29 (100)	19 (100)	12 (100)

## Discussion

### Principal Findings

In this study, we developed a codebook to conduct a content analysis of vaccination apps and explored the use of computational approaches to identify the feature importance of vaccination apps, reduce the dimensionality of our feature space, and categorize vaccination apps using k-means clustering in an unsupervised case approach. When examining the feature importance of the 119 vaccination apps and 42 features, we found that the most important features could be categorized and explained through PC1 and PC2. For PC1, the top features found in this component were predominately associated with the *privacy and readability* category from the codebook. The category of *information exchange* had the most prominent features associated with PC2. On the basis of these results, incorporating information exchange functions and improving the readability of policy-related information should include expert involvement in vaccination app design (as denoted by clusters 1 and 2 in Table 3). Among the selected features in the cluster analysis, cluster 1 had the highest percentage of vaccination apps that provided a privacy policy, allowed a customized vaccination schedule, and targeted parents with regard to app use. Some apps that were not designed to track child vaccination information targeted parents (ie, KnowAsYouGo). Studies have detailed the lack of a government regulatory presence in the app market [6] as it relates to data privacy. Our work shows the lack of HIPAA compliance in vaccination-related apps (Table 2), although it is crucial for designers of vaccination apps in the United States to ensure agreement with HIPAA laws [41]. A transdisciplinary research approach in vaccination app design would allow for greater use by mHealth app users and opportunities to improve users' health literacy related to vaccines. Ultimately, this would result in an overall improvement of potential information exchange with public health providers.

### App Development and Feature Analysis

mHealth technology has the potential to improve the efficiency and convenience of health care information exchange. Our findings can be categorized into two major themes: (1) features that limit the functionality of apps and (2) features that impede the overall user experience. Although most apps are moderately received by their users, based on the app rating feature, there were salient weaknesses identified through the use of PCA. This further suggests that the limitations within the reviewed vaccination apps must be addressed. On the basis of the k-means cluster analysis and the selected features, only 1 cluster of vaccination apps did not provide evidence for user vaccine schedule customization. Functionality improvements to mHealth apps could allow for a connection between patients and medical professionals to provide timely care. Systematic incorporation of information exchange features and improving policy readability would result in notable enhancements to future apps, as well as those that are currently on the market and fail to incorporate these features.

We concluded that most vaccination apps were not developed alongside health professionals. There is no standard for expert involvement in app development for any sector, and integrating medical experts in the development of mHealth apps is important, considering the increased use of mHealth technology in health care spaces [42]. Specifically related to vaccination apps, they serve as a potential tool for vaccine advocacy, administration, documentation, and monitoring success within vaccination programs. Previous research has shown a lack of engagement from public health agencies, who might have benefited from a better estimation of immunization coverage and preparedness for incoming epidemics [43]. The apps we studied were absent of any data-sharing features with public health departments, although vaccine tracking is important when monitoring vaccine programs.

## Health Literacy and Health Communication

Health literacy involves the ability of individuals to find, understand, and use services to educate themselves to make health-related decisions [44]. One of the most correlated features of the evaluated apps based on the PC1 results was the readability tests. Improved readability in mHealth apps allows for increased use among consumers and helps individuals personally educate themselves to make healthier decisions in their lives [45]. Although our readability was focused on important user privacy content, these findings also have implications for other areas of the app that require high literacy skills to operate. Our results also revealed opportunities to redesign how privacy policy information and HIPAA compliance are communicated within vaccination apps [46]. Although results from the readability measures showed that vaccination apps scored an average of 6.51 on the Flesch-Kincaid Grade readability scale, other audio and video approaches may be leveraged to improve understanding of the policy information. In reference to the information exchange theory, we see that it is crucial to have users' personal information secured to ensure credibility. The development and redesign of the information exchange process within the apps prove to be an essential feature to adhere to policies such as HIPAA. Through these developmental improvements, we may experience an increase in vaccination app use across multiple public health sectors [47,48].

## Privacy and Security

Transferring vaccination records from paper to digital requires strict data standards and interoperability to ensure security [49]. Interoperability describes the extent to which systems can exchange and interpret shared data based on standards across health care settings. Interoperability allows for the secure exchange of medical information, which is essential for successful technological advances in health care. Less than half of the apps analyzed contained features that allowed data to be shared for personal recommendations. "Some information exchange methods involve 'rolling out' the electronic health records (EHRs) to unaffiliated health care organizations, creating an interface between different EHRs, or sharing a portal that allows others to view their information" [50]. Opportunities exist to develop evidence-based apps with regard to health data security and privacy concerns [51]. Credibility is a major concern of mHealth apps and may influence consumer use. This could lead to creating a systematic approach to mHealth vaccination app development and how these apps securely connect patient information with EHR systems [48].

## Strengths and Limitations

The validity of our research is upheld through a diligent acquisition and analysis of the 119 Android and iOS apps. We used 2 computational approaches to reduce the feature space and cluster our apps. Furthermore, PCA allows for the identification of specific features correlated to the larger PCs. Following the use of PCA and k-means clustering, our data provide a visual representation that is palatable for diverse audiences. The method used in our work has implications for other domain areas to examine the most important features when considering app design.

Despite a rigorous procurement and analysis of the 119 apps, our research contains several limitations. First, the apps that did not meet the 12-month time frame of representation on their respective platforms were removed from the analysis [22]. Although analyzing these apps independently was not the primary focus of this study, if included in our study, they could have affected the outcome of specific features, particularly the apps on the Android operating system. Future work should systematically evaluate apps that were discontinued during the study and compare their impact on study results. We did not observe the same issue with iOS apps. This yields potential complications for the replicability of our research in accordance with the obtained data. As a result of selecting apps exclusively from the Android and iOS app stores, there is potential for vaccination-related apps in other marketplaces to be excluded, affecting the study results. Another limitation involves bias related to the data selection process. Oversampling the Android apps creates an imbalance in the feature representation that may already be inherent to the data.

Second, this study was started in 2019, before the COVID-19 pandemic. Vaccination hesitancy along with misinformation has exacerbated vaccination uptake concerns. The landscape related to vaccination campaigns and the use of vaccination apps has changed significantly since this study started. Therefore, changes in apps that address misinformation, vaccine hesitancy, and telehealth services should be considered in future studies. Third, we used 2 exploratory machine learning approaches that can be affected by the data set size, number of features, and number of clusters. Instead of k-means clustering, the use of a hierarchical clustering method can account for grouping concerns during the cluster assignment step. Future work may incorporate other computational techniques to analyze these nuanced differences.

Finally, the researchers conducting this study are a US-based team; therefore, this research is intended to facilitate future app development. This research is also intended to supplement the further improvement of vaccination apps currently used in the United States. Not all the 119 apps featured in our research are based in the United States; this adds to the limitations of the research as it may complicate health recommendations that adhere to government and regional guidelines. Per the variation in countries where the apps are based, HIPAA compliance may not apply to other nations, and this may additionally complicate comparisons. Despite some apps being based in other countries, many internationally focused apps have followed or referenced the Centers for Disease Control and Prevention recommendations for vaccination schedules.

## Future Implications

The use of vaccines as a tool in personal and public health remains a cornerstone of disease prevention. Despite the advancement of vaccine technology and the promotion of vaccines as safe and effective, vaccine hesitancy has led to the resurgence of preventable childhood diseases. This resurgence threatens the effectiveness of vaccines as a public health tool. Technology, particularly mHealth apps, enables the intersection of public health and IT to potentially manifest positive vaccine health behaviors in individuals. Understanding the descriptive,

usability, information exchange, and privacy features of these 119 mHealth apps has the potential to provide researchers and health care professionals information concerning features that should be considered when designing vaccination apps as a public health instrument.

There is conflicting literature on the overall effectiveness of mHealth apps to assist with improving vaccination coverage; however, our research yields recommendations for mHealth vaccination apps developed in the future. One recommendation is to incorporate a transdisciplinary research approach to mHealth app development, in which medical professionals, app developers, public health experts, and users can collaborate throughout the app development process. This ensures engagement from multiple stakeholders and reliable information exchange between agencies and users. As noted in the previous section, although our study was conducted before the COVID-19 pandemic, our findings could prove relevant for the ongoing monitoring of COVID-19 metrics, vaccination documentation, and beyond. One such example for mHealth apps is contact tracing for COVID-19 or serving as a liaison for information exchange between experts and users. A recent study described the most frequently installed features of contact-tracking apps as alert systems and government accountability [52]. However, the need for the exchange of information for public health

purposes in contact tracing diminishes the data protection of the users. This affects users' uptake of these mHealth apps, and prior work has shown that many apps do not include participatory user involvement with contact-tracing apps [53]. Future directions for this research include the development of a sustainable bidirectional information exchange framework for vaccination mHealth apps.

## Conclusions

We conclude that our computational method was able to identify important features of vaccination apps correlating with end user experience and categorize those apps through cluster analysis (Multimedia Appendix 1). Results from PC1 show that the top 5 features correlated with readability, and results from PC2 show that most of the top 5 features correlated with user customization. Results from our computational method provide evidence that data information exchange among different health care entities should be leveraged to provide patient-centric health care. Privacy and security concerns around the collection, storage, and sharing of health data should be addressed during the app design development process. Collaboration among multiple health stakeholders during design and development can improve the overall functionality of vaccination-related apps.

---

## Acknowledgments

The authors would like to acknowledge the Office of Undergraduate Research online pilot, Paper Chase project, at the University of North Carolina Charlotte. The authors would also like to acknowledge Drs Jessamyn Bowling and Alicia Dahl, and Ms Lisa Krinner for supporting the students during this project.

---

## Authors' Contributions

GSJ designed and conducted the computational experiments and analyzed the results of the study. GSJ, DN, EP, RS, ML, and RA assisted with the writing and editing of every section of the manuscript. PA completed the descriptive statistics used in the study. QX assisted with the development of the codebook. SC assisted with the development of the codebook and provided overall direction for the project codebook implementation.

---

## Conflicts of Interest

None declared.

---

### Multimedia Appendix 1

Completed k-means analysis using 5 clusters.

[DOCX File , 136 KB - [formative\\_v6i10e36818\\_app1.docx](#) ]

---

### Multimedia Appendix 2

List of names for the apps that are represented in each cluster using the k-means clustering method.

[DOCX File , 25 KB - [formative\\_v6i10e36818\\_app2.docx](#) ]

---

## References

1. Awad A, Trenfield SJ, Pollard TD, Ong JJ, Elbadawi M, McCoubrey LE, et al. Connected healthcare: Improving patient care using digital health technologies. *Adv Drug Deliv Rev* 2021 Nov;178:113958. [doi: [10.1016/j.addr.2021.113958](#)] [Medline: [34478781](#)]
2. Mueller RC. Exploring family nurse practitioners' practices in recommending mHealth apps to patients. *Comput Inform Nurs* 2020 Feb;38(2):71-79. [doi: [10.1097/CIN.0000000000000580](#)] [Medline: [31804242](#)]



3. Ng YC, Alexander S, Frith KH. Integration of mobile health applications in health information technology initiatives: expanding opportunities for nurse participation in population health. *Comput Inform Nurs* 2018 May;36(5):209-213. [doi: [10.1097/CIN.0000000000000445](https://doi.org/10.1097/CIN.0000000000000445)] [Medline: [29734236](https://pubmed.ncbi.nlm.nih.gov/29734236/)]
4. Bhuyan SS, Lu N, Chandak A, Kim H, Wyant D, Bhatt J, et al. Use of mobile health applications for health-seeking behavior among US adults. *J Med Syst* 2016 Jun;40(6):153. [doi: [10.1007/s10916-016-0492-7](https://doi.org/10.1007/s10916-016-0492-7)] [Medline: [27147516](https://pubmed.ncbi.nlm.nih.gov/27147516/)]
5. Istepanian R, Jovanov E, Zhang YT. Introduction to the special section on M-Health: beyond seamless mobility and global wireless health-care connectivity. *IEEE Trans Inf Technol Biomed* 2004 Dec;8(4):405-414. [doi: [10.1109/titb.2004.840019](https://doi.org/10.1109/titb.2004.840019)] [Medline: [15615031](https://pubmed.ncbi.nlm.nih.gov/15615031/)]
6. Facciola A, Visalli G, Orlando A, Bertuccio MP, Spataro P, Squeri R, et al. Vaccine hesitancy: an overview on parents' opinions about vaccination and possible reasons of vaccine refusal. *J Public Health Res* 2019 Mar 11;8(1):1436 [FREE Full text] [doi: [10.4081/jphr.2019.1436](https://doi.org/10.4081/jphr.2019.1436)] [Medline: [30997357](https://pubmed.ncbi.nlm.nih.gov/30997357/)]
7. Glenn T, Monteith S. Privacy in the digital world: medical and health data outside of HIPAA protections. *Curr Psychiatry Rep* 2014 Nov;16(11):494. [doi: [10.1007/s11920-014-0494-4](https://doi.org/10.1007/s11920-014-0494-4)] [Medline: [25218603](https://pubmed.ncbi.nlm.nih.gov/25218603/)]
8. Bianco A, Mascaro V, Zucco R, Pavia M. Parent perspectives on childhood vaccination: how to deal with vaccine hesitancy and refusal? *Vaccine* 2019 Feb 08;37(7):984-990. [doi: [10.1016/j.vaccine.2018.12.062](https://doi.org/10.1016/j.vaccine.2018.12.062)] [Medline: [30655175](https://pubmed.ncbi.nlm.nih.gov/30655175/)]
9. Dubé E, Laberge C, Guay M, Bramadat P, Roy R, Bettinger JA. Vaccine hesitancy: an overview. *Hum Vaccin Immunother* 2013 Aug;9(8):1763-1773 [FREE Full text] [doi: [10.4161/hv.24657](https://doi.org/10.4161/hv.24657)] [Medline: [23584253](https://pubmed.ncbi.nlm.nih.gov/23584253/)]
10. MacDonald NE, SAGE Working Group on Vaccine Hesitancy. Vaccine hesitancy: definition, scope and determinants. *Vaccine* 2015 Aug 14;33(34):4161-4164 [FREE Full text] [doi: [10.1016/j.vaccine.2015.04.036](https://doi.org/10.1016/j.vaccine.2015.04.036)] [Medline: [25896383](https://pubmed.ncbi.nlm.nih.gov/25896383/)]
11. Harrison EA, Wu JW. Vaccine confidence in the time of COVID-19. *Eur J Epidemiol* 2020 Apr;35(4):325-330 [FREE Full text] [doi: [10.1007/s10654-020-00634-3](https://doi.org/10.1007/s10654-020-00634-3)] [Medline: [32318915](https://pubmed.ncbi.nlm.nih.gov/32318915/)]
12. 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/)]
13. Latella LE, McAuley RJ, Rabinowitz M. Beliefs about vaccinations: comparing a sample from a medical school to that from the general population. *Int J Environ Res Public Health* 2018 Mar 28;15(4):620 [FREE Full text] [doi: [10.3390/ijerph15040620](https://doi.org/10.3390/ijerph15040620)] [Medline: [29597319](https://pubmed.ncbi.nlm.nih.gov/29597319/)]
14. Simeoni R, Maccioni G, Giansanti D. The vaccination process against the COVID-19: opportunities, problems and mHealth support. *Healthcare (Basel)* 2021 Sep 06;9(9):1165 [FREE Full text] [doi: [10.3390/healthcare9091165](https://doi.org/10.3390/healthcare9091165)] [Medline: [34574939](https://pubmed.ncbi.nlm.nih.gov/34574939/)]
15. de Cock C, van Velthoven M, Milne-Ives M, Mooney M, Meinert E. Use of apps to promote childhood vaccination: systematic review. *JMIR Mhealth Uhealth* 2020 May 18;8(5):e17371 [FREE Full text] [doi: [10.2196/17371](https://doi.org/10.2196/17371)] [Medline: [32421684](https://pubmed.ncbi.nlm.nih.gov/32421684/)]
16. Balzarini F, Frascella B, Oradini-Alacreu A, Gaetti G, Lopalco PL, Edelstein M, et al. Does the use of personal electronic health records increase vaccine uptake? A systematic review. *Vaccine* 2020 Aug 27;38(38):5966-5978. [doi: [10.1016/j.vaccine.2020.05.083](https://doi.org/10.1016/j.vaccine.2020.05.083)] [Medline: [32620374](https://pubmed.ncbi.nlm.nih.gov/32620374/)]
17. Bender JL, Yue RY, 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](https://doi.org/10.2196/jmir.2661)] [Medline: [24366061](https://pubmed.ncbi.nlm.nih.gov/24366061/)]
18. Shen N, Levitan MJ, Johnson A, Bender JL, Hamilton-Page M, Jadad AA, et al. Finding a depression app: a review and content analysis of the depression app marketplace. *JMIR Mhealth Uhealth* 2015 Feb 16;3(1):e16 [FREE Full text] [doi: [10.2196/mhealth.3713](https://doi.org/10.2196/mhealth.3713)] [Medline: [25689790](https://pubmed.ncbi.nlm.nih.gov/25689790/)]
19. Lewis TL, Boissaud-Cooke MA, Aungst TD, Eysenbach G. Consensus on use of the term "App" versus "Application" for reporting of mHealth research. *J Med Internet Res* 2014 Jul 17;16(7):e174 [FREE Full text] [doi: [10.2196/jmir.3460](https://doi.org/10.2196/jmir.3460)] [Medline: [25033233](https://pubmed.ncbi.nlm.nih.gov/25033233/)]
20. Ling CX, Li C. Data mining for direct marketing: problems and solutions. In: *Proceedings of the Third International Conference on Knowledge Discovery and Data Mining*. 1998 Presented at: KDD '98; August 27-31, 1998; New York, NY, USA p. 73-79 URL: <https://www.aaai.org/Papers/KDD/1998/KDD98-011.pdf>
21. He H, Garcia EA. Learning from imbalanced data. *IEEE Trans Knowl Data Eng* 2009 Sep;21(9):1263-1284. [doi: [10.1109/tkde.2008.239](https://doi.org/10.1109/tkde.2008.239)]
22. Tabi K, Randhawa AS, Choi F, Mithani Z, Albers F, Schnieder M, et al. Mobile apps for medication management: review and analysis. *JMIR Mhealth Uhealth* 2019 Sep 11;7(9):e13608 [FREE Full text] [doi: [10.2196/13608](https://doi.org/10.2196/13608)] [Medline: [31512580](https://pubmed.ncbi.nlm.nih.gov/31512580/)]
23. Wang RJ. Branded mobile application adoption and customer engagement behavior. *Comput Human Behav* 2020 May;106:106245. [doi: [10.1016/j.chb.2020.106245](https://doi.org/10.1016/j.chb.2020.106245)]
24. Ebone A, Tan Y, Jia X. A performance evaluation of cross-platform mobile application development approaches. In: *Proceedings of the IEEE/ACM 5th International Conference on Mobile Software Engineering and Systems*. 2018 Presented at: MOBILESoft '18; May 27-June 3, 2018; Gothenburg, Sweden p. 92-93. [doi: [10.1145/3197231.3197252](https://doi.org/10.1145/3197231.3197252)]
25. McHugh ML. Interrater reliability: the kappa statistic. *Biochem Med (Zagreb)* 2012;22(3):276-282 [FREE Full text] [Medline: [23092060](https://pubmed.ncbi.nlm.nih.gov/23092060/)]

26. Jimenez G, Lum E, Car J. Examining diabetes management apps recommended from a Google search: content analysis. *JMIR Mhealth Uhealth* 2019 Jan 16;7(1):e11848 [FREE Full text] [doi: [10.2196/11848](https://doi.org/10.2196/11848)] [Medline: [30303485](https://pubmed.ncbi.nlm.nih.gov/30303485/)]
27. Yin H, Wardenaar KJ, Wang Y, Wang N, Chen W, Zhang Y, et al. Mobile mental health apps in China: systematic app store search. *J Med Internet Res* 2020 Jul 27;22(7):e14915 [FREE Full text] [doi: [10.2196/14915](https://doi.org/10.2196/14915)] [Medline: [32716301](https://pubmed.ncbi.nlm.nih.gov/32716301/)]
28. Robillard JM, Feng TL, Sporn AB, Lai JA, Lo C, Ta M, et al. Availability, readability, and content of privacy policies and terms of agreements of mental health apps. *Internet Interv* 2019 Sep;17:100243 [FREE Full text] [doi: [10.1016/j.invent.2019.100243](https://doi.org/10.1016/j.invent.2019.100243)] [Medline: [30949436](https://pubmed.ncbi.nlm.nih.gov/30949436/)]
29. Arora S, Yttri J, Nilse W. Privacy and security in mobile health (mHealth) research. *Alcohol Res* 2014;36(1):143-151 [FREE Full text] [Medline: [26259009](https://pubmed.ncbi.nlm.nih.gov/26259009/)]
30. Aljedaani B, Babar MA. Challenges with developing secure mobile health applications: systematic review. *JMIR Mhealth Uhealth* 2021 Jun 21;9(6):e15654 [FREE Full text] [doi: [10.2196/15654](https://doi.org/10.2196/15654)] [Medline: [34152277](https://pubmed.ncbi.nlm.nih.gov/34152277/)]
31. Crescenzi-Lanna L. Emotions, private speech, involvement and other aspects of young children's interactions with educational apps. *Comput Human Behav* 2020 Oct;111:106430. [doi: [10.1016/j.chb.2020.106430](https://doi.org/10.1016/j.chb.2020.106430)]
32. Sarker IH, Abushark YB, Khan AI. ContextPCA: predicting context-aware smartphone apps usage based on machine learning techniques. *Symmetry* 2020 Apr 01;12(4):499. [doi: [10.3390/sym12040499](https://doi.org/10.3390/sym12040499)]
33. Jolliffe IT, Cadima J. Principal component analysis: a review and recent developments. *Philos Trans A Math Phys Eng Sci* 2016 Apr 13;374(2065):20150202 [FREE Full text] [doi: [10.1098/rsta.2015.0202](https://doi.org/10.1098/rsta.2015.0202)] [Medline: [26953178](https://pubmed.ncbi.nlm.nih.gov/26953178/)]
34. Steinley D. K-means clustering: a half-century synthesis. *Br J Math Stat Psychol* 2006 May;59(Pt 1):1-34. [doi: [10.1348/000711005X48266](https://doi.org/10.1348/000711005X48266)] [Medline: [16709277](https://pubmed.ncbi.nlm.nih.gov/16709277/)]
35. Kruse J, Toledo P, Belton TB, Testani EJ, Evans CT, Grobman WA, et al. Readability, content, and quality of COVID-19 patient education materials from academic medical centers in the United States. *Am J Infect Control* 2021 Jun;49(6):690-693 [FREE Full text] [doi: [10.1016/j.ajic.2020.11.023](https://doi.org/10.1016/j.ajic.2020.11.023)] [Medline: [33259825](https://pubmed.ncbi.nlm.nih.gov/33259825/)]
36. Dobbins C, Rawassizadeh R. Clustering of physical activities for quantified self and mHealth applications. In: Proceedings of the 2015 IEEE International Conference on Computer and Information Technology; Ubiquitous Computing and Communications; Dependable, Autonomic and Secure Computing; Pervasive Intelligence and Computing. 2015 Presented at: CIT '15; October 26-28, 2015; Liverpool, UK p. 1423-1428. [doi: [10.1109/cit/iucc/dasc/picom.2015.213](https://doi.org/10.1109/cit/iucc/dasc/picom.2015.213)]
37. Syakur MA, Khotimah BK, Rochman EM, Satoto BD. Integration K-means clustering method and elbow method for identification of the best customer profile cluster. In: IOP Conference Series: Materials Science and Engineering, Volume 336, The 2nd International Conference on Vocational Education and Electrical Engineering. 2018 Apr 16 Presented at: ICVEE '18; November 9, 2017; Surabaya, Indonesia. [doi: [10.1088/1757-899x/336/1/012017](https://doi.org/10.1088/1757-899x/336/1/012017)]
38. Rousseeuw PJ. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *J Comput Appl Math* 1987 Nov 23;20:53-65 [FREE Full text] [doi: [10.1016/0377-0427\(87\)90125-7](https://doi.org/10.1016/0377-0427(87)90125-7)]
39. Burger SV. Introduction to Machine Learning with R. Sebastopol, CA, USA: O'Reilly Media; Mar 2018.
40. Seeber L, Conrad T, Hoppe C, Obermeier P, Chen X, Karsch K, et al. Educating parents about the vaccination status of their children: a user-centered mobile application. *Prev Med Rep* 2017 Mar;5:241-250 [FREE Full text] [doi: [10.1016/j.pmedr.2017.01.002](https://doi.org/10.1016/j.pmedr.2017.01.002)] [Medline: [28127527](https://pubmed.ncbi.nlm.nih.gov/28127527/)]
41. Al Ayubi SU, Pelletier A, Sunthara G, Gujral N, Mittal V, Bourgeois FC. A mobile app development guideline for hospital settings: maximizing the use of and minimizing the security risks of "bring your own devices" policies. *JMIR Mhealth Uhealth* 2016 May 11;4(2):e50 [FREE Full text] [doi: [10.2196/mhealth.4424](https://doi.org/10.2196/mhealth.4424)] [Medline: [27169345](https://pubmed.ncbi.nlm.nih.gov/27169345/)]
42. Pagoto S, Schneider K, Jovic M, DeBiase M, Mann D. Evidence-based strategies in weight-loss mobile apps. *Am J Prev Med* 2013 Nov;45(5):576-582. [doi: [10.1016/j.amepre.2013.04.025](https://doi.org/10.1016/j.amepre.2013.04.025)] [Medline: [24139770](https://pubmed.ncbi.nlm.nih.gov/24139770/)]
43. Fadda M, Galimberti E, Fiordelli M, Romanò L, Zanetti A, Schulz PJ. Effectiveness of a smartphone app to increase parents' knowledge and empowerment in the MMR vaccination decision: a randomized controlled trial. *Hum Vaccin Immunother* 2017 Nov 02;13(11):2512-2521 [FREE Full text] [doi: [10.1080/21645515.2017.1360456](https://doi.org/10.1080/21645515.2017.1360456)] [Medline: [29125783](https://pubmed.ncbi.nlm.nih.gov/29125783/)]
44. Liu C, Wang D, Liu C, Jiang J, Wang X, Chen H, et al. What is the meaning of health literacy? A systematic review and qualitative synthesis. *Fam Med Community Health* 2020 May;8(2):e000351 [FREE Full text] [doi: [10.1136/fmch-2020-000351](https://doi.org/10.1136/fmch-2020-000351)] [Medline: [32414834](https://pubmed.ncbi.nlm.nih.gov/32414834/)]
45. Dunn Lopez K, Chae S, Michele G, Fraczkowski D, Habibi P, Chattopadhyay D, et al. Improved readability and functions needed for mHealth apps targeting patients with heart failure: an app store review. *Res Nurs Health* 2021 Feb;44(1):71-80 [FREE Full text] [doi: [10.1002/nur.22078](https://doi.org/10.1002/nur.22078)] [Medline: [33107056](https://pubmed.ncbi.nlm.nih.gov/33107056/)]
46. Powell AC, Singh P, Torous J. The complexity of mental health app privacy policies: a potential barrier to privacy. *JMIR Mhealth Uhealth* 2018 Jul 30;6(7):e158 [FREE Full text] [doi: [10.2196/mhealth.9871](https://doi.org/10.2196/mhealth.9871)] [Medline: [30061090](https://pubmed.ncbi.nlm.nih.gov/30061090/)]
47. Willcox JC, Dobson R, Whittaker R. Old-fashioned technology in the era of "bling": is there a future for text messaging in health care? *J Med Internet Res* 2019 Dec 20;21(12):e16630 [FREE Full text] [doi: [10.2196/16630](https://doi.org/10.2196/16630)] [Medline: [31859678](https://pubmed.ncbi.nlm.nih.gov/31859678/)]
48. Al-Azzam MK, Bader Alazzam M, Khalid al-Manasra M. MHealth for decision making support: a case study of eHealth in the public sector. *Int J Adv Comput Sci Appl* 2019;10(5):381-387. [doi: [10.14569/ijacsa.2019.0100547](https://doi.org/10.14569/ijacsa.2019.0100547)]
49. Maurer W, Seeber L, Rundblad G, Kochhar S, Trusko B, Kisler B, Vienna Vaccine Safety Initiative. Standardization and simplification of vaccination records. *Expert Rev Vaccines* 2014 Apr;13(4):545-559. [doi: [10.1586/14760584.2014.892833](https://doi.org/10.1586/14760584.2014.892833)] [Medline: [24597495](https://pubmed.ncbi.nlm.nih.gov/24597495/)]

50. Everson J. The implications and impact of 3 approaches to health information exchange: community, enterprise, and vendor-mediated health information exchange. *Learn Health Syst* 2017 Apr;1(2):e10021 [FREE Full text] [doi: [10.1002/lrh2.10021](https://doi.org/10.1002/lrh2.10021)] [Medline: [31245558](https://pubmed.ncbi.nlm.nih.gov/31245558/)]
51. Mosa AS, Yoo I, Sheets L. A systematic review of healthcare applications for smartphones. *BMC Med Inform Decis Mak* 2012 Jul 10;12:67 [FREE Full text] [doi: [10.1186/1472-6947-12-67](https://doi.org/10.1186/1472-6947-12-67)] [Medline: [22781312](https://pubmed.ncbi.nlm.nih.gov/22781312/)]
52. Kolasa K, Mazzi F, Leszczuk-Czubkowska E, Zrubka Z, Péntek M. State of the art in adoption of contact tracing apps and recommendations regarding privacy protection and public health: systematic review. *JMIR Mhealth Uhealth* 2021 Jun 10;9(6):e23250 [FREE Full text] [doi: [10.2196/23250](https://doi.org/10.2196/23250)] [Medline: [34033581](https://pubmed.ncbi.nlm.nih.gov/34033581/)]
53. Osmanliu E, Rafie E, Bédard S, Paquette J, Gore G, Pomey MP. Considerations for the design and implementation of COVID-19 contact tracing apps: scoping review. *JMIR Mhealth Uhealth* 2021 Jun 09;9(6):e27102 [FREE Full text] [doi: [10.2196/27102](https://doi.org/10.2196/27102)] [Medline: [34038376](https://pubmed.ncbi.nlm.nih.gov/34038376/)]

## Abbreviations

**EHR:** electronic health record

**HIPAA:** Health Insurance Portability and Accountability Act

**mHealth:** mobile health

**PC:** principal component

**PCA:** principal component analysis

*Edited by A Mavragani; submitted 28.01.22; peer-reviewed by T Giles-Vernick, P Nelson, J Ropero; comments to author 24.03.22; revised version received 31.05.22; accepted 16.06.22; published 11.10.22.*

*Please cite as:*

*Shaw Jr G, Nadkarni D, Phann E, Sielaty R, Ledenyi M, Abnowf R, Xu Q, Arredondo P, Chen S*

*Separating Features From Functionality in Vaccination Apps: Computational Analysis*

*JMIR Form Res* 2022;6(10):e36818

URL: <https://formative.jmir.org/2022/10/e36818>

doi: [10.2196/36818](https://doi.org/10.2196/36818)

PMID: [36222791](https://pubmed.ncbi.nlm.nih.gov/36222791/)

©George Shaw Jr, Devaki Nadkarni, Eric Phann, Rachel Sielaty, Madeleine Ledenyi, Razaan Abnowf, Qian Xu, Paul Arredondo, Shi Chen. Originally published in *JMIR Formative Research* (<https://formative.jmir.org>), 11.10.2022. 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

# Social Media Mining of Long-COVID Self-Medication Reported by Reddit Users: Feasibility Study to Support Drug Repurposing

Jonathan Koss<sup>1</sup>, MSc; Sabine Bohnet-Joschko<sup>1</sup>, Prof Dr

Department of Management and Entrepreneurship, Faculty of Management, Economics and Society, Witten/Herdecke University, Witten, Germany

**Corresponding Author:**

Jonathan Koss, MSc

Department of Management and Entrepreneurship

Faculty of Management, Economics and Society

Witten/Herdecke University

Alfred-Herrhausen-Str. 50

Witten, 58455

Germany

Phone: 49 2302926475

Email: [jonathan.koss@uni-wh.de](mailto:jonathan.koss@uni-wh.de)

## Abstract

**Background:** Since the beginning of the COVID-19 pandemic, over 480 million people have been infected and more than 6 million people have died from COVID-19 worldwide. In some patients with acute COVID-19, symptoms manifest over a longer period, which is also called “long-COVID.” Unmet medical needs related to long-COVID are high, since there are no treatments approved. Patients experiment with various medications and supplements hoping to alleviate their suffering. They often share their experiences on social media.

**Objective:** The aim of this study was to explore the feasibility of social media mining methods to extract important compounds from the perspective of patients. The goal is to provide an overview of different medication strategies and important agents mentioned in Reddit users’ self-reports to support hypothesis generation for drug repurposing, by incorporating patients’ experiences.

**Methods:** We used named-entity recognition to extract substances representing medications or supplements used to treat long-COVID from almost 70,000 posts on the “/r/covidlonghaulers” subreddit. We analyzed substances by frequency, co-occurrences, and network analysis to identify important substances and substance clusters.

**Results:** The named-entity recognition algorithm achieved an F1 score of 0.67. A total of 28,447 substance entities and 5789 word co-occurrence pairs were extracted. “Histamine antagonists,” “famotidine,” “magnesium,” “vitamins,” and “steroids” were the most frequently mentioned substances. Network analysis revealed three clusters of substances, indicating certain medication patterns.

**Conclusions:** This feasibility study indicates that network analysis can be used to characterize the medication strategies discussed in social media. Comparison with existing literature shows that this approach identifies substances that are promising candidates for drug repurposing, such as antihistamines, steroids, or antidepressants. In the context of a pandemic, the proposed method could be used to support drug repurposing hypothesis development by prioritizing substances that are important to users.

(*JMIR Form Res* 2022;6(10):e39582) doi:[10.2196/39582](https://doi.org/10.2196/39582)

## KEYWORDS

social media mining; drug repurposing; long-COVID; crowdsourcing; COVID-19; Reddit; social media; content analysis; network analysis; recognition algorithm; treatment

## Introduction

### Background

Since the beginning of the COVID-19 pandemic, over 480 million people have been infected and more than 6 million

people have died from COVID-19 worldwide [1]. In some patients with acute COVID-19, symptoms manifest over a longer period of time [2]. Owing to this phenomenon, the term “long-COVID” (LC) has emerged [3]. LC refers to both postacute (lasting longer than 4 weeks) and chronic (lasting longer than 12 weeks) symptoms [3,4]. At least one symptom

persists in 32%-87% of previously hospitalized patients [4]. Furthermore, the incidence of LC is estimated to range between 10% and 35% in individuals who have not been hospitalized [5]. The economic costs associated with LC symptomatology could be significant. Reynolds et al [6] stated that chronic fatigue syndrome, which has similar characteristics to LC [7], leads to a 37% reduction in household productivity and a 54% reduction in labor force productivity in the United States. Unmet medical needs have motivated immense research activities. ClinicalTrials.gov lists more than 7000 studies in the field of COVID-19, including more than 600 LC-specific studies [8].

### Retrospective Clinical Analysis

The large number of ongoing studies highlights a key challenge in drug development. There are numerous substances that are potentially effective. It is therefore essential to identify promising substances and narrow down the number of potential drug candidates to those showing the most promise. Given the urgency, scarcity of financial resources, and the high risk of failure in pharmaceutical research [9], drug repurposing (DR) appears to be a promising strategy for LC drug development. The exploitation of existing drugs for new therapeutic purposes usually leads to shorter development cycles with lower costs [10]. For example, existing drugs have proven to be safe for use in humans. Accordingly, phase I clinical trials are not required [10]. From a historical perspective, DR has often been serendipitous [10,11], but systematic approaches also exist to identify promising target leads [12]. One of these approaches is retrospective clinical analysis [12], which has already been used in the context of the COVID-19 pandemic [12]. Retrospective clinical analysis involves learning from real-world experience (eg, evaluating clinical case reports) to hypothesize applications of existing drugs for new indications [13].

### Mining Patients' Experiences From Social Media

Traditionally, retrospective clinical analysis is based on information that is collected and stored in databases, explicitly dedicated for health care system-related applications. Signals for potential DR are subsequently generated by professionals analyzing the data. Nowadays, there is a growing awareness in medical research that the collective intelligence of the affected patients, jointly searching for a solution to improve a medical condition, can be a driver of innovation [14,15] that is leveraged in the innovation process known as "crowdsourcing" [14,16]. While traditional crowdsourcing aims at active collaboration (eg, between a pharmaceutical company and an external patient group), online forums enable a passive approach to collecting real-world data by offering relevant content for analysis, which is also called "passive crowdsourcing" [17,18]. For instance, researchers have analyzed data from disease-specific social media platforms to identify medications that are used outside the approved indication (off-label use) [10,19]. Off-label use provides information to support hypotheses regarding DR [10,19]. These approaches save time and costs associated with data collection, while incorporating patients' real-world experiences. However, social media mining (SMM) [19], a term

that refers to the collection of methods used for conducting passive crowdsourcing, poses significant risks in terms of bias [10,19]. For instance, owing to the age-related user behavior on social media platforms, the data might not be generalizable to the whole population [10].

### Generating Hypotheses on DR From Discussions Among Long-Haulers

In this study, we aimed to capture substances such as medications and supplements that are relevant to the coping strategies of patients with LC. Accordingly, we applied the principle of retrospective clinical analysis using passive crowdsourcing by applying SMM. Since there is no approved drug for LC, promising LC candidates based on off-label properties could not be determined. Instead, we used an exploratory method, mainly consisting of the application of named-entity recognition (NER) and network analysis, aiming to provide an overview of different treatment strategies and important compounds from the patient's perspective for DR hypothesis generation. Methodologically similar approaches have previously been used to identify substances used in self-medication regarding opioid withdrawal [20] or to monitor potential drug interactions and reactions [21]. Furthermore, network analysis has been used to explore discussions related to certain diseases [22] or to explore the public perspective on vaccines [23,24]. For example, Lewis et al [22] used network analysis to analyze reasons for older adults to join a diabetes online community. Luo et al [24] used network analysis to explore public perceptions of the COVID-19 vaccine.

To our best knowledge, this study is the first to explore treatment strategies and important medications to support DR hypothesis generation by applying network analysis.

### Research Objectives

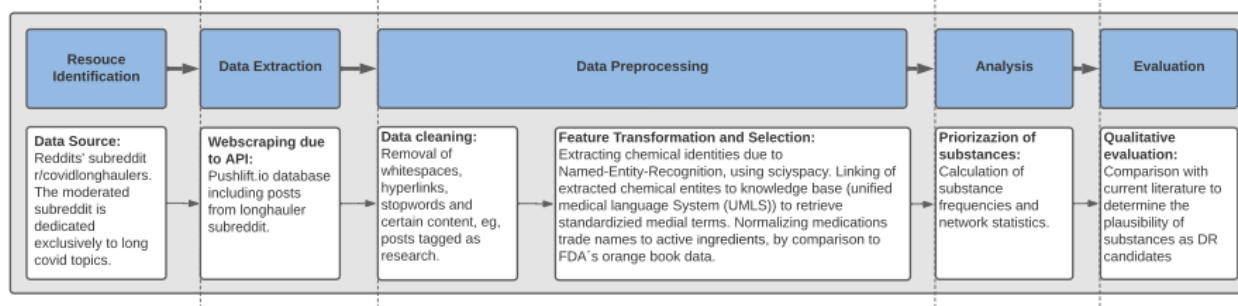
The aim of this feasibility study was to evaluate whether the proposed method can be employed to support DR hypothesis generation based on the experiences of affected individuals shared on Reddit. To this end, we first explored which substances are mentioned in LC online discussions regarding self-medication. Second, we investigated whether there are clusters of substances often discussed together, indicating treatment strategies. Third, we attempted to identify the most important substances in these clusters to indicate respective treatment strategies.

## Methods

### Overview

The methodology used in this study consists of the following steps: (1) extraction of appropriate data, (2) detection of substance entities mentioned in users' posts using NER, and (3) analysis of substance frequencies and co-occurrence networks of substance entities. Figure 1 outlines the end-to-end workflow in detail.

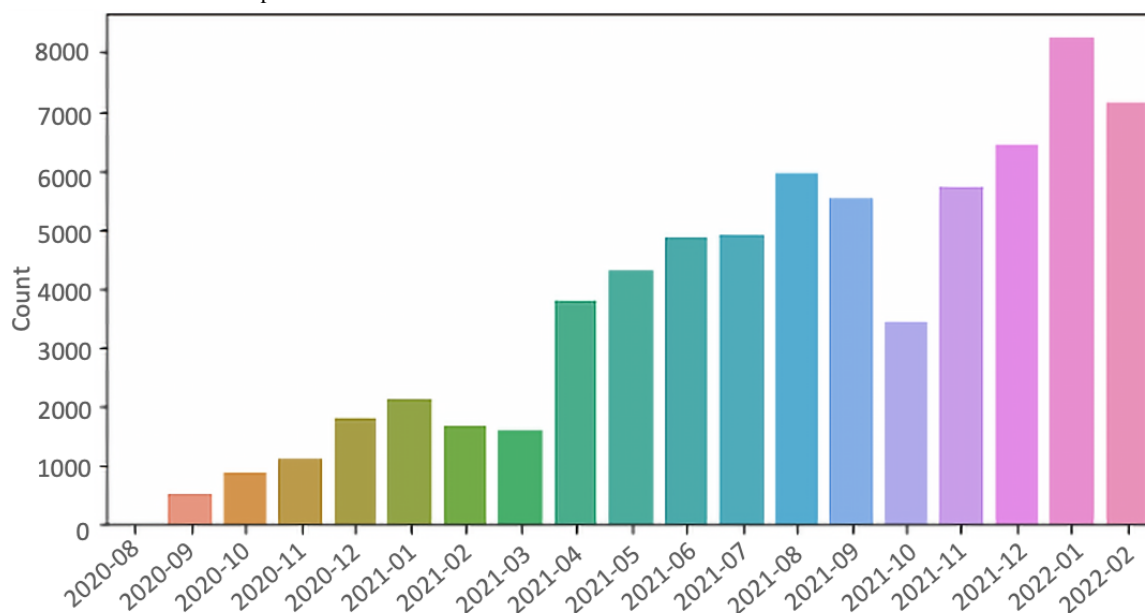
**Figure 1.** End-to-end detailed study workflow. The workflow can be divided into the following steps: resource identification, data extraction, data preprocessing, analysis, and evaluation. API: application programming interface; DR: drug repositioning; FDA: Food and Drug Administration.



## Data Source and Extraction

Reddit is a social media platform that is organized in theme-specific forums called “subreddits” [20]. The data extraction process was performed using Pushshift [25], which is a platform that collects Reddit data and has been available to researchers since 2015 [25]. The extracted data consist of posts and metadata from the subreddit “/r/covidlonghaulers,” which has already been used to explore LC symptoms [26,27]. This subreddit is actively moderated by specific users and provides a medium for LC-related discussions. The content is subject to strict rules prohibiting the promotion of alternative treatment, misinformation, and conspiracy theories. As of January 3, 2022, the subreddit had over 24,000 subscribers and 20,000 threads.

**Figure 2.** Overview of the number of posts at different dates.



## Substance Entity Extraction

First, the text was preprocessed to improve the data quality. For example, hyperlinks, tabs, and blank lines were removed. The substances of interest mentioned by patients in posts needed to be extracted and structured for subsequent analysis [19] using NER [28]. We defined substances of interest as explicitly mentioned substances or groups of substances that can be considered as treatments. For instance, we captured conventional supplements (eg, vitamin supplements) or prescription medications (eg, antidepressants) that were discussed by users.

Users self-report their LC experiences such as discussing symptoms [26] and medications.

Beyond the extraction of posts, metadata such as the username, date of the post, or the so-called “link flair text” can be extracted. Link flair text represent thematic tags that are used to associate posts (initial and subsequent posts) with specific categories. This provides researchers with the ability to exclude data unrelated to the analysis. For example, posts tagged as articles, research articles, or humor posts were excluded from the analysis (see [Multimedia Appendix 1](#)). Additionally, posts without tags were excluded. The analyzed data included 68,268 posts written by 8717 users between August 31, 2020, and March 1, 2022 (Figure 2).

ScispaCy provides several NER models related to medical issues [29] and was used to extract symptoms from LC Reddit posts [27]. In principle, two ScispaCy models are applicable for this purpose. The first model is the en\_ner\_bc5cdr\_md model [30], which can detect chemical substances and diseases. Since the module covers chemical substances in general and not specific medications or dietary supplements, we defined stop words such as “ethanol” to narrow the focus of the analysis to substances of interest. The second model is the Med7 model [31], which specifically focuses on drug extraction. Running both models together yielded the best results [27]. Negated substances were

excluded from the extraction by considering negotiation terms. For this purpose, we added the Negex algorithm [32] to the NER ScispaCy pipeline. Negex identifies different forms of negation patterns and was initially developed for application to clinical texts [32]. Subsequently, the named entities were normalized and filtered to improve data quality for subsequent analyses. Therefore, the extracted entities were matched against an external knowledge base and were either replaced with standard medical vocabulary or discarded if no match was found. For this purpose, we used the ScispaCy entity linker, which matches entities with the unified medical language system (UMLS) knowledge base [33]. The EntityLinker pipeline performs a string nearest-neighbor search for entities to match them with the UMLS concepts [29]. We considered 0.85 as a threshold value for the overlap with UMLS concepts [29]. To evaluate the entity extraction performance, 500 randomly selected posts from the entire corpus were manually annotated. For data annotation, two annotators were involved. The intercoder reliability was 0.94. The F1 score was used as the evaluation metric [34]. By applying inexact string matching between data from the Food and Drug Administration's Orange Book [35] and the extracted entities, brand names were normalized by their active ingredient [36]. For instance, "zyrtec" was replaced with "ceterizine hydrochloride."

## Network Analysis

The network analysis method offers the possibility of visualizing and evaluating relationships in text. In this study, we used network analysis to obtain an overview of the spectrum of substances and identify potential substance clusters. Similar approaches have been used to identify substances and their effects, including self-medication in opioid withdrawal [20] or pharmacovigilance settings [21].

Features for network analysis consist of the nodes (represented by the extracted entities) and the "edges," which represent the relationship between the nodes as weight based on the co-occurrence of entities. A co-occurrence was defined as the mention of two or more substances in one post [37,38]. Duplicates of entities within one post were removed to avoid assigning more weight to longer posts that mention specific entities more frequently. Subsequently, the information on the co-occurrence of substances was converted into a pointwise mutual information (PMI) matrix [20,39,40]. We only considered associations between entities that co-occurred more frequently than expected based on their overall frequency, also called positive PMI, which was proven to be beneficial for extracting semantic representations [20,41]. To improve the quality of the visualization and analysis, substances occurring less than 10 times and node pairs below the average PMI weight were excluded [22,42]. False-positive nodes were manually removed. Using the PMI matrix, an undirected graph was created and analyzed using Gephi software. Gephi is an open-source software for network analysis, which allows spatialization, filtering, navigation, manipulation, and clustering of entities [43].

## Community Detection

We used clustering, also referred to as "community detection," to identify potential drug and/or supplement strategies for LC.

Community detection describes the clustering of nodes (in our case, substances) that are strongly associated with each other according to their edges. Hence, a cluster consists of substances that are strongly associated and discussed with one another. Communities can be determined using various clustering algorithms. For this purpose, the relatively new Leiden algorithm was used [44]. In contrast to the Louvain algorithm, which has been widely used in network analysis in the past, the Leiden algorithm has several advantages [45] such as having more meaningful partitioning [44]. The modularity value (Q) was used as the quality function [44]. A Q value of at least 0.3 implies meaningful clustering [44]. The result of clustering, and consequently Q, was significantly determined by the preselected resolution [46]. Following an iterative process, we aimed to find a proper balance between the number and relevance of the discovered communities and the resulting modularity by applying different resolution values [46]. To analyze the most important substances in the network, we calculated the degree centrality (number of linkages of a node) [47]; the higher the centrality, the more important the substance is in the network [47,48].

## Results

### Substances in Self-Reports

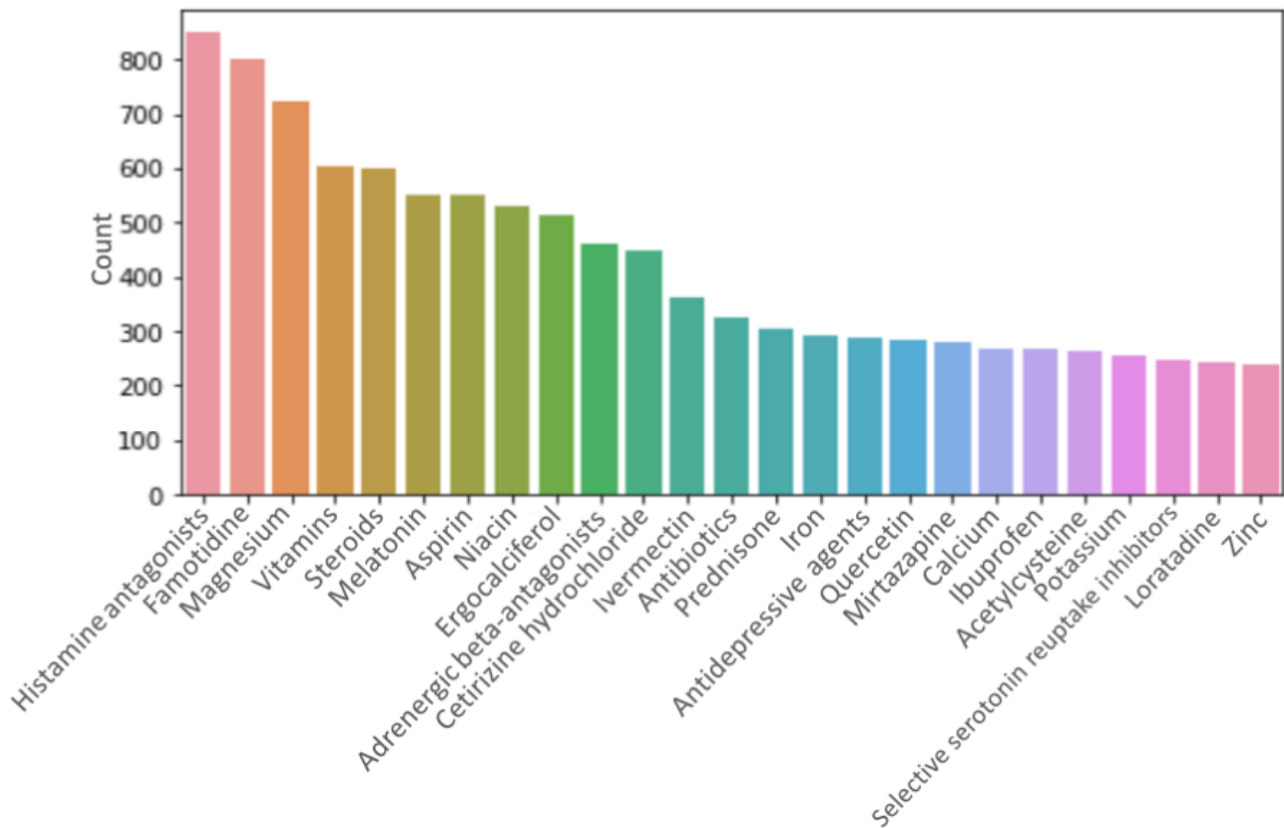
The NER algorithm achieved an F1 score of 0.67 (precision=0.69, recall=0.66). Error analysis was performed on the incorrectly labeled entities. Errors were classified into lexical and dictionary errors [49]. A lexical error (38.5%) refers to the case in which users employ a variety of terms when referring to the substances they use. For example, our model failed to detect and extract the term "benzos," since it is a slang abbreviation for the drug "benzodiazepine" and was not indexed in the model. Another example of expressions that our model did not recognize are compound terms of more than a single word (eg, "anti histamines") with the algorithm only extracting "histamines" and was thus missing the preceding word "anti," giving the extracted entity a different meaning.

A dictionary error (61.5%) refers to certain terms that are not specifying a concrete substance but rather a substance group; for example, "electrolytes" were not captured, whereas explicitly mentioned substances representing electrolytes such as "magnesium" were reliably detected. Furthermore, the algorithm extracted substances that are not considered as treatments, such as "chlorine."

A total of 28,447 substance entities and 5789 word-co-occurrence pairs were extracted. "Histamine antagonists," "famotidine," "magnesium," "vitamins," and "steroids" were the most frequently mentioned substances that appeared at least once in a post (duplicate mentions of a substance in a post were disregarded) (Figure 3). A list of all substances can be found in [Multimedia Appendix 1](#).

The most frequent word pairs are listed in [Table 1](#). For example, the pairing that occurred most frequently with 218 mentions was cetirizine hydrochloride–famotidine.

**Figure 3.** The 25 most frequently mentioned substances that appeared at least once in a post. For example, “histamine antagonists” were discussed in more than 800 different posts.



**Table 1.** The most frequent co-occurrences.

Rank	Substance–Substance pair	Frequency (number of mentions)
1	Cetirizine Hydrochloride–Famotidine	218
2	Famotidine–Histamine Antagonists	135
3	Potassium–Magnesium	106
4	Famotidine–Loratadine	98
5	Ergocalciferol–Magnesium	96
6	Cetirizine Hydrochloride–Histamine Antagonists	95
7	Aspirin–Famotidine	88
8	Loratadine–Histamine Antagonists	82
9	Zinc–Ascorbic Acid	78
10	Famotidine–Melatonin	78

## Substance Clusters

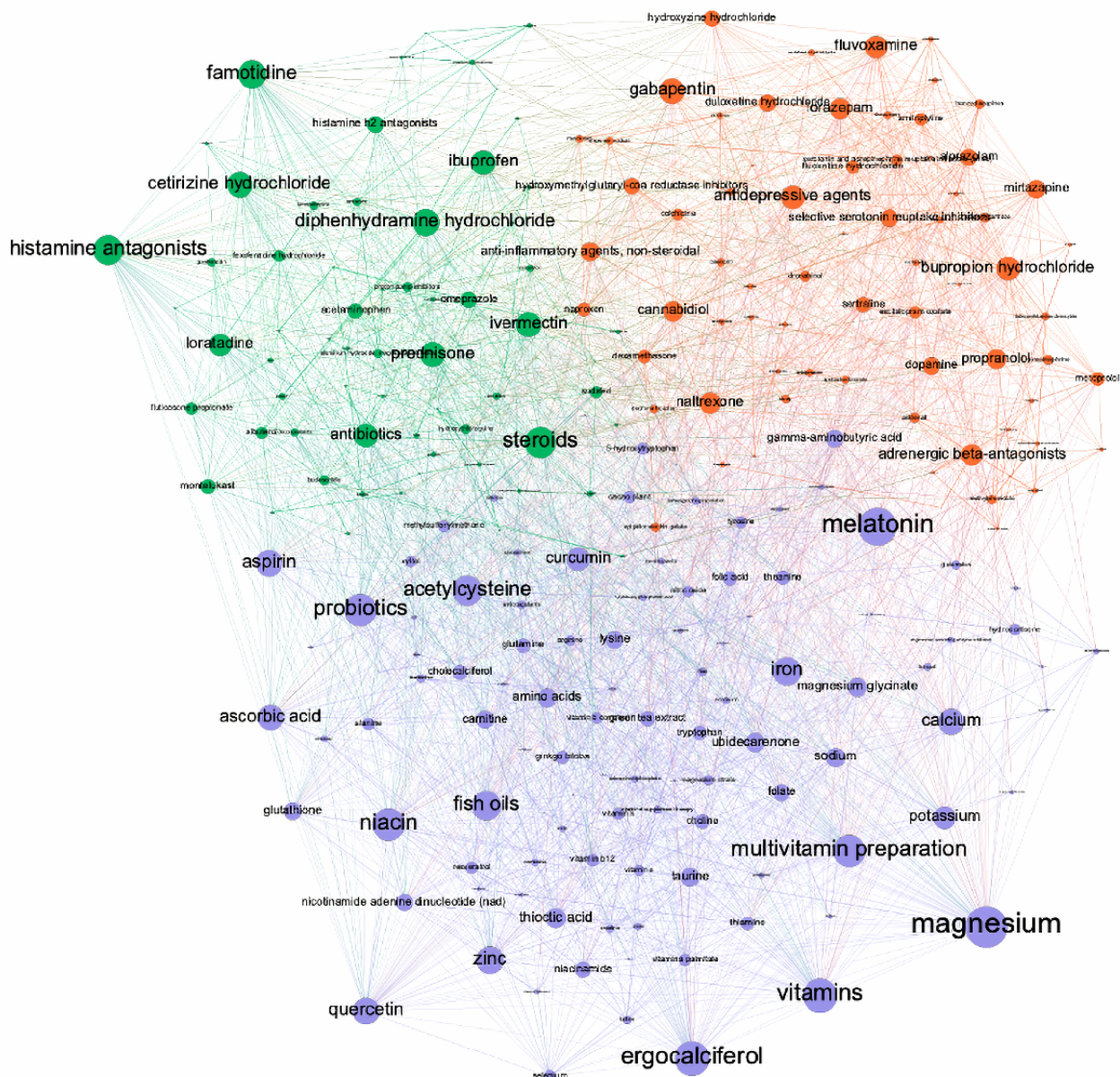
### Overview

Using a resolution of 0.6, three clusters were found. They consisted of 244 nodes and 3570 edges. The modularity value

was 0.48, indicating a reasonable partitioning of communities [50,51]. The average clustering coefficient was 0.414. Overall, these scores indicated that the network (Figure 4) had no random structure [47]. Coloring in the network indicates the community and node size degree centrality.



**Figure 4.** Substance network and clusters. Substances are presented by nodes; the larger the size of a node, the higher degree centrality. Coloring refers to detected communities; violet represents cluster 1, orange refers to cluster 2, and green highlights substances of cluster 3.



### Cluster 1

Cluster 1 mainly consisted of supplements and several over-the-counter (OTC) medications, which are often used in the context of flu-like diseases (Figure 5). The top 10 most important substances and the respective substance classes

measured by degree centrality are displayed in Table 2. The retrieved entities belong to the drug classes of electrolyte/mineral replacement, vitamins, respiratory tract agents such as acetylcysteine, and nutritional supplements such as fish oil or probiotics.

Figure 5. Examples of posts containing typical substance co-occurrences of clusters.



Table 2. Characteristics of clusters.

Cluster	Total share of nodes	Ten most important substances (by degree centrality)
1	42.85%	magnesium, melatonin, ergocalciferol, vitamin, multivitamin preparation, niacin, probiotics, acetylsteine, fish oils, zinc
2	29.51%	gabapentin, bupropion hydrochloride, antidepressive agents, fluvoxamine, adrenergic beta-antagonists, naltrexone, lorazepam, cannabidiol, propranolol, nonsteroidal anti-inflammatory agents
3	26.64%	steroids, histamine antagonists, famotidine, diphenhydramine hydrochloride, cetirizine hydrochloride, prednisone, ibuprofen, antibiotics, loratadine, ivermectin

### Cluster 2

Cluster 2 mostly included prescription medicines such as those used for the treatment of psychological, mental, or neurological disorders (Figure 4). Examples of the 10 most important substances (Table 2) include the anticonvulsant drug gabapentin, antidepressants such as bupropion hydrochloride, and adrenergic  $\beta$ -antagonists such as propranolol. Furthermore, opioid antagonists such as naltrexone, anxiolytics such as lorazepam, and nonsteroidal anti-inflammatory agents such as naproxen incorporated high degree centrality.

### Cluster 3

Cluster 3 mainly included prescription and OTC medicines that are often used to treat allergic reactions and inflammation (Figure 4). The top 10 important substances (Table 2) belonged to the drug classes of steroids such as prednisone, antihistamines such as famotidine, antibiotics, and nonsteroidal anti-inflammatory agents such as ibuprofen.

## Discussion

### Principal Results

Posts were extracted from an LC-specific subreddit to analyze the discussed substances with the aim of evaluating whether the proposed method could be used to support hypothesis development for DR. In the absence of approved medications for LC (which will also be the case in future pandemic situations), all substances can be considered for off-label use, which makes it difficult to evaluate the potential candidates that apply traditional SMM DR approaches. Instead of filtering substances for off-label use, we considered frequencies and

network analysis to facilitate the identification of important substances from the patients' point of view.

The substances mentioned the most frequently in our feasibility study were antihistamines in general and famotidine, followed by supplements such as magnesium. Moreover, vitamins and steroids were frequently discussed substances. To analyze the strength of the substance-substance combination, the PMI was used to compare the strength of the associations with random associations considering the overall frequency of substances. These substances and their associations formed a nonrandom network consisting of three substance clusters, identified by community detection, implying systematic discussion and usage of substances. For instance, the most frequently mentioned class of substances, histamine antagonists, was found to be highly associated with other inflammatory substances such as steroids.

The clusters, mainly consisting of anti-inflammatory agents and supplements, incorporate the most often mentioned entities. Moreover, medications used for the treatment of psychological, mental, or neurological disorders also formed a cluster. The latter cluster can be assumed to be a less prevalent treatment regime in our sample because its substances occur less frequently. Nevertheless, all clusters reflect the treatment approaches described by the users.

### Supporting DR Hypothesis Generation by Analyzing Patients' Self-Reports

The results of our feasibility study highlight drugs and self-treatment strategies discussed by long-haulers on Reddit. We were able to successfully identify substance communities (representing different treatment strategies) in the substance network and drugs of high(er) importance to users (based on

degree centrality) within these communities. Comparing the results to the current literature, our findings are supported by the successful identification of promising drug candidates already discussed by the scientific community. For instance, according to Crook et al [52], antihistamines are considered potential DR candidates. Significant improvements in long-term symptoms have been reported in case reports [53] and observational studies [54]. This seems to be backed by discussions from long-haulers: antihistamines are the most frequently discussed substances and are important (measured by degree centrality) to the cluster of anti-inflammatory drugs.

Similarly, Crook et al [52] concluded in their review that antidepressants such as serotonin-norepinephrine reuptake inhibitors and selective serotonin reuptake inhibitors could be repurposed for the treatment of LC, as they have been associated with a reduced risk of death or intubation in acute COVID-19 cases [52,55] and a reduction in peripheral inflammatory markers [52,56]. In April 2021, Sukhatme et al [57] conducted review on the mechanisms of action of fluvoxamine and its role in acute COVID-19 treatment. The authors concluded that it “is also tempting to speculate on a role for fluvoxamine in COVID-19 long-haulers.” In April 2022, Khani et al [58] expressed the hypothesis that the majority of LC symptoms might not be directly due to COVID-19 but likely result from COVID-19-associated inflammation and Epstein-Barr virus (EBV) reactivation. The authors argue that fluvoxamine might have beneficial effects in reducing LC symptoms due to the modulatory effects on central mechanisms (eg, reduction of endoplasmic reticulum stress and inflammation). In medication cluster 2, fluvoxamine was the second most important antidepressant, as measured by degree centrality (Table 2). Interestingly, our data are based on posts up to March 2022, and clearly indicate that users already consider this drug to be important in their LC treatment strategies while the scientific community is still working on hypothesis generation.

The other antidepressant that is most frequently mentioned and most central to long-hauler discussions was bupropion hydrochloride, representing a norepinephrine/dopamine reuptake inhibitor (NDRI) [59]. This drug is currently barely discussed in the research community. There seems to be a significant discrepancy in the perception of importance of NDRI of long-haulers and the scientific community.

Additional drugs that appear to be important to long-haulers but that are barely discussed in clinical research include naltrexone, adrenergic  $\beta$ -antagonists, and prednisone. For instance, prednisone, as a corticosteroid, emerged as the most central steroid used for self-treatment by long-haulers in our study. Chen et al [60] found a high incidence of EBV coinfection in acute COVID-19 cases and concluded that patients may be advised to use a corticosteroid. Following the hypothesis of Khani et al [58] that EBV is also of central importance in LC, corticosteroid could be useful in the treatment of LC, and thus explain the identified central position in the long-haulers discussion identified in this study. In fact, Goel et al [61] reported that systemic steroids are helpful in hastening the recovery of a selected subset of patients with LC [61]. Another recently published single-center interventional pre-post study demonstrated that low-dose naltrexone is safe in patients with

LC and may improve well-being (eg, reduce symptomatology) [62].

In summary, reports on clinical research indicate that our proposed method might support the early identification of promising DR candidates. When incorporating patients' experiences regarding specific drugs, network analysis has proven to be especially useful to slice the data in a meaningful way. In particular, network analysis enables the identification of different (self-) treatment strategies and corresponding drugs beyond raw frequencies. For example, medications used for the treatment of psychological, mental, or neurological disorders represent a medication cluster; however, these drugs are generally reported at lower frequency. This might result from the fact that most of these drugs are only available on prescription and therefore access is limited in contrast to OTCs. Substance community detection and degree centrality highlight the strategies and drugs that otherwise could be overlooked. Furthermore, network analysis allows to distinguish between systematic and random discussions of substances. This is important information, as a systematic discussion might be an indicator for data quality and the knowledge of the crowd. Our results imply that patients' experiences shared on social media influence others' self-treatment decisions [63]. Positive experiences reported by users will lead to other users adopting the same approaches, leading to the increase of discussion of potentially helpful substances [64].

However, it is beyond the scope of this study to review all of the identified compounds for their potential relevance to LC DR. We encourage professionals to consider the findings as a starting point for hypothesis generation by narrowing down potential DR candidates to drugs that are important for long-haulers. Clearly, we cannot derive any conclusions about effectiveness based on this analysis. Nevertheless, these drugs appear to be frequently used by long-haulers as treatments. Therefore, these drugs should be further evaluated by the scientific community to determine whether they might be effective or even harmful, which would also be of relevance to communicate from a public health perspective.

### Limitations and Future Work

Our study has several limitations in line with comparable studies based on similar methods. As our NER algorithm relies on pretrained models, the error analysis we performed implies that a custom model trained on annotated examples of the posts used in this study would increase the accuracy of the results. To avoid missing entities in the normalization process (eg, due to the use of slang), a customized dictionary could be defined, which links slang terms of substances used by patients to their medical terminology. Further, substances that fail in the normalization process could be revised and normalized manually. However, the performance of the NER algorithm can be considered to be appropriate. Other studies applying pretrained NER algorithms to extract medical entities from Reddit data showed similar F1 scores. Foufi et al [49] used PKDE4J [65] to identify biomedical substances from disease-specific subreddits, and 71.48% of the extracted entities were correctly labeled. Šćepanović et al [66] evaluated six state-of-the-art pretrained language models in a bidirectional long short-term memory-conditional random

field (BiLSTM-CRF) model to identify medical entities in disease-specific subreddits; the F1 scores ranged from 0.64 to 0.73. Approaches combining different pretrained language models can improve performance. For example, Šćepanović et al [66] built a customized NER system by combining a BiLSTM-CRF sequence labeling architecture with contextual embeddings, which scored 0.71 on symptoms and 0.77 on drugs.

### Determining the Outcome of Substances

We analyzed the importance of substances in the long-haulers' discussions, but we did not analyze whether the substances were helpful in terms of the outcomes, which should be considered in future studies. Several approaches can be applied to approximate outcomes. The gold standard is the manual examination of medication-specific posts by medical professionals incorporating domain knowledge. Further programmatic approaches could include analyses of average sentiments regarding drugs to determine whether a treatment is perceived to be useful by users [67]. Another possibility could be to capture correlations between substances and effects mentioned in a sentence by the application of dependency parsing [68] or by implementing an observational study design [69]. However, reliable assessment of causality of underlying treatment effects is impossible because of various limitations [19]; for instance, inaccurate use of medical terminology by users would bias the results, even if machine-learning algorithms perform with perfect accuracy. Moreover, our analysis indicates a trend of polymedication, which would confound the analysis of single substances. Furthermore, the data quality is low compared with data obtained in traditional study designs, and several confounders such as demographic variables are unknown. In general, even if holistic patient information is available, social media data should be interpreted with caution. For instance, ivermectin was considered a potential DR candidate for acute COVID-19. However, this is suspected to be supported by flawed research [70] and it was later found to be ineffective in randomized clinical trials [71]. While we cannot evaluate the efficacy of ivermectin in patients with LC, it can be assumed that users' treatment decisions may also be influenced by potentially flawed information.

### Generalizability of Results and Inference of Indications

Even though important substances were found from the users' perspective, the results cannot be generalized for all patients

with LC since demographics or symptom distributions were not analyzed. Although previous studies [26,27] indicate that the spectrum of symptoms in the subreddit broadly agrees with recently published studies on LC [26,27], it can be assumed that the treatment choice of users depends on their symptoms. Therefore, the implementation of a bimodal network to reveal correlations between symptoms and medications could be useful. Moreover, the demographic distribution of users might not be representative of the entire long-hauler population. In fact, research suggests that the distribution of Reddit is skewed by age and gender [72], which is a major limitation to SMM in general [19], indicating that the data source represents a young male subpopulation [72]. There are two possible improvements in future SMM studies: (1) combining data from multiple platforms could lower user bias depending on the specific platform; and (2) algorithms can be applied to infer demographic variables and analyze diverse user data, including metadata [19].

### Conclusions

In this feasibility study, we tested the application of SMM methods to support the development of hypotheses on LC DR. To this end, we extracted substance entities to analyze frequencies and co-occurrences and subsequently used them to identify substance clusters. Our results highlight certain approaches to DR, such as antihistamines, steroids, or antidepressants, while also indicating that patients experiment with a wide range of substances in a systematic manner. This feasibility study demonstrates that network analysis can be used to characterize the medication strategies discussed. Comparison with existing literature indicates that the approach identifies substances that are plausible candidates for drug repurposing. The comparison also showed that some substances are important from the users' point of view while they are barely discussed in the scientific community. These substances should be reviewed by experts to assess their potential efficacy or harmfulness. The result could either lead to a DR hypothesis or underline the need to communicate the potential risks of the drug in the community of long-haulers. In the context of a pandemic, the proposed method might be used to support DR hypothesis development by prioritizing substances that are important to users in a cost- and time-effective manner.

---

### Acknowledgments

This research was supported as part of the ATLAS project "Innovation and digital transformation in healthcare" by the state of North Rhine-Westphalia, Germany (grant number: ITG-1-1).

---

### Conflicts of Interest

None declared.

---

### Multimedia Appendix 1

Link flair text (LFT) tags included and excluded from the analysis, and total substances extracted from posts. [[DOCX File, 25 KB - formative\\_v6i10e39582\\_app1.docx](#)]

---

### References

1. WHO Coronavirus (COVID-19) Dashboard. World Health Organization. URL: <https://covid19.who.int> [accessed 2022-08-30]
2. Lopez-Leon S, Wegman-Ostrosky T, Perelman C, Sepulveda R, Rebolledo P, Cuapio A, et al. More than 50 long-term effects of COVID-19: a systematic review and meta-analysis. *Sci Rep* 2021 Aug 09;11(1):16144. [doi: [10.1038/s41598-021-95565-8](https://doi.org/10.1038/s41598-021-95565-8)] [Medline: [34373540](https://pubmed.ncbi.nlm.nih.gov/34373540/)]
3. Michelen M, Manoharan L, Elkheir N, Cheng V, Dagens A, Hastie C, et al. Characterising long COVID: a living systematic review. *BMJ Glob Health* 2021 Sep 27;6(9):e005427 [FREE Full text] [doi: [10.1136/bmjgh-2021-005427](https://doi.org/10.1136/bmjgh-2021-005427)] [Medline: [34580069](https://pubmed.ncbi.nlm.nih.gov/34580069/)]
4. Nalbandian A, Sehgal K, Gupta A, Madhavan MV, McGroder C, Stevens JS, et al. Post-acute COVID-19 syndrome. *Nat Med* 2021 Apr 22;27(4):601-615 [FREE Full text] [doi: [10.1038/s41591-021-01283-z](https://doi.org/10.1038/s41591-021-01283-z)] [Medline: [33753937](https://pubmed.ncbi.nlm.nih.gov/33753937/)]
5. Pavli A, Theodoridou M, Maltezou HC. Post-COVID syndrome: incidence, clinical spectrum, and challenges for primary healthcare professionals. *Arch Med Res* 2021 Aug;52(6):575-581 [FREE Full text] [doi: [10.1016/j.arcmed.2021.03.010](https://doi.org/10.1016/j.arcmed.2021.03.010)] [Medline: [33962805](https://pubmed.ncbi.nlm.nih.gov/33962805/)]
6. Reynolds KJ, Vernon SD, Bouchery E, Reeves WC. The economic impact of chronic fatigue syndrome. *Cost Eff Resour Alloc* 2004 Jul 21;2(1):4 [FREE Full text] [doi: [10.1186/1478-7547-2-4](https://doi.org/10.1186/1478-7547-2-4)] [Medline: [15210053](https://pubmed.ncbi.nlm.nih.gov/15210053/)]
7. Wong TL, Weitzer DJ. Long COVID and myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS)-a systemic review and comparison of clinical presentation and symptomatology. *Medicina* 2021 May 26;57(5):418 [FREE Full text] [doi: [10.3390/medicina57050418](https://doi.org/10.3390/medicina57050418)] [Medline: [33925784](https://pubmed.ncbi.nlm.nih.gov/33925784/)]
8. Clinical studies related to COVID-19. *ClinicalTrials.gov*. URL: <https://clinicaltrials.gov/ct2/results?cond=COVID-19> [accessed 2022-08-30]
9. Pammolli F, Magazzini L, Riccaboni M. *Nat Rev Drug Discov* 2011 Jul 1;10(6):428-438. [doi: [10.1038/nrd3405](https://doi.org/10.1038/nrd3405)] [Medline: [21629293](https://pubmed.ncbi.nlm.nih.gov/21629293/)]
10. Pushpakom S, Iorio F, Eyers PA, Escott KJ, Hopper S, Wells A, et al. Drug repurposing: progress, challenges and recommendations. *Nat Rev Drug Discov* 2019 Jan 12;18(1):41-58. [doi: [10.1038/nrd.2018.168](https://doi.org/10.1038/nrd.2018.168)] [Medline: [30310233](https://pubmed.ncbi.nlm.nih.gov/30310233/)]
11. Cha Y, Erez T, Reynolds IJ, Kumar D, Ross J, Koytiger G, et al. Drug repurposing from the perspective of pharmaceutical companies. *Br J Pharmacol* 2018 Jan 18;175(2):168-180. [doi: [10.1111/bph.13798](https://doi.org/10.1111/bph.13798)] [Medline: [28369768](https://pubmed.ncbi.nlm.nih.gov/28369768/)]
12. Ng YL, Salim CK, Chu JJH. Drug repurposing for COVID-19: approaches, challenges and promising candidates. *Pharmacol Ther* 2021 Dec;228:107930 [FREE Full text] [doi: [10.1016/j.pharmthera.2021.107930](https://doi.org/10.1016/j.pharmthera.2021.107930)] [Medline: [34174275](https://pubmed.ncbi.nlm.nih.gov/34174275/)]
13. Cavalla D. Using human experience to identify drug repurposing opportunities: theory and practice. *Br J Clin Pharmacol* 2019 Apr 03;85(4):680-689. [doi: [10.1111/bcp.13851](https://doi.org/10.1111/bcp.13851)] [Medline: [30648285](https://pubmed.ncbi.nlm.nih.gov/30648285/)]
14. Tucker J, Day S, Tang W, Bayus B. Crowdsourcing in medical research: concepts and applications. *PeerJ* 2019;7:e6762. [doi: [10.7717/peerj.6762](https://doi.org/10.7717/peerj.6762)] [Medline: [30997295](https://pubmed.ncbi.nlm.nih.gov/30997295/)]
15. Sharma N. Patient centric approach for clinical trials: current trend and new opportunities. *Perspect Clin Res* 2015;6(3):134-138 [FREE Full text] [doi: [10.4103/2229-3485.159936](https://doi.org/10.4103/2229-3485.159936)] [Medline: [26229748](https://pubmed.ncbi.nlm.nih.gov/26229748/)]
16. Boudreau K, Lakhani K. Using the crowd as an innovation partner. *Harv Bus Rev* 2013 May;91(4):60-9, 140 [FREE Full text] [Medline: [23593768](https://pubmed.ncbi.nlm.nih.gov/23593768/)]
17. Charalabidis Y, Loukis E, Androutsopoulou A, Karkaletsis V, Triantafillou A. Passive crowdsourcing in government using social media. *Transform Gov: People, Process Policy* 2014;8(2):283-306. [doi: [10.1108/TG-09-2013-0035](https://doi.org/10.1108/TG-09-2013-0035)]
18. Ahmed S, Rajput AE, Sarirete A, Aljaberi A, Alghanem O, Alsheraigi A. Studying unemployment effects on mental health: social media versus the traditional approach. *Sustainability* 2020 Oct 02;12(19):8130. [doi: [10.3390/su12198130](https://doi.org/10.3390/su12198130)]
19. Koss J, Rheinlaender A, Truebel H, Bohnet-Joschko S. Social media mining in drug development-Fundamentals and use cases. *Drug Discov Today* 2021 Dec;26(12):2871-2880 [FREE Full text] [doi: [10.1016/j.drudis.2021.08.012](https://doi.org/10.1016/j.drudis.2021.08.012)] [Medline: [34481080](https://pubmed.ncbi.nlm.nih.gov/34481080/)]
20. Preiss A, Baumgartner P, Edlund MJ, Bobashev GV. Using named entity recognition to identify substances used in the self-medication of opioid withdrawal: natural language processing study of Reddit data. *JMIR Form Res* 2022 Mar 30;6(3):e33919 [FREE Full text] [doi: [10.2196/33919](https://doi.org/10.2196/33919)] [Medline: [35353047](https://pubmed.ncbi.nlm.nih.gov/35353047/)]
21. Correia R, Li L, Rocha L. Monitoring potential drug interactions and reactions via network analysis of Instagram user timelines. *Pac Symp Biocomput* 2016;21:492-503 [FREE Full text] [Medline: [26776212](https://pubmed.ncbi.nlm.nih.gov/26776212/)]
22. Lewis JA, Gee PM, Ho CL, Miller LMS. Understanding why older adults with type 2 diabetes join diabetes online communities: semantic network analyses. *JMIR Aging* 2018 Jul 28;1(1):e10649 [FREE Full text] [doi: [10.2196/10649](https://doi.org/10.2196/10649)] [Medline: [31518243](https://pubmed.ncbi.nlm.nih.gov/31518243/)]
23. Wang Q, Zhang W, Cai H, Cao Y. Understanding the perceptions of Chinese women of the commercially available domestic and imported HPV vaccine: A semantic network analysis. *Vaccine* 2020 Dec 14;38(52):8334-8342. [doi: [10.1016/j.vaccine.2020.11.016](https://doi.org/10.1016/j.vaccine.2020.11.016)] [Medline: [33190947](https://pubmed.ncbi.nlm.nih.gov/33190947/)]
24. Luo C, Chen A, Cui B, Liao W. Exploring public perceptions of the COVID-19 vaccine online from a cultural perspective: semantic network analysis of two social media platforms in the United States and China. *Telemat Inform* 2021 Dec;65:101712 [FREE Full text] [doi: [10.1016/j.tele.2021.101712](https://doi.org/10.1016/j.tele.2021.101712)] [Medline: [34887618](https://pubmed.ncbi.nlm.nih.gov/34887618/)]
25. Baumgartner J, Zannettou S, Keegan B, Squire M, Blackburn J. The pushshift Reddit dataset. 2020 Presented at: Fourteenth International AAI Conference on Web and Social Media (ICWSM 2020); June 8-11, 2020; Atlanta, GA.

26. Sarker A, Ge Y. Mining long-COVID symptoms from Reddit: characterizing post-COVID syndrome from patient reports. *JAMIA Open* 2021 Jul;4(3):ooab075 [FREE Full text] [doi: [10.1093/jamiaopen/ooab075](https://doi.org/10.1093/jamiaopen/ooab075)] [Medline: [34485849](https://pubmed.ncbi.nlm.nih.gov/34485849/)]
27. Sarabadani S, Baruah G, Fossat Y, Jeon J. Longitudinal changes of COVID-19 symptoms in social media: observational study. *J Med Internet Res* 2022 Feb 16;24(2):e33959 [FREE Full text] [doi: [10.2196/33959](https://doi.org/10.2196/33959)] [Medline: [35076400](https://pubmed.ncbi.nlm.nih.gov/35076400/)]
28. Mansouri A, Affendey L, Mamat A. Named entity recognition approaches. *Int J Comput Sci Netw Secur* 2008;8(2):339-344.
29. Neumann M, King D, Beltagy I, Ammar W. ScispaCy: fast and robust models for biomedical natural language processing. arXiv. 2019. URL: <https://arxiv.org/abs/1902.07669> [accessed 2022-08-30]
30. Li J, Sun Y, Johnson R, Sciaky D, Wei C, Leaman R, et al. BioCreative V CDR task corpus: a resource for chemical disease relation extraction. *Database* 2016;2016:baw068 [FREE Full text] [doi: [10.1093/database/baw068](https://doi.org/10.1093/database/baw068)] [Medline: [27161011](https://pubmed.ncbi.nlm.nih.gov/27161011/)]
31. Kormilitzin A, Vaci N, Liu Q, Nevado-Holgado A. Med7: a transferable clinical natural language processing model for electronic health records. *Artif Intell Med* 2021 Aug;118:102086. [doi: [10.1016/j.artmed.2021.102086](https://doi.org/10.1016/j.artmed.2021.102086)] [Medline: [34412834](https://pubmed.ncbi.nlm.nih.gov/34412834/)]
32. Chapman WW, Bridewell W, Hanbury P, Cooper GF, Buchanan BG. A simple algorithm for identifying negated findings and diseases in discharge summaries. *J Biomed Inform* 2001 Oct;34(5):301-310 [FREE Full text] [doi: [10.1006/jbin.2001.1029](https://doi.org/10.1006/jbin.2001.1029)] [Medline: [12123149](https://pubmed.ncbi.nlm.nih.gov/12123149/)]
33. Bodenreider O. The Unified Medical Language System (UMLS): integrating biomedical terminology. *Nucleic Acids Res* 2004 Jan 01;32(Database issue):D267-D270 [FREE Full text] [doi: [10.1093/nar/gkh061](https://doi.org/10.1093/nar/gkh061)] [Medline: [14681409](https://pubmed.ncbi.nlm.nih.gov/14681409/)]
34. Hripcsak G. Agreement, the F-measure, and reliability in information retrieval. *J Am Med Inform Assoc* 2005 Jan 31;12(3):296-298. [doi: [10.1197/jamia.m1733](https://doi.org/10.1197/jamia.m1733)]
35. Orange book: approved drug products with therapeutic equivalence evaluations. US Food and Drug Administration. URL: <https://www.fda.gov/drugs/drug-approvals-and-databases/approved-drug-products-therapeutic-equivalence-evaluations-orange-book> [accessed 2022-08-30]
36. Ma'ayan A, Jenkins S, Goldfarb J, Iyengar R. Network analysis of FDA approved drugs and their targets. *Mt Sinai J Med* 2007 May;74(1):27-32 [FREE Full text] [doi: [10.1002/msj.20002](https://doi.org/10.1002/msj.20002)] [Medline: [17516560](https://pubmed.ncbi.nlm.nih.gov/17516560/)]
37. Marshall S, Yang C, Ping Q, Zhao M, Avis N, Ip E. Symptom clusters in women with breast cancer: an analysis of data from social media and a research study. *Qual Life Res* 2016 Mar;25(3):547-557 [FREE Full text] [doi: [10.1007/s11136-015-1156-7](https://doi.org/10.1007/s11136-015-1156-7)] [Medline: [26476836](https://pubmed.ncbi.nlm.nih.gov/26476836/)]
38. Vitte J, Gao F, Coppola G, Judkins AR, Giovannini M. Timing of Smarcb1 and Nf2 inactivation determines schwannoma versus rhabdoid tumor development. *Nat Commun* 2017 Aug 21;8(1):300. [doi: [10.1038/s41467-017-00346-5](https://doi.org/10.1038/s41467-017-00346-5)] [Medline: [28824165](https://pubmed.ncbi.nlm.nih.gov/28824165/)]
39. Liu C, Lu X. Analyzing hidden populations online: topic, emotion, and social network of HIV-related users in the largest Chinese online community. *BMC Med Inform Decis Mak* 2018 Jan 05;18(1):2 [FREE Full text] [doi: [10.1186/s12911-017-0579-1](https://doi.org/10.1186/s12911-017-0579-1)] [Medline: [29304788](https://pubmed.ncbi.nlm.nih.gov/29304788/)]
40. Teng C, Lin Y, Adamic L. Recipe recommendation using ingredient networks. 2012 Presented at: 4th annual ACM Web Science Conference (WebSci '12); June 22-24, 2012; Evanston, IL. [doi: [10.1145/2380718.2380757](https://doi.org/10.1145/2380718.2380757)]
41. Bullinaria JA, Levy JP. Extracting semantic representations from word co-occurrence statistics: stop-lists, stemming, and SVD. *Behav Res Methods* 2012 Oct 19;44(3):890-907. [doi: [10.3758/s13428-011-0183-8](https://doi.org/10.3758/s13428-011-0183-8)] [Medline: [22258891](https://pubmed.ncbi.nlm.nih.gov/22258891/)]
42. Calabrese C, Ding J, Millam B, Barnett GA. The uproar over gene-edited babies: a semantic network analysis of CRISPR on Twitter. *Env Commun* 2019 Dec 19;14(7):954-970. [doi: [10.1080/17524032.2019.1699135](https://doi.org/10.1080/17524032.2019.1699135)]
43. Bastian M, Heymann S, Jacomy M. Gephi: an open source software for exploring and manipulating networks. 2009 Presented at: Third International AAAI Conference on Weblogs and Social Media; May 17-20, 2009; San Jose, CA.
44. Traag VA, Waltman L, van Eck NJ. From Louvain to Leiden: guaranteeing well-connected communities. *Sci Rep* 2019 Mar 26;9(1):5233. [doi: [10.1038/s41598-019-41695-z](https://doi.org/10.1038/s41598-019-41695-z)] [Medline: [30914743](https://pubmed.ncbi.nlm.nih.gov/30914743/)]
45. Hairol Anuar SH, Abas ZA, Yunus NM, Mohd Zaki NH, Hashim NA, Mokhtar MF, et al. Comparison between Louvain and Leiden Algorithm for Network Structure: A Review. 2021 Presented at: 1st International Conference on Material Processing and Technology (ICMProTech 2021); July 14-15, 2021; Perlis, Malaysia. [doi: [10.1088/1742-6596/2129/1/012028](https://doi.org/10.1088/1742-6596/2129/1/012028)]
46. Arroyo-Machado W, Torres-Salinas D, Robinson-Garcia N. Identifying and characterizing social media communities: a socio-semantic network approach to altmetrics. *Scientometrics* 2021 Oct 12;126(11):9267-9289 [FREE Full text] [doi: [10.1007/s11192-021-04167-8](https://doi.org/10.1007/s11192-021-04167-8)] [Medline: [34658460](https://pubmed.ncbi.nlm.nih.gov/34658460/)]
47. Dotsika F, Watkins A. Identifying potentially disruptive trends by means of keyword network analysis. *Technol Forecast Soc Change* 2017 Jun;119:114-127. [doi: [10.1016/j.techfore.2017.03.020](https://doi.org/10.1016/j.techfore.2017.03.020)]
48. Hoser B, Hotho A, Jäschke R, Schmitz C, Stumme G. Semantic network analysis of ontologies. In: Sure Y, Domingue J, editors. *The Semantic Web: Research and Applications*. ESWC 2006. Lecture Notes in Computer Science, vol 4011. Berlin, Heidelberg: Springer; 2006.
49. Foufi V, Timakum T, Gaudet-Blavignac C, Lovis C, Song M. Mining of textual health information from Reddit: analysis of chronic diseases with extracted entities and their relations. *J Med Internet Res* 2019 Jun 13;21(6):e12876 [FREE Full text] [doi: [10.2196/12876](https://doi.org/10.2196/12876)] [Medline: [31199327](https://pubmed.ncbi.nlm.nih.gov/31199327/)]
50. Varsha K, Patil K. An overview of community detection algorithms in social networks. 2020 Presented at: International Conference on Inventive Computation Technologies (ICICT), 2020; February 26-28, 2020; Coimbatore, Tamilnadu, India. [doi: [10.1109/iciict48043.2020.9112563](https://doi.org/10.1109/iciict48043.2020.9112563)]

51. Newman MEJ. Modularity and community structure in networks. *Proc Natl Acad Sci U S A* 2006 Jul 06;103(23):8577-8582 [FREE Full text] [doi: [10.1073/pnas.0601602103](https://doi.org/10.1073/pnas.0601602103)] [Medline: [16723398](https://pubmed.ncbi.nlm.nih.gov/16723398/)]
52. Crook H, Raza S, Nowell J, Young M, Edison P. Long covid-mechanisms, risk factors, and management. *BMJ* 2021 Jul 26;374:n1648. [doi: [10.1136/bmj.n1648](https://doi.org/10.1136/bmj.n1648)] [Medline: [34312178](https://pubmed.ncbi.nlm.nih.gov/34312178/)]
53. Pinto MD, Lambert N, Downs CA, Abraham H, Hughes TD, Rahmani AM, et al. Antihistamines for postacute sequelae of SARS-CoV-2 infection. *J Nurse Pract* 2022 Mar;18(3):335-338 [FREE Full text] [doi: [10.1016/j.nurpra.2021.12.016](https://doi.org/10.1016/j.nurpra.2021.12.016)] [Medline: [35153633](https://pubmed.ncbi.nlm.nih.gov/35153633/)]
54. Glynne P, Tahmasebi N, Gant V, Gupta R. Long COVID following mild SARS-CoV-2 infection: characteristic T cell alterations and response to antihistamines. *J Investig Med* 2022 Jan 05;70(1):61-67 [FREE Full text] [doi: [10.1136/jim-2021-002051](https://doi.org/10.1136/jim-2021-002051)] [Medline: [34611034](https://pubmed.ncbi.nlm.nih.gov/34611034/)]
55. Hoertel N, Sánchez-Rico M, Vernet R, Beeker N, Jannot A, Neuraz A, AP-HP/Universities/INSERM COVID-19 Research Collaboration AP-HP COVID CDR Initiative. Association between antidepressant use and reduced risk of intubation or death in hospitalized patients with COVID-19: results from an observational study. *Mol Psychiatry* 2021 Sep 04;26(9):5199-5212. [doi: [10.1038/s41380-021-01021-4](https://doi.org/10.1038/s41380-021-01021-4)] [Medline: [33536545](https://pubmed.ncbi.nlm.nih.gov/33536545/)]
56. Köhler CA, Freitas TH, Stubbs B, Maes M, Solmi M, Veronese N, et al. Peripheral alterations in cytokine and chemokine levels after antidepressant drug treatment for major depressive disorder: systematic review and meta-analysis. *Mol Neurobiol* 2018 May 13;55(5):4195-4206. [doi: [10.1007/s12035-017-0632-1](https://doi.org/10.1007/s12035-017-0632-1)] [Medline: [28612257](https://pubmed.ncbi.nlm.nih.gov/28612257/)]
57. Sukhatme VP, Reiersen AM, Vayttaden SJ, Sukhatme VV. Fluvoxamine: a review of its mechanism of action and its role in COVID-19. *Front Pharmacol* 2021 Apr 20;12:652688. [doi: [10.3389/fphar.2021.652688](https://doi.org/10.3389/fphar.2021.652688)] [Medline: [33959018](https://pubmed.ncbi.nlm.nih.gov/33959018/)]
58. Khani E, Entezari-Maleki T. Fluvoxamine and long COVID-19; a new role for sigma-1 receptor (S1R) agonists. *Mol Psychiatry* 2022 May 06:online ahead of print [FREE Full text] [doi: [10.1038/s41380-022-01545-3](https://doi.org/10.1038/s41380-022-01545-3)] [Medline: [35388182](https://pubmed.ncbi.nlm.nih.gov/35388182/)]
59. Krystal AD, Thase ME, Tucker VL, Goodale EP. Bupropion HCL and sleep in patients with depression. *Curr Psych Rev* 2007 May 01;3(2):123-128. [doi: [10.2174/157340007780599096](https://doi.org/10.2174/157340007780599096)]
60. Chen T, Song J, Liu H, Zheng H, Chen C. Positive Epstein-Barr virus detection in coronavirus disease 2019 (COVID-19) patients. *Sci Rep* 2021 May 25;11(1):10902. [doi: [10.1038/s41598-021-90351-y](https://doi.org/10.1038/s41598-021-90351-y)] [Medline: [34035353](https://pubmed.ncbi.nlm.nih.gov/34035353/)]
61. Goel N, Goyal N, Nagaraja R, Kumar R. Systemic corticosteroids for management of 'long-COVID': an evaluation after 3 months of treatment. *Monaldi Arch Chest Dis* 2021 Dec 03;92(2):1981. [doi: [10.4081/monaldi.2021.1981](https://doi.org/10.4081/monaldi.2021.1981)] [Medline: [34730322](https://pubmed.ncbi.nlm.nih.gov/34730322/)]
62. O'Kelly B, Vidal L, McHugh T, Woo J, Avramovic G, Lambert JS. Safety and efficacy of low dose naltrexone in a long covid cohort; an interventional pre-post study. *Brain Behav Immun Health* 2022 Oct;24:100485 [FREE Full text] [doi: [10.1016/j.bbih.2022.100485](https://doi.org/10.1016/j.bbih.2022.100485)] [Medline: [35814187](https://pubmed.ncbi.nlm.nih.gov/35814187/)]
63. Benetoli A, Chen T, Aslani P. How patients' use of social media impacts their interactions with healthcare professionals. *Patient Educ Couns* 2018 Mar;101(3):439-444. [doi: [10.1016/j.pec.2017.08.015](https://doi.org/10.1016/j.pec.2017.08.015)] [Medline: [28882545](https://pubmed.ncbi.nlm.nih.gov/28882545/)]
64. Ogink T, Dong JQ. Stimulating innovation by user feedback on social media: The case of an online user innovation community. *Technol Forecast Soc Change* 2019 Jul;144:295-302. [doi: [10.1016/j.techfore.2017.07.029](https://doi.org/10.1016/j.techfore.2017.07.029)]
65. Song M, Kim WC, Lee D, Heo GE, Kang KY. PKDE4J: entity and relation extraction for public knowledge discovery. *J Biomed Inform* 2015 Oct;57:320-332 [FREE Full text] [doi: [10.1016/j.jbi.2015.08.008](https://doi.org/10.1016/j.jbi.2015.08.008)] [Medline: [26277115](https://pubmed.ncbi.nlm.nih.gov/26277115/)]
66. Scepanovic S, Martin-Lopez E, Quercia D, Baykaner K. Extracting medical entities from social media. 2020 Presented at: CHIL '20: ACM Conference on Health, Inference, and Learning; April 2-4, 2020; Toronto, ON. [doi: [10.1145/3368555.3384467](https://doi.org/10.1145/3368555.3384467)]
67. Gräßer F, Kallumadi S, Malberg H, Zaunseder S. Aspect-based sentiment analysis of drug reviews applying cross-domain and cross-data learning. 2018 Presented at: DH '18: 2018 International Conference on Digital Health; April 23-26, 2018; Lyon, France. [doi: [10.1145/3194658.3194677](https://doi.org/10.1145/3194658.3194677)]
68. Doan S, Yang EW, Tilak SS, Li PW, Zisook DS, Torii M. Extracting health-related causality from twitter messages using natural language processing. *BMC Med Inform Decis Mak* 2019 Apr 04;19(Suppl 3):79 [FREE Full text] [doi: [10.1186/s12911-019-0785-0](https://doi.org/10.1186/s12911-019-0785-0)] [Medline: [30943954](https://pubmed.ncbi.nlm.nih.gov/30943954/)]
69. Saha K, Sugar B, Torous J, Abrahao B, Kıcıman E, De Choudhury M. A social media study on the effects of psychiatric medication use. *Proc Int AAAI Conf Weblogs Soc Media* 2019 Jul 07;13:440-451 [FREE Full text] [Medline: [32280562](https://pubmed.ncbi.nlm.nih.gov/32280562/)]
70. Meyerowitz-Katz G, Wieten S, Medina Arellano MDJ, Yamey G. Unethical studies of ivermectin for covid-19. *BMJ* 2022 Apr 14;377:o917. [doi: [10.1136/bmj.o917](https://doi.org/10.1136/bmj.o917)] [Medline: [35422425](https://pubmed.ncbi.nlm.nih.gov/35422425/)]
71. Roman Y, Burela P, Pasupuleti V, Piscocoya A, Vidal J, Hernandez A. Ivermectin for the treatment of coronavirus disease 2019: a systematic review and meta-analysis of randomized controlled trials. *Clin Infect Dis* 2022 Mar 23;74(6):1022-1029 [FREE Full text] [doi: [10.1093/cid/ciab591](https://doi.org/10.1093/cid/ciab591)] [Medline: [34181716](https://pubmed.ncbi.nlm.nih.gov/34181716/)]
72. Amaya A, Bach R, Keusch F, Kreuter F. New data sources in social science research: things to know before working with Reddit data. *Soc Sci Comput Rev* 2019 Dec 18;39(5):943-960. [doi: [10.1177/0894439319893305](https://doi.org/10.1177/0894439319893305)]

## Abbreviations

**BiLSTM-CRF:** bidirectional long short-term memory–conditional random field

**DR:** drug repurposing  
**EBV:** Epstein-Barr virus  
**LC:** long COVID  
**NDRI:** norepinephrine/dopamine reuptake inhibitor  
**NER:** named-entity recognition  
**OTC:** over the counter  
**PMI:** pointwise mutual information  
**SMM:** social media mining  
**UMLS:** unified medical language system

*Edited by A Mavragani; submitted 16.05.22; peer-reviewed by A Sarker, X Wang; comments to author 06.06.22; revised version received 27.06.22; accepted 09.08.22; published 03.10.22.*

*Please cite as:*

*Koss J, Bohnet-Joschko S*

*Social Media Mining of Long-COVID Self-Medication Reported by Reddit Users: Feasibility Study to Support Drug Repurposing*

*JMIR Form Res 2022;6(10):e39582*

*URL: <https://formative.jmir.org/2022/10/e39582>*

*doi: [10.2196/39582](https://doi.org/10.2196/39582)*

*PMID: [36007131](https://pubmed.ncbi.nlm.nih.gov/36007131/)*

©Jonathan Koss, Sabine Bohnet-Joschko. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 03.10.2022. 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

# Collective Value Promotes the Willingness to Share Provaccination Messages on Social Media in China: Randomized Controlled Trial

Chunye Fu<sup>1</sup>, PhD; Xiaokang Lyu<sup>1</sup>, PhD; Mingdi Mi<sup>2</sup>, MD

<sup>1</sup>Department of Social Psychology, Zhou Enlai School of Government, Nankai University, Tianjin, China

<sup>2</sup>Students' Affairs Division, Weinan Vocational & Technical College, Weinan, China

**Corresponding Author:**

Xiaokang Lyu, PhD

Department of Social Psychology

Zhou Enlai School of Government

Nankai University

38 Tongyan Road

Haihe Education Park

Tianjin, 300350

China

Phone: 86 13920214514

Email: [luxk@nankai.edu.cn](mailto:luxk@nankai.edu.cn)

## Abstract

**Background:** The proliferation of vaccine misinformation on social media has seriously corrupted the public's confidence in vaccination. Proactively sharing provaccination messages on social media is a cost-effective way to enhance global vaccination rates and resist vaccine misinformation. However, few strategies for encouraging the public to proactively share vaccine-related knowledge on social media have been developed.

**Objective:** This research examines the effect of value type (individual vs collective) and message framing (gain vs loss) on influenza vaccination intention (experiment 1) and the willingness to share provaccination messages (experiment 2) among Chinese adults during the COVID-19 pandemic. The primary aim was to evaluate whether messages that emphasized collective value were more effective in increasing the willingness to share than messages that emphasized individual value.

**Methods:** We enrolled 450 Chinese adults for experiment 1 (n=250, 55.6%) and experiment 2 (n=200, 44.4%). Participants were randomly assigned to individual-gain, individual-loss, collective-gain, or collective-loss conditions with regard to the message in each experiment using the online survey platform's randomization function. Experiment 1 also included a control group. The primary outcome was influenza vaccination intention in experiment 1 and the willingness to share provaccination messages in experiment 2.

**Results:** The valid sample included 213 adults in experiment 1 (females: n=151, 70.9%; mean age 29 [SD 9] years; at least some college education: n=202, 94.8%; single: n=131, 61.5%) and 171 adults in experiment 2 (females: n=106, 62.0%; mean age 28 [SD 7] years; at least some college education: n=163, 95.3%; single: n=95, 55.6%). Influenza vaccination intention was stronger in the individual-value conditions than in the collective-value conditions ( $F_{3,166}=4.96$ ,  $P=.03$ ,  $\eta^2=0.03$ ). The reverse result was found for the willingness to share provaccination messages ( $F_{3,165}=6.87$ ,  $P=.01$ ,  $\eta^2=0.04$ ). Specifically, participants who received a message emphasizing collective value had a higher intention to share the message than participants who read a message emphasizing individual value ( $F_{3,165}=6.87$ ,  $P=.01$ ,  $\eta^2=0.04$ ), and the perceived responsibility for message sharing played a mediating role (indirect effect=0.23, 95% lower limit confidence interval [LLCI] 0.41, 95% upper limit confidence interval [ULCI] 0.07). In addition, gain framing facilitated influenza vaccination intention more than loss framing ( $F_{3,166}=5.96$ ,  $P=.02$ ,  $\eta^2=0.04$ ). However, experiment 2 did not find that message framing affected message-sharing willingness. Neither experiment found an interaction between value type and message framing.

**Conclusions:** Strengthened individual value rather than collective value is more likely to persuade Chinese adults to vaccinate. However, these adults are more likely to share a message that emphasizes collective rather than individual value, and the perceived responsibility for message sharing plays a mediating role.

**KEYWORDS**

individual value; collective value; vaccination; message-sharing willingness; perceived responsibility; misinformation; vaccine misinformation; public health; influenza vaccine; social media; COVID-19

## Introduction

### Background

The proliferation of vaccine misinformation on social media has seriously corrupted the public's confidence in vaccination [1-3]. Extreme antivaxers create vaccine misinformation online [4-6], and ordinary users often spread vaccine misinformation voluntarily on social media [7,8] because of the emotional and attractive nature of misinformation. Fortunately, social media platforms are increasingly implementing positive interventions to resist vaccination misinformation [9,10].

These endeavors mainly focus on 2 different approaches. The first approach relies on information technology, such as fact-checking labels [9] or debiasing strategies [11], to reduce the impact of existing vaccine misinformation on its audience. The other approach provides more vaccination knowledge on social media to suppress the diffusion of vaccine misinformation [4,8,12,13].

However, additional strategies to mobilize the public to debunk vaccine misinformation on social media are needed to successfully achieve public health goals. Just as vaccine misinformation can drive sharing intentions, provaccination information, if presented properly, can motivate the public's willingness to share on social media. Inspiring the public's willingness to share provaccination information can improve the current situation in which antivaccination messages occupy a more controversial space than provaccination messages on social media [8,14], and prevent the adverse effects of misinformation [15]. Hence, the main goal of this study is to explore strategies to present effective messages to improve the public's willingness to share provaccination messages on social media.

Previous studies have confirmed that emphasizing the collective benefits of vaccination, such as the long-term benefits of herd immunity for society, could positively affect participants' vaccination intention [8,16,17]. In particular, according to the "sharing for the community" model [18], community interest serves as a more influential motivation for the online sharing of health-related knowledge than personal gain does [19], and information-sharing behaviors driven by community interest are based on reciprocity norms [20]. Therefore, it is logical to hypothesize that provaccination messages that emphasize collective value can promote the willingness to share provaccination messages better than those that emphasize individual value.

In addition, message framing plays a key role in health communication [21]. Based on gain versus loss message framing [22,23], the value of vaccination can be framed by highlighting either desirable consequences (eg, health benefits) or undesirable consequences (eg, disease risk). Although a meta-analytic review

showed that gain-framing messages are significantly more likely to encourage illness prevention behaviors compared to loss-framing messages [22], another study showed that framing effects are significant only when individuals perceive the issue to have high personal relevance [24]. Therefore, we assume that the framing effects may be less obvious when messages are described with collective value, since personal relevance is diminished when messages are described with collective value compared to individual value. Thus, we used message framing as an independent variable to explore whether message framing and value type synergistically affect vaccination message-sharing intention.

### Goal of This Study

We conducted 2 survey experiments during the COVID-19 pandemic outbreak in mainland China. Experiment 1 explored whether a message that emphasized collective value improved vaccination intention more than one that emphasized individual value during the COVID-19 pandemic outbreak. Experiment 2 tested our primary hypothesis of whether collective value significantly facilitates the dissemination of provaccination messages online. We also measured the perceived importance and responsibility of message sharing as mediating variables in experiment 2 to examine the mechanism of collective value on the willingness to share messages.

## Methods

### Design

Two 2 (value type: collective vs individual) × 2 (message framing: gain vs loss) between-participant experiments were conducted using Sojump [25], an online study portal. Experiment 1 explored the main effect of value type and message framing on vaccination intention for the influenza vaccine (including a no-message control group), while experiment 2 focused on message-sharing willingness.

### Participants

The sample size was estimated using G\*Power 3.1 [26], assuming a statistical power of 90%, a significance level of .05, and an effect size of 0.25. The estimated minimum sample size for 2-factor ANOVA was 172 (n=43, 25%, for each condition). Accordingly, we recruited 450 volunteers in total (experiment 1: n=250, 55.6%; experiment 2: n=200, 44.4%) for monetary compensation.

All participants were native Chinese readers; none had a medical education background or were from Wuhan, China. Manipulation checks in the survey resulted in a total of 384 valid participants in the experiments (experiment 1: n=213, 85.2%; experiment 2: n=171, 85.5%). Table 1 presents the descriptive information about the participants.

**Table 1.** Descriptive information about participants in experiments 1 and 2.

Variables	Experiment 1 (N=213)	Experiment 2 (N=171)
<b>Age (years)</b>		
Range	18-58	18-56
Mean (SD)	29 (9)	28 (7)
<b>Gender, n (%)</b>		
Female	151 (70.9)	106 (62.0)
Male	62 (20.1)	65 (38.0)
<b>Marital status, n (%)</b>		
Single	131 (61.5)	95 (55.6)
Married	82 (38.5)	76 (44.4)
<b>Education, n (%)</b>		
Middle school graduate and below	2 (0.9)	2 (1.2)
High school graduate	9 (4.2)	6 (3.5)
College degree holder	129 (60.6)	124 (72.5)
Graduate degree holder and above	73 (34.3)	39 (22.8)

## Message Stimuli and Pretest

### *Message Stimuli*

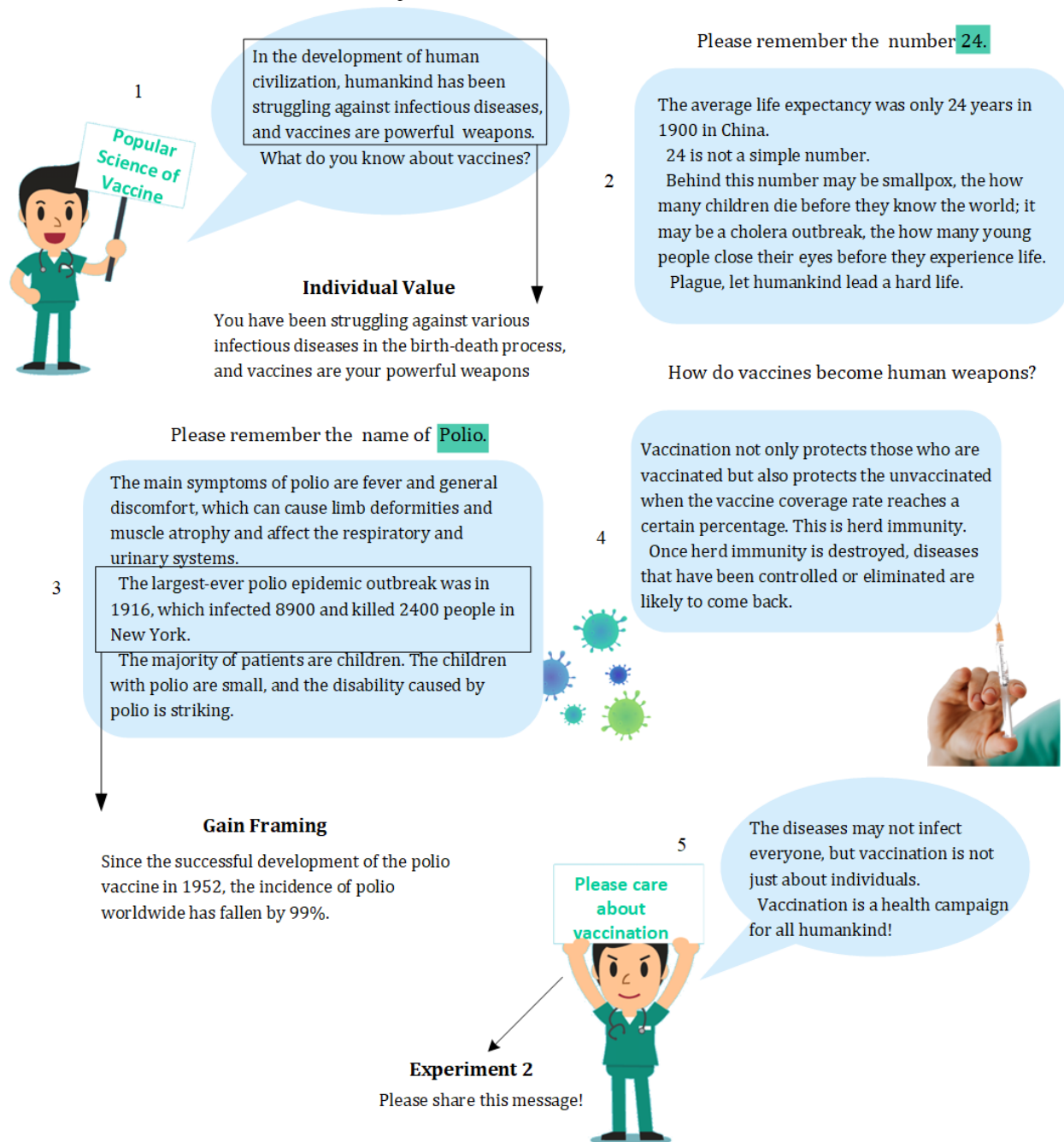
The content of the messages was developed based on considerable scientific research published in vaccination-related fields [8,27,28]. We wrote 5 consecutive paragraphs, with an associated theme in each message. In approximately 450 Chinese characters, each of the messages described a story about immunization and stated the substantial value of vaccination. To increase the vividness of the messages, we used pictures from a royalty-free image web archive.

Figure 1 presents the condition of collective loss in experiment 1. Value type was manipulated by emphasizing collective or

individual value under different conditions. For example, herd immunity was explained in the condition of collective value, while individual immunity was used in the condition of individual value. In addition, the subjects of the sentences were changed from “human” (collective value) to “you” (individual value).

In line with prior framing-related literature [29], message framing was described separately by strengthening the benefit of vaccination in gain framing or the harm of not being vaccinated in loss framing. We also used specific examples of vaccination to convey the messages (eg, the polio vaccine has saved human lives, or polio has killed humans).

**Figure 1.** Material for the collective-loss condition for experiment 1.



### Pretest of Messages

We tested the validity of the messages with 120 participants (the manipulation check resulted in 111 valid participants) who did not participate in the formal experiments. The 111 participants (females:  $n=83$ , 74.8%; mean age 33 [SD 12] years; at least some college education:  $n=93$ , 83.7%; single:  $n=58$ , 52.3%) were native Chinese readers; none had a medical education background or were from Wuhan, China. Each participant was randomly assigned to 1 of the 4 materials using the online survey platform's randomization function.

First, we used a single 7-point item to measure participants' credibility ratings of the messages. The responses showed no significant differences between the credibility ratings for collective and individual value ( $F_{3,107}=1.45$ ,  $P=.23$ ) or loss and

gain framing ( $F_{3,107}=0.003$ ,  $P=.95$ ). Second, we asked participants to choose the message they read with regard to collective or individual value describing the benefits of vaccination or the risk of not vaccinating. The results revealed that more than 90% of the participants could accurately distinguish collective or individual value and gain or loss framing. Thus, the materials were considered valid for use in subsequent experiments.

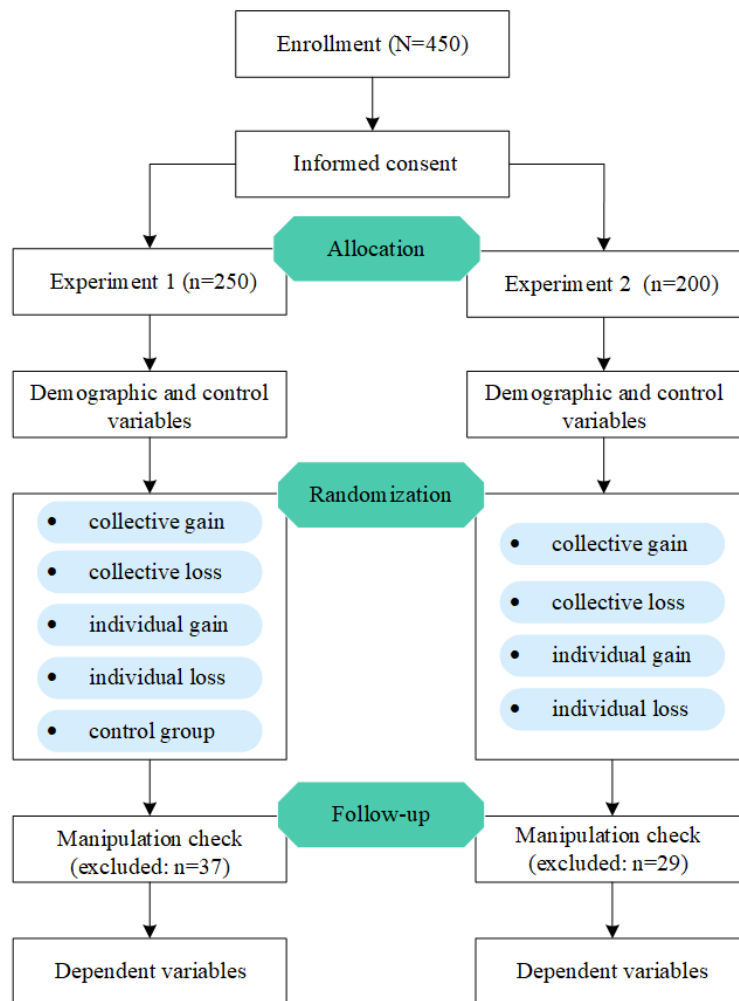
### Procedure

The system flowchart can be seen in Figure 2. After obtaining informed consent, the survey gathered information about the participants' demographics (gender, age, education, and marital status) and control variables. The participants were randomly assigned to 1 of the conditions using the online survey platform's randomization function. Those assigned to 1 of the

4 message conditions proceeded to a page that showed them a message and required them to answer 2 manipulation check questions. They were then instructed to respond to questions

related to the variables described in the following section. Those in the control condition in experiment 1 proceeded directly to a similar survey without reading a message.

**Figure 2.** Schematic representation of the experimental flow.



## Measures

All control and dependent variables used in our experiments are presented in [Table 2](#). The details of all measures used in experiments are provided in [Multimedia Appendix 1](#).

**Table 2.** Control and dependent variables used in experiments 1 and 2.

Variables	Experiment(s)	Measures	Items, n	Sample item	Score range	Cronbach $\alpha$
Control	1 and 2	Vaccine confidence	3	Overall, I think vaccines are safe.	1=extremely disagree to 5=extremely agree	.85 (experiment 1) and .72 (experiment 2)
Control	2	Propensity to share on social media	3	I share new messages on social media.	1=rarely to 5=always	.72
Dependent	1	Vaccine acceptance	10	Vaccines are effective at preventing diseases.	1=strongly disagree to 7=strongly agree	.76
Dependent	1	Influenza vaccination intention	3	How likely is it that you are going to receive the flu vaccination this winter?	1=extremely unlikely to 5=extremely likely	.82
Dependent	2	Perceived importance of message	3	The message contains crucial knowledge about the vaccine.	1=extremely disagree to 5=extremely agree	.74
Dependent	2	Perceived responsibility for message sharing	6	At this moment, I believe I have the responsibility for sharing this message to improve the influenza vaccination rate.	1=strongly disagree to 7=strongly agree	.89
Dependent	2	Message-sharing willingness	3	I will share this message using different social networking tools.	1=extremely disagree to 5=extremely agree	.72

### Control Variables

Vaccine confidence, assessed by modifying the 4-item vaccine confidence scale to 3 items by removing the item about religion [30], was measured as a control variable in experiments 1 and 2. The propensity to share on social media [31] served as a control variable in experiment 2.

### Dependent Variables

We focused on vaccination intention for the 2020-2021 influenza vaccine. A self-administered scale was used. Specifically, participants were provided with medical news. The news read, "Once COVID-19 infection occurs together with influenza, the difficulty of diagnosis will be increased..." In addition, the intention to receive the influenza vaccination was measured by how likely the participants were to accept the influenza vaccination for themselves or recommend it to their families in the winter. The secondary dependent variable in experiment 1

was vaccine acceptance, which captured the entire conceptual domain of vaccine acceptance [32].

Experiment 2 comprised 3 dependent variables: the perceived importance of messages [33], the perceived responsibility for message sharing [34], and message-sharing willingness [33].

### Ethical Considerations

The study was conducted according to the guidelines of the Declaration of Helsinki and approved by the Institutional Review Board of Nankai University (#NKUIRB2020023, approval date February 25, 2020).

## Results

### Experiment 1

The descriptive and inferential statistics of experiment 1 are shown in Table 3. Demographic variables were also included in the analysis as control variables.

**Table 3.** Means (SDs) for each condition and inferential statistics of each condition compared to the control group of experiment 1.

Value and framing	Sample size (N=213), n (%)	Vaccine acceptance		Vaccination intention	
		Mean (SD)	P value	Mean (SD)	P value
<b>Collective</b>					
Gain	43 (20.2)	5.46 (0.79)	.02	3.91(0.87)	.01
Loss	41 (19.2)	5.43 (0.65)	.06	3.57(0.89)	.48
<b>Individual</b>					
Gain	44 (20.7)	5.53 (0.71)	.02	4.11(0.65)	<.001
Loss	42 (19.7)	5.60 (0.65)	.01	3.94(0.63)	<.001
Control group	43 (20.2)	5.06 (0.74)	N/A <sup>a</sup>	3.35(1.12)	N/A

<sup>a</sup>N/A: not applicable.

### Vaccine Acceptance

The main effects of value type ( $F_{3,166}=0.16$ ,  $P=.69$ ), message framing ( $F_{3,166}=0.001$ ,  $P=.97$ ) and 2-way interactions between value type and message framing ( $F_{3,166}=0.49$ ,  $P=.49$ ) were not statistically significant for vaccine acceptance. Further pairwise comparisons showed that participants who had received a message on any condition expressed significantly higher vaccine acceptance than those in the control group. Except for the collective-loss condition, the difference was only borderline significant.

### Influenza Vaccination Intention

The results of the influenza vaccination intention index revealed a significant main effect of value type ( $F_{3,166}=4.96$ ,  $P=.03$ ,

$\eta^2=0.03$ ). Vaccination intention was stronger in the individual-value conditions than in the collective-value conditions. Messaging framing also exhibited a significant main effect ( $F_{3,166}=5.96$ ,  $P=.02$ ,  $\eta^2=0.04$ ). Vaccination intention was stronger in the gain-framing condition than in the loss-framing condition. No interaction effects were significant ( $F_{3,166}=0.83$ ,  $P=.36$ ). We also found that the participants in all conditions except for the collective-loss condition had significantly higher vaccination intention than those in the control group.

### Experiment 2

The descriptive statistics of experiment 2 are shown in [Table 4](#).

**Table 4.** Means (SDs) for each condition of experiment 2.

Value and framing	Sample size (N=171), n (%)	Perceived importance, mean (SD)	Perceived responsibility, mean (SD)	Message-sharing willingness, mean (SD)
<b>Collective</b>				
Gain	43 (25.1)	4.54 (0.43)	6.19 (0.62)	3.96 (0.62)
Loss	42 (24.6)	4.61 (0.42)	6.23 (0.67)	4.04 (0.68)
<b>Individual</b>				
Gain	44 (25.7)	4.44 (0.51)	5.87 (0.77)	3.75 (0.81)
Loss	42 (24.6)	4.21 (0.56)	5.62 (1.07)	3.44 (1.01)

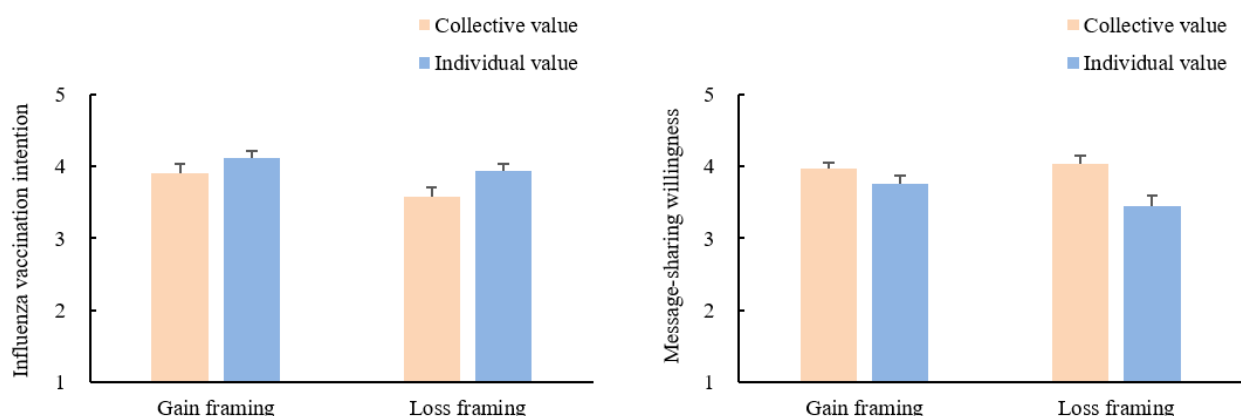
### Perceived Importance, Perceived Responsibility, and Message-Sharing Willingness

Significant differences were found between collective and individual value. Participants who received a message emphasizing collective value had higher perceived importance ( $F_{3,165}=4.72$ ,  $P=.03$ ,  $\eta^2=0.03$ ), higher perceived responsibility ( $F_{3,165}=9.31$ ,  $P=.003$ ,  $\eta^2=0.06$ ), and a higher willingness to

share the message ( $F_{3,165}=6.87$ ,  $P=.01$ ,  $\eta^2=0.04$ ) than participants who read a message emphasizing individual value. There was no significant main effect of message framing or 2-way interactions between value type and message framing on each dependent variable.

Influenza vaccination intention and message-sharing willingness were the 2 primary outcomes in experiments 1 and 2, respectively. Therefore, we present the 2 dependent variables together in [Figure 3](#) to facilitate a comparison of trends.

**Figure 3.** Effect of value type on influenza vaccination intention (experiment 1) and message-sharing willingness (experiment 2).



### Mediating Effect Analysis

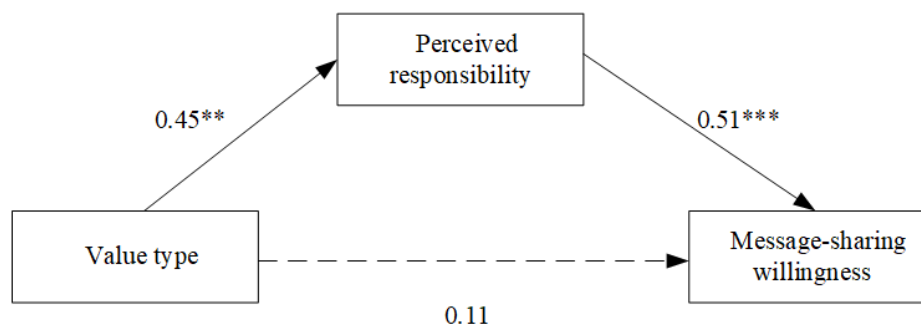
A mediating effect analysis was conducted to determine whether the effect of collective value was mediated by perceived importance and responsibility. The mediational model was

assessed using the SPSS PROCESS macro (model 4) with a 95% bias-corrected CI based on 5000 bootstrap samples [35]. The value type (collective value=1, individual value=0) was used as the independent variable, perceived importance and

responsibility were parallel mediators, and message-sharing intention was the outcome variable.

The analysis yielded a significant model ( $F_{8,162}=6.88$ ,  $P<.001$ ,  $R^2=0.25$ ). The direct effect of value type on the intention to share was not significant after introducing the indirect effects of perceived importance and responsibility into the model

**Figure 4.** Statistical model of the meditational model. \*\* $P<.01$ ; \*\*\* $P<.001$ .



## Discussion

### Principal Findings

In experiment 1, we found that individual value had a more substantial effect than collective value on influenza vaccination intention. Interestingly, the results from experiment 2 indicated that the message describing collective value promoted message-sharing willingness better than the message related to individual value; the perceived responsibility for message sharing mediated the effect of collective value on message-sharing willingness but not on perceived importance.

In general, individual-focused messages increased vaccination intention, but collective-focused messages increased the likelihood of sharing. These results seem to produce conundrums for message designers and reveal that designers should finely construct specific messages for different targets. For example, if the goal is to make provaccination messages occupy more space on social media, then collective value should be prioritized to describe messages. In another circumstance, if the improvement of vaccination rates is more important and urgent, especially under circumstances with high risks, individual value should be considered first. Moreover, the message of collective value described as gain framing also enhanced vaccination intention compared to the control group. Therefore, collective-gain messages effectively increase vaccination intention, while enhancing message-sharing willingness.

This study also obtained some other results. First, there were no differences between collective and individual value or gain and loss framing on vaccine acceptance; however, vaccine acceptance improved in all intervention groups compared to the control group. Second, gain framing was more effective than loss framing in influenza vaccination intention. Finally, we did not find an interaction effect between value type and message framing on any dependent variables.

These results contribute to the literature on online health knowledge sharing and provide implications for practice.

(Figure 4). One indirect effect of perceived importance was also not significant (indirect effect=0.03, 95% lower limit confidence interval [LLCI] 0.11, 95% upper limit confidence interval [ULCI] -0.01). In contrast, the other indirect effect of perceived responsibility was significant (indirect effect=0.23, 95% LLCI 0.41, 95% ULCI 0.07).

### Limitations

There are some limitations of this study. First, we did not design a condition that combined collective and individual value and that tested whether such messages can promote both vaccination intention and message-sharing willingness. This limitation can be addressed through follow-up research.

Second, the external validity of this study is limited. In a realistic network environment, other types of information with high attractiveness or interesting themes, such as stories about celebrities or sports news, may cause the public to ignore provaccination messages [36]. Therefore, it is necessary to simulate the real network context to verify the robustness of the results of this study.

Third, we conducted the research in the early stages of COVID-19. During this period, the epidemic in Wuhan, China, was more serious, while people in other regions were mostly in a state of unknown risk. As the form of the epidemic continues to change in China, the public's perception of the risk will also change dynamically. This may further lead to changes in the way messages function in different severe situations. Subsequent studies could consider collecting data at different time points and drawing on studies that use time series models [37,38] in an attempt to optimize health strategies to resist vaccine misinformation and boost vaccination.

Finally, the participants were mainly female, younger, and single and had a college degree or above. Although we controlled for these demographic variables in the analysis, the external validity of the experimental results is limited. This study can be further improved to examine whether the function of the messages varies among populations.

Despite the limitations, the study complements and extends previous studies that focused on collective benefits in personal vaccination decisions.



## Comparison With Prior Work

The effect of value type on vaccination intention contradicts a previous study that found that collective benefit promotes vaccination intention more than individual benefit does [16]. Notably, a study supports our finding, showing that individuals self-categorized as being in a high-risk group are more likely to adopt vaccination behavior following self-benefit messages than social benefit messages [39]. Our study was conducted during the COVID-19 outbreak, and high risk perception might have resulted in the public's heightened sensitivity to individual value messages.

In contrast, message-sharing willingness was promoted by collective value more than individual value, and perceived responsibility played a mediatory role. These findings are consistent with a similar study that found that social integrative

benefits significantly and positively influence knowledge-sharing intention in virtual communities [40].

Finally, the results of message framing coincide with a previous study that suggested that gain-framing messages are more likely to encourage illness prevention behaviors compared to loss-framing messages [22].

## Conclusion

Online messages emphasizing individual value promote the intention to vaccinate more strongly than those emphasizing collective value, but this effect is reversed in message-sharing willingness. Furthermore, perceived responsibility for message sharing plays a mediating role between collective value and message-sharing willingness. Our findings have practical implications for constructing and providing effective, targeted provaccination messages.

---

## Acknowledgments

This research was supported by grants from the National Social Science Fund of China (grant number 20ASH015) and the science research project of the Liberal Arts Development Fund of Nankai University (grant number ZB22BZ0109).

---

## Conflicts of Interest

None declared.

---

## Editorial Notice

This randomized study was not registered. The editor granted an exception from ICMJE rules mandating prospective registration of randomized trials because the risk of bias appears low and the study was considered formative. However, readers are advised to carefully assess the validity of any potential explicit or implicit claims related to primary outcomes or effectiveness.

---

### Multimedia Appendix 1

Details of all measures used in experiments.

[[DOCX File, 18 KB - formative\\_v6i10e35744\\_app1.docx](#)]

---

### Multimedia Appendix 2

CONSORT checklist.

[[PDF File \(Adobe PDF File\), 65 KB - formative\\_v6i10e35744\\_app2.pdf](#)]

---

## References

1. Smith N, Graham T. Mapping the anti-vaccination movement on Facebook. *Inf Commun Soc* 2017 Dec 27;22(9):1310-1327. [doi: [10.1080/1369118x.2017.1418406](https://doi.org/10.1080/1369118x.2017.1418406)]
2. Larson HJ. Blocking information on COVID-19 can fuel the spread of misinformation. *Nature* 2020 Apr;580(7803):306. [doi: [10.1038/d41586-020-00920-w](https://doi.org/10.1038/d41586-020-00920-w)] [Medline: [32231320](https://pubmed.ncbi.nlm.nih.gov/32231320/)]
3. Catalan-Matamoros D, Peñafiel-Saiz C. Exploring the relationship between newspaper coverage of vaccines and childhood vaccination rates in Spain. *Hum Vaccin Immunother* 2020 May 03;16(5):1055-1061 [FREE Full text] [doi: [10.1080/21645515.2019.1708163](https://doi.org/10.1080/21645515.2019.1708163)] [Medline: [32017659](https://pubmed.ncbi.nlm.nih.gov/32017659/)]
4. Hoffman BL, Felter EM, Chu K, Shensa A, Hermann C, Wolynn T, et al. It's not all about autism: the emerging landscape of anti-vaccination sentiment on Facebook. *Vaccine* 2019 Apr 10;37(16):2216-2223. [doi: [10.1016/j.vaccine.2019.03.003](https://doi.org/10.1016/j.vaccine.2019.03.003)] [Medline: [30905530](https://pubmed.ncbi.nlm.nih.gov/30905530/)]
5. Harvey AM, Thompson S, Lac A, Coolidge FL. Fear and derision: a quantitative content analysis of provaccine and antivaccine internet memes. *Health Educ Behav* 2019 Dec;46(6):1012-1023. [doi: [10.1177/1090198119866886](https://doi.org/10.1177/1090198119866886)] [Medline: [31789076](https://pubmed.ncbi.nlm.nih.gov/31789076/)]
6. Broniatowski DA, Reyna VF. To illuminate and motivate: a fuzzy-trace model of the spread of information online. *Comput Math Organ Theory* 2020 Dec;26:431-464 [FREE Full text] [doi: [10.1007/s10588-019-09297-2](https://doi.org/10.1007/s10588-019-09297-2)] [Medline: [33737859](https://pubmed.ncbi.nlm.nih.gov/33737859/)]
7. Ngai CSB, Singh RG, Yao L. Impact of COVID-19 vaccine misinformation on social media virality: content analysis of message themes and writing strategies. *J Med Internet Res* 2022 Jul 06;24(7):e37806 [FREE Full text] [doi: [10.2196/37806](https://doi.org/10.2196/37806)] [Medline: [35731969](https://pubmed.ncbi.nlm.nih.gov/35731969/)]

8. Basch CH, Zybert P, Reeves R, Basch CE. What do popular YouTube videos say about vaccines? *Child Care Health Dev* 2017 Jul;43(4):499-503. [doi: [10.1111/cch.12442](https://doi.org/10.1111/cch.12442)] [Medline: [28105642](https://pubmed.ncbi.nlm.nih.gov/28105642/)]
9. Zhang J, Featherstone JD, Calabrese C, Wojcieszak M. Effects of fact-checking social media vaccine misinformation on attitudes toward vaccines. *Prev Med* 2021 Apr;145:106408. [doi: [10.1016/j.ypmed.2020.106408](https://doi.org/10.1016/j.ypmed.2020.106408)] [Medline: [33388335](https://pubmed.ncbi.nlm.nih.gov/33388335/)]
10. Shah Z, Surian D, Dyda A, Coiera E, Mandl KD, Dunn AG. Automatically appraising the credibility of vaccine-related web pages shared on social media: a Twitter surveillance study. *J Med Internet Res* 2019 Nov 04;21(11):e14007 [FREE Full text] [doi: [10.2196/14007](https://doi.org/10.2196/14007)] [Medline: [31682571](https://pubmed.ncbi.nlm.nih.gov/31682571/)]
11. Ludolph R, Allam A, Schulz PJ. Manipulating Google's knowledge graph box to counter biased information processing during an online search on vaccination: application of a technological debiasing strategy. *J Med Internet Res* 2016 Jun 02;18(6):e137 [FREE Full text] [doi: [10.2196/jmir.5430](https://doi.org/10.2196/jmir.5430)] [Medline: [27255736](https://pubmed.ncbi.nlm.nih.gov/27255736/)]
12. Wong LP, Wong PF, AbuBakar S. Vaccine hesitancy and the resurgence of vaccine preventable diseases: the way forward for Malaysia, a Southeast Asian country. *Hum Vaccin Immunother* 2020 Jul 02;16(7):1511-1520 [FREE Full text] [doi: [10.1080/21645515.2019.1706935](https://doi.org/10.1080/21645515.2019.1706935)] [Medline: [31977285](https://pubmed.ncbi.nlm.nih.gov/31977285/)]
13. Suzuki Y, Sukegawa A, Ueda Y, Sekine M, Enomoto T, Miyagi E. Effect of a brief web-based educational intervention on willingness to consider human papillomavirus vaccination for children in Japan: randomized controlled trial. *J Med Internet Res* 2021 Sep 27;23(9):e28355 [FREE Full text] [doi: [10.2196/28355](https://doi.org/10.2196/28355)] [Medline: [34569941](https://pubmed.ncbi.nlm.nih.gov/34569941/)]
14. Skafle I, Nordahl-Hansen A, Quintana DS, Wynn R, Gabarron E. Misinformation about COVID-19 vaccines on social media: rapid review. *J Med Internet Res* 2022 Aug 04;24(8):e37367 [FREE Full text] [doi: [10.2196/37367](https://doi.org/10.2196/37367)] [Medline: [35816685](https://pubmed.ncbi.nlm.nih.gov/35816685/)]
15. Jolley D, Douglas KM. Prevention is better than cure: addressing anti-vaccine conspiracy theories. *J Appl Soc Psychol* 2017 Jun 28;47(8):459-469. [doi: [10.1111/jasp.12453](https://doi.org/10.1111/jasp.12453)]
16. Arnesen S, Bærøe K, Cappelen C, Carlsen B. Could information about herd immunity help us achieve herd immunity? Evidence from a population representative survey experiment. *Scand J Public Health* 2018 Dec;46(8):854-858. [doi: [10.1177/1403494818770298](https://doi.org/10.1177/1403494818770298)] [Medline: [29741455](https://pubmed.ncbi.nlm.nih.gov/29741455/)]
17. Hakim H, Provencher T, Chambers CT, Driedger SM, Dube E, Gavaruzzi T, et al. Interventions to help people understand community immunity: a systematic review. *Vaccine* 2019 Jan 07;37(2):235-247. [doi: [10.1016/j.vaccine.2018.11.016](https://doi.org/10.1016/j.vaccine.2018.11.016)] [Medline: [30528593](https://pubmed.ncbi.nlm.nih.gov/30528593/)]
18. Mazzuchelli A, Gurioli M, Graziano D, Quacquarelli B, Aouina-Mejri C. How to fight against food waste in the digital era: key factors for a successful food sharing platform. *J Bus Res* 2021 Jan;124:47-58. [doi: [10.1016/j.jbusres.2020.11.055](https://doi.org/10.1016/j.jbusres.2020.11.055)]
19. Oh S. The characteristics and motivations of health answerers for sharing information, knowledge, and experiences in online environments. *J Am Soc Inf Sci* 2011 Nov 01;63(3):543-557. [doi: [10.1002/asi.21676](https://doi.org/10.1002/asi.21676)]
20. Pai P, Tsai H. Reciprocity norms and information-sharing behavior in online consumption communities: an empirical investigation of antecedents and moderators. *Inf Manag* 2016 Jan;53(1):38-52. [doi: [10.1016/j.im.2015.08.002](https://doi.org/10.1016/j.im.2015.08.002)]
21. Jun J, Wen TJ, Wu L. The effects of self vs. group affirmation and message framing on college students' vape-free campus policy support. *Health Commun* 2021 Oct;36(11):1441-1451. [doi: [10.1080/10410236.2020.1767446](https://doi.org/10.1080/10410236.2020.1767446)] [Medline: [32438843](https://pubmed.ncbi.nlm.nih.gov/32438843/)]
22. Gallagher KM, Updegraff JA. Health message framing effects on attitudes, intentions, and behavior: a meta-analytic review. *Ann Behav Med* 2012 Feb;43(1):101-116. [doi: [10.1007/s12160-011-9308-7](https://doi.org/10.1007/s12160-011-9308-7)] [Medline: [21993844](https://pubmed.ncbi.nlm.nih.gov/21993844/)]
23. Van't Riet J, Cox AD, Cox D, Zimet GD, De Bruijn G, Van den Putte B, et al. Does perceived risk influence the effects of message framing? A new investigation of a widely held notion. *Psychol Health* 2014;29(8):933-949 [FREE Full text] [doi: [10.1080/08870446.2014.896916](https://doi.org/10.1080/08870446.2014.896916)] [Medline: [24579986](https://pubmed.ncbi.nlm.nih.gov/24579986/)]
24. Bosone L, Martinez F. When, how and why is loss-framing more effective than gain- and non-gain-framing in the promotion of detection behaviors? *Int Rev Soc Psychol* 2017 Aug 02;30(1):184-192. [doi: [10.5334/irsp.15](https://doi.org/10.5334/irsp.15)]
25. Questionnaire Star Can Meet Your Various Questionnaire/Form Needs. URL: <https://www.wjx.cn> [accessed 2022-09-12]
26. Faul F, Erdfelder E, Lang A, Buchner A. G\*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res Methods* 2007 May;39(2):175-191. [doi: [10.3758/bf03193146](https://doi.org/10.3758/bf03193146)] [Medline: [17695343](https://pubmed.ncbi.nlm.nih.gov/17695343/)]
27. Betsch C, Böhm R, Korn L, Holtmann C. On the benefits of explaining herd immunity in vaccine advocacy. *Nat Hum Behav* 2017 Mar 06;1(3):0056. [doi: [10.1038/s41562-017-0056](https://doi.org/10.1038/s41562-017-0056)]
28. Doherty M, Buchy P, Standaert B, Giaquinto C, Prado-Cohrs D. Vaccine impact: benefits for human health. *Vaccine* 2016 Dec 20;34(52):6707-6714 [FREE Full text] [doi: [10.1016/j.vaccine.2016.10.025](https://doi.org/10.1016/j.vaccine.2016.10.025)] [Medline: [27773475](https://pubmed.ncbi.nlm.nih.gov/27773475/)]
29. Dixon GN. Making vaccine messaging stick: perceived causal instability as a barrier to effective vaccine messaging. *J Health Commun* 2017 Aug;22(8):631-637. [doi: [10.1080/10810730.2017.1337832](https://doi.org/10.1080/10810730.2017.1337832)] [Medline: [28682187](https://pubmed.ncbi.nlm.nih.gov/28682187/)]
30. Larson HJ, de Figueiredo A, Xiaohong Z, Schulz WS, Verger P, Johnston IG, et al. The state of vaccine confidence 2016: global insights through a 67-country survey. *EBioMedicine* 2016 Oct;12:295-301 [FREE Full text] [doi: [10.1016/j.ebiom.2016.08.042](https://doi.org/10.1016/j.ebiom.2016.08.042)] [Medline: [27658738](https://pubmed.ncbi.nlm.nih.gov/27658738/)]
31. Pal A, Chua AY, Hoe-Lian Goh D. Debunking rumors on social media: the use of denials. *Comput Hum Behav* 2019 Jul;96:110-122. [doi: [10.1016/j.chb.2019.02.022](https://doi.org/10.1016/j.chb.2019.02.022)]
32. Sarathchandra D, Navin MC, Largent MA, McCright AM. A survey instrument for measuring vaccine acceptance. *Prev Med* 2018 Apr;109:1-7. [doi: [10.1016/j.ypmed.2018.01.006](https://doi.org/10.1016/j.ypmed.2018.01.006)] [Medline: [29337069](https://pubmed.ncbi.nlm.nih.gov/29337069/)]

33. Oh HJ, Lee H. When do people verify and share health rumors on social media? The effects of message importance, health anxiety, and health literacy. *J Health Commun* 2019;24(11):837-847. [doi: [10.1080/10810730.2019.1677824](https://doi.org/10.1080/10810730.2019.1677824)] [Medline: [31609678](https://pubmed.ncbi.nlm.nih.gov/31609678/)]
34. Wu B, Yang Z. The impact of moral identity on consumers' green consumption tendency: the role of perceived responsibility for environmental damage. *J Environ Psychol* 2018 Oct;59:74-84. [doi: [10.1016/j.jenvp.2018.08.011](https://doi.org/10.1016/j.jenvp.2018.08.011)]
35. Bolin JH. *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach*. New York, NY: Guilford Press; 2014.
36. Wojcieszak M, Garrett R. Social identity, selective exposure, and affective polarization : how priming national identity shapes attitudes toward immigrants via news selection. *Hum Commun Res* 2018;44:247-273. [doi: [10.1093/hcr/hqx010](https://doi.org/10.1093/hcr/hqx010)]
37. Maleki M, Mahmoudi MR, Heydari MH, Pho K. Modeling and forecasting the spread and death rate of coronavirus (COVID-19) in the world using time series models. *Chaos Solitons Fractals* 2020 Nov;140:110151 [FREE Full text] [doi: [10.1016/j.chaos.2020.110151](https://doi.org/10.1016/j.chaos.2020.110151)] [Medline: [32834639](https://pubmed.ncbi.nlm.nih.gov/32834639/)]
38. Maleki M, Mahmoudi MR, Wraith D, Pho K. Time series modelling to forecast the confirmed and recovered cases of COVID-19. *Travel Med Infect Dis* 2020;37:101742. [doi: [10.1016/j.tmaid.2020.101742](https://doi.org/10.1016/j.tmaid.2020.101742)] [Medline: [33081974](https://pubmed.ncbi.nlm.nih.gov/33081974/)]
39. Isler O, Isler B, Kopsacheilis O, Ferguson E. Limits of the social-benefit motive among high-risk patients: a field experiment on influenza vaccination behaviour. *BMC Public Health* 2020 Feb 17;20(1):240 [FREE Full text] [doi: [10.1186/s12889-020-8246-3](https://doi.org/10.1186/s12889-020-8246-3)] [Medline: [32066407](https://pubmed.ncbi.nlm.nih.gov/32066407/)]
40. Liao T. Developing an antecedent model of knowledge sharing intention in virtual communities. *Univ Access Inf Soc* 2016 Feb 8;16(1):215-224. [doi: [10.1007/s10209-016-0452-5](https://doi.org/10.1007/s10209-016-0452-5)]

## Abbreviations

**LLCI:** lower limit confidence interval

**ULCI:** upper limit confidence interval

*Edited by A Mavragani; submitted 16.12.21; peer-reviewed by Z Jin, M Mahmoudi; comments to author 09.06.22; revised version received 30.07.22; accepted 10.08.22; published 04.10.22.*

*Please cite as:*

*Fu C, Lyu X, Mi M*

*Collective Value Promotes the Willingness to Share Pro Vaccination Messages on Social Media in China: Randomized Controlled Trial*

*JMIR Form Res* 2022;6(10):e35744

URL: <https://formative.jmir.org/2022/10/e35744>

doi: [10.2196/35744](https://doi.org/10.2196/35744)

PMID: [36067417](https://pubmed.ncbi.nlm.nih.gov/36067417/)

©Chunye Fu, Xiaokang Lyu, Mingdi Mi. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 04.10.2022. 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

# COVID-Washing in US Food and Beverage Marketing on Twitter: Content Analysis

Krystle A Tsai<sup>1</sup>, MA; Omni Cassidy<sup>1</sup>, PhD; Josh Arshonsky<sup>1</sup>, BA; Sara Bond<sup>1</sup>, MPH; Inés M Del Giudice<sup>1</sup>, MPH; Marie A Bragg<sup>1</sup>, PhD

Department of Population Health, New York University Grossman School of Medicine, New York City, NY, United States

**Corresponding Author:**

Marie A Bragg, PhD

Department of Population Health

New York University Grossman School of Medicine

180 Madison Ave, 3rd Fl

New York City, NY, 10016

United States

Phone: 1 646 501 2717

Email: [marie.bragg@nyulangone.org](mailto:marie.bragg@nyulangone.org)

## Abstract

**Background:** Food companies have increased digital and social media ad expenditures during the COVID-19 pandemic, capitalizing on the coinciding increase in social media use during the pandemic. The extent of pandemic-related social media advertising and marketing tactics have been previously reported. No studies, however, have evaluated how food and beverage companies used COVID-washing on social media posts in the United States or analyzed the nutritional content of advertised food and beverage products. This study was designed to address these gaps by evaluating how food and beverage companies capitalize on the COVID-19 pandemic to promote unhealthy foods and sugary beverages.

**Objective:** We aimed to document the types and frequencies of COVID-19–related themes in US food and beverage companies' Twitter posts during the first wave of the pandemic in the United States, and assess the nutritional quality of food and beverage products featured in these tweets.

**Methods:** Research assistants visited the Twitter accounts of the most-marketed food and beverage brands, and screen-captured all tweets posted between March 1 and May 31, 2020. Researchers documented the date of the tweet; the number of likes, views, comments, and “retweets”; and the type of food and beverage products. We coded tweets for the following 10 COVID-19 themes: (1) social distancing, staying home, or working remotely; (2) contactless delivery or pick-up; (3) handwashing or sanitizing; (4) masks; (5) safety or protection; (6) staying connected with others; (7) staying active; (8) frontline or essential workers; (9) monetary relief, donations, or unemployment; and (10) pandemic, unprecedented, or difficult times. Researchers calculated the nutrient profile index scores for featured foods and sorted beverages into categories based on sugar content.

**Results:** Our final sample included 874 COVID-19–themed tweets from 52 food and beverage brands. Social distancing themes appeared most frequently (n=367, 42%), followed by pandemic, unprecedented, or difficult times (n=246, 28.2%), and contactless delivery (n=237, 27.1%). The majority of tweets (n=682, 78%) promoted foods and beverages. Among those tweets featuring foods and beverages, 89.6% (n=611) promoted unhealthy products, whereas 17.2% (n=117) promoted healthy products.

**Conclusions:** Our findings point to a concerning marketing tactic in which major food and beverage companies promote unhealthy foods and sugary beverages during the COVID-19 pandemic. Given that nutrition-related diseases such as obesity and diabetes are risk factors for COVID-19 morbidity and mortality, food and beverage companies should reduce the promotion of unhealthy products to help decrease the prevalence of health conditions that place people at higher risk for severe illness and death due to COVID-19.

(*JMIR Form Res* 2022;6(10):e37642) doi:[10.2196/37642](https://doi.org/10.2196/37642)

**KEYWORDS**

COVID-19; food and beverage marketing; Twitter; obesity; social media; nutrition; Twitter; risk factor; health advertising; marketing

## Introduction

Health is determined not only by biology and individual choices but also by corporate practices and their influence on the social environments in which people live and work [1]. These “corporate determinants of health” were most recently defined as “strategies and approaches used by the private sector to promote products and choices that are detrimental to health” [2,3]. One example of a corporate determinant of poor health is unhealthy food and beverage marketing. Exposure to food advertising increases calorie consumption [4] and is particularly concerning during the current COVID-19 pandemic, given that obesity and diabetes are risk factors for COVID-19 morbidity and mortality [5-8].

Food marketing exerts a powerful influence on children’s and adults’ diets. To date, the majority of food advertising promotes products that are energy dense and nutrient poor [9]. Many research studies substantiate that such marketing of unhealthy foods affects young consumers’ preferences and purchasing behaviors, and increases their consumption of those marketed foods [4,9-14].

Social media use increased during the COVID-19 pandemic [15-19], coinciding with food and beverage companies’ growth in marketing efforts on social media. Companies that shifted their marketing dollars to digital media-centric platforms saw significant profit [20]: ad revenue from companies like YouTube, Facebook, and Twitter grew by up to 49% during the first quarter of 2021 [21]. The extent of pandemic-related social media advertising and marketing tactics have been previously reported [22-25]. Despite companies’ efforts to align with public health initiatives, COVID-washing may contribute to obesogenic environments and increase health risk. A content analysis of Australian-based social media posts from Australian food and beverage companies during the pandemic showed that 100% of unhealthy food and beverage parent companies included “COVID-washing” marketing techniques—the use of COVID-19-related themes (eg, social distancing) to market products [25]. The use of COVID-19-related marketing by food and beverage companies has also been documented in Latin America. A content analysis of Facebook posts of 5 fast food chains in Argentina, Bolivia, Guatemala, and Peru also documented the use of COVID-19-related marketing strategies during the pandemic [22], and a study in Brazil documented a high prevalence of unhealthy food advertising on the largest online food delivery platform in Latin America during the fourth month of the COVID-19 pandemic [23]. No studies, however, have evaluated how food and beverage companies used COVID-washing on social media posts in the United States or analyzed the nutritional content of advertised food and beverage products. Most of the world’s largest food companies are from the United States (eg, 5 of the top 10 largest food companies are from the United States) [26], and those US-based companies are increasingly targeting other countries with their unhealthy products [27], making it urgent to understand their domestic and international marketing practices. Conversely, Latin American brands do not have a large advertising presence in the United States—the most-marketed brands in the United States largely include US-based brands along with a few brands

from outside of the United States (eg, Nestle based in Switzerland and Danone based in France) [26-30].

This study was designed to address these gaps by evaluating how food and beverage companies capitalize on the COVID-19 pandemic to promote unhealthy foods and sugary beverages. We aimed to document the types and frequencies of COVID-19-related themes in US food and beverage companies’ Twitter posts (ie, “tweets”), and assess the nutritional quality of featured products in these tweets.

## Methods

### Selection of Food and Beverage Brands

We conducted a content analysis of food and beverage brands’ US Twitter accounts during the first wave of the COVID-19 pandemic. We chose Twitter as the social media platform to analyze in our study because companies spend US \$2.9 billion of their advertising budget on Twitter [31]. Additionally, our previous research shows that food and beverage brands use diverse and powerful advertising strategies on Twitter, including adopting personalities that drive consumer engagement and interaction on the platform [32]. To determine our sample of food and beverage brands, we selected a random subset of 29 of the most-marketed brands from the following categories, using the Rudd Center for Food Policy and Obesity’s Food Advertising to Children and Teens Scores (FACTS) reports: fast food [30], sugary drinks [33], cereals [34], and snacks [28]. Five research assistants then visited Twitter and searched for each brand’s official Twitter account page. On each account page, Twitter suggests three similar accounts in a “You might like” list. We also included these suggested accounts in our sample. Twitter uses an algorithm to populate these “You might like” lists, suggesting additional brands that are likely relevant and of interest to users who follow and interact with the initial brand account they are viewing. Our choice to include brands from these lists was because users who “follow” a particular brand account on Twitter are likely to also follow the suggested accounts in the “you might like” list on their account page. By also including these suggested brands, we aimed to produce a robust sample of popular food brands on Twitter. This search process yielded a sample of 52 brands—29 brands from the FACTS reports and 23 brands from Twitter’s “You might like” lists.

### Collection of Tweets From Brands’ Twitter Accounts

The lead author divided the 52 brands between seven researchers, who then visited their assigned brands’ Twitter accounts and recorded the number of “followers” each brand had. To collect our sample of tweets, the researchers then scrolled through each brand’s reverse-chronological feed of posted tweets, until they reached a tweet dated May 31, 2020. They then screen-captured each tweet posted between March 1 and May 31, 2020, and documented the following data for each captured tweet in an Excel (Microsoft Corporation) spreadsheet: date of posted tweet, number of likes, number of views, number of comments, and number of “retweets.” We only collected tweets and retweets generated by the brands, and did not collect any user-generated content. Finally, the researcher identified and documented each food and beverage product featured in

the tweet. Products were included in the data collection if they were either referenced in the text of the tweet or depicted with an image of the product.

### Codebook Development and Qualitative Coding of Tweets

To create a codebook, the lead author and two senior researchers reviewed a subset of tweets in our sample to identify featured COVID-19-related themes. These identified themes were then documented and defined in a codebook to be used in a content analysis of the full sample of tweets. The lead author then trained the seven researchers on how to use the codebook to perform the qualitative coding. If throughout the coding process, a researcher identified an additional COVID-19-related theme, they notified the lead author who then updated the codebook. The final codebook included 10 distinct COVID-19-related themes: (1) social distancing, staying home, or working remotely; (2) contactless delivery or pick-up; (3) handwashing or sanitizing; (4) masks; (5) safety/protection; (6) staying connected with others; (7) staying active; (8) frontline or essential workers; (9) monetary relief, donations, or unemployment; and (10) pandemic, unprecedented, or difficult times (Table S1 in [Multimedia Appendix 1](#)).

Each tweet was then reviewed and coded for the 10 established COVID-19-related themes by one of the seven researchers. If there was a lack of clarity on how a tweet should be coded, the researcher raised the question with the research team, and through a team discussion, a consensus was reached. To account for coding consistency, the lead author reviewed 20% of the sample from each of the seven researchers. If a coding discrepancy was identified, the lead author discussed the discrepancy with the initial coder, and they reached a consensus. Researchers coded tweets for each theme that was featured, so some tweets were coded with more than one theme.

### Nutritional Analysis of Food and Beverages

We calculated nutrition scores for each featured food product using the validated Nutrient Profile Model (NPM) [35], with higher scores representing less healthy products and lower scores representing healthier products. We then converted the NPM scores to a more interpretable nutrient profile index (NPI) [34] score using the following formula:  $NPI\ score = -2 \times NPM\ score + 70$ . NPI scores ranged from 1 to 100, with 1 being the worst nutrition score and 100 being the best score. We classified food items with an NPI score  $\geq 64$  as “healthy,” as this is the threshold for products that can be advertised to children in the United Kingdom [36]. The NPM is limited in its use with beverages and codes some sugary sweetened beverages as “healthy,” so we sorted featured beverages into the categories based on the sugar content outlined in the Rudd Center’s Sugary Drink FACTS Report [33]. Food and beverage products that were referenced in the text of the tweet or depicted with an image of the product were included in the nutritional analysis. In instances where a featured food or beverage product was not easily identifiable (eg, image of a brand’s take-out bag), we calculated and averaged the NPI scores of the brand’s top five marketed products, as described by Bragg et al [37].

## Results

After excluding brands without active Twitter accounts, we identified 52 unique food and beverage brands. During the study period, the 52 food and beverage brands collectively posted 2307 tweets, and 874 of those tweets featured COVID-19-related themes. Brands mentioned social distancing ( $n=367$ , 42% of COVID-19-themed tweets) most frequently in their tweets, followed by pandemic, unprecedented, or difficult times ( $n=246$ , 28.2%), contactless delivery ( $n=237$ , 27.1%), and frontline or essential workers ( $n=193$ , 22.1%; [Figure 1](#)). Half of the COVID-19-themed tweets ( $n=440$ , 50.1%) featured more than one COVID-19 theme. Engagement by followers (likes, comments, and retweets) varied between COVID-19 themes ([Table 1](#)) and brands ([Table 2](#)). Tweets featuring social distancing generated the most interactions (likes, comments, and retweets) of the 10 themes, with 118,838 likes; 11,268 comments; and 21,643 retweets ( $n=151,749$ , 23.2% of the total 655,551 interactions). Tweets referencing pandemic, unprecedented, or difficult times generated the next greatest number of interactions ( $n=124,632$ , 19% of the total interactions) with 96,423 likes; 6417 comments; and 21,792 retweets. Tweets referencing staying active generated the least number of interactions in our sample ( $n=1673$ , 0.3%).

Fast-food companies comprised more than one-quarter of the 52 brands in our sample ( $n=14$ , 27%) and generated more than half of the 874 COVID-19-themed tweets in our sample ( $n=461$ , 52.7%). Taco Bell (47/461, 10.2%), Little Caesars (46/461, 10%), and Chick-fil-A (41/461, 8.9%) posted the most COVID-19-themed tweets among fast-food brands, comprising 15.3% (134/874) of the total COVID-19-themed tweets in our sample. Fast-food companies had the majority of followers in our sample (27,274,600/37,063,870, 73.6%), and their tweets had the highest level of engagement, accruing the majority of the total number of “likes” (205,026/270,697, 75.7%), comments (20,333/27,633, 73.6%), and retweets (43,530/55,683, 78.2%). Contactless delivery was the most frequently referenced theme in the 461 tweets by fast-food brands ( $n=208$ , 45.1%), followed by social distancing ( $n=135$ , 29.3%), and pandemic ( $n=113$ , 24.5%).

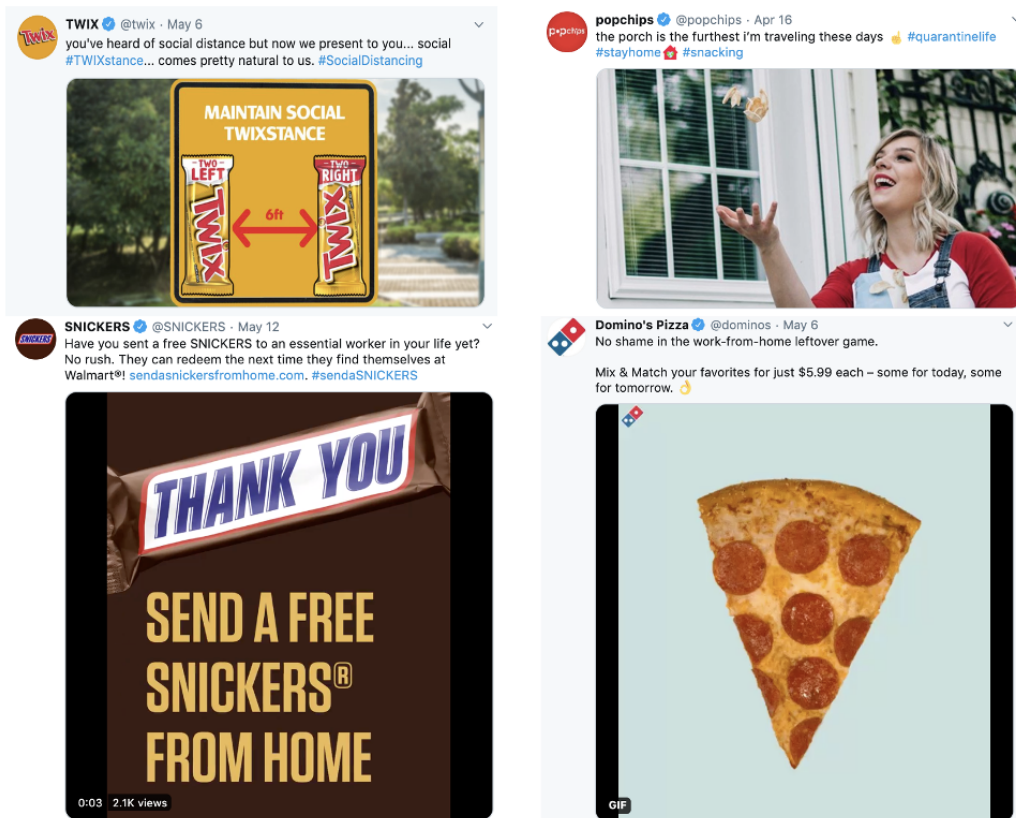
Snack brands comprised over one-third of the 52 brands in our sample ( $n=19$ , 37%) and generated almost one-third of the 874 COVID-19-themed tweets in our sample ( $n=260$ , 29.7%). Snack brands had a total of 2,343,412 (6.3%) followers, and their tweets generated a total of 16,618 (6.1%) likes; 2454 (8.9%) comments; and 4187 (7.5%) retweets. Social distancing was the most frequently referenced theme in tweets by snack brands (132/260, 50.8%), followed by pandemic (93/260, 35.8%).

Beverage brands comprised 23% ( $n=12$ ) of the 52 brands in our sample and had the second highest number of followers in our sample ( $n=7,104,432$ , 19.2%). Beverage brands generated 9.5% ( $n=83$ ) of the COVID-19-themed tweets in our sample but had the second highest level of engagement, accruing a total of 47,691 (17.6%) likes; 4692 (17%) comments; and 7711 (13.9%) retweets. Social distancing was the most frequently referenced theme in the 83 beverage brand tweets ( $n=50$ , 60%), followed by staying connected ( $n=26$ , 31%) and pandemic ( $n=24$ , 29%).

A total of 78% (n=682) of the 874 tweets that featured COVID-19–related themes also promoted foods and beverages. Of these 682 COVID-19–related tweets, 89.6% (n=611) promoted unhealthy products, whereas 17.2% (n=117) promoted healthy products. Unhealthy foods were promoted more frequently (n=556, 81.5%) than sugary beverages (n=100, 14.7%). Healthy food and beverages were promoted in only 8.4% (n=57) and 9.4% (n=64) of tweets, respectively. A total of 10.4% (n=71) of tweets promoted only healthy products,

whereas 82.8% (n=565) promoted only unhealthy products. A total of 38% (n=20) of the 52 brands in our sample promoted both healthy and unhealthy products in their COVID-19–themed tweets. Fast-food and snack brands promoted unhealthy products (n=367, 89.7% and n=168, 92.8%, respectively) more often than healthy products (n=59, 14.4% and n=24, 13.3%, respectively). Beverage brands promoted healthy products more often than unhealthy products (n=32, 62.8% vs n=27, 52.9%).

**Figure 1.** Four tweets in our sample that use COVID-19 themes (social distancing/staying home/working remotely; frontline/essential workers) to promote unhealthy food products.



**Table 1.** Summary data for COVID-19–related themes featured across all tweets (N=874).

COVID-19 theme <sup>a</sup>	Frequency (N=1526), n (%)	Likes (N=503,649), n (%)	Comments (N=46,966), n (%)	Retweets (N=104,936), n (%)	Video views <sup>b</sup> (N=32,065,091,897), n (%)
Social distancing, staying home, working remotely	367 (24.1)	118,838 (23.6)	11,268 (24.0)	21,643 (20.6)	25,378,029 (0.1)
Pandemic, unprecedented or difficult times	246 (16.1)	96,423 (19.1)	6417 (13.7)	21,792 (20.8)	16,000,164,700 (49.9)
Contactless delivery and pick-up	237 (15.5)	60,363 (12.0)	8802 (18.7)	15,085 (14.4)	3,996,010 (0.0)
Frontline and essential workers	193 (12.7)	63,689 (12.7)	5435 (11.8)	11,817 (11.3)	3,719,072 (0.0)
Monetary relief, donations, unemployment	192 (12.6)	37,860 (7.5)	3911 (8.3)	9260 (8.8)	16,005,367,147 (49.9)
Staying connected with others	125 (8.2)	63,414 (12.6)	4322 (9.2)	10,979 (10.5)	24,731,677 (0.0)
Safety, protection	74 (4.9)	37,531 (7.5)	4046 (8.6)	9077 (8.7)	128,128 (0.0)
Masks	51 (3.3)	10,271 (2.0)	794 (1.7)	1758 (1.7)	1,576,670 (0.0)
Handwashing, sanitizing	23 (1.5)	13,903 (2.8)	1894 (4.0)	3286 (3.1)	23,364 (0.0)
Staying active	18 (1.2)	1357 (0.3)	77 (0.2)	239 (0.2)	7100 (0.0)

<sup>a</sup>Some tweets included more than one theme, so the total number of theme instances is greater than the total number of tweets.

<sup>b</sup>Data only includes tweets that feature video media (n=192), as Twitter only reports the number of views for videos.



**Table 2.** Summary data for food and beverage brands in sample (N=52).

Brands	Tweets (N=874), n	Total followers (N=37,063,870), n	Total likes (N=270,697), n	Total com- ments (N=27,633), n	Total retweets (N=55,683), n	Total videos (N=167), n	Total video views (N=94,390,779), n	Category
3 Musketeers	17	3646	187	14	33	0	N/A <sup>a</sup>	Snacks
Altoids	3	1072	71	5	4	0	N/A	Snacks
Ben's Original	3	7126	11	6	4	0	N/A	Other
Breyer's	3	123,800	47	3	18	0	N/A	Snacks
Brisk Tea	2	34,700	18	3	4	1	270	Beverages
Cheerios	1	88,100	4	0	1	0	N/A	Cereal
Chick-fil-A	41	1,100,000	17,093	1316	2,818	7	111,200	Fast food
Chobani	15	103,300	633	29	160	1	16,900,000	Snacks
Cinnabon	30	169,600	2018	270	569	4	10,300	Fast food
Coca-Cola	9	3,300,000	3466	435	666	1	11,900	Beverages
Coffee Mate	1	59,400	67	8	8	0	N/A	Other
Combos	4	5712	84	12	22	0	N/A	Snacks
Domino's	27	1,300,000	5448	1838	948	2	15,900	Fast food
Dunkin Donuts	39	1,200,000	10,851	802	1785	10	78,600	Fast food
Five Gum	2	22,200	13	7	2	1	261	Snacks
Gatorade	3	330,400	69	13	15	0	N/A	Beverages
Hellmann's	5	36,400	51	4	18	2	695	Other
Kashi Foods	10	11,800	19	3	4	0	N/A	Snacks
Kellogg	63	77,300	1202	65	466	12	8256	Snacks
KFC	25	1,400,000	5306	1105	1165	9	106,800	Fast food
KitKat	11	409,300	995	43	207	5	10,700	Snacks
Kool-Aid	6	74,200	26,112	1122	4070	1	961,700	Beverages
Krispy Kreme	22	346,000	14,492	477	3070	1	5500	Fast food
Lay's	6	457,500	606	254	173	2	2461	Snacks
Lipton Tea	4	55,000	135	5	33	0	N/A	Beverages
Little Caesars	46	316,200	9917	915	2679	22	5,830,600	Fast food
Lunchables	4	9982	70	21	19	0	N/A	Snacks
M&Ms	9	118,400	811	50	135	2	2192	Snacks
Magnum Ice Cream	4	56,500	53	14	6	2	932	Snacks
McDonald's	22	3,600,000	23,946	3430	3098	2	159,700	Fast food
Miracle Whip	4	17,100	37	16	6	0	N/A	Other
Oscar Meyer	35	38,100	2243	192	631	6	74,342	Snacks
Panera Bread	39	460,800	8249	632	1797	9	1,534,900	Fast food
Pepsi	9	3,000,000	15,607	2857	2567	2	68,091,000	Beverages
Perrier USA	6	14,300	28	1	9	1	149	Beverages
Pizza Hut	38	1,600,000	4022	1127	923	17	94,000	Fast Food
Poland Spring	7	7932	56	2	13	3	397	Beverages
Popchips	38	48,800	89	17	46	10	1529	Snacks
Skittles	7	407,000	6505	1340	1519	2	116,800	Snacks
Snickers	10	373,100	2207	245	643	3	6300	Snacks

Brands	Tweets (N=874), n	Total followers (N=37,063,870), n	Total likes (N=270,697), n	Total com- ments (N=27,633), n	Total retweets (N=55,683), n	Total videos (N=167), n	Total video views (N=94,390,779), n	Category
Sonic Drive-In	32	382,000	2608	451	603	10	27,900	Fast food
Starbucks	34	11,100,000	66,380	4360	13,381	3	101,800	Fast food
Subway	19	2,300,000	2014	391	513	2	8700	Fast food
Taco Bell	47	1,900,000	32,682	3219	10,181	3	101,800	Fast food
TGI Fridays	26	214,200	464	57	135	6	5100	Other
Topo Chico	10	10,600	107	9	22	0	N/A	Beverages
Tostitos	7	40,000	128	16	19	0	N/A	Snacks
Tropicana	1	185,800	9	1	2	0	N/A	Beverages
Twix	12	34,900	655	122	80	1	898	Snacks
Velveeta	30	20,100	728	65	83	0	N/A	Other
Vita Coco	15	57,900	2060	243	304	1	7100	Beverages
Zico Coconut Water	11	33,600	24	1	6	1	97	Beverages

<sup>a</sup>N/A: not applicable.

## Discussion

### Principal Findings

We identified a wide range of COVID-19–related themes. Brands used social distancing themes most frequently, often with a play-on-words using the brand name or slogan (eg, “Maintain Social TWIXstance” or “Lunchables from home”). During the first wave of the COVID-19 pandemic, between March and May 2020, we found that 78% (n=682) of our sampled tweets with COVID-19–related themes promoted food and beverage products, of which the majority were unhealthy.

Our findings are consistent with previous studies that have analyzed food and beverage marketing during the COVID-19 pandemic, showing that major food and beverage brands worldwide frequently posted on social media using COVID-19–related themes [25,38,39]. In our sample of tweets, companies most frequently referenced social distancing themes (eg, “OH \_\_\_ 6 ft \_\_\_ YEAH!” [Kool-Aid], “No shame in the work-from-home leftover game” [Domino’s Pizza], “having your breaks at a distance” [KitKat], or “Quarantine Cuisine” [Popchips]). This finding is consistent with previous studies that examined other social media platforms like Facebook and Instagram where the majority of posts from major food and beverage brands referenced “isolation activities” (eg, #stayhome) [25]—activities that are included under our “social distancing/staying home/working remotely” theme. Discussing social distancing and isolation activities may be useful for food and beverage brands. With millions of families remaining in their homes, people have turned to the internet to order food and beverages. The four major US food delivery apps—Grubhub, DoorDash, Uber Eats, and Postmates—reported a collective US \$3 billion increase in revenues in the second and third quarters of 2020 due to the pandemic’s shelter-in-place restrictions [40]. Particularly concerning is the increase in online ordering of fast food that continued to increase through 2021. A recent report by UpMenu, an online food ordering software

service provider, reported that Chick-fil-A experienced a 590% increase in online ordering during the pandemic, followed closely behind by Burger King and Little Caesars with increases of 407% and 279%, respectively [40,41]. Our findings that Little Caesars and Chick-fil-A were in the top 4 most active brands posting COVID-19–themed content in our sample, suggests that COVID-washing may have played a role in this increase in online ordering during the pandemic.

Using the NPM, this study is the first to also examine the nutritional quality of products promoted in companies’ COVID-19–related tweets. The analysis revealed that most products were classified as unhealthy, which is consistent with other analyses of food and beverage advertisements [42-47]. However, COVID-19 raises new ethical considerations for food and beverage companies’ marketing practices because diet-related diseases like obesity and diabetes are major risk factors for severe illness and death due to COVID-19.

These findings point to a concerning yet typical marketing tactic of food and beverage companies. Corporate marketers taking advantage of trending social or health issues for profit is a textbook strategy from the corporate marketing playbook. However, using the uncertainty and disruptions of the COVID-19 pandemic for promotion is particularly more egregious than say seasonal or holiday food marketing campaigns because overconsuming unhealthy food and drink can indirectly increase the risk of more severe COVID-19–related outcomes.

Posting on Twitter and other social media platforms has enabled companies to market themselves to a wide audience, as most of the US population during our study period followed stay-at-home orders and spent more time online than ever before. Scholars refer to this kind of corporate activity as a type of cause marketing or corporate social responsibility [48] strategy, whereby companies align themselves with certain social or health issues to enhance their own image. COVID-washing

portrays food and beverage companies as empathetic and responsive to the pandemic. In reality, however, this is a classic corporate strategy that may contribute to poor diet and exacerbate poor outcomes related to COVID-19.

### Limitations

This study has several limitations. One limitation is that it is possible we did not capture every food and beverage brand's Twitter account in our sample. We included a subset of the top marketed brands from the Rudd Center for Food Policy and Obesity's FACTS reports in our sample, which likely omitted smaller food and beverage brands. While we also included the additional brands suggested in Twitter's "You might like" list on each brand's Twitter account, these suggested lists were not comprehensive, as they only include three brands each. It is therefore possible that there are additional food and beverage brands that we did not capture in our sample. Further, as only Twitter was included in our sample, we only captured a subset of these food and beverage brands' COVID-19-themed marketing. It is possible that some food and beverage companies may maintain a more active marketing presence on other social media platforms such as Instagram, TikTok, and Facebook. Future studies should examine COVID-19-themed marketing across all social media platforms.

A second potential limitation is that we selected the brands in our sample from unhealthier food categories (fast food, sugary drinks, snacks, and cereal) and did not focus on brands in healthier categories. Our sample did, however, include some brands with healthy products (eg, Perrier, Poland Spring, Chobani, or Subway), and many of the unhealthy brands in our sample also offer some items that are healthy. Over one-third of the brands ( $n=20$ , 39%) in our sample featured a mix of

healthy and unhealthy products in their tweets; however, the healthier options were not products that appeared as frequently in the ads in our sample. Future research could examine the prevalence of COVID-washing by healthier food and beverage brands; however, we predict this would likely comprise a tiny proportion of the overall food marketing landscape because we know from previous research that the unhealthy brands spend more on advertising than healthy brands [28,30,33] and that food companies mostly promote their unhealthy products on their social media accounts [49-51]. Because of this discrepancy, we were less interested in comparing healthy versus unhealthy brands and more interested in documenting the extent to which these unhealthy brands are promoting junk food. This phenomenon is problematic because unhealthy brands have the most money to spend and should not be capitalizing on the pandemic to sell products given the risk factors of obesity and diabetes related to COVID-19.

Another limitation is that we were unable to measure the actual exposure and impressions of these Twitter posts. Finally, we did not examine how these COVID-19-themed ads may directly affect the dietary and purchasing choices of consumers, which is an area of future research.

### Conclusions

Our findings document the prevalence of food and beverage companies' use of COVID-19 themes in tweets that promote unhealthy foods and sugary beverages. Given that obesity and diabetes are risk factors for COVID-19 morbidity and mortality, food and beverage companies should reduce the promotion of unhealthy products to help decrease the prevalence of health conditions that place people at higher risk for severe illness and death due to COVID-19.

---

### Acknowledgments

We would like to thank the following research assistants for their contributions to the data extraction and coding: Yuanqi Gu, MS; Melanie Lowenstein, BA; Emma Gutstein, MPH; Chioma Anugwom, MPH; and Sumaita Mahmood, BS. The aforementioned contributors have no conflicts of interest to report. This study was supported, in part, by a National Institutes of Health Early Independence Award (DP5OD021373-05; principal investigator MAB) from the National Institutes of Health Office of the Director and an Agency for Healthcare Research and Quality fellowship grant (1T32HS026120-01; OC). These funding sources had no role in the study design; in the collection, analysis, or interpretation of data; in the writing of the report; nor in the decision to submit the article for publication.

---

### Authors' Contributions

KAT conceived of and designed the study; led the acquisition, analysis, and interpretation of data; and generated tables and figures. SB and IMDG participated in the acquisition of the data. All authors were involved in drafting the manuscript. MAB and OC provided critical revision of the manuscript. All authors had final approval of the submitted and published versions of the manuscript.

---

### Conflicts of Interest

None declared.

---

Multimedia Appendix 1

Table S1.

[[DOCX File, 18 KB - formative\\_v6i10e37642\\_app1.docx](#) ]

---

### References

<https://formative.jmir.org/2022/10/e37642>

JMIR Form Res 2022 | vol. 6 | iss. 10 | e37642 | p.619  
(page number not for citation purposes)

1. McKee M, Stuckler D. Revisiting the corporate and commercial determinants of health. *Am J Public Health* 2018 Sep;108(9):1167-1170. [doi: [10.2105/AJPH.2018.304510](https://doi.org/10.2105/AJPH.2018.304510)] [Medline: [30024808](https://pubmed.ncbi.nlm.nih.gov/30024808/)]
2. Kickbusch I, Allen L, Franz C. The commercial determinants of health. *Lancet Global Health* 2016 Dec;4(12):e895-e896. [doi: [10.1016/s2214-109x\(16\)30217-0](https://doi.org/10.1016/s2214-109x(16)30217-0)]
3. Millar JS. The corporate determinants of health: how big business affects our health, and the need for government action!. *Can J Public Health* 2013 May 14;104(4):e327-e329 [FREE Full text] [doi: [10.17269/cjph.104.3849](https://doi.org/10.17269/cjph.104.3849)] [Medline: [24044474](https://pubmed.ncbi.nlm.nih.gov/24044474/)]
4. Qutteina Y, De Backer C, Smits T. Media food marketing and eating outcomes among pre-adolescents and adolescents: a systematic review and meta-analysis. *Obes Rev* 2019 Dec;20(12):1708-1719. [doi: [10.1111/obr.12929](https://doi.org/10.1111/obr.12929)] [Medline: [31468652](https://pubmed.ncbi.nlm.nih.gov/31468652/)]
5. Rajpal A, Rahimi L, Ismail-Beigi F. Factors leading to high morbidity and mortality of COVID-19 in patients with type 2 diabetes. *J Diabetes* 2020 Dec;12(12):895-908 [FREE Full text] [doi: [10.1111/1753-0407.13085](https://doi.org/10.1111/1753-0407.13085)] [Medline: [32671936](https://pubmed.ncbi.nlm.nih.gov/32671936/)]
6. Richardson S, Hirsch J, Narasimhan M, Crawford JM, McGinn T, Davidson KW, the Northwell COVID-19 Research Consortium, et al. Presenting characteristics, comorbidities, and outcomes among 5700 patients hospitalized with COVID-19 in the New York City Area. *JAMA* 2020 May 26;323(20):2052-2059 [FREE Full text] [doi: [10.1001/jama.2020.6775](https://doi.org/10.1001/jama.2020.6775)] [Medline: [32320003](https://pubmed.ncbi.nlm.nih.gov/32320003/)]
7. Palaiodimos L, Kokkinidis DG, Li W, Karamanis D, Ognibene J, Arora S, et al. Severe obesity, increasing age and male sex are independently associated with worse in-hospital outcomes, and higher in-hospital mortality, in a cohort of patients with COVID-19 in the Bronx, New York. *Metabolism* 2020 Jul;108:154262 [FREE Full text] [doi: [10.1016/j.metabol.2020.154262](https://doi.org/10.1016/j.metabol.2020.154262)] [Medline: [32422233](https://pubmed.ncbi.nlm.nih.gov/32422233/)]
8. Harris J, Schwartz M, Munsell C. Fast Food FACTS 2013: measuring progress in nutrition and marketing to children and teens. *Fast Food FACTS*. 2013. URL: [http://www.fastfoodmarketing.org/media/fastfoodfacts\\_report.pdf](http://www.fastfoodmarketing.org/media/fastfoodfacts_report.pdf) [accessed 2022-08-31]
9. Kraak VI, Rincón-Gallardo Patiño S, Sacks G. An accountability evaluation for the International Food & Beverage Alliance's Global Policy on Marketing Communications to Children to reduce obesity: A narrative review to inform policy. *Obes Rev* 2019 Nov;20 Suppl 2:90-106. [doi: [10.1111/obr.12859](https://doi.org/10.1111/obr.12859)] [Medline: [31034139](https://pubmed.ncbi.nlm.nih.gov/31034139/)]
10. Smith R, Kelly B, Yeatman H, Boyland E. Food marketing influences children's attitudes, preferences and consumption: a systematic critical review. *Nutrients* 2019 May 18;11(4):875 [FREE Full text] [doi: [10.3390/nu11040875](https://doi.org/10.3390/nu11040875)] [Medline: [31003489](https://pubmed.ncbi.nlm.nih.gov/31003489/)]
11. Hallez L, Qutteina Y, Raedschelders M, Boen F, Smits T. That's my cue to eat: a systematic review of the persuasiveness of front-of-pack cues on food packages for children vs. adults. *Nutrients* 2020 May 11;12(4):1062 [FREE Full text] [doi: [10.3390/nu12041062](https://doi.org/10.3390/nu12041062)] [Medline: [32290522](https://pubmed.ncbi.nlm.nih.gov/32290522/)]
12. Scully M, Wakefield M, Niven P, Chapman K, Crawford D, Pratt IS, NaSSDA Study Team. Association between food marketing exposure and adolescents' food choices and eating behaviors. *Appetite* 2012 Mar;58(1):1-5. [doi: [10.1016/j.appet.2011.09.020](https://doi.org/10.1016/j.appet.2011.09.020)] [Medline: [22001023](https://pubmed.ncbi.nlm.nih.gov/22001023/)]
13. McClure AC, Tanski SE, Gilbert-Diamond D, Adachi-Mejia AM, Li Z, Li Z, et al. Receptivity to television fast-food restaurant marketing and obesity among U.S. youth. *Am J Prev Med* 2013 Dec;45(5):560-568 [FREE Full text] [doi: [10.1016/j.amepre.2013.06.011](https://doi.org/10.1016/j.amepre.2013.06.011)] [Medline: [24139768](https://pubmed.ncbi.nlm.nih.gov/24139768/)]
14. Story M, Neumark-Sztainer D, French S. Individual and environmental influences on adolescent eating behaviors. *J Am Diet Assoc* 2002 Mar;102(3 Suppl):S40-S51. [doi: [10.1016/s0002-8223\(02\)90421-9](https://doi.org/10.1016/s0002-8223(02)90421-9)] [Medline: [11902388](https://pubmed.ncbi.nlm.nih.gov/11902388/)]
15. Drouin M, McDaniel BT, Pater J, Toscos T. How parents and their children used social media and technology at the beginning of the COVID-19 pandemic and associations with anxiety. *Cyberpsychol Behav Soc Netw* 2020 Dec;23(11):727-736. [doi: [10.1089/cyber.2020.0284](https://doi.org/10.1089/cyber.2020.0284)] [Medline: [32726144](https://pubmed.ncbi.nlm.nih.gov/32726144/)]
16. Williamson DA. US consumers are flocking to TikTok. *Insider Intelligence*. 2020 Apr 27. URL: <https://www.insiderintelligence.com/content/us-consumers-are-flocking-to-tiktok> [accessed 2022-08-19]
17. Kemp S. Digital 2020: July global statshot. *DataReportal*. 2020 Jul 21. URL: <https://datareportal.com/reports/digital-2020-july-global-statshot> [accessed 2022-08-01]
18. COVID-19 barometer: consumer attitudes, media habits and expectations. *Kantar*. URL: <https://www.kantar.com/Inspiration/Coronavirus/COVID-19-Barometer-Consumer-attitudes-media-habits-and-expectations> [accessed 2022-08-01]
19. Dixon S. Increased time spent on social by U.S. users during COVID-19 pandemic 2020. *Statista*. 2022 Apr 28. URL: <https://www.statista.com/statistics/1116148/more-time-spent-social-media-platforms-users-usa-coronavirus/> [accessed 2022-08-01]
20. Hagey K, Vranica S. How Covid-19 supercharged the advertising “Triopoly” of Google, Facebook and Amazon. *The Wall Street Journal*. 2021 Mar 19. URL: <https://www.wsj.com/articles/how-covid-19-supercharged-the-advertising-triopoly-of-google-facebook-and-amazon-11616163738> [accessed 2021-09-22]
21. Alphabet Q1 2021 Earnings Call. *Alphabet*. 2020 Apr 27. URL: [https://abc.xyz/investor/static/pdf/2021\\_Q1\\_Earnings\\_Transcript.pdf?cache=aa96ea2](https://abc.xyz/investor/static/pdf/2021_Q1_Earnings_Transcript.pdf?cache=aa96ea2) [accessed 2020-09-20]
22. Rozas L, Castronuovo L, Busse P, Mus S, Barnoya J, Garrón A, et al. Data on the Facebook marketing strategies used by fast-food chains in four Latin American countries during the COVID-19 lockdowns. *BMC Res Notes* 2021 Dec 20;14(1):463 [FREE Full text] [doi: [10.1186/s13104-021-05870-8](https://doi.org/10.1186/s13104-021-05870-8)] [Medline: [34930460](https://pubmed.ncbi.nlm.nih.gov/34930460/)]

23. Horta P, Matos JDP, Mendes L. Digital food environment during the coronavirus disease 2019 (COVID-19) pandemic in Brazil: an analysis of food advertising in an online food delivery platform. *Br J Nutr* 2021 Sep 14;126(5):767-772 [FREE Full text] [doi: [10.1017/S0007114520004560](https://doi.org/10.1017/S0007114520004560)] [Medline: [33208203](https://pubmed.ncbi.nlm.nih.gov/33208203/)]
24. Limb M. Covid-19: food and drink companies are exploiting pandemic to further their brands, analysis finds. *BMJ* 2020 Sep 11;370:m3548. [doi: [10.1136/bmj.m3548](https://doi.org/10.1136/bmj.m3548)] [Medline: [32917584](https://pubmed.ncbi.nlm.nih.gov/32917584/)]
25. Martino F, Brooks R, Browne J, Carah N, Zorbas C, Corben K, et al. The nature and extent of online marketing by big food and big alcohol during the COVID-19 pandemic in Australia: content analysis study. *JMIR Public Health Surveill* 2021 Mar 12;7(3):e25202 [FREE Full text] [doi: [10.2196/25202](https://doi.org/10.2196/25202)] [Medline: [33709935](https://pubmed.ncbi.nlm.nih.gov/33709935/)]
26. Sorvino C. Forbes Global 2000: the world's largest food companies in 2022. *Forbes*. 2022 May 12. URL: <https://www.forbes.com/sites/chloesorvino/2022/05/12/the-worlds-largest-food-companies-in-2022/> [accessed 2022-08-19]
27. Cassidy O, Shin H, Song E, Jiang E, Harri R, Cano C, et al. Comparing McDonald's food marketing practices on official Instagram accounts across 15 countries. *BMJ Nutr Prev Health* 2021;4(2):510-518 [FREE Full text] [doi: [10.1136/bmjnph-2021-000229](https://doi.org/10.1136/bmjnph-2021-000229)] [Medline: [35028520](https://pubmed.ncbi.nlm.nih.gov/35028520/)]
28. Harris JL, Schwartz MB, Shehan C, Hyary M, Appel J, Haraghey K, et al. Snack FACTS 2015: evaluating snack food nutrition and marketing to youth. UConn Rudd Center for Food Policy and Health. 2015. URL: [https://uconnruddcenter.org/wp-content/uploads/sites/2909/2020/11/SnackFACTS\\_2015\\_Fulldraft03.pdf](https://uconnruddcenter.org/wp-content/uploads/sites/2909/2020/11/SnackFACTS_2015_Fulldraft03.pdf) [accessed 2022-08-31]
29. Fleming-Milici F, Phaneuf L, Elmi H, Harris JL. Fast Food FACTS 2021: social media supplement: fast-food restaurant social media accounts and examples of campaigns targeted to teens and Hispanic and Black youth. UConn Rudd Center for Food Policy and Health. 2021. URL: <https://media.ruddcenter.uconn.edu/PDFs/SocialMedia2021.pdf> [accessed 2022-08-31]
30. Harris JL, Fleming-Milici F, Phaneuf L, Jensen M, Choi YY, McCann M, et al. Fast Food Facts 2021: fast food advertising: billions in spending, continued high exposure by youth. UConn Rudd Center for Food Policy and Health. 2021. URL: <https://media.ruddcenter.uconn.edu/PDFs/FACTS2021.pdf> [accessed 2022-08-31]
31. Q4 and fiscal year 2020: letter to shareholders. @TwitterIR. 2021. URL: [https://s22.q4cdn.com/826641620/files/doc\\_financials/2020/q4/FINAL-Q4'20-TWTR-Shareholder-Letter.pdf](https://s22.q4cdn.com/826641620/files/doc_financials/2020/q4/FINAL-Q4'20-TWTR-Shareholder-Letter.pdf) [accessed 2022-08-31]
32. Greene T, Seet C, Rodríguez Barrio A, McIntyre D, Kelly B, Bragg MA. Brands with personalities - good for businesses, but bad for public health? A content analysis of how food and beverage brands personify themselves on Twitter. *Public Health Nutr* 2022 Jan;25(1):51-60 [FREE Full text] [doi: [10.1017/S1368980021001439](https://doi.org/10.1017/S1368980021001439)] [Medline: [33820575](https://pubmed.ncbi.nlm.nih.gov/33820575/)]
33. Harris JL, Fleming-Milici F, Kibwana-Jaff A, Phaneuf L. Sugary drink FACTS 2020: sugary drink advertising to youth: continued barrier to public health progress. *Sugary Drink FACTS*. 2020. URL: [https://www.sugarydrinkfacts.org/resources/Sugary%20Drink%20FACTS%202020/Sugary\\_Drink\\_FACTS\\_Full%20Report\\_final.pdf](https://www.sugarydrinkfacts.org/resources/Sugary%20Drink%20FACTS%202020/Sugary_Drink_FACTS_Full%20Report_final.pdf) [accessed 2022-08-31]
34. Harris JL, Schwartz MB, Brownell KD, Sarda V, Dembek C, Munsell C, et al. Cereal FACTS 2012: limited progress in the nutrition quality and marketing of children's cereals. *Cereal FACTS*. 2012. URL: [http://www.cerealfacts.org/media/cereal\\_facts\\_report\\_2012\\_7.12.pdf](http://www.cerealfacts.org/media/cereal_facts_report_2012_7.12.pdf) [accessed 2022-08-31]
35. Rayner M, Scarborough P, Boxer A, Stockley L. Nutrient profiles: development of final model. *ResearchGate*. 2005. URL: [https://www.researchgate.net/profile/Peter\\_Scarborough/publication/266447771\\_Nutrient\\_profiles\\_Development\\_of\\_Final\\_Model\\_Final\\_Report/links/5440d4fe0cf218719077d82d.pdf](https://www.researchgate.net/profile/Peter_Scarborough/publication/266447771_Nutrient_profiles_Development_of_Final_Model_Final_Report/links/5440d4fe0cf218719077d82d.pdf) [accessed 2022-08-31]
36. Harris JL, Schwartz MB, Brownell KD, Sarda V, Weinberg ME, Speers S, et al. Cereal FACTS: evaluating the nutrition quality and marketing of children's cereals. *Cereal FACTS*. 2009. URL: [https://www.cerealfacts.org/media/Cereal\\_FACTS\\_Report\\_2009.pdf](https://www.cerealfacts.org/media/Cereal_FACTS_Report_2009.pdf) [accessed 2022-08-31]
37. Bragg M, Yanamadala S, Roberto C, Harris J, Brownell K. Athlete endorsements in food marketing. *Pediatrics* 2013 Dec;132(5):805-810. [doi: [10.1542/peds.2013-0093](https://doi.org/10.1542/peds.2013-0093)] [Medline: [24101762](https://pubmed.ncbi.nlm.nih.gov/24101762/)]
38. Gerritsen S, Sing F, Lin K, Martino F, Backholer K, Culpin A, et al. The timing, nature and extent of social media marketing by unhealthy food and drinks brands during the COVID-19 pandemic in New Zealand. *Front Nutr* 2021;8:645349. [doi: [10.3389/fnut.2021.645349](https://doi.org/10.3389/fnut.2021.645349)] [Medline: [33748177](https://pubmed.ncbi.nlm.nih.gov/33748177/)]
39. Antúnez L, Alcaire F, Brunet G, Bove I, Ares G. COVID-washing of ultra-processed products: the content of digital marketing on Facebook during the COVID-19 pandemic in Uruguay. *Public Health Nutr* 2021 Apr;24(5):1142-1152 [FREE Full text] [doi: [10.1017/S1368980021000306](https://doi.org/10.1017/S1368980021000306)] [Medline: [33494846](https://pubmed.ncbi.nlm.nih.gov/33494846/)]
40. Sumagaysay L. The pandemic has more than doubled food-delivery apps' business. Now what? *MarketWatch*. 2020 Nov 25. URL: <https://tinyurl.com/4tvd67wa> [accessed 2021-09-22]
41. UpMenu. URL: <http://upmenu.com> [accessed 2021-09-22]
42. Kelly B, Halford JC, Boyland EJ, Chapman K, Bautista-Castaño I, Berg C, et al. Television food advertising to children: a global perspective. *Am J Public Health* 2010 Oct;100(9):1730-1736. [doi: [10.2105/AJPH.2009.179267](https://doi.org/10.2105/AJPH.2009.179267)] [Medline: [20634464](https://pubmed.ncbi.nlm.nih.gov/20634464/)]
43. Bragg MA, Liu PJ, Roberto CA, Sarda V, Harris JL, Brownell KD. The use of sports references in marketing of food and beverage products in supermarkets. *Public Health Nutr* 2013 May;16(4):738-742. [doi: [10.1017/S1368980012003163](https://doi.org/10.1017/S1368980012003163)] [Medline: [22874497](https://pubmed.ncbi.nlm.nih.gov/22874497/)]

44. Bragg MA, Pageot YK, Amico A, Miller AN, Gasbarre A, Rummo PE, et al. Fast food, beverage, and snack brands on social media in the United States: an examination of marketing techniques utilized in 2000 brand posts. *Pediatr Obes* 2020 May;15(5):e12606. [doi: [10.1111/ijpo.12606](https://doi.org/10.1111/ijpo.12606)] [Medline: [31875654](https://pubmed.ncbi.nlm.nih.gov/31875654/)]
45. Boelsen-Robinson T, Backholer K, Peeters A. Digital marketing of unhealthy foods to Australian children and adolescents. *Health Promot Int* 2016 Oct;31(3):523-533. [doi: [10.1093/heapro/dav008](https://doi.org/10.1093/heapro/dav008)] [Medline: [25769977](https://pubmed.ncbi.nlm.nih.gov/25769977/)]
46. Vandevijvere S, Kelly B, Swinburn B. Food references and marketing in popular magazines for children and adolescents in New Zealand: a content analysis. *Obes Res Clin Pract* 2014 Dec;8:109-181. [doi: [10.1016/j.orcp.2014.10.197](https://doi.org/10.1016/j.orcp.2014.10.197)]
47. Kumar G, Onufrak S, Zytynick D, Kingsley B, Park S. Self-reported advertising exposure to sugar-sweetened beverages among US youth. *Public Health Nutr* 2015 May;18(7):1173-1179. [doi: [10.1017/S1368980014001785](https://doi.org/10.1017/S1368980014001785)] [Medline: [25166512](https://pubmed.ncbi.nlm.nih.gov/25166512/)]
48. Aguinis H, Glavas A. What we know and don't know about corporate social responsibility. *J Manage* 2012 Mar 01;38(4):932-968. [doi: [10.1177/0149206311436079](https://doi.org/10.1177/0149206311436079)]
49. Freeman B, Kelly B, Baur L, Chapman K, Chapman S, Gill T, et al. Digital junk: food and beverage marketing on Facebook. *Am J Public Health* 2014 Dec;104(12):e56-e64. [doi: [10.2105/AJPH.2014.302167](https://doi.org/10.2105/AJPH.2014.302167)] [Medline: [25322294](https://pubmed.ncbi.nlm.nih.gov/25322294/)]
50. Laestadius LI, Wahl MM. Mobilizing social media users to become advertisers: corporate hashtag campaigns as a public health concern. *Digit Health* 2017;3:2055207617710802 [FREE Full text] [doi: [10.1177/2055207617710802](https://doi.org/10.1177/2055207617710802)] [Medline: [29942600](https://pubmed.ncbi.nlm.nih.gov/29942600/)]
51. Dunlop S, Freeman B, Jones S. Marketing to youth in the digital age: the promotion of unhealthy products and health promoting behaviours on social media. *Media Commun* 2016 Jun 16;4(3):35-49. [doi: [10.17645/mac.v4i3.522](https://doi.org/10.17645/mac.v4i3.522)]

## Abbreviations

**FACTS:** Food Advertising to Children and Teens Scores

**NPI:** nutrient profile index

**NPM:** Nutrient Profile Model

*Edited by A Mavragani; submitted 01.03.22; peer-reviewed by F Martino, S Zwier; comments to author 22.06.22; revised version received 19.08.22; accepted 29.08.22; published 14.10.22.*

*Please cite as:*

*Tsai KA, Cassidy O, Arshonsky J, Bond S, Del Giudice IM, Bragg MA*

*COVID-Washing in US Food and Beverage Marketing on Twitter: Content Analysis*

*JMIR Form Res* 2022;6(10):e37642

URL: <https://formative.jmir.org/2022/10/e37642>

doi: [10.2196/37642](https://doi.org/10.2196/37642)

PMID: [36040957](https://pubmed.ncbi.nlm.nih.gov/36040957/)

©Krystle A Tsai, Omni Cassidy, Josh Arshonsky, Sara Bond, Inés M Del Giudice, Marie A Bragg. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 14.10.2022. 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

# Population Health Metrics During the Early Stages of the COVID-19 Pandemic: Correlative Pilot Study

Marie A Severson<sup>1</sup>, MSc; David A Cassada<sup>2</sup>, MSc; Victor C Huber<sup>1</sup>, PhD; Daniel D Snow<sup>2</sup>, PhD; Lisa M McFadden<sup>1</sup>, PhD

<sup>1</sup>Division of Basic Biomedical Sciences, University of South Dakota, Vermillion, SD, United States

<sup>2</sup>Water Sciences Laboratory, University of Nebraska–Lincoln, Lincoln, NE, United States

**Corresponding Author:**

Lisa M McFadden, PhD

Division of Basic Biomedical Sciences

University of South Dakota

414 East Clark Street

Vermillion, SD, 57069

United States

Phone: 1 605 658 6476

Email: [lisa.mcfadden@usd.edu](mailto:lisa.mcfadden@usd.edu)

## Abstract

**Background:** COVID-19 has caused nearly 1 million deaths in the United States, not to mention job losses, business and school closures, stay-at-home orders, and mask mandates. Many people have suffered increased anxiety and depression since the pandemic began. Not only have mental health symptoms become more prevalent, but alcohol consumption has also increased during this time. Helplines offer important insight into both physical and mental wellness of a population by offering immediate, anonymous, cheap, and accessible resources for health and substance use disorders (SUD) that was unobstructed by many of the mandates of the pandemic. Further, the pandemic also launched the use of wastewater surveillance, which has the potential for tracking not only population infections but also consumption of substances such as alcohol.

**Objective:** This study assessed the feasibility of using multiple public surveillance metrics, such as helpline calls, COVID-19 cases, and alcohol metabolites in wastewater, to better understand the need for interventions or public health programs in the time of a public health emergency.

**Methods:** Ethanol metabolites were analyzed from wastewater collected twice weekly from September 29 to December 4, 2020, in a Midwestern state. Calls made to the helpline regarding housing, health care, and mental health/SUD were correlated with ethanol metabolites analyzed from wastewater samples, as well as the number of COVID-19 cases during the sampling period.

**Results:** Correlations were observed between COVID-19 cases and helpline calls regarding housing and health care needs. No correlation was observed between the number of COVID-19 cases and mental health/SUD calls. COVID-19 cases on Tuesdays were correlated with the alcohol metabolite ethyl glucuronide (EtG). Finally, EtG levels were negatively associated with mental health/SUD helpline calls.

**Conclusions:** Although helpline calls provided critical services for health care and housing-related concerns early in the pandemic, evidence suggests helpline calls for mental health/SUD-related concerns were unrelated to COVID-19 metrics. Instead, COVID metrics were associated with alcohol metabolites in wastewater. Although this research was formative, with continued and expanded monitoring of population metrics, such as helpline usage, COVID-19 metrics, and wastewater, strategies can be implemented to create precision programs to address the needs of the population.

(*JMIR Form Res* 2022;6(10):e40215) doi:[10.2196/40215](https://doi.org/10.2196/40215)

**KEYWORDS**

COVID-19; ethyl glucuronide; wastewater; stress; helpline

## Introduction

The SARS-CoV-2 pandemic has caused tremendous stress in the United States. Widespread unemployment and financial strain [1,2], orders to shelter in place [3], health care infrastructure stress [4], and disruptions to education caused significant distress and uncertainty for many people in the United States. National surveys have found that anxiety and depression increased during the pandemic [5]. Early in the pandemic, serious psychological distress was 3.5 times higher in April 2020 compared to 2018 in a nationally representative survey of adults [6]. Similarly, a national survey by the US Census Bureau found that anxiety and depression were 4 times higher in 2020 compared to 2019 [7]. These levels of anxiety and depressive symptoms persisted into the fall of 2020, when COVID-19 cases began to increase. Overall, the negative impacts of COVID-19 on mental health have been well documented [8-11].

Although psychological distress persisted, people changed their activities. Emergency department employees reported decreased exercise and increased alcohol consumption [4]. Increased alcohol consumption was reported in other studies as well. A 51% increase in alcohol sales occurred during the early weeks of the pandemic [12]. Across the United States, beer, wine, and liquor store sales increased by 21.5% in October 2020 compared to the previous 3 years. Elevated alcohol sales continued into November and December 2020 when compared to the average sales of the 3 prior years [13]. Similarly, alcohol consumption increased in 2 independent surveys [8,10]. This increase was greater in females compared to males [12,14]. Urban college students also reported more alcohol consumption following the start of the pandemic [15]. The shift in mental health and the rise in alcohol use since the beginning of the pandemic may highlight the critical need for connecting people with resources for mental health concerns. However, given the stress of the pandemic on the health care infrastructure and subsequent restrictions in face-to-face appointments, population metrics of alcohol use that are less influenced by pandemic-associated restrictions are critically needed. Wastewater analysis of alcohol metabolites provides a community-wide assessment of alcohol use, which does not depend upon restrictions imposed by the pandemic.

Helplines are a critical source of information during times of crisis. These helplines provide immediate, anonymous, cheap, and accessible resources for mental health and substance use disorders (SUD) [16]. Moreover, during the SARS-CoV-2 pandemic, helplines overcame many of the restrictions imposed with traditional face-to-face services, thus allowing them to function unobstructed during these times. Within the United States, few studies have examined the use of helplines during a national emergency, and most have focused on suicidality or depression [17,18]. One study by Brühlhart et al [16] examined helpline calls during the SARS-CoV-2 pandemic in multiple countries, including the United States. Findings revealed a significant increase in call volume 6 weeks after the first major outbreak [16]. Subsequent analysis from France and Germany found increased call volumes related to infections and loneliness but only a small decrease in call volumes related to SUD.

Subsequent waves revealed no significant change in call volumes related to SUD. However, understanding helpline use during the pandemic in the United States remains underinvestigated.

The purpose of this study was to examine the relationships among population health metrics during a public health emergency. Specifically, we compared COVID-19 cases with calls to a regional helpline, including inquiries related to housing, health care, and mental health/SUD needs. We also compared these variables with alcohol metabolites found in wastewater samples during this time to estimate population alcohol use. Findings provide the feasibility of these public health-monitoring systems to improve insights into the specific needs for programs to address population health during a public health emergency.

## Methods

### Ethical Considerations

To reduce the risk of identifying communities, metrics as suggested by Prichard et al [19] were used. Only the broad geographical location (state) was given. These efforts aligned with ethical guidelines for wastewater-based epidemiology set forth by the Sewage Analysis Core group Europe [20] to help minimize risk to participating communities and their citizens. Similarly, given that county-level metrics would provide identifiable cities, only correlations among metrics were included. Helpline calls and COVID-19 case metrics were considered exempt from human subjects' ethical review under the federal regulations for human subjects (45 CFR Part 46) because (1) the data came from sources that are publicly available and (2) the data were deidentified and uncoded and stripped of identifiers [21].

### Wastewater Samples

Twenty-four-hour composite wastewater samples were collected at the inlet of a wastewater treatment plant at approximately 9:00 a.m. on Tuesdays and Fridays, September 29-December 4, 2020. Approximately 500 mL of wastewater was frozen until analysis. Samples were shipped to the Water Sciences Laboratory at the University of Nebraska-Lincoln for analysis of ethyl glucuronide (EtG) and ethyl sulfate (EtS) by multiple reaction monitoring using ultrahigh-pressure liquid chromatography coupled with tandem quadrupole mass spectrometry (UHPLC-MS-MS).

### Analysis of Ethanol Metabolites

Sulfate (EtS) and glucuronide (EtG) metabolites of ethanol were analyzed by isotope dilution using a Xevo TQ-S micro triple quadrupole mass analyzer interfaced with an Acquity UHPLC and UniSpray™ ionization source in a negative-ion-detection-mode spectrometer (Waters Corporation, Milford, MA). An Acquity BEH C18 column (2.1 mm × 50 mm, 1.7 μm) was used for separation at a flow rate of 0.6 mL/minute and a temperature of 40 °C. In addition, 0.1% (v/v) formic acid with 1.0 g/L of ammonium formate in water (solvent A) and a 50/50 (v/v) methanol/acetonitrile (solvent B) binary gradient was used, as listed in Table 1. The injection volume was 2.0 μL. Mass spectrometer settings included the following:



desolvation gas temperature=600 °C, desolvation gas flow=1000 L/hour (N<sub>2</sub>), cone gas flow=50 L/hour (N<sub>2</sub>), impactor voltage=1.0 kV, LM1 resolution=9.2, HM1 resolution=15.1, ion energy 1=0.1 V, LM2 resolution=9.0, HM2 resolution=15.1, and ion energy 2=0.6 V. Retention times and mass scan segments are listed in Table 2. The UHPLC-MS-MS equipment was controlled by MassLynx software (version 4.2 SCN 1017, Waters Corporation). Aqueous standards were prepared by

dilution of analytical standards for deuterium-labeled and nonlabeled EtG and EtS (certified reference material, Cerilliant, Sigma-Aldrich, St Louis, MO, USA) in purified organic-free reagent water over the range of 1.0-200 ng/mL. Instrument detection limits, estimated as 3 times the SD of the lowest calibration standard, were 0.93 and 0.23 for EtG and EtS, respectively. Syringe-filtered (0.2 µm pore size) samples were spiked with internal standards and analyzed without additional preparation.

**Table 1.** Gradient setting for UHPLC-MS-MS<sup>a</sup> analysis at a flow rate of 0.6 mL/minute. All solvent transitions were linear.

Time (minutes)	0.3% (v/v) formic acid/1.0 g/L of ammonium formate/H <sub>2</sub> O	50/50 (v/v) methanol/acetonitrile
0	100%	0%
1.0	100%	0%
2.0	50%	50%
2.5	0%	100%
5.75	0%	100%
6.0	100%	0%
9.0	100%	0%

<sup>a</sup>UHPLC-MS-MS: ultrahigh-pressure liquid chromatography coupled with tandem quadrupole mass spectrometry.

**Table 2.** Retention times and mass scan settings.

Compound	Retention time (minutes)	Precursor ion (amu)	Product ion (amu)	Dwell time (ms)	Fragmentation voltage (V)	Collision energy (V)
EtS <sup>a</sup>	0.34	124.9	96.9	23	40	11
d5-EtS	0.34	129.9	97.9	23	40	11
EtG <sup>b</sup>	0.55	220.9	84.9	23	40	15
d5-EtG	0.54	225.9	84.9	23	40	15

<sup>a</sup>EtS: ethyl sulfate.

<sup>b</sup>EtG: ethyl glucuronide.

## Helpline Call Statistics

The South Dakota Helpline Center is a free resource for those in need of assistance. People may call, text, or email the helpline center about their needs 24 hours a day, 7 days a week. Each contact is classified by the needs expressed during the call, including but not limited to needs for housing, health care, food/meals, and mental health/SUD assistance. Data on daily calls and call categories are posted on the interactive website [22]. To ensure caller privacy, categories of calls below a minimum threshold are suppressed. To ensure that data exceeded the threshold to be categorized, calls were aggregated to align with wastewater sampling (sum of calls from Saturday to Tuesday, sum of calls from Wednesday to Friday). Calls during the sampling period were included in the analysis described.

## COVID-19 Cases

The South Dakota Department of Health reported positive COVID-19 clinical cases daily during the sampling period. Positive cases included only South Dakota residents regardless of where they were tested. Cases were reported by the county of residence. Case counts and 7-day averages were reported daily [23].

## Alcohol Metabolites

Sulfate (EtS) and glucuronide (EtG) metabolites of ethanol were used as a population estimate of alcohol use. Wastewater epidemiology is a well-established metric of human drug consumption [24]. Chen et al [25] have previously shown that wastewater estimates of alcohol consumption were in good agreement with US survey data based on sampling from 3 communities over a 1-year period.

## Statistics

SAS Studio was used for bivariate Pearson 2-tailed correlational analysis. Violations of assumptions were not observed; therefore, parametric analysis was used. Variables (helpline: total calls, health care, housing, and mental health/SUD related; COVID-19: county cases, county 7-day average cases, and state 7-day average cases; alcohol metabolites: EtS and EtG) were first correlated across all dates (N=19) and then by day of the week (Tuesday: n=10, 53%; Friday: n=9, 47%). Significance was set at  $P < .05$ .

## Results

### Correlations Among COVID-19 Cases, Alcohol Metabolites, and Helpline Calls

When all sampling dates were included (Table 3), calls to the helpline were significantly positively correlated with measures of COVID-19 cases, including new county cases ( $t_{17}=2.90$ ,  $P=.01$ ), county 7-day moving average ( $t_{17}=6.87$ ,  $P<.001$ ), and state 7-day moving average ( $t_{17}=9.83$ ,  $P<.001$ ). More specifically, measures of COVID-19 cases were significantly correlated with helpline calls related to health care (county 7-day average:  $t_{17}=3.13$ ,  $P=.01$ ; state 7-day average:  $t_{17}=3.07$ ,  $P=.01$ ) and housing (county 7-day average:  $t_{17}=3.28$ ,  $P=.004$ ; state 7-day average:  $t_{17}=3.56$ ,  $P=.002$ ) but not mental health/SUD (county 7-day average:  $t_{17}=-1.08$ ,  $P=.29$ ; state 7-day average:  $t_{17}=-1.06$ ,  $P=.30$ ). In contrast, helpline calls related to mental health/SUD were significantly negatively correlated with wastewater EtG ( $t_{17}=2.29$ ,  $P=.03$ ) but not EtS ( $t_{17}=0.10$ ,  $P=.92$ ) levels.

When examining measures by day of the week, measures of COVID-19 cases were still significantly correlated with total helpline calls on Tuesdays (county cases:  $t_8=2.70$ ,  $P=.03$ ; county 7-day average:  $t_8=4.70$ ,  $P=.002$ ; state 7-day average:  $t_8=6.73$ ,

$P<.001$ ; Table 4) and Fridays (county 7-day average:  $t_7=4.43$ ,  $P=.003$ ; state 7-day average:  $t_7=7.10$ ,  $P<.001$ ; Table 5) except for new county cases (Friday:  $t_7=1.06$ ,  $P=.33$ ; Table 5). On Fridays, county COVID cases ( $t_7=3.06$ ,  $P=.02$ ), county moving average ( $t_7=5.47$ ,  $P<.001$ ), and state moving average ( $t_7=3.72$ ,  $P=.01$ ) were significantly positively correlated with calls related to health care but not housing (county cases:  $t_7=0.22$ ,  $P=.83$ ; county 7-day average:  $t_7=1.47$ ,  $P=.18$ ; state 7-day average:  $t_7=2.09$ ,  $P=.08$ ) or mental health/SUD (county cases:  $t_7=0.43$ ,  $P=.68$ ; county 7-day average:  $t_7=0.99$ ,  $P=.36$ ; state 7-day average:  $t_7=1.36$ ,  $P=.21$ ). In contrast, on Tuesdays, county and state moving average COVID-19 cases were significantly correlated with helpline calls for housing (county 7-day average:  $t_8=2.85$ ,  $P=.02$ ; state 7-day average:  $t_8=2.71$ ,  $P=.03$ ) but not health care (county 7-day average:  $t_8=1.58$ ,  $P=.15$ ; state 7-day average:  $t_8=1.79$ ,  $P=.11$ ) or mental health/SUD (county 7-day average:  $t_8=0.57$ ,  $P=.58$ ; state 7-day average:  $t_8=0.32$ ,  $P=.76$ ). Further, county cases were significantly positively correlated with EtG ( $t_8=2.63$ ,  $P=.03$ ). EtG was negatively correlated with helpline calls related to mental health/SUD ( $t_8=2.82$ ,  $P=.02$ ) on Tuesdays. Similarly, on Fridays, EtG was negatively correlated with helpline calls related to mental health/SUD but was only marginally associated ( $t_7=-2.13$ ,  $P=.07$ ).

**Table 3.** Correlations among COVID-19 cases, alcohol metabolites, and helpline calls (N=19).

Measures	County cases	County 7-day average	SD 7-day average	EtG <sup>a</sup>	EtS <sup>b</sup>	SD helpline total calls	SD helpline-health care	SD helpline-housing	SD helpline- mental health/SUD <sup>c</sup>
<b>County cases</b>									
r	1.00	0.78 <sup>d</sup>	0.68 <sup>d</sup>	0.03	-0.07	0.58 <sup>d</sup>	0.45	0.34	-0.13
t <sub>17</sub>	N/A <sup>e</sup>	5.07	3.86	0.13	-0.30	2.90	2.07	1.49	-0.53
P value	N/A	<.001	.001	.90	.77	.01	.05	.16	.60
<b>County 7-day average</b>									
r	0.78 <sup>d</sup>	1.00	0.93 <sup>d</sup>	0.24	-0.09	0.86 <sup>d</sup>	0.61 <sup>d</sup>	0.62 <sup>d</sup>	-0.25
t <sub>17</sub>	5.07	N/A	10.38	1.00	-0.35	6.87	3.13	3.28	-1.08
P value	<.001	N/A	<.001	.33	.73	<.001	.01	.004	.29
<b>SD 7-day average</b>									
r	0.68 <sup>d</sup>	0.93 <sup>d</sup>	1.00	0.28	-0.06	0.92 <sup>d</sup>	0.60 <sup>d</sup>	0.65 <sup>d</sup>	-0.25
t <sub>17</sub>	3.86	10.38	N/A	1.20	-0.26	9.83	3.07	3.56	-1.06
P value	.001	<.001	N/A	.25	.80	<.001	.01	.002	.30
<b>EtG</b>									
r	0.03	0.24	0.28	1.00	0.34	0.19	-0.01	0.08	-0.49 <sup>d</sup>
t <sub>17</sub>	0.13	1.00	1.20	N/A	1.51	0.78	-0.05	0.31	-2.29
P value	.90	.33	.25	N/A	.15	.44	.96	.76	.03
<b>EtS</b>									
r	-0.07	-0.09	-0.06	0.34	1.00	-0.05	-0.11	-0.12	0.02
t <sub>17</sub>	-0.30	-0.35	-0.26	1.51	N/A	-0.19	-0.46	-0.49	0.10
P value	.77	.73	.80	.15	N/A	.85	.65	.63	.92
<b>SD helpline total calls</b>									
r	0.58 <sup>d</sup>	0.86 <sup>d</sup>	0.92 <sup>d</sup>	0.19	-0.05	1.00	0.72 <sup>d</sup>	0.67 <sup>d</sup>	-0.21
t <sub>17</sub>	2.90	6.87	9.83	0.78	-0.19	N/A	4.24	3.68	-0.89
P value	.01	<.001	<.001	.44	.85	N/A	<.001	.002	.39
<b>SD helpline-health care</b>									
r	0.45	0.61 <sup>d</sup>	0.60 <sup>d</sup>	-0.01	-0.11	0.72 <sup>d</sup>	1.00	0.16	-0.34
t <sub>17</sub>	2.07	3.13	3.07	-0.05	-0.46	4.24	N/A	0.66	-1.50
P value	.05	.01	.01	.96	.65	<.001	N/A	.52	.15
<b>SD helpline-housing</b>									
r	0.34	0.62 <sup>d</sup>	0.65 <sup>d</sup>	0.08	-0.12	0.67 <sup>d</sup>	0.16	1.00	0.05
t <sub>17</sub>	1.49	3.28	3.56	0.31	-0.49	3.68	0.66	N/A	0.20
P value	.16	.004	.002	.76	.63	.002	.52	N/A	.84
<b>SD helpline-mental health/SUD</b>									
r	-0.13	-0.25	-0.25	-0.49 <sup>d</sup>	0.02	-0.21	-0.34	0.05	1.00
t <sub>17</sub>	-0.53	-1.08	-1.06	-2.29	0.10	-0.89	-1.50	0.20	N/A
P value	.60	.29	.30	.03	.92	.39	.15	.84	N/A

<sup>a</sup>EtG: ethyl glucuronide.<sup>b</sup>EtS: ethyl sulfate.<sup>c</sup>SUD: substance use disorders.

<sup>d</sup> $P < .05$

<sup>e</sup>N/A: not applicable.

**Table 4.** Tuesday correlations among COVID-19 cases, alcohol metabolites, and helpline calls (n=10).

Measures	County cases	County 7-day average	SD 7-day average	EtG <sup>a</sup>	EtS <sup>b</sup>	SD helpline total calls	SD helpline-health care	SD helpline-housing	SD helpline- mental health/SUD <sup>c</sup>
<b>County cases</b>									
r	1.00	0.89 <sup>d</sup>	0.84 <sup>d</sup>	0.68 <sup>d</sup>	0.04	0.69 <sup>d</sup>	0.32	0.54	-0.27
<i>t</i> <sub>8</sub>	N/A <sup>e</sup>	5.39	4.30	2.63	0.11	2.70	0.97	1.83	-0.80
<i>P</i> value	N/A	<.001	.003	.03	.92	.03	.36	.11	.44
<b>County 7-day average</b>									
r	0.89 <sup>d</sup>	1.00	0.91 <sup>d</sup>	0.60	0.03	0.86 <sup>d</sup>	0.49	0.71 <sup>d</sup>	-0.20
<i>t</i> <sub>8</sub>	5.39	N/A	6.02	2.10	0.09	4.70	1.58	2.85	-0.57
<i>P</i> value	<.001	N/A	<.001	.07	.93	.002	.15	.02	.58
<b>SD 7-day average</b>									
r	0.84 <sup>d</sup>	0.91 <sup>d</sup>	1.00	0.42	-0.04	0.92 <sup>d</sup>	0.54	0.69 <sup>d</sup>	-0.11
<i>t</i> <sub>8</sub>	4.30	6.02	N/A	1.30	-0.11	6.73	1.79	2.71	-0.32
<i>P</i> value	.003	<.001	N/A	.23	.92	<.001	.11	.03	.76
<b>EtG</b>									
r	0.68 <sup>d</sup>	0.60	0.42	1.00	-0.03	0.32	0.26	0.12	-0.71 <sup>d</sup>
<i>t</i> <sub>8</sub>	2.63	2.10	1.30	N/A	-0.07	0.95	0.75	0.34	-2.82
<i>P</i> value	.03	.07	.23	N/A	.94	.37	.48	.75	.02
<b>EtS</b>									
r	0.04	0.03	-0.04	-0.03	1.00	0.06	0.22	-0.29	-0.39
<i>t</i> <sub>8</sub>	0.11	0.09	-0.11	-0.07	N/A	0.17	0.63	-0.86	-1.20
<i>P</i> value	.92	.93	.92	.94	N/A	.87	.55	.41	.27
<b>SD helpline total calls</b>									
r	0.69 <sup>d</sup>	0.86 <sup>d</sup>	0.92 <sup>d</sup>	0.32	0.06	1.00	0.74 <sup>d</sup>	0.60	-0.14
<i>t</i> <sub>8</sub>	2.70	4.70	6.73	0.95	0.17	N/A	3.11	2.13	-0.39
<i>P</i> value	.03	.002	<.001	.37	.87	N/A	.01	.07	.71
<b>SD helpline-health care</b>									
r	0.32	0.49	0.54	0.26	0.22	0.74 <sup>d</sup>	1.00	0.14	0.46
<i>t</i> <sub>8</sub>	0.97	1.58	1.79	0.75	0.63	3.11	N/A	0.40	-1.45
<i>P</i> value	.36	.15	.11	.48	.55	.01	N/A	.70	.18
<b>SD helpline-housing</b>									
r	0.54	0.71 <sup>d</sup>	0.69 <sup>d</sup>	0.12	0.29	0.60	0.14	1.00	0.21
<i>t</i> <sub>8</sub>	1.83	2.85	2.71	0.34	-0.86	2.13	0.40	N/A	0.62
<i>P</i> value	.11	.02	.03	.75	.41	.07	.70	N/A	.55
<b>SD helpline-mental health/SUD</b>									
r	0.27	0.20	0.11	0.71 <sup>d</sup>	0.39	0.14	0.46	0.21	1.00
<i>t</i> <sub>8</sub>	-0.80	-0.57	-0.32	-2.82	-1.20	-0.39	-1.45	0.62	N/A
<i>P</i> value	.44	.58	.76	.02	.27	.71	.18	.55	N/A

<sup>a</sup>EtG: ethyl glucuronide.<sup>b</sup>EtS: ethyl sulfate.<sup>c</sup>SUD: substance use disorders.

<sup>d</sup> $P < .05$

<sup>e</sup>N/A: not applicable.

**Table 5.** Friday correlations among COVID-19 cases, alcohol metabolites and helpline calls (n=9).

Measures	County cases	County 7-day average	SD 7-day average	EtG <sup>a</sup>	EtS <sup>b</sup>	SD helpline total calls	SD helpline-health care	SD helpline-housing	SD helpline- mental health/SUD <sup>c</sup>
<b>County cases</b>									
r	1.00	0.60	0.48	0.39	0.02	0.37	0.76 <sup>d</sup>	0.08	0.16
<i>t</i> <sub>7</sub>	N/A <sup>e</sup>	1.99	1.46	-1.11	-0.04	1.06	3.06	-0.22	0.43
<i>P</i> value	N/A	.09	.19	.30	.97	.33	.02	.83	.68
<b>County 7-day average</b>									
r	0.60	1.00	0.97 <sup>d</sup>	0.18	0.15	0.86 <sup>d</sup>	0.90 <sup>d</sup>	0.49	0.35
<i>t</i> <sub>7</sub>	1.99	N/A	10.02	0.50	-0.40	4.43	5.47	1.47	-0.99
<i>P</i> value	.09	N/A	<.001	.63	.70	.003	<.001	.18	.36
<b>SD 7-day average</b>									
r	0.48	0.97 <sup>d</sup>	1.00	0.31	0.19	0.94 <sup>d</sup>	0.82 <sup>d</sup>	0.62	0.46
<i>t</i> <sub>7</sub>	1.46	10.02	N/A	0.87	-0.50	7.10	3.72	2.09	-1.36
<i>P</i> value	.19	<.001	N/A	.41	.63	<.001	.01	.08	.21
<b>EtG</b>									
r	0.39	0.18	0.31	1.00	0.37	0.32	0.06	0.22	0.63
<i>t</i> <sub>7</sub>	-1.11	0.50	0.87	N/A	-1.06	0.90	0.15	0.59	-2.13
<i>P</i> value	.30	.63	.41	N/A	.32	.40	.88	.58	.07
<b>EtS</b>									
r	0.02	0.15	0.19	0.37	1.00	0.01	0.06	0.03	0.33
<i>t</i> <sub>7</sub>	-0.04	-0.40	-0.50	-1.06	N/A	-0.04	-0.15	-0.09	0.94
<i>P</i> value	.97	.70	.63	.32	N/A	.97	.88	.93	.38
<b>SD helpline total calls</b>									
r	0.37	0.86 <sup>d</sup>	0.94 <sup>d</sup>	0.32	0.01	1.00	0.71 <sup>d</sup>	0.77 <sup>d</sup>	0.34
<i>t</i> <sub>7</sub>	1.06	4.43	7.10	0.90	-0.04	N/A	2.67	3.22	-0.97
<i>P</i> value	.33	.003	<.001	.40	.97	N/A	.03	.01	.36
<b>SD helpline-health care</b>									
r	0.76 <sup>d</sup>	0.90 <sup>d</sup>	0.82 <sup>d</sup>	0.06	0.06	0.71 <sup>d</sup>	1.00	0.17	0.10
<i>t</i> <sub>7</sub>	3.06	5.47	3.72	0.15	-0.15	2.67	N/A	0.47	-0.25
<i>P</i> value	.02	<.001	.01	.88	.88	.03	N/A	.65	.81
<b>SD helpline-housing</b>									
r	0.08	0.49	0.62	0.22	0.03	0.77 <sup>d</sup>	0.17	1.00	0.27
<i>t</i> <sub>7</sub>	-0.22	1.47	2.09	0.59	-0.09	3.22	0.47	N/A	-0.75
<i>P</i> value	.83	.18	.08	.58	.93	.01	.65	N/A	.48
<b>SD helpline-mental health/SUD</b>									
r	0.16	0.35	0.46	0.63	0.33	0.34	0.10	0.27	1.00
<i>t</i> <sub>7</sub>	0.43	-0.99	-1.36	-2.13	0.94	-0.97	-0.25	-0.75	N/A
<i>P</i> value	.68	.36	.21	.07	.38	.36	.81	.48	N/A

<sup>a</sup>EtG: ethyl glucuronide.<sup>b</sup>EtS: ethyl sulfate.<sup>c</sup>SUD: substance use disorders.

<sup>d</sup> $P < .05$ <sup>e</sup>N/A: not applicable.

## Discussion

### Principal Findings

This study assessed the feasibility of using multiple public surveillance metrics, such as helpline calls, COVID-19 cases, and wastewater alcohol metabolites, to better understand public health and the need for interventions at the time of a public health emergency. Evidence supports that the helpline provided critical support for people navigating the pandemic. Multiple metrics of clinical COVID-19 cases were positively correlated with total helpline calls and helpline calls regarding health care. Moreover, the helpline was an essential resource to assist people experiencing housing insecurities that may have been associated with economic downturns related to the pandemic. However, COVID-19 cases were not significantly correlated with helpline calls for mental health and SUD resources. Instead, county COVID-19 cases were significantly associated with the alcohol metabolite EtG. Further, on Tuesdays, helpline calls regarding mental health/SUD were negatively correlated with EtG levels. We speculate that these findings suggest that people may have been using alcohol to cope with the stressors associated with the pandemic, especially during the week.

Nationwide, studies have suggested increases in fear, anxiety, and depression since the beginning of the pandemic. According to the US Census Household Pulse Survey 2020-2022, a national average of 40.14% of participants reported symptoms of either anxiety or depression during a similar sampling period [5]. Nationally reported symptoms peaked on November 11-23, 2020, at 42.6%, but persisted until mid-February 2021. During the sampling period of this study, an average of 33.4% of South Dakotan participants reported symptoms of either anxiety or depression, peaking on November 11-23, 2020, at 42.3%. The South Dakota data align with national trends, suggesting widespread mental health concerns.

Extensive research has shown the cyclical relationship between stress and alcohol use. Acute alcohol consumption, through modulation of the neurotransmitter  $\gamma$ -aminobutyric acid (GABA), temporarily alleviates symptoms associated with stress and anxiety [26]. Chronic alcohol use, however, increases negative affect, especially during withdrawal, leading to a preoccupation with the seeking of more alcohol [27]. Despite negative consequences, the cycle of alcohol use may continue into alcohol use disorder [28]. The consequences of this are devastating, leading to serious health complications and death [28]. Developing public health surveillance systems that can monitor problematic alcohol use rapidly enough to develop intervention programs before these serious consequences occur is critical for improving population health.

Finally, alcohol-induced mortality increased by 41% from 2019 to 2020 in the state of South Dakota according to the Centers for Disease Control and Prevention (CDC) [29]. According to the State of South Dakota Department of Health, chronic liver disease and cirrhosis, 1 of the leading causes of death attributed to alcohol misuse, increased by 53% from 2019 to 2020 [30].

Unfortunately, deaths due to chronic liver disease and cirrhosis have more than doubled from 2019 to 2021. In 2019, chronic liver disease and cirrhosis were the tenth-leading cause of death in the state; by 2021, they jumped to the seventh-leading cause of death. Although full statistics detailing alcohol-related deaths are not yet available for 2021, alcohol misuse was associated with increased mortality in the state since the beginning of the pandemic. The establishment of wastewater monitoring for drug and alcohol misuse integrated with other metrics, such as helpline call volumes, may be a critical tool for alerting public health officials of the need for SUD treatments, including connecting people with resources, such as call helplines.

### Comparison With Prior Work

The peak in anxiety or depressive symptoms in South Dakotans also corresponds with the peak number of COVID-19 cases during this period. The number of COVID-19 cases reported to the South Dakota Department of Health during the study period peaked on November 12, 2020, at 1960 cases [23]. Of note, peak EtG levels were found approximately 1 week after the peak in COVID-19 cases. Prior research has found that depression and anxiety levels peak at similar times as COVID-19 cases [7]. Finally, others have found that pandemic-associated distress was also associated with problematic alcohol consumption [31,32]. For example, Gilley et al [33] found that alcohol consumption was associated with COVID-19 positivity. These surveys, combined with wastewater analysis, could provide critical information about stress and other potentially harmful behaviors due to the pandemic or other health crises. The information obtained from wastewater analysis and other metrics could help determine spatial and temporal patterns of alcohol use such that targeted campaigns could be created to connect people most in need of mental health and SUD resources.

### Limitations

The stability of alcohol metabolites in wastewater must be considered. Previous studies suggest that EtS is a more stable biomarker of ethanol consumption than EtG [34,35]. This property may play an important role in our findings and must be considered. Given its lack of stability, EtG was more likely to be excreted during the early morning hours, especially on Tuesdays. This may represent a population with a greater risk of developing problematic alcohol consumption behaviors. In contrast, EtS levels may be more inclusive of all alcohol consumers, including social drinkers as well as those with the potential for problematic drinking.

Indeed, this study found that Tuesday levels of EtG but not EtS were significantly correlated with COVID-19 cases and negatively correlated with mental health-related calls to the helpline. No such correlates between COVID-19 cases and alcohol metabolites occurred on Fridays, when social drinking is more common. Continued research is needed to explore these metabolites with indicators of misuse.

Finally, more complex analysis was not conducted in this study. Given the preliminary nature of the research, complex analysis,



such as time series with lags incorporated, would be premature. Future research is necessary to develop more complex modeling.

### Conclusion

Although this research was formative, findings from this study suggest that helpline calls during the observed period were associated with housing and health care inquires but not mental health/SUD concerns. However, alcohol consumption metrics were negatively associated with mental health/SUD inquires to the helpline. These findings highlight the critical need for using

multiple public surveillance metrics of population health to provide a better understanding of community needs. Tailoring public health announcements around these metrics to direct people to mental health resources would provide a precision medicine approach to promoting positive coping strategies. Finally, although more research is needed, these findings suggest that wastewater surveillance may be helpful in monitoring and impacting population health metrics beyond that of SARS-CoV-2.

### Acknowledgments

This work was supported by General Medical Sciences (grant nos. GM115458 and GM121341) and the National Science Foundation (NSF; DGE-1633213). The content, study design, collection, analysis, interpretation of data, writing of the report, and the decision to submit the paper for publication are solely the responsibility of the authors and do not necessarily represent the views of the National Institutes of Health (NIH) or the NSF.

### Data Availability

The COVID-19 and helpline data analyzed in the study are on the South Dakota Department of Health website and helpline website. Alcohol wastewater data for the study are not publicly available due to community privacy reasons but are available from the corresponding author upon reasonable request.

### Authors' Contributions

MAS performed writing—original draft, visualization, and funding acquisition; DAC, methodology, validation, and investigation; VCH, investigation and writing—review and editing; DDS, methodology, validation, investigation, and writing—review and editing; and LMM, conceptualization, formal analysis, data curation, investigation, writing—original draft, supervision, and funding acquisition. All authors have reviewed and approve of the final paper.

### Conflicts of Interest

None declared.

### References

1. U.S. Bureau of Labor Statistics. Labor Force Statistics from the Current Population Survey: Supplemental Data Measuring the Effects of the Coronavirus (COVID-19) Pandemic on the Labor Market. URL: <https://www.bls.gov/cps/effects-of-the-coronavirus-covid-19-pandemic.htm> [accessed 2022-10-11]
2. Lee J, Kapteyn A, Clomax A, Jin H. Estimating influences of unemployment and underemployment on mental health during the COVID-19 pandemic: who suffers the most? *Public Health* 2021 Dec;201:48-54 [FREE Full text] [doi: [10.1016/j.puhe.2021.09.038](https://doi.org/10.1016/j.puhe.2021.09.038)] [Medline: [34781158](https://pubmed.ncbi.nlm.nih.gov/34781158/)]
3. Moreland A, Herlihy C, Tynan M, Sunshine G, McCord R, Hilton C, CDC Public Health Law Program, CDC COVID-19 Response Team, Mitigation Policy Analysis Unit. Timing of state and territorial covid-19 stay-at-home orders and changes in population movement - United States, March 1-May 31, 2020. *MMWR Morb Mortal Wkly Rep* 2020 Sep 04;69(35):1198-1203 [FREE Full text] [doi: [10.15585/mmwr.mm6935a2](https://doi.org/10.15585/mmwr.mm6935a2)] [Medline: [32881851](https://pubmed.ncbi.nlm.nih.gov/32881851/)]
4. Dehon E, Zachrisson K, Peltzer-Jones J, Tabatabai R, Clair E, Puskarich M, et al. Sources of distress and coping strategies among emergency physicians during covid-19. *West J Emerg Med* 2021 Oct 27;22(6):1240-1252 [FREE Full text] [doi: [10.5811/westjem.2021.9.53406](https://doi.org/10.5811/westjem.2021.9.53406)] [Medline: [34787546](https://pubmed.ncbi.nlm.nih.gov/34787546/)]
5. Centers for Disease Control and Prevention (CDC). Household Pulse Survey. URL: <https://www.cdc.gov/nchs/covid19/pulse/mental-health.htm> [accessed 2022-03-31]
6. McGinty E, Presskreischer R, Han H, Barry C. Psychological distress and loneliness reported by US Adults in 2018 and April 2020. *JAMA* 2020 Jul 07;324(1):93-94 [FREE Full text] [doi: [10.1001/jama.2020.9740](https://doi.org/10.1001/jama.2020.9740)] [Medline: [32492088](https://pubmed.ncbi.nlm.nih.gov/32492088/)]
7. Twenge J, McAllister C, Joiner T. Anxiety and depressive symptoms in U.S. Census Bureau assessments of adults: trends from 2019 to fall 2020 across demographic groups. *J Anxiety Disord* 2021 Oct;83:102455 [FREE Full text] [doi: [10.1016/j.janxdis.2021.102455](https://doi.org/10.1016/j.janxdis.2021.102455)] [Medline: [34332230](https://pubmed.ncbi.nlm.nih.gov/34332230/)]
8. Johnson R, Marker K, Mayer D, Shortt J, Kao D, Barnes K, et al. COVID-19 surveillance in the biobank at the Colorado Center for Personalized Medicine: observational study. *JMIR Public Health Surveill* 2022 Jun 13;8(6):e37327 [FREE Full text] [doi: [10.2196/37327](https://doi.org/10.2196/37327)] [Medline: [35486493](https://pubmed.ncbi.nlm.nih.gov/35486493/)]

9. Haucke M, Heinz A, Liu S, Heinzl S. The impact of covid-19 lockdown on daily activities, cognitions, and stress in a lonely and distressed population: temporal dynamic network analysis. *J Med Internet Res* 2022 Mar 17;24(3):e32598 [FREE Full text] [doi: [10.2196/32598](https://doi.org/10.2196/32598)] [Medline: [35191843](https://pubmed.ncbi.nlm.nih.gov/35191843/)]
10. Šorgo A, Crnkovič N, Gabrovec B, Cesar K, Selak Š. Influence of forced online distance education during the covid-19 pandemic on the perceived stress of postsecondary students: cross-sectional study. *J Med Internet Res* 2022 Mar 15;24(3):e30778 [FREE Full text] [doi: [10.2196/30778](https://doi.org/10.2196/30778)] [Medline: [35171098](https://pubmed.ncbi.nlm.nih.gov/35171098/)]
11. El Gindi H, Shalaby R, Gusnowski A, Vuong W, Suroid S, Hrabok M, et al. The mental health impact of the covid-19 pandemic among physicians, nurses, and other health care providers in Alberta: cross-sectional survey. *JMIR Form Res* 2022 Mar 09;6(3):e27469 [FREE Full text] [doi: [10.2196/27469](https://doi.org/10.2196/27469)] [Medline: [34995203](https://pubmed.ncbi.nlm.nih.gov/34995203/)]
12. Pollard M, Tucker J, Green HJ. Changes in adult alcohol use and consequences during the covid-19 pandemic in the US. *JAMA Netw Open* 2020 Sep 01;3(9):e2022942 [FREE Full text] [doi: [10.1001/jamanetworkopen.2020.22942](https://doi.org/10.1001/jamanetworkopen.2020.22942)] [Medline: [32990735](https://pubmed.ncbi.nlm.nih.gov/32990735/)]
13. Federal Reserve Bank of St. Louis. Retail Sales: Beer, Wine, and Liquor Stores (MRTSSM4453USN). URL: <https://fred.stlouisfed.org/series/MRTSSM4453USN> [accessed 2022-10-11]
14. Rodriguez L, Litt D, Stewart S. Drinking to cope with the pandemic: the unique associations of COVID-19-related perceived threat and psychological distress to drinking behaviors in American men and women. *Addict Behav* 2020 Nov;110:106532 [FREE Full text] [doi: [10.1016/j.addbeh.2020.106532](https://doi.org/10.1016/j.addbeh.2020.106532)] [Medline: [32652385](https://pubmed.ncbi.nlm.nih.gov/32652385/)]
15. Lechner W, Laurene K, Patel S, Anderson M, Grega C, Kenne D. Changes in alcohol use as a function of psychological distress and social support following COVID-19 related university closings. *Addict Behav* 2020 Nov;110:106527 [FREE Full text] [doi: [10.1016/j.addbeh.2020.106527](https://doi.org/10.1016/j.addbeh.2020.106527)] [Medline: [32679435](https://pubmed.ncbi.nlm.nih.gov/32679435/)]
16. Brillhart M, Klotzbücher V, Lalive R, Reich S. Mental health concerns during the COVID-19 pandemic as revealed by helpline calls. *Nature* 2021 Dec;600(7887):121-126 [FREE Full text] [doi: [10.1038/s41586-021-04099-6](https://doi.org/10.1038/s41586-021-04099-6)] [Medline: [34789873](https://pubmed.ncbi.nlm.nih.gov/34789873/)]
17. Gould M, Lake A, Galfalvy H, Kleinman M, Munfakh J, Wright J, et al. Follow-up with callers to the National Suicide Prevention Lifeline: evaluation of callers' perceptions of care. *Suicide Life Threat Behav* 2018 Feb;48(1):75-86 [FREE Full text] [doi: [10.1111/sltb.12339](https://doi.org/10.1111/sltb.12339)] [Medline: [28261860](https://pubmed.ncbi.nlm.nih.gov/28261860/)]
18. Scerri J, Sammut A, Cilia Vincenti S, Grech P, Galea M, Scerri C, et al. Reaching out for help: calls to a mental health helpline prior to and during the covid-19 pandemic. *Int J Environ Res Public Health* 2021 Apr 23;18(9):4505 [FREE Full text] [doi: [10.3390/ijerph18094505](https://doi.org/10.3390/ijerph18094505)] [Medline: [33922749](https://pubmed.ncbi.nlm.nih.gov/33922749/)]
19. Prichard J, Hall W, de Voogt P, Zuccato E. Sewage epidemiology and illicit drug research: the development of ethical research guidelines. *Sci Total Environ* 2014 Feb 15;472:550-555 [FREE Full text] [doi: [10.1016/j.scitotenv.2013.11.039](https://doi.org/10.1016/j.scitotenv.2013.11.039)] [Medline: [24317162](https://pubmed.ncbi.nlm.nih.gov/24317162/)]
20. Sewage Analysis Core group Europe (SCORE). Ethical Research Guidelines for Sewage Epidemiology. URL: [https://www.emcdda.europa.eu/drugs-library/ethical-research-guidelines-wastewater-based-epidemiology-and-related-fields\\_en](https://www.emcdda.europa.eu/drugs-library/ethical-research-guidelines-wastewater-based-epidemiology-and-related-fields_en) [accessed 2020-07-20]
21. University of Maryland. Use of Publicly Available OR Identifiable Private Sources of Information. URL: <https://tinyurl.com/4254f398> [accessed 2022-09-20]
22. Helpline Center. Overview. URL: <https://www.helplinecenter.org/2-1-1/data/> [accessed 2022-03-31]
23. South Dakota Department of Health (SDDOH). South Dakota Covid-19 Dashboard. URL: <https://doh.sd.gov/COVID/Dashboard.aspx> [accessed 2022-10-11]
24. Reinstadler V, Ausweger V, Grabher A, Kreidl M, Huber S, Grandner J, et al. Monitoring drug consumption in Innsbruck during coronavirus disease 2019 (COVID-19) lockdown by wastewater analysis. *Sci Total Environ* 2021 Feb 25;757:144006 [FREE Full text] [doi: [10.1016/j.scitotenv.2020.144006](https://doi.org/10.1016/j.scitotenv.2020.144006)] [Medline: [33310574](https://pubmed.ncbi.nlm.nih.gov/33310574/)]
25. Chen J, Venkatesan A, Halden R. Alcohol and nicotine consumption trends in three U.S. communities determined by wastewater-based epidemiology. *Sci Total Environ* 2019 Mar 15;656:174-183 [FREE Full text] [doi: [10.1016/j.scitotenv.2018.11.350](https://doi.org/10.1016/j.scitotenv.2018.11.350)] [Medline: [30504019](https://pubmed.ncbi.nlm.nih.gov/30504019/)]
26. Olsen R, Liang J. Role of GABA receptors in alcohol use disorders suggested by chronic intermittent ethanol (CIE) rodent model. *Mol Brain* 2017 Sep 20;10(1):45 [FREE Full text] [doi: [10.1186/s13041-017-0325-8](https://doi.org/10.1186/s13041-017-0325-8)] [Medline: [28931433](https://pubmed.ncbi.nlm.nih.gov/28931433/)]
27. Koob G. Drug addiction: hyperkatifeia/negative reinforcement as a framework for medications development. *Pharmacol Rev* 2021 Jan;73(1):163-201 [FREE Full text] [doi: [10.1124/pharmrev.120.000083](https://doi.org/10.1124/pharmrev.120.000083)] [Medline: [33318153](https://pubmed.ncbi.nlm.nih.gov/33318153/)]
28. Yang W, Singla R, Maheshwari O, Fontaine C, Gil-Mohapel J. Alcohol use disorder: neurobiology and therapeutics. *Biomedicines* 2022 May 21;10(5):1192 [FREE Full text] [doi: [10.3390/biomedicines10051192](https://doi.org/10.3390/biomedicines10051192)] [Medline: [35625928](https://pubmed.ncbi.nlm.nih.gov/35625928/)]
29. Centers for Disease Control and Prevention (CDC). About Underlying Cause of Death, 1999-2020. URL: <https://wonder.cdc.gov/ucd-icd10.html> [accessed 2022-10-11]
30. South Dakota Department of Health, Office of Health Statistics. South Dakota Mortality Report, 2021 (Final). URL: <https://doh.sd.gov/documents/statistics/Mortality-Report-2022-Provisional.pdf> [accessed 2022-09-21]
31. Martinez P, Karriker-Jaffe K, Ye Y, Patterson D, Greenfield T, Mulia N, et al. Mental health and drinking to cope in the early COVID period: data from the 2019-2020 US National Alcohol Survey. *Addict Behav* 2022 May;128:107247 [FREE Full text] [doi: [10.1016/j.addbeh.2022.107247](https://doi.org/10.1016/j.addbeh.2022.107247)]

32. McPhee M, Keough M, Rundle S, Heath L, Wardell J, Hendershot C. Depression, environmental reward, coping motives and alcohol consumption during the covid-19 pandemic. *Front Psychiatry* 2020;11:574676 [FREE Full text] [doi: [10.3389/fpsy.2020.574676](https://doi.org/10.3389/fpsy.2020.574676)] [Medline: [33192708](https://pubmed.ncbi.nlm.nih.gov/33192708/)]
33. Gilley K, Baroudi L, Yu M, Gainsburg I, Reddy N, Bradley C, et al. Risk factors for covid-19 in college students identified by physical, mental, and social health reported during the fall 2020 semester: observational study using the Roadmap app and Fitbit wearable sensors. *JMIR Ment Health* 2022 Feb 10;9(2):e34645 [FREE Full text] [doi: [10.2196/34645](https://doi.org/10.2196/34645)] [Medline: [34992051](https://pubmed.ncbi.nlm.nih.gov/34992051/)]
34. Banks A, Lai F, Mueller J, Jiang G, Carter S, Thai P. Potential impact of the sewer system on the applicability of alcohol and tobacco biomarkers in wastewater-based epidemiology. *Drug Test Anal* 2018 Mar;10(3):530-538 [FREE Full text] [doi: [10.1002/dta.2246](https://doi.org/10.1002/dta.2246)] [Medline: [28688172](https://pubmed.ncbi.nlm.nih.gov/28688172/)]
35. Gao J, Li J, Jiang G, Yuan Z, Eaglesham G, Covaci A, et al. Stability of alcohol and tobacco consumption biomarkers in a real rising main sewer. *Water Res* 2018 Jul 01;138:19-26 [FREE Full text] [doi: [10.1016/j.watres.2018.03.036](https://doi.org/10.1016/j.watres.2018.03.036)] [Medline: [29571085](https://pubmed.ncbi.nlm.nih.gov/29571085/)]

## Abbreviations

**EtG:** ethanol glucuronide

**EtS:** ethanol sulfate

**SUD:** substance use disorders

**UHPLC-MS-MS:** ultrahigh-pressure liquid chromatography coupled with tandem quadrupole mass spectrometry

*Edited by A Mavragani; submitted 10.06.22; peer-reviewed by D Goebert, J Huckins; comments to author 08.09.22; revised version received 28.09.22; accepted 03.10.22; published 17.10.22.*

*Please cite as:*

*Severson MA, Cassada DA, Huber VC, Snow DD, McFadden LM*

*Population Health Metrics During the Early Stages of the COVID-19 Pandemic: Correlative Pilot Study*

*JMIR Form Res* 2022;6(10):e40215

URL: <https://formative.jmir.org/2022/10/e40215>

doi: [10.2196/40215](https://doi.org/10.2196/40215)

PMID: [36219745](https://pubmed.ncbi.nlm.nih.gov/36219745/)

©Marie A Severson, David A Cassada, Victor C Huber, Daniel D Snow, Lisa M McFadden. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 17.10.2022. 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 At-Home Serial Testing Using Over-the-Counter SARS-CoV-2 Tests With a Digital Smartphone App for Assistance: Longitudinal Cohort Study

Carly Herbert<sup>1</sup>, BA; John Broach<sup>2,3</sup>, MPH, MD; William Heetderks<sup>4</sup>, MD, PhD; Felicia Qashu<sup>5</sup>, PhD; Laura Gibson<sup>6</sup>, MD; Caitlin Pretz<sup>1</sup>, MSc; Kelsey Woods<sup>1</sup>, BA; Vik Kheterpal<sup>7</sup>, MD; Thejas Suvarna<sup>7</sup>, BBA, BSE; Christopher Nowak<sup>7</sup>, BA; Peter Lazar<sup>8</sup>, BA; Didem Ayturk<sup>8</sup>, MS; Bruce Barton<sup>8</sup>, PhD; Chad Achenbach<sup>9</sup>, MPH, MD; Robert Murphy<sup>9</sup>, MD; David McManus<sup>1,10</sup>, MD; Apurv Soni<sup>1,8</sup>, MD, PhD

<sup>1</sup>Program in Digital Medicine, Department of Medicine, University of Massachusetts Chan Medical School, Worcester, MA, United States

<sup>2</sup>Department of Emergency Medicine, University of Massachusetts Chan Medical School, Worcester, MA, United States

<sup>3</sup>UMass Memorial Medical Center, Worcester, MA, United States

<sup>4</sup>National Institute of Biomedical Imaging and Bioengineering, Bethesda, MD, United States

<sup>5</sup>National Institutes of Health, Bethesda, MD, United States

<sup>6</sup>Division of Infectious Disease and Immunology, Departments of Medicine and Pediatrics, University of Massachusetts Chan Medical School, Worcester, MA, United States

<sup>7</sup>CareEvolution, LLC, Ann Arbor, MI, United States

<sup>8</sup>Department of Population and Quantitative Health Sciences, University of Massachusetts Chan Medical School, Worcester, MA, United States

<sup>9</sup>Division of Infectious Disease, Department of Medicine, Feinberg School of Medicine, Northwestern University, Chicago, IL, United States

<sup>10</sup>Department of Cardiology, University of Massachusetts Chan Medical School, Worcester, MA, United States

**Corresponding Author:**

Carly Herbert, BA

Program in Digital Medicine

Department of Medicine

University of Massachusetts Chan Medical School

55 N Lake Ave

Worcester, MA, 01655

United States

Phone: 1 4047180676

Email: [carly.herbert@umassmed.edu](mailto:carly.herbert@umassmed.edu)

## Abstract

**Background:** The ongoing SARS-CoV-2 pandemic necessitates the development of accurate, rapid, and affordable diagnostics to help curb disease transmission, morbidity, and mortality. Rapid antigen tests are important tools for scaling up testing for SARS-CoV-2; however, little is known about individuals' use of rapid antigen tests at home and how to facilitate the user experience.

**Objective:** This study aimed to describe the feasibility and acceptability of serial self-testing with rapid antigen tests for SARS-CoV-2, including need for assistance and the reliability of self-interpretation.

**Methods:** A total of 206 adults in the United States with smartphones were enrolled in this single-arm feasibility study in February and March 2021. All participants were asked to self-test for COVID-19 at home using rapid antigen tests daily for 14 days and use a smartphone app for testing assistance and to report their results. The main outcomes were adherence to the testing schedule, the acceptability of testing and smartphone app experiences, and the reliability of participants versus study team's interpretation of test results. Descriptive statistics were used to report the acceptability, adherence, overall rating, and experience of using the at-home test and MyDataHelps app. The usability, acceptability, adherence, and quality of at-home testing were analyzed across different sociodemographic, age, and educational attainment groups.

**Results:** Of the 206 enrolled participants, 189 (91.7%) and 159 (77.2%) completed testing and follow-up surveys, respectively. In total, 51.3% (97/189) of study participants were women, the average age was 40.7 years, 34.4% (65/189) were non-White, and 82% (155/189) had a bachelor's degree or higher. Most (n=133/206, 64.6%) participants showed high testing adherence,

meaning they completed over 75% of the assigned tests. Participants' interpretations of test results demonstrated high agreement (2106/2130, 98.9%) with the study verified results, with a  $\kappa$  score of 0.29 ( $P < .001$ ). Participants reported high satisfaction with self-testing and the smartphone app, with 98.7% (157/159) reporting that they would recommend the self-test and smartphone app to others. These results were consistent across age, race/ethnicity, and gender.

**Conclusions:** Participants' high adherence to the recommended testing schedule, significant reliability between participants and study staff's test interpretation, and the acceptability of the smartphone app and self-test indicate that self-tests for SARS-CoV-2 with a smartphone app for assistance and reporting is a highly feasible testing modality among a diverse population of adults in the United States.

(*JMIR Form Res* 2022;6(10):e35426) doi:[10.2196/35426](https://doi.org/10.2196/35426)

## KEYWORDS

COVID-19; SARS-CoV-2; rapid tests; MyDataHelps smartphone app; mHealth; mobile health; serial self-testing; digital health; pandemic; self test

## Introduction

Since the emergence of the SARS-CoV-2 pandemic in late 2019, more than 590 million cases and 6.5 million deaths from COVID-19 have been reported worldwide [1]. Over 2 years into the pandemic, the United States continues to face waves of increasing SARS-CoV-2 cases. The ongoing pandemic necessitates the development of accurate, rapid, and affordable diagnostics to help curb SARS-CoV-2 disease transmission, morbidity, and mortality, as well as safely navigate social re-engagement [2].

Antigen-based rapid diagnostic tests (Ag-RDTs) detect SARS-CoV-2 viral proteins in multiple specimen types and facilitate opportunities for large-scale, cost-effective testing solutions [3,4]. Ag-RDTs are preferred by over two-thirds of SARS-CoV-2 test users, especially in comparison to molecular tests, which can take days to receive the result [5,6]. Numerous Ag-RDTs for SARS-CoV-2 have received Emergency Use Authorization by the US Food and Drug Administration for point-of-care testing in the health care setting and, more recently, for at-home use, with evidence consistently showing the validity of self-collected specimens for SARS-CoV-2 testing [7,8]. Self-testing at home offers great opportunity for scaling up and implementing regular testing of both asymptomatic and symptomatic individuals, a key step toward controlling the COVID-19 pandemic [2,9-11]. Furthermore, self-testing offers the opportunity to increase testing access across geographic, sociodemographic, and socioeconomic groups to improve health outcomes and reduce health care disparities [12-14]. However, little is known about individuals' use of rapid antigen tests at home and how to facilitate the user experience [5,15].

The objectives of this study were to describe the feasibility of the at-home use of rapid antigen tests for SARS-CoV-2, as well as participants' use of the MyDataHelps smartphone app (CareEvolution) to support at-home testing. This study aimed to describe the usability and acceptability of self-tests for SARS-CoV-2, variation in use across different sociodemographic and socioeconomic groups, and how participants interact with the MyDataHelps smartphone app to report symptoms and test results. We hypothesized that the acceptability and usability of the rapid antigen tests and smartphone app would be consistent across sociodemographic and socioeconomic groups.

## Methods

### Study Population

Participants were recruited from the University of Massachusetts (UMass) Chan Medical School and Northwestern University using best practices developed by the RADx Tech Community Health Equity and Engagement Team to maximize the representation of diverse age, sex, race, ethnicity, education, and socioeconomic groups [16]. Participants were enrolled in the study during February and March 2021. For inclusion in the study, individuals were required to be aged  $\geq 18$  years, be willing to use their own smartphone device and download the MyDataHelps app, have reported no symptoms attributable to COVID-19 within 48 hours prior to screening, and be proficient in English.

### Ethics Approval

Details of the study procedures were explained to participants, and written informed consent was obtained from all participants. This study was approved by the Institutional Review Board (H00022342) at UMass Chan Medical School and Northwestern University, which had a reliance agreement with the UMass Chan Medical School Institutional Review Board.

### Study Procedures

All participants were asked to self-test for COVID-19 at home daily over a consecutive 14-day period. Participants were mailed a QuickVue test kit (Quidel) containing testing supplies for 25 tests; written testing instructions; and a prepaid, pre-addressed return box for test strips with return instructions. All test kits used anterior nasal swabs, and instructions directed participants on how to properly swab their nasal cavity. Participants were given access to the MyDataHelps app to support testing. The MyDataHelps app allowed participants to view testing instructions, report test results, verify test results with the study team, track their testing history, respond to surveys, and access the study team's contact information (Multimedia Appendix 1).

Participants were informed that they could report their test results to the study team either digitally through the MyDataHelps app or through manual written recording. If participants opted to use the MyDataHelps app for reporting test results, they were asked to report their interpretation of the test results—positive or negative—and upload an image of the

test strip ([Multimedia Appendix 1](#)). Study coordinators validated all test results using the test strip image, with digital verification occurring within 24 hours of reporting. Written test results were mailed to study coordinators along with all test strips for verification at the end of the study. Participants were instructed to contact study coordinators with any questions during the study period. All interactions between study coordinators and participants were recorded in a contact log.

If a participant showed COVID-19 symptoms, reported close exposure to a person positive for SARS-CoV-2, or tested positive on a home test, the study team contacted the participant and scheduled confirmatory SARS-CoV-2 polymerase chain reaction testing by trained personnel using established protocols and procedures. If an individual tested positive for SARS-CoV-2 on confirmatory testing, participants were removed from the study and received an exit survey on the day of the positive results out of safety for the participants.

### Study Questionnaires

Questionnaires were administered to participants through the MyDataHelps app. Eligible and consenting study participants were given a baseline survey, daily surveys, a midpoint survey, and an exit survey. The baseline survey assessed their sociodemographic characteristics, anthropometrics, health status, and social engagement. Health status included questions regarding disability, pregnancy, current alcohol and cigarette use, chronic conditions, and report of common COVID-19 symptoms. Daily surveys asked for self-interpretation of the test results (positive or negative), image upload of the test strip, and symptom reporting. The midpoint survey was given to participants on day 7, and participants were asked to self-report the total number of completed tests and their social engagement practices. Lastly, the exit survey, on day 14 of the study, asked for a self-report of the total number of tests completed, acceptability and experience of using the MyDataHelps app, acceptability of the at-home test, social engagement, insurance status (no insurance, private, or public insurance), and health status. The acceptability of the at-home test was assessed by asking participants if they would recommend the self-test to someone else using the Net Promotor Scale. The number of tests reported and number of daily image uploads over the 14-day testing period determined adherence to the testing schedule. Adherence to the testing schedule was classified into

4 categories: no (0%) adherence, low (<50%) adherence, moderate (50% to 75%) adherence, and high (>75%) adherence. Using a Likert scale, the experience and acceptability of the MyDataHelps app was assessed by asking for the participants' overall rating of the app (1=one of the worst apps I've used to 5=one of the best apps I've used). Participants were asked how often they had difficulties using the app (1=all the time to 5=never), the usefulness of different features of the app (1=really useful to 5=really useless), whether they would recommend the app to another person to help them perform an at-home test for COVID-19 (1=definitely yes to 5=definitely not), and if they would continue using the app to keep testing themselves at home for COVID-19 (1=definitely yes to 5=definitely not).

### Data Analysis

Descriptive statistics were used to report the acceptability, adherence, overall rating, and experience of using the at-home test and MyDataHelps app. The usability, acceptability, adherence, and quality of at-home testing were analyzed across different sociodemographic, age, and educational attainment groups to evaluate the impact on existing socioeconomic disparities, using ANOVA to evaluate significance. The study coordinator contact log was text mined and used to generate a word cloud in R statistical software (version 4.2.1; R Foundation for Statistical Computing) to characterize participant interactions with study staff.

## Results

### Participant Characteristics

A total of 206 participants enrolled in the study during February and March 2021. There were 5 (5/206, 2.4%) participants who tested positive for SARS-CoV-2 during the study period, and they were removed from the study. Of the 206 participants, 189 (91.7%) and 159 (77.2%) completed testing and follow-up surveys, respectively. Among participants who completed testing, slightly more than half (97/189, 51.3%) were women, the average age of the study population was 40.7 years, 34.4% (65/189) were non-White, and 82% (155/189) had a bachelor's degree or higher ([Table 1](#)). At the time of study enrollment (February 2021), only 2.5% of the US population were fully vaccinated for SARS-CoV-2 [[17](#)].

**Table 1.** Participant characteristics stratified by testing adherence.

Characteristic	Adherence to daily testing		
	Low (n=19), n (%)	Moderate (n=37), n (%)	High (n=133), n (%)
<b>Age (years)</b>			
18-39	11 (57.9)	24 (64.9)	62 (46.6)
40-64	3 (15.8)	13 (35.1)	49 (36.8)
≥65	2 (10.5)	0 (0)	18 (13.5)
No answer	3 (15.8)	0 (0)	4 (3)
<b>Gender</b>			
Male	8 (42.1)	12 (32.4)	62 (46.6)
Female	8 (42.1)	24 (64.9)	65 (48.9)
Nonbinary or transgender	1 (5.3)	1 (2.7)	3 (2.3)
No answer	2 (10.5)	0 (0)	3 (2.3)
<b>Race/ethnicity</b>			
Hispanic	2 (10.5)	8 (21.6)	12 (9)
Non-Hispanic Asian	1 (5.3)	5 (13.5)	13 (9.8)
Non-Hispanic Black	2 (10.5)	4 (10.8)	15 (11.3)
Non-Hispanic Other	1 (5.3)	0 (0)	2 (1.5)
Non-Hispanic White	11 (57.9)	20 (54.1)	89 (66.9)
No answer	2 (10.5)	0 (0)	2 (1.5)
<b>Education level</b>			
Master's degree or higher	7 (36.8)	7 (18.9)	34 (25.6)
Bachelor's degree or equivalent	6 (31.6)	24 (64.9)	77 (57.9)
High school or lower	4 (21.1)	5 (13.5)	19 (14.3)
No answer	2 (10.5)	1 (2.7)	3 (2.3)
<b>Employment status</b>			
Working now	10 (52.6)	28 (75.7)	94 (70.7)
Student	2 (10.5)	5 (13.5)	10 (7.5)
Retired	3 (15.8)	1 (2.7)	19 (14.3)
Other	4 (21.1)	3 (8.1)	10 (7.5)

### Patient-Reported Usability and Acceptability of At-Home Testing

In all, 91.7% (189/206) of the participants performed 1 or more tests during the study period (Table 2). The majority (133/206, 64.6%) of the participants showed high adherence to testing and picture upload, characterized as testing and uploading the picture of the test strip to the app on more than 75% of the indicated days (Table 1). Participants aged 18-39 years comprised the majority of the moderate (24/37, 64.9%) and low (11/19, 57.9%) adherence groups, whereas 90% (18/20) of the

participants aged ≥65 years reported high adherence ( $P=.03$ ; Table 1). Comparatively, only 63.9% (62/97) and 75.4% (49/65) of participants aged 18-39 years and 40-64 years demonstrated high adherence, respectively. Participants' interpretations of test results demonstrated high agreement (2106/2130, 98.9%) with the study verified results, with a  $\kappa$  score of 0.29 ( $P<.001$ ; Table 3). Overall, participants reported high satisfaction with at-home testing, with 98.7% (157/159) of the participants reporting that they would definitely or likely recommend the self-test to others (Figure 1).

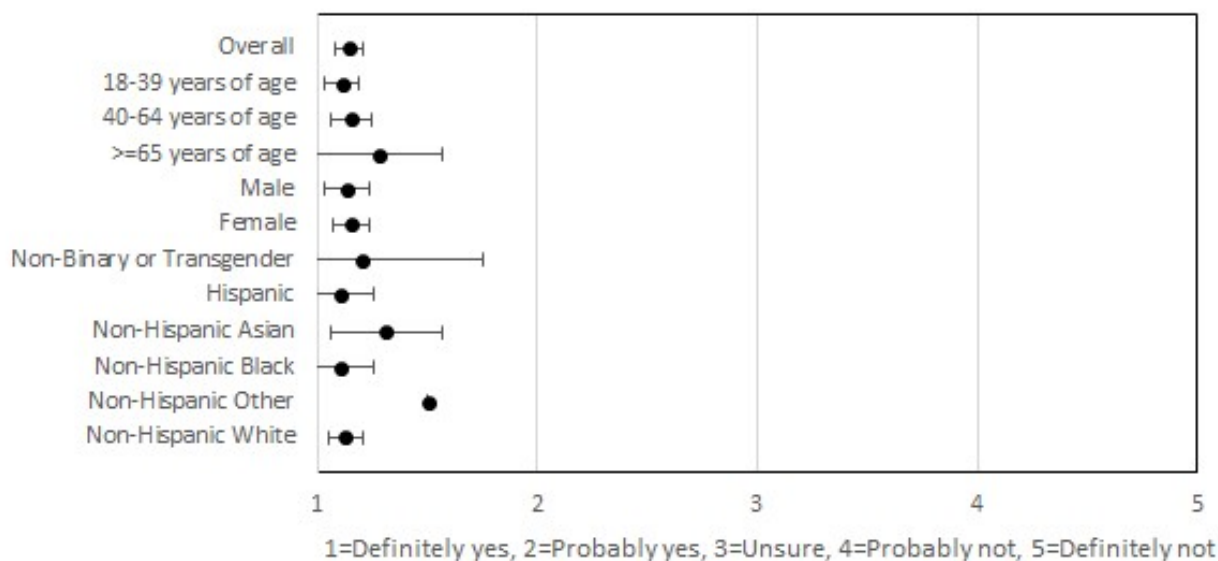
**Table 2.** Total number of tests performed in the 14-day study period.

Total number tests performed in the 14-day study period	Participants (N=206), n (%)
0	17 (8.3)
1	2 (1)
2	3 (1.5)
3	2 (1)
4	2 (1)
5	2 (1)
6	2 (1)
7	6 (2.9)
8	7 (3.4)
9	9 (4.4)
10	22 (10.7)
11	24 (11.7)
12	32 (15.5)
13	35 (17)
14	41 (19.9)

**Table 3.** Reliability of self-interpretation versus study verification of at-home antigen-based rapid diagnostic tests.

Self-interpretation	Study verification			Total
	Negative	Positive	Invalid	
Negative	2102	5	11	2118
Positive	4	4	2	10
Invalid	2	0	0	2
Total	2108	9	13	2130

**Figure 1.** Usability and acceptability of self-tests.



**MyDataHelps Participant Usability**

Participants also reported high satisfaction with the MyDataHelps app. In all, 98.7% (157/159) of the participants indicated that they would definitely or probably recommend

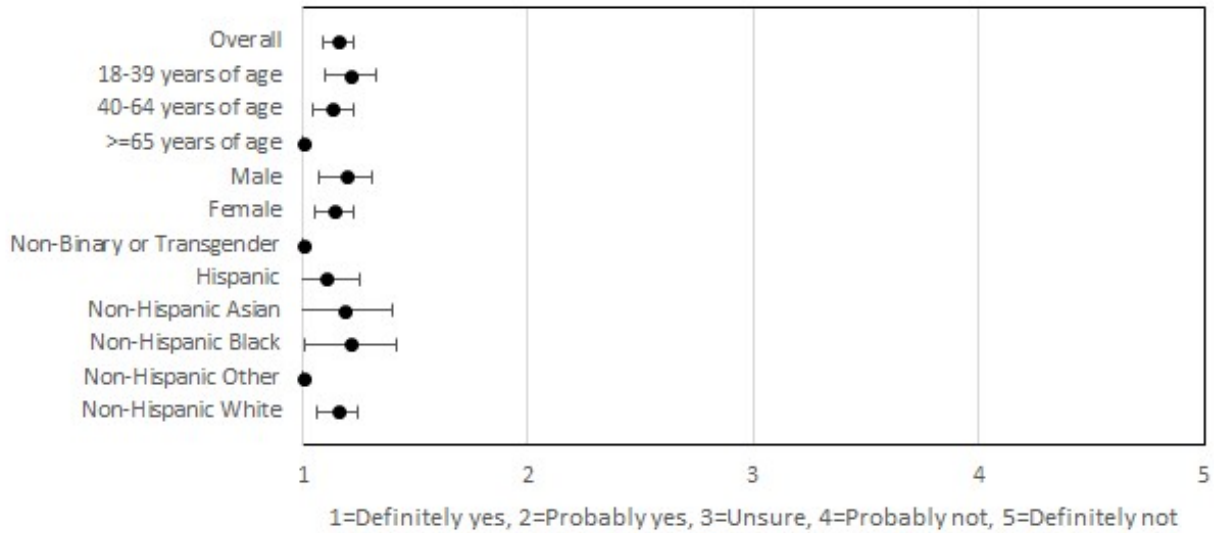
the app to others, with 91.8% (146/159) indicating that they would continue using the app for at-home testing if possible (Figures 2 and 3). These results were consistent across all age, race/ethnicity, and gender groups. In all, 77.4% (123/159) of



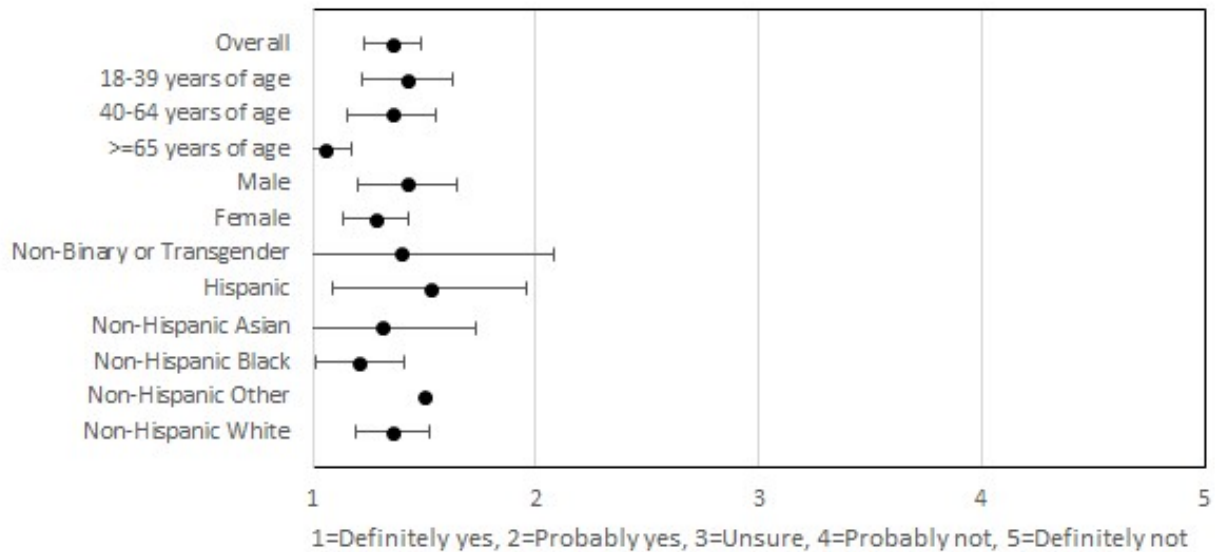
the participants reported never having difficulties using the app, and 3.8% (6/159) reported having difficulties with the app most or all of the time. Among participants who reported difficulties, internet connection issues (5/6, 83%) were the most common reason. Participants on average found the COVID-19 testing

instructions to be the most useful feature of the app, with 88.1% (140/159) of the participants finding this feature “very useful.” The overall rating of the app was 4.4 out of 5, and the overall rating did not differ by age, gender, or race/ethnicity (Figure 4).

**Figure 2.** Participants’ willingness to recommend smartphone app to others.



**Figure 3.** Participants’ interest in continuing to use the smartphone app after the study period.





were very capable of administering, reading, and reporting test results at home without clinical assistance.

### Comparison With Prior Work

Although many previous studies have analyzed the performance of rapid antigen tests for the detection of SARS-CoV-2, few studies have looked at users' testing behavior. It is important to understand who uses rapid antigen tests, when people use rapid antigen tests, and how people test to facilitate the development of effective testing interventions. Nguyen et al [5] found that, among 31 employees of a large company, a daily serial rapid antigen testing intervention with an associated mobile app was highly acceptable, with mean adherence of 88% over a 21-day period. This finding is similar to our own findings of adherence, with over 60% of participants displaying high adherence (>75%) during the 14-day study period. Although the study by Nguyen et al [5] was nested in an employer testing program, with weekly COVID-19 testing required as a condition for employment, our study was based among households residing in 2 large metropolitan cities. The consistency of daily testing adherence across these 2 populations is notable and adds to the external validity of these results. Nguyen et al [5] also found that the acceptability of daily testing was related to the perceived threat of COVID-19, and participants were more likely to find daily testing acceptable in times of high SARS-CoV-2 prevalence. Our study was conducted prior to the widespread distribution of vaccines for SARS-CoV-2; therefore, it is possible that the perceived threat of COVID-19 was generally high throughout the population, contributing to high acceptability and adherence. It is important to reassess COVID-19 testing behaviors as the pandemic continues to evolve to understand the motivations and challenges with testing.

Additionally, the COVID-19 pandemic has escalated pre-pandemic health care disparities within the United States, and geographic inequities in COVID-19 incidence and testing availability persist [19-21]. Non-English speakers, persons of color, and those of lower socioeconomic status are less likely to have access to testing for SARS-CoV-2 than their counterparts, despite simultaneously having an increased proportion of positive cases and mortality [22,23]. Bringing health care services outside the traditional clinical environment offers solutions to accessibility, as well as bridging the gap to populations who have been systematically exploited by the health care system. In this study, we found that the acceptability and usability of at-home testing was consistent across all

race/ethnicity categorizations, indicating that at-home testing could be a promising tool in addressing COVID-19 disparities.

As individuals navigate the return to work and school in the age of COVID-19, it is important that individuals have access to frequent and rapid testing to guide social engagement [24]. However, more information is needed on the diagnostic capabilities and limitations of these tools to ensure that individuals interpret the implications of their test results properly [25]. Additionally, although participants were asked to adhere to a 14-day continuous testing schedule for the purpose of this study, it is important to investigate further optimal testing schedules for SARS-CoV-2 detection [11]. We must also continue to evaluate the accessibility of at-home testing and the MyDataHelps app among diverse communities, including non-English speakers [26].

### Strengths and Limitations

This is the first study to look at the feasibility of at-home Ag-RDTs, filling an important gap in the literature. The strengths of this study include the longitudinal design, which allowed us to analyze adherence over time, and the use of a digital app for testing assistance and survey administration. The use of a digital app allowed participants to engage in the study from their homes, decreasing the burden of participation. Additionally, the wide inclusion criteria allowed the enrollment of a diverse cohort. However, this study is not without limitations. Study participants were required to speak English and have access to a smartphone, which limits the generalizability of our findings. Over 85% of adults in the United States own a smartphone; however, smartphone users vary from nonsmartphone users in terms of education, income, and age [27]. Additionally, only rapid antigen tests using nasal specimen collection were analyzed in this study; therefore, additional work may be necessary to evaluate the feasibility and acceptability of alternative SARS-CoV-2 testing modalities.

### Conclusions

As society establishes a new normal amid an ongoing pandemic, the development of accurate and rapid diagnostics is necessary to help curb SARS-CoV-2 disease transmission and safely navigate social re-engagement. The use of self-tests for COVID-19 with the MyDataHelps app for testing assistance was shown to be a feasible and accessible testing modality across gender, age, and racial groups, and more investigations into the efficacy of these testing modalities is indicated.

---

### Acknowledgments

This work was supported by the National Heart, Lung, and Blood Institute at the National Institutes of Health (3U54HL143541-02S1 and 3U54HL143541-02S2) through the RADx-Tech program.

---

### Conflicts of Interest

CN and VK are employees of CareEvolution. DM has received personal fees and grants from Bristol Myers Squibb, Pfizer, Flexcon, and Fitbit; personal fees for editorial work from Heart Rhythm Society and from Avania for work on Data Safety Monitoring Boards; financial research support from Boehringer Ingelheim; and non-financial research support from Apple, Samsung, and Fitbit.

## Multimedia Appendix 1

Screenshots of the MyDataHelps app.

[\[PPTX File , 705 KB - formative\\_v6i10e35426\\_app1.pptx \]](#)**References**

1. Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. *Lancet Infect Dis* 2020 May;20(5):533-534 [[FREE Full text](#)] [doi: [10.1016/S1473-3099\(20\)30120-1](https://doi.org/10.1016/S1473-3099(20)30120-1)] [Medline: [32087114](https://pubmed.ncbi.nlm.nih.gov/32087114/)]
2. Tromberg BJ, Schwetz TA, Pérez-Stable EJ, Hodes RJ, Woychik RP, Bright RA, et al. Rapid scaling up of COVID-19 diagnostic testing in the United States - the NIH RADx Initiative. *N Engl J Med* 2020 Sep 10;383(11):1071-1077 [[FREE Full text](#)] [doi: [10.1056/NEJMSr2022263](https://doi.org/10.1056/NEJMSr2022263)] [Medline: [32706958](https://pubmed.ncbi.nlm.nih.gov/32706958/)]
3. Ricks S, Kendall EA, Dowdy DW, Sacks JA, Schumacher SG, Arinaminpathy N. Quantifying the potential value of antigen-detection rapid diagnostic tests for COVID-19: a modelling analysis. *BMC Med* 2021 Mar 09;19(1):75 [[FREE Full text](#)] [doi: [10.1186/s12916-021-01948-z](https://doi.org/10.1186/s12916-021-01948-z)] [Medline: [33685466](https://pubmed.ncbi.nlm.nih.gov/33685466/)]
4. Dinnes J, Deeks JJ, Berhane S, Taylor M, Adriano A, Davenport C, Cochrane COVID-19 Diagnostic Test Accuracy Group. Rapid, point-of-care antigen and molecular-based tests for diagnosis of SARS-CoV-2 infection. *Cochrane Database Syst Rev* 2021 Mar 24;3:CD013705 [[FREE Full text](#)] [doi: [10.1002/14651858.CD013705.pub2](https://doi.org/10.1002/14651858.CD013705.pub2)] [Medline: [33760236](https://pubmed.ncbi.nlm.nih.gov/33760236/)]
5. Nguyen N, Lane B, Lee IS, Gorman S, Wu Y, Li A, et al. A mixed methods study evaluating acceptability of a daily COVID-19 testing regimen with a mobile-app connected, at-home, rapid antigen test: implications for current and future pandemics. *PLoS One* 2022 Aug 08;17(8):e0267766 [[FREE Full text](#)] [doi: [10.1371/journal.pone.0267766](https://doi.org/10.1371/journal.pone.0267766)] [Medline: [35939422](https://pubmed.ncbi.nlm.nih.gov/35939422/)]
6. Møller IJB, Utke AR, Rysgaard UK, Østergaard LJ, Jespersen S. Diagnostic performance, user acceptability, and safety of unsupervised SARS-CoV-2 rapid antigen-detecting tests performed at home. *Int J Infect Dis* 2022 Mar;116:358-364 [[FREE Full text](#)] [doi: [10.1016/j.ijid.2022.01.019](https://doi.org/10.1016/j.ijid.2022.01.019)] [Medline: [35038598](https://pubmed.ncbi.nlm.nih.gov/35038598/)]
7. Tu YP, Jennings R, Hart B, Cangelosi GA, Wood RC, Wehber K, et al. Swabs collected by patients or health care workers for SARS-CoV-2 testing. *N Engl J Med* 2020 Jul 30;383(5):494-496 [[FREE Full text](#)] [doi: [10.1056/NEJMc2016321](https://doi.org/10.1056/NEJMc2016321)] [Medline: [32492294](https://pubmed.ncbi.nlm.nih.gov/32492294/)]
8. Hanson KE, Barker AP, Hillyard DR, Gilmore N, Barrett JW, Orlandi RR, et al. Self-collected anterior nasal and saliva specimens versus health care worker-collected nasopharyngeal swabs for the molecular detection of SARS-CoV-2. *J Clin Microbiol* 2020 Oct 21;58(11):1824-1844 [[FREE Full text](#)] [doi: [10.1128/JCM.01824-20](https://doi.org/10.1128/JCM.01824-20)] [Medline: [32817233](https://pubmed.ncbi.nlm.nih.gov/32817233/)]
9. Subramanian R, He Q, Pascual M. Quantifying asymptomatic infection and transmission of COVID-19 in New York City using observed cases, serology, and testing capacity. *Proc Natl Acad Sci U S A* 2021 Mar 02;118(9):e2019716118 [[FREE Full text](#)] [doi: [10.1073/pnas.2019716118](https://doi.org/10.1073/pnas.2019716118)] [Medline: [33571106](https://pubmed.ncbi.nlm.nih.gov/33571106/)]
10. Vermund SH, Pitzer VE. Asymptomatic transmission and the infection fatality risk for COVID-19: implications for school reopening. *Clin Infect Dis* 2021 May 04;72(9):1493-1496 [[FREE Full text](#)] [doi: [10.1093/cid/ciaa855](https://doi.org/10.1093/cid/ciaa855)] [Medline: [32584967](https://pubmed.ncbi.nlm.nih.gov/32584967/)]
11. Soni A, Herbert C, Lin H, Pretz C, Stamegna P, Orwig T, et al. Performance of screening for SARS-CoV-2 using rapid antigen tests to detect incidence of symptomatic and asymptomatic SARS-CoV-2 infection: findings from the Test Us at Home prospective cohort study. medRxiv Preprint posted online on August 06, 2022 [[FREE Full text](#)] [doi: [10.1101/2022.08.05.22278466](https://doi.org/10.1101/2022.08.05.22278466)] [Medline: [35982680](https://pubmed.ncbi.nlm.nih.gov/35982680/)]
12. Hu T, Yue H, Wang C, She B, Ye X, Liu R, et al. Racial segregation, testing site access, and COVID-19 incidence rate in Massachusetts, USA. *Int J Environ Res Public Health* 2020 Dec 19;17(24):9528 [[FREE Full text](#)] [doi: [10.3390/ijerph17249528](https://doi.org/10.3390/ijerph17249528)] [Medline: [33352650](https://pubmed.ncbi.nlm.nih.gov/33352650/)]
13. Hernandez JH, Karletsos D, Avegno J, Reed CH. Is COVID-19 community level testing effective in reaching at-risk populations? evidence from spatial analysis of New Orleans patient data at walk-up sites. *BMC Public Health* 2021 Apr 01;21(1):632 [[FREE Full text](#)] [doi: [10.1186/s12889-021-10717-9](https://doi.org/10.1186/s12889-021-10717-9)] [Medline: [33789647](https://pubmed.ncbi.nlm.nih.gov/33789647/)]
14. Dalva-Baird NP, Alobuia WM, Bendavid E, Bhattacharya J. Racial and ethnic inequities in the early distribution of U.S. COVID-19 testing sites and mortality. *Eur J Clin Invest* 2021 Nov 27;51(11):e13669 [[FREE Full text](#)] [doi: [10.1111/eci.13669](https://doi.org/10.1111/eci.13669)] [Medline: [34390487](https://pubmed.ncbi.nlm.nih.gov/34390487/)]
15. Herbert C, Khetarpal V, Suvarna T, Broach J, Marquez JL, Gerber B, et al. Design and preliminary findings of adherence to the self-testing for our protection from COVID-19 (STOP COVID-19) risk-based testing protocol: prospective digital study. *JMIR Form Res* 2022 Jun 16;6(6):e38113 [[FREE Full text](#)] [doi: [10.2196/38113](https://doi.org/10.2196/38113)] [Medline: [35649180](https://pubmed.ncbi.nlm.nih.gov/35649180/)]
16. Gibson LL, Fahey NM, Hafer N, Buchholz B, Dunlap DR, Murphy RL, et al. The RADx Tech Clinical Studies Core: a model for academic based clinical studies. *IEEE Open J Eng Med Biol* 2021 Apr 28;2:152-157 [[FREE Full text](#)] [doi: [10.1109/OJEMB.2021.3070830](https://doi.org/10.1109/OJEMB.2021.3070830)] [Medline: [34192287](https://pubmed.ncbi.nlm.nih.gov/34192287/)]
17. COVID data tracker. Centers for Disease Control and Prevention. URL: <https://covid.cdc.gov/covid-data-tracker/#variant-proportions> [accessed 2021-05-12]
18. Yanez ND, Weiss NS, Romand JA, Treggiari MM. COVID-19 mortality risk for older men and women. *BMC Public Health* 2020 Nov 19;20(1):1742 [[FREE Full text](#)] [doi: [10.1186/s12889-020-09826-8](https://doi.org/10.1186/s12889-020-09826-8)] [Medline: [33213391](https://pubmed.ncbi.nlm.nih.gov/33213391/)]

19. Zalla LC, Martin CL, Edwards JK, Gartner DR, Noppert GA. A geography of risk: structural racism and coronavirus disease 2019 mortality in the United States. *Am J Epidemiol* 2021 Aug 01;190(8):1439-1446 [FREE Full text] [doi: [10.1093/aje/kwab059](https://doi.org/10.1093/aje/kwab059)] [Medline: [33710272](https://pubmed.ncbi.nlm.nih.gov/33710272/)]
20. Hsiao CJ, Patel AGM, Fasanya HO, Stoffel MR, Beal SG, Winston-McPherson GN, et al. The lines that held us: assessing racial and socioeconomic disparities in SARS-CoV-2 testing. *J Appl Lab Med* 2021 Sep 01;6(5):1143-1154 [FREE Full text] [doi: [10.1093/jalm/jfab059](https://doi.org/10.1093/jalm/jfab059)] [Medline: [34240171](https://pubmed.ncbi.nlm.nih.gov/34240171/)]
21. Bilal U, Tabb LP, Barber S, Diez Roux AV. Spatial inequities in COVID-19 testing, positivity, confirmed cases, and mortality in 3 U.S. cities : an ecological study. *Ann Intern Med* 2021 Jul;174(7):936-944 [FREE Full text] [doi: [10.7326/M20-3936](https://doi.org/10.7326/M20-3936)] [Medline: [33780289](https://pubmed.ncbi.nlm.nih.gov/33780289/)]
22. Grigsby-Toussaint DS, Shin JC, Jones A. Disparities in the distribution of COVID-19 testing sites in black and Latino areas in new York City. *Prev Med* 2021 Jun;147:106463. [doi: [10.1016/j.ypmed.2021.106463](https://doi.org/10.1016/j.ypmed.2021.106463)] [Medline: [33647352](https://pubmed.ncbi.nlm.nih.gov/33647352/)]
23. Johansson MA, Quandelacy TM, Kada S, Prasad PV, Steele M, Brooks JT, et al. SARS-CoV-2 transmission from people without COVID-19 symptoms. *JAMA Netw Open* 2021 Jan 04;4(1):e2035057 [FREE Full text] [doi: [10.1001/jamanetworkopen.2020.35057](https://doi.org/10.1001/jamanetworkopen.2020.35057)] [Medline: [33410879](https://pubmed.ncbi.nlm.nih.gov/33410879/)]
24. Mina MJ, Andersen KG. COVID-19 testing: one size does not fit all. *Science* 2021 Jan 08;371(6525):126-127. [doi: [10.1126/science.abe9187](https://doi.org/10.1126/science.abe9187)] [Medline: [33414210](https://pubmed.ncbi.nlm.nih.gov/33414210/)]
25. Woloshin S, Dewitt B, Krishnamurti T, Fischhoff B. Assessing how consumers interpret and act on results from at-home COVID-19 self-test kits: a randomized clinical trial. *JAMA Intern Med* 2022 Mar 01;182(3):332-341. [doi: [10.1001/jamainternmed.2021.8075](https://doi.org/10.1001/jamainternmed.2021.8075)] [Medline: [35099501](https://pubmed.ncbi.nlm.nih.gov/35099501/)]
26. Fulmer AA, Abboud GA, Wallace LS. Health literacy characteristics of over-the-counter rapid antigen COVID-19 test materials. *Res Social Adm Pharm* 2022 Aug 15;S1551-7411(22):00269-00268 [FREE Full text] [doi: [10.1016/j.sapharm.2022.08.003](https://doi.org/10.1016/j.sapharm.2022.08.003)] [Medline: [35987673](https://pubmed.ncbi.nlm.nih.gov/35987673/)]
27. Demographics of mobile device ownership and adoption in the United States. Pew Research Center. 2021 Apr 07. URL: <https://www.pewresearch.org/internet/fact-sheet/mobile/> [accessed 2022-08-14]

## Abbreviations

**Ag-RDTs:** antigen-based rapid diagnostic tests

**UMass:** University of Massachusetts

*Edited by A Mavragani; submitted 03.12.21; peer-reviewed by C Memering, S Badawy; comments to author 29.06.22; revised version received 24.08.22; accepted 25.08.22; published 18.10.22.*

*Please cite as:*

*Herbert C, Broach J, Heetderks W, Qashu F, Gibson L, Pretz C, Woods K, Kheterpal V, Suvarna T, Nowak C, Lazar P, Ayturk D, Barton B, Achenbach C, Murphy R, McManus D, Soni A*

*Feasibility of At-Home Serial Testing Using Over-the-Counter SARS-CoV-2 Tests With a Digital Smartphone App for Assistance: Longitudinal Cohort Study*

*JMIR Form Res* 2022;6(10):e35426

URL: <https://formative.jmir.org/2022/10/e35426>

doi: [10.2196/35426](https://doi.org/10.2196/35426)

PMID: [36041004](https://pubmed.ncbi.nlm.nih.gov/36041004/)

©Carly Herbert, John Broach, William Heetderks, Felicia Qashu, Laura Gibson, Caitlin Pretz, Kelsey Woods, Vik Kheterpal, Thejas Suvarna, Christopher Nowak, Peter Lazar, Didem Ayturk, Bruce Barton, Chad Achenbach, Robert Murphy, David McManus, Apurv Soni. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 18.10.2022. 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

# Anxiety, Post–COVID-19 Syndrome-Related Depression, and Suicidal Thoughts and Behaviors in COVID-19 Survivors: Cross-sectional Study

Sean F Woodward<sup>1,2</sup>, PhD; Sumra Bari<sup>1</sup>, PhD; Nicole Vike<sup>1</sup>, PhD; Shamal Lalvani<sup>3</sup>, MS; Khrystyna Stetsiv<sup>1</sup>, BA; Byoung Woo Kim<sup>1</sup>, MA; Leandros Stefanopoulos<sup>4</sup>, MS; Nicos Maglaveras<sup>4</sup>, PhD; Hans Breiter<sup>1,5</sup>, MD; Aggelos K Katsaggelos<sup>3,6,7</sup>, PhD

<sup>1</sup>Department of Psychiatry and Behavioral Sciences, Northwestern University, Chicago, IL, United States

<sup>2</sup>Feinberg School of Medicine, Northwestern University, Chicago, IL, United States

<sup>3</sup>Department of Electrical and Computer Engineering, Northwestern University, Evanston, IL, United States

<sup>4</sup>Laboratory of Medical Informatics, Aristotle University of Thessaloniki, Thessaloniki, Greece

<sup>5</sup>Laboratory of Neuroimaging and Genetics, Department of Psychiatry, Massachusetts General Hospital and Harvard University, Boston, MA, United States

<sup>6</sup>Department of Computer Science, Northwestern University, Evanston, IL, United States

<sup>7</sup>Department of Radiology, Northwestern University, Chicago, IL, United States

**Corresponding Author:**

Aggelos K Katsaggelos, PhD

Department of Electrical and Computer Engineering

Northwestern University

2145 Sheridan Road

Evanston, IL, 60208

United States

Phone: 1 847 491 3451

Email: [a-katsaggelos@northwestern.edu](mailto:a-katsaggelos@northwestern.edu)

## Abstract

**Background:** Although the mental health impacts of COVID-19 on the general population have been well studied, studies of the long-term impacts of COVID-19 on infected individuals are relatively new. To date, depression, anxiety, and neurological symptoms associated with post–COVID-19 syndrome (PCS) have been observed in the months following COVID-19 recovery. Suicidal thoughts and behavior (STB) have also been preliminarily proposed as sequelae of COVID-19.

**Objective:** We asked 3 questions. First, do participants reporting a history of COVID-19 diagnosis or a close relative having severe COVID-19 symptoms score higher on depression (Patient Health Questionnaire-9 [PHQ-9]) or state anxiety (State Trait Anxiety Index) screens than those who do not? Second, do participants reporting a COVID-19 diagnosis score higher on PCS-related PHQ-9 items? Third, do participants reporting a COVID-19 diagnosis or a close relative having severe COVID-19 symptoms score higher in STB before, during, or after the first year of the pandemic?

**Methods:** This preliminary study analyzed responses to a COVID-19 and mental health questionnaire obtained from a US population sample, whose data were collected between February 2021 and March 2021. We used the Mann-Whitney *U* test to detect differences in the medians of the total PHQ-9 scores, PHQ-9 component scores, and several STB scores between participants claiming a past clinician diagnosis of COVID-19 and those denying one, as well as between participants claiming severe COVID-19 symptoms in a close relative and those denying them. Where significant differences existed, we created linear regression models to predict the scores based on COVID-19 response as well as demographics to identify potential confounding factors in the Mann-Whitney relationships. Moreover, for STB scores, which corresponded to 5 questions asking about 3 different time intervals (i.e., past 1 year or more, past 1 month to 1 year, and past 1 month), we developed repeated-measures ANOVAs to determine whether scores tended to vary over time.

**Results:** We found greater total depression (PHQ-9) and state anxiety (State Trait Anxiety Index) scores in those with COVID-19 history than those without (Bonferroni  $P=.001$  and Bonferroni  $P=.004$ ) despite a similar history of diagnosed depression and anxiety. Greater scores were noted for a subset of depression symptoms (PHQ-9 items) that overlapped with the symptoms of

PCS (all Bonferroni  $P_s < .05$ ). Moreover, we found greater overall STB scores in those with COVID-19 history, equally in time windows preceding, during, and proceeding infection (all Bonferroni  $P_s < .05$ ).

**Conclusions:** We confirm previous studies linking depression and anxiety diagnoses to COVID-19 recovery. Moreover, our findings suggest that depression diagnoses associated with COVID-19 history relate to PCS symptoms, and that STB associated with COVID-19 in some cases precede infection.

(*JMIR Form Res* 2022;6(10):e36656) doi:[10.2196/36656](https://doi.org/10.2196/36656)

## KEYWORDS

COVID-19; post-COVID-19 syndrome; suicidality; depression; Patient Health Questionnaire-9; PHQ-9; State Trait Anxiety Index; STAI

## Introduction

### Background

More than 60 million COVID-19 cases, directly linked to over 830,000 deaths, have been reported in the United States at the time of this writing [1]. Although our understanding of the disease pathophysiology continues to develop, COVID-19 is largely considered a respiratory and cardiovascular disease caused by infection by the respiratory virus SARS-CoV-2, with highly variable presentation [2]. Symptoms including fatigue, dyspnea, anosmia, and ageusia have been reported to persist for  $\geq 7$  months after infection [3-5] and are now attributed to post-COVID-19 syndrome (PCS), otherwise known as long-haul syndrome.

COVID-19 and its pandemic have also been associated with psychological changes; however, most research on this topic has focused on the COVID-19-free population. In this population, general distress [6] as well as increased levels of anxiety, depression [7,8], and posttraumatic stress disorder symptoms [9] have been observed. These symptoms have been specifically attributed to general fears of infection, interpersonal and economic burdens of social distancing measures [10,11], and downstream sleep disturbances [7,12]. Increased levels of suicidal ideation [13,14] and suicide rates [15,16] in the general population have also been noted, although whether suicide rates have truly increased during the COVID-19 pandemic remains debated [17]. Speculative causes for increased suicidal thoughts and behavior (STB) during the pandemic include increased domestic violence [18], social distancing and decreased interpersonal support [13], and specific psychological phenomena exacerbated by the pandemic such as burdensomeness, thwarted belongingness, and stress sensitivity [19].

Studies of the direct psychological and psychiatric impacts of COVID-19 are relatively sparse. Increased levels of depression, anxiety, and posttraumatic stress disorder symptoms have been observed both during COVID-19 hospitalization [20] and up to 3 months after infection [21-23]. One study observed not only increased anxiety and mood disorders but also increased psychotic and substance abuse disorders at 3 and 6 months after COVID-19 infection [24]. Studies on the long-term psychiatric consequences of SARS-CoV-1 infection forewarn that such sequelae may persist for as many as 50 months after infection [25,26]. Moreover, neurological symptoms such as headaches, concentration difficulties, subjective memory loss, and reduced

attention span are reported to accompany more common symptoms of PCS [27].

These psychiatric and neurological sequelae have been proposed to fundamentally predispose patients with COVID-19 infection and COVID-19 survivors to increased STB [28,29]. In particular, the literature features several case studies of suicide attempts among hospitalized patients with COVID-19, whose motives may reflect universal experiences of COVID-19 hospitalization [30,31]. However, only a small number of studies have quantitatively addressed STB in patients with COVID-19; one reported increased suicidal ideation in patients 1 month after COVID-19 hospitalization [32], and another characterized suicides in patients testing positive for COVID-19 infection reported in the media [33]. Interestingly, a study of the general population linked increased suicidal ideation with a desire to seek COVID-19 exposure [13], implying a potential reversal of the assumption that COVID-19 precedes an increase in STB.

### Objectives

This study examined responses to a questionnaire digitally distributed to a population sample of 506 adults in early March 2021. The questionnaire included a depression screener (Patient Health Questionnaire-9 [PHQ-9]) [34], a state anxiety screener (State Trait Anxiety Index [STAI]) [35], a custom set of questions about STB experienced in 3 periods (1 month ago, 1-12 months ago, and  $>12$  months ago), a question asking whether the participant ever received a COVID-19 diagnosis, and a question asking whether the participant ever had a friend and/or family member who had severe or fatal COVID-19 symptoms. Data analysis focused on 3 questions. First, assuming a similar prevalence of depression and anxiety history among participants with and without COVID-19 history, we asked whether COVID-19 infection (a personal diagnosis or severe symptomology in a close friend or relative) was linked to—and explicitly causative of—increased depression (PHQ-9) or anxiety (state component of the STAI) screening scores. Second, we asked whether COVID-19 infection was linked to—and implicitly causative of—increases in specific-item PHQ-9 scores and specifically in PCS-related scores (in particular, scores corresponding to fatigue, problems with concentration, psychomotor retardation or agitation, and altered appetite) for patients admitting a COVID-19 diagnosis. Lastly, we asked whether COVID-19 infection was linked to increased STB before or after infection and whether any causality in either direction was suggested. In all cases, we considered confounding factors from demographic information, as well as previous diagnoses of depression and anxiety. This study sought not only

to confirm known associations between the experience of COVID-19 and greater depression and anxiety scores but also to expose the drivers of greater depression scores with respect to known PCS symptomology. Moreover, it sought to clarify the effect of COVID-19 on STB in a population reporting pre-COVID-19 STB.

## Methods

### Participant Recruitment and Demographics

Study participants were recruited by Gold Research Inc from multiple vendors. Gold Research's vendors recruit the emails of willing participants in multiple ways. Some are recruited "by invitation only" from customer databases of large companies in revenue-sharing agreements, some are recruited from social media, some are recruited via direct mail, and others sign up voluntarily to participate in research studies in lieu of monetary or other incentives such as coupons for everyday household purchases. During recruitment, all survey respondents also go through a double opt-in process to indicate the types of research studies they would like to participate in along with providing their profiles on different demographic attributes like age, race, and gender. This information is then used to reflect representation against US Census metrics. In this process, respondents are also asked multiple test questions to screen out those providing random and illogical responses or showing flatline or speeder behavior. Along with having cohort demographics balanced to meet the demographic criteria established by the US Census, Gold Research also oversampled 15% of the sample for mental health conditions.

Gold Research reported that over 50,000 respondents were contacted for questionnaire completion. They estimated that over 37,500 (75%) either did not respond or said no. Of the remaining 12,500 who did click on the survey link, more than 50% did not complete the questionnaire. Of the >6000 who completed the survey, those who did not clear the data integrity assessments were omitted to obtain the final number of completed surveys.

We assessed multiple mental health conditions including depression symptoms and STB. In this study, we focused only on depression and STB. The request for participation stated Gold Research was administering the study on behalf of its client, Northwestern University, to study emotional health (see text at end of this section for detailed instructions in the survey about the solicitation, study description, and opt-in procedures). Participants meeting quality assurance procedures (including completion of the survey) were studied, up to a limit of 500 to 520 participants, resulting in 506 participants in the final cohort, of which 379 (74.9%) met all quality assurance criteria.

Questionnaire responses were digitally collected between the end of February 2021 and the first week of March 2021, approximately 1 year following the official pandemic declaration in the United States (March 11, 2020) [36]. Final filtering for data quality (see the Data Quality Assurance section) reduced the final sample size to 379 questionnaires or participants (Figure S1 in [Multimedia Appendix 1](#)). All analyses centered on this subset of 379 participants, which, based on the nature of the databases sampled by Gold Research Inc, as well as the demographic statistics shown below, was determined to be an approximate, random sample of the adult US population. Assessment of location data by state showed that the sample of 379 participants had a broad geographic distribution (Figure S2 in [Multimedia Appendix 1](#)). Participants reported age, gender ("What gender do you consider yourself?"), ethnicity ("What ethnicity do you consider yourself?"), annual household income, employment status, years of school, and highest level of education ([Table 1](#)).

The final sample comprised 379 participants (age: mean 43, SD 15 years; range 18-70 years) ([Table 1](#)). Most participants identified as female (215/379, 56.7%), White (251/379, 66.2%), having an annual household income of US \$25,000 to US \$50,000 (98/379, 25.9%), employed full-time (161/379, 42.5%), and having completed some college (105/379, 27.7%). The gender, ethnicity, and education percentages approximated US Census Bureau figures at the time of data collection [37].

For initial recruitment, potential participants received the communication displayed in [Textbox 1](#).

If potential participants responded with "Accept," they were then sent the message in [Textbox 2](#).

The survey would then begin if they pressed "Next."

The participants were asked questions related to their COVID-19 status. They were asked (1) to report whether they were ever diagnosed with COVID-19 by a medical clinician (variable COVID-DIAGNOSIS: POS/NEG indicates a positive or nondiagnosis), (2) whether they ever received a positive COVID-19 laboratory test (variable COVID-TEST: POS/NEG), and (3) whether a family member or close friend ever experienced serious symptoms or died of COVID-19 (variable COVID-FAMILY: POS/NEG). Because 30 participants were members of both the COVID-DIAGNOSIS POS and COVID-TEST POS groups, which corresponded to 77% (30/39) of the participants of the total COVID-DIAGNOSIS POS group and 73% (30/41) of the participants of the total COVID-TEST POS group, our subsequent analysis featured only the COVID-DIAGNOSIS variable given its collinearity with COVID-TEST. A similar analysis, yielding similar results, for the COVID-TEST variable can be found in [Multimedia Appendix 1](#).



**Table 1.** Participant demographic information with breakdown of questionnaire responses.

Characteristic	Participants <sup>a</sup>
<b>Age (years)</b>	
Count, n (%)	379 (100)
Mean (SD); range	43.21 (15.38); 18.0-70.0
<b>Gender, n (%)</b>	
Male	163 (43)
Female	215 (56.7)
Other	1 (0.3)
<b>Ethnicity, n (%)</b>	
White	251 (66.2)
African American	52 (13.7)
Hispanic	33 (8.7)
Asian or Pacific Islander	13 (3.4)
Native American or Alaska Native	7 (1.8)
Other	3 (0.8)
Mixed (≥2 ethnicities)	17 (4.5)
Prefer not to answer	3 (0.8)
<b>Annual household income<sup>b</sup> (US \$), n (%)</b>	
<25,000	88 (23.2)
25,000-50,000	98 (25.9)
50,000-75,000	73 (19.3)
75,000-100,000	48 (12.7)
100,000-150,000	38 (10)
150,000-300,000	26 (6.9)
>300,000	8 (2.1)
<b>Employment status, n (%)</b>	
Unemployed	61 (16.1)
Full-time	161 (42.5)
Part-time	43 (11.3)
Self-employed	26 (6.9)
More than one job	3 (0.3)
Retired	61 (16.1)
Other	24 (6.3)
<b>Years of school</b>	
Count, n (%)	379 (100)
Mean (SD); range	13.18 (5.15); 1.0-30.0
<b>Highest education level, n (%)</b>	
Some high school	7 (1.8)
High school graduate	99 (26.1)
Some college	105 (27.7)
Bachelor's degree	86 (22.7)
Some graduate school	13 (3.4)

Characteristic	Participants <sup>a</sup>
Graduate degree	30 (7.9)
Postgrad or doctorate	39 (10.3)

<sup>a</sup>Response counts across all participants.

<sup>b</sup>Income and education are considered ordinal variables, and their integer values (prepending the long-form responses in the second column) were used for statistical computations.

**Textbox 1.** Communication sent to potential participants during initial recruitment.

Gold Research Inc, a national market research firm, and its client, Northwestern University, request your participation in this study of emotional health. We will be evaluating how different emotions and experiences are connected and may relate to our emotional health. The information you provide will be kept confidential, coded to be anonymous so it cannot be connected back to you and will be used only for research purposes. Researchers will not be able to contact you or restudy you after this survey. We will not share your information with any other third party. We will also not use your information to identify you individually or use your responses to market or sell other services or products to you. As part of this effort, you will not be asked to provide any personal identifiers such as your name, email, phone number, address or social media handles. A unique identifier will be generated for you and each survey participant to enhance privacy. As part of the survey process, we will be able to tell if you completed the survey, but we will not be able to tell which answers were yours. For this study, we are going to ask you some questions about yourself and how much you like or dislike a set of pictures. You may discontinue this study at any time. We appreciate your help with this study, given the serious challenges facing many people regarding emotional health at this time. We thank you in advance.

1. Accept
2. Decline

**Textbox 2.** Follow-up communication sent to participants.

Thank you for participating in our survey. All responses during this survey are anonymous and confidential. We will be able to tell if you completed the survey, but we will not be able to tell which answers were yours. In this study, we aim to understand how different emotions and experiences relate to visual processing.

We are going to:

- Ask you some questions about yourself
- Have you rate how much you like or dislike a set of pictures

For this study, your identity is protected and your answers are anonymous and confidential. Press “Next” to proceed.

## Ethics Approval

Participation was offered noting that Gold Research was administering an emotional health questionnaire on behalf of Northwestern University, with the phrasing, “We will be evaluating how different emotions and experiences are connected and may relate to our emotional health.” The complete text about the solicitation, study description, and opt-in procedures can be found in the Participant Recruitment subsection of the Methods section. All participants provided informed consent following oversight by the Northwestern University Institutional Review Board, which reviewed and approved the project (approval number STU00213665). The participants were guaranteed anonymity and confidentiality, and the researchers possessed no protected health information.

## Survey Questions and Scoring

Our survey questionnaire consisted of the PHQ-9, the 20 “state” questions from the STAI Form Y, behavioral neurology and mental health questions including those pertaining to STB from the Massachusetts General Hospital Subjective Question screener of the Phenotype Genotype Project in Addiction and Mood Disorder [38], and questions pertaining to COVID-19 history. A cognitive task related to picture ratings was also included (data not reported here) and incorporated into general quality assurance. The questions pertinent to this study are outlined in Table 2, with a breakdown of responses to ancillary demographic questions given in Table 1. The questionnaire is not outlined in its entirety here.

**Table 2.** Questionnaire items.

Variable <sup>a</sup>	Question preamble	Question <sup>b</sup>	Value space <sup>c</sup>
COVID-TEST	N/A <sup>d</sup>	Have you ever tested positive for COVID-19?	Binary
COVID-DIAGNOSIS	N/A	Have you ever been diagnosed with COVID-19 by a medical clinician?	Binary
COVID-FAMILY	N/A	Has anyone in your family or group of friends had serious symptoms or died of COVID-19?	Binary
PHQ9-1	Over the last two weeks, how often have you been bothered by any of the following problems? 1 (not at all), 2 (several days), 3 (more than half the days), 4 (nearly every day)	Little interest or pleasure in doing things	[0,3]
PHQ9-2	Same as above	Feeling down, hopeless, or depressed	[0,3]
PHQ9-3	Same as above	Trouble falling/staying asleep, sleeping too much	[0,3]
PHQ9-4	Same as above	Feeling tired or having little energy	[0,3]
PHQ9-5	Same as above	Poor appetite or overeating	[0,3]
PHQ9-6	Same as above	Feeling bad about yourself or that you are a failure or have let yourself or your family down	[0,3]
PHQ9-7	Same as above	Trouble concentrating on things, such as reading the newspaper or watching television	[0,3]
PHQ9-8	Same as above	Moving or speaking so slowly that other people could have noticed. Or the opposite, being so fidgety or restless that you have been moving around a lot more than usual	[0,3]
PHQ9-9	Same as above	Thoughts that you would be better off dead or of hurting yourself in some way	[0,3]
PHQ9-SUM	Same as above	Sum of the 9 PHQ-9 <sup>e</sup> scores	[0,27]
STAI-SUM	N/A	Sum of the 20 STAI <sup>f</sup> Form Y State scores	[0,60]
DEPRESSION-YRS	For how many years have you had the following diagnosis by a medical professional, psychologist, or physician?	Depression?	[0,60]
DEPRESSION-HX	Same as above	Depression?	Binary
ANXIETY-YRS	Same as above	Anxiety disorder?	[0,60]
ANXIETY-HX	Same as above	Anxiety disorder?	Binary
S-PASSIVE-LT	Please rate how often you have experienced each of the following (in the following time periods): 1 (never), 2 (rarely), 3 or 4 (sometimes), 5 or 6 (often), 7 (always)	Wish to go to sleep and not wake up—past 1 year or more	[1,7]
S-PASSIVE-MT	Same as above	Wish to go to sleep and not wake up—past 1 month to 1 year	[1,7]
S-PASSIVE-ST	Same as above	Wish to go to sleep and not wake up—past 1 month	[1,7]
S-ACTIVE-LT	Same as above	Wanting to hurt yourself or take your own life—past 1 year or more	[1,7]
S-ACTIVE-MT	Same as above	Wanting to hurt yourself or take your own life—past 1 month to 1 year	[1,7]
S-ACTIVE-ST	Same as above	Wanting to hurt yourself or take your own life—past 1 month	[1,7]
S-PLAN-LT	Same as above	Having a plan to take your own life—past 1 year or more	[1,7]
S-PLAN-MT	Same as above	Having a plan to take your own life—past 1 month to 1 year	[1,7]

Variable <sup>a</sup>	Question preamble	Question <sup>b</sup>	Value space <sup>c</sup>
S-PLAN-ST	Same as above	Having a plan to take your own life—past 1 month	[1,7]
S-HISTORY-LT	Same as above	Prior attempts at hurting yourself, or taking your own life—past 1 year or more	[1,7]
S-HISTORY-MT	Same as above	Prior attempts at hurting yourself, or taking your own life—past 1 month to 1 year	[1,7]
S-HISTORY-ST	Same as above	Prior attempts at hurting yourself, or taking your own life—past 1 month	[1,7]
S-SAFETY-LT	Same as above	Having a safety plan for not hurting yourself for when those feelings arise—past 1 year or more	[1,7]
S-SAFETY-MT	Same as above	Having a safety plan for not hurting yourself for when those feelings arise—past 1 month to 1 year	[1,7]
S-SAFETY-ST	Same as above	Having a safety plan for not hurting yourself for when those feelings arise—past 1 month	[1,7]

<sup>a</sup>Designations of the scores or variables obtained.

<sup>b</sup>Corresponding questions as phrased in the questionnaire.

<sup>c</sup>Range of possible responses.

<sup>d</sup>N/A: not applicable.

<sup>e</sup>PHQ-9: Patient Health Questionnaire-9.

<sup>f</sup>STAI: State Trait Anxiety Index.

The PHQ-9 includes 9 questions, as outlined in Table 2 [34]. PHQ-9 is a self-report-based screening tool for depression. A numeric response to a single question can take any integer value from 0 to 4. Across 9 questions, a cumulative score  $\geq 5$  is associated with mild depression, and a score  $\geq 10$  is associated with moderate depression. For statistical analyses, responses to these questions were pooled to obtain cumulative scores and considered independent scores.

The STAI Form Y includes 20 “state” anxiety questions [35]. This questionnaire subset is a tool for assessing a respondent’s current anxiety in response to a perceived danger, based on self-reporting. A numeric response to a single question can take any integer value from 0 to 3, with a greater cumulative score associated with greater anxiety. We considered only cumulative scores.

Five questions pertaining to current and historic STB were included in the questionnaire; these were adapted from the Massachusetts General Hospital Phenotype Genotype Project [38], with direct antecedents in the *Manual of Psychiatric Emergencies 4th Edition* [39]. These are outlined in their entirety in Table 2. The 5 questions were each repeated 3 times for 3 blocks of time asking how often participants experienced symptoms in the (1) past 1 year or more, (2) past 1 month to 1 year, and (3) past 1 month on a 1 to 7 Likert scale (1=never, 2-3=rarely, 4-5=sometimes, 6=often, and 7=always). Responses to these questions were considered independent and were not pooled to obtain cumulative scores.

As noted previously, the questionnaire also asked whether participants were ever diagnosed with COVID-19 by a medical clinician (COVID-DIAGNOSIS: POS/NEG), whether they had received a positive COVID-19 laboratory test (COVID-TEST: POS/NEG), and whether a family member or close friend had ever experienced serious symptoms or died of COVID-19 (COVID-FAMILY: POS/NEG).

During the analysis, we assumed that the vast majority of the participants who had COVID-19 (either by clinician diagnosis or test) or had a family member affected would experience symptoms during the year preceding survey participation, that is, between March 5, 2020 (1 year prior to the end of the survey) and February 12, 2021 (the day before which a person could be diagnosed with COVID-19 and be expected to have their acute symptoms resolved before the beginning of the survey, 2 weeks later). We justify this assumption as follows. On March 5, 2020, a total of 41.45 cumulative cases of COVID-19 infection were reported in the United States; on February 12, 2021, a total of 27,504,011 cumulative cases were reported; and on March 5, 2021 (the closing of the survey), 28,866,565 cumulative cases were reported [1]. We, therefore, expect that 95% of the participants who experienced acute COVID-19 in the assumed interval were likely resolved at the time of the survey. The remaining 5% may have experienced acute COVID-19 symptoms during the survey period.

### Data Quality Assurance

An initial set of 506 participants was obtained from Gold Research, Inc, based on the following inclusion criteria: age 18 to 70 years, capacity to complete the questionnaire, and enrollment from multiple databases of potential participants.

The subsequent quality assessment used 4 exclusion criteria: (1) participants with the same responses throughout any section of the questionnaire (eg, “1” for all questions), (2) participants indicating they had been diagnosed by a clinician with 10 or more illnesses (data outside depression and anxiety not described here), (3) participants with minimal variance in a picture rating task (all pictures were rated the same or varied only by 1 point; data not described here), and (4) participants reporting inconsistent education level and years of education and participants who completed the questionnaire in less than 500 seconds. After applying the aforementioned exclusion criteria,

74.9% (379/506) of participants qualified for the statistical analysis.

### Statistical Analyses

The aims of this study were divided into 3 categories. First, assuming a similar prevalence of depression and anxiety history among participants with and without COVID-19 history, we asked whether COVID-19 infection (a personal diagnosis or severe symptomology in a close friend or relative) was linked to—and explicitly causative of—increased depression (PHQ-9) or anxiety (state STAI) screening scores. Second, we asked whether COVID-19 infection was linked to—and implicitly

causative of—increases in specific-item PHQ-9 scores and specifically in PCS-related scores (in particular, scores corresponding to fatigue, problems with concentration, psychomotor retardation or agitation, and altered appetite) for those patients admitting a COVID-19 diagnosis. Finally, we asked whether COVID-19 infection was linked to increased STB before or after infection and whether any causality in either direction was suggested. In all cases, we considered confounding factors from demographic information, as well as from previous diagnoses of depression and anxiety. [Table 3](#) gives a broad framework of the a priori study design.

**Table 3.** Overview of hypotheses and analyses underlying this study.

Hypothesis <sup>a</sup>	Expected causality <sup>b</sup>	A priori covariates	Analysis tools
COVID-19 diagnosis → total PHQ-9 <sup>c</sup>	Explicit	Group A covariates <sup>d</sup> ; COVID-19 in relation	Group A analyses <sup>e</sup>
COVID-19 in relation → total PHQ-9	Explicit	Group A covariates; COVID-19 diagnosis	Group A analyses
COVID-19 diagnosis → total STAI <sup>f</sup>	Explicit	Group A covariates; COVID-19 in relation	Group A analyses
COVID-19 in relation → total STAI	Explicit	Group A covariates; COVID-19 diagnosis	Group A analyses
COVID-19 diagnosis → PHQ-9	Implicit	Group A covariates; COVID-19 in relation	Group A analyses
COVID-19 in relation → PHQ-9	Implicit	Group A covariates; COVID-19 diagnosis	Group A analyses
COVID-19 diagnosis ↔ STB <sup>g</sup> items (3 time windows)	Exploratory	Group A covariates; COVID-19 in relation	Group A analyses; RM-ANOVA <sup>h</sup>
COVID-19 in relation ↔ STB items (3 time windows)	Exploratory	Group A covariates; COVID-19 diagnosis	Group A analyses; RM-ANOVA

<sup>a</sup>Arrows indicate the expected direction of causality (if any).

<sup>b</sup>When expected causality is “explicit,” we possess a baseline measure of the dependent variable, and when “implicit,” causality is believed to be implied by context.

<sup>c</sup>PHQ-9: Patient Health Questionnaire-9.

<sup>d</sup>“Group A covariates” include demographics, depression history, and anxiety history.

<sup>e</sup>“Group A analyses” include Mann-Whitney *U* testing and linear regression covariate analysis.

<sup>f</sup>STAI: State Trait Anxiety Index.

<sup>g</sup>STB: suicidal thoughts and behavior.

<sup>h</sup>RM-ANOVA: repeated-measures ANOVA.

Data analysis was done in 4 parts and numbered 1 to 4 in the following text; the Results section is also organized using this framework. In part 1, we determined differences in participant demographics and depression or anxiety history variables based on their COVID-19 status (ie, COVID-DIAGNOSIS and COVID-FAMILY) to identify potential confounding variables. In part 2, we assessed differences in cumulative PHQ-9 (PHQ9-SUM) and state STAI (STAI-SUM) scores based on COVID-19 status. In part 3, we examined differences in scores for individual PHQ-9 questions (PHQ9-1 through PHQ9-9) based on COVID-19 status. In part 4, we quantified differences in STB (listed in [Table 2](#)) scores based on COVID-19 status. Given that STB was also reported for 3 time frames (ie, past 1 year or more, past 1 month to 1 year, and past 1 month), 30 repeated-measures ANOVAs were run (ie, 3 COVID-19 tests × 2 COVID-19 statuses × 5 questions) to assess potential differences between time frames.

For this study, only the COVID-DIAGNOSIS and COVID-FAMILY analyses are reported in the main text. Similar analyses based on a third COVID-19 status variable, positive

self-reported COVID-19 laboratory testing (COVID-TEST), appear in [Multimedia Appendix 1](#), given that there was significant overlap between the COVID-DIAGNOSIS and COVID-TEST responses (see the Participant Demographics subsection).

Part 1 specifically involved comparing the medians of demographic variables (reported in the Results section) and 4 depression or anxiety history variables (DEPRESSION-YRS [years carrying depression diagnosis], DEPRESSION-HX [any history of depression diagnosis], ANXIETY-YRS [years carrying anxiety diagnosis], and ANXIETY-HX [any history of anxiety diagnosis] from [Table 2](#)) between 2 COVID-19 status groups (ie, COVID-DIAGNOSIS POS vs NEG group; COVID-FAMILY POS vs NEG group) using Mann-Whitney *U* testing ( $\alpha=.05$ ). *P* values were adjusted for multiple comparisons using Bonferroni correction. Significant differences in median scores (Bonferroni  $P<.05$ ) indicated potential confounders to be controlled in parts 2 to 4.

The analysis for parts 2 to 4 each involved 2 steps. First, the analysis began by comparing the medians of the mental health variables (as grouped in the first paragraph of this section) between COVID-19 status groups by using the Mann-Whitney *U* test ( $\alpha=.05$ ) with subsequent Bonferroni correction ( $P<.05$ ). Second, to confirm the robustness of these differences to the potential demographic confounders identified in part 1, linear regressions were run to separately model the mental health variables from COVID-DIAGNOSIS and COVID-FAMILY. For each unique pair of mental health (dependent variable) and COVID-19 status variables (independent variable), we built 1 univariate model and 1 multivariate model. The multivariate model included potential confounders identified in part 1 as covariates. To prevent discrepancy between the sample sizes of the POS and NEG groups from causing linear regression models to preferentially fit NEG group data, we applied the synthetic minority oversampling technique (SMOTE) to the POS group data to equalize the number of POS and NEG samples before fitting. SMOTE is considered a standard oversampling technique [40], and we emphasize that it was not implemented until after initial Mann-Whitney *U* testing. For each univariate or multivariate model pair, we compared the crude and adjusted nonstandardized regression coefficients (denoted as B) for the COVID-19 status variable. A difference of >10% was interpreted to indicate confounding and prompted us to report the adjusted regression coefficient [41]; we otherwise reported the crude regression coefficient. [Multimedia Appendix 1](#) provides more details on this approach for assessing potential confounding factors in our main analyses.

Based on our observation that 22.3% (23/103) of the participants in the COVID-FAMILY POS group were also in the

COVID-DIAGNOSIS POS group, we included COVID-DIAGNOSIS as a covariate in all adjusted linear regression models involving COVID-FAMILY. However, COVID-FAMILY was not used as a covariate in the models involving COVID-DIAGNOSIS. This decision was based on the observation that 59% (23/39) of the participants in the COVID-DIAGNOSIS POS group were also in the COVID-FAMILY POS group, suggesting that the consequences of a personal COVID-19 diagnosis could not be meaningfully separated from experiencing severe symptoms in a friend or family member.

## Results

### Overview

A small fraction of the participants were members of the COVID-DIAGNOSIS POS (39/379, 10.3%) or COVID-FAMILY POS (28/103, 27.2%) groups (Table 4), which is broadly consistent with reports of illness incidence during the first year of the COVID-19 pandemic (ie, 10%). A total of 23 participants were members of both COVID-DIAGNOSIS POS and COVID-FAMILY POS groups. Chi-square testing indicated higher levels of COVID-DIAGNOSIS POS status in the COVID-FAMILY POS group than in the overall sample (23/103, 22.3%;  $P=.001$ ) and similarly higher levels of COVID-FAMILY status in the COVID-DIAGNOSIS POS group as when compared to the overall sample (23/39, 59%;  $P<.001$ ). Although not presented as a main finding of this paper, these influenced the construction of our regression models, detailed in the Statistical Analyses subsection.

**Table 4.** Participant demographic statistics<sup>a</sup>.

Variable	Sample description		Mann-Whitney <i>U</i> test		
	NEG group, median (IQR)	POS group, median (IQR)	<i>U</i> statistic	<i>P</i> value	Bonferroni <i>P</i> value
<b>COVID-DIAGNOSIS (NEG: n=340; POS: n=39)</b>					
Age (years)	45.0 (28.0)	36.0 (19.5)	4644.0	.001	.004
Gender	N/A <sup>a</sup>	N/A	5253.0	.007	.03
Income	2.0 (3.0)	4.0 (2.0)	4026.5	<.001	<.001
Education	3.0 (2.0)	4.0 (3.5)	5315.5	.02	.08
<b>COVID-FAMILY (NEG: n=276; POS: n=103)</b>					
Age (years)	44.0 (27.25)	39.0 (31.0)	12,976.0	.10	.38
Gender	N/A	N/A	13,674.0	.25	>.99
Income	3.0 (3.0)	2.0 (2.0)	13,730.0	.30	>.99
Education	3.0 (2.0)	3.0 (1.5)	13,661.0	.28	>.99

<sup>a</sup>N/A: not applicable.

The results are presented according to the 4 parts outlined in the Statistical Analyses section. First, we present the results of testing for differences in demographic and depression or anxiety history values based on COVID-19 status; this analysis was done to identify potential confounding variables for subsequent analyses. These variables were then included in relevant

multivariate regression analyses in the second, third, and fourth analysis sections to quantify confounding effects. Second, we determined differences in cumulative depression (PHQ-9) and anxiety (STAI) scores based on COVID-19 status, including univariate and multivariate linear regressions intended to highlight potential confounding effects indicated previously.

Third, we performed the same analysis as in part 2 but for each question in the PHQ-9, as opposed to the cumulative score. Fourth, we performed the same analysis as in the previous 2 parts but for STB scores. Note that although this analysis considers the COVID-DIAGNOSIS and COVID-FAMILY variables, an analogous analysis for COVID-TEST is included in Tables S1-S4 in [Multimedia Appendix 1](#). Please refer to [Table 2](#) for descriptions and abbreviations of all the survey questions considered here.

### Analysis of Demographics and Depression or Anxiety History Against COVID-19 Status

This analysis sought to identify covariates for subsequent analyses. Participants in the COVID-DIAGNOSIS POS group reported a lower median age than that reported in the COVID-DIAGNOSIS NEG group (Bonferroni  $P < .05$ ; [Table 4](#)), greater median annual household income, and identified as female (gender) more frequently than participants in the NEG group (Bonferroni  $P < .05$ ). Median age, gender, income, and

highest education level did not differ between POS and NEG groups for COVID-19 (all Bonferroni  $P_s > .05$ ). Although data for the demographic variables ethnicity and employment were also collected, there were  $< 5$  samples in most categories for the COVID-DIAGNOSIS POS group, and we were, therefore, unable to perform valid comparisons between the POS and NEG groups for these 2 variables.

Participants in the COVID-DIAGNOSIS POS group did not report a different median DEPRESSION-HX, DEPRESSION-YRS, ANXIETY-HX, or ANXIETY-YRS score than those in the NEG group (all Bonferroni  $P_s > .05$ ; [Table 5](#)). By contrast, the median DEPRESSION-HX and DEPRESSION-YRS scores were greater among participants in the COVID-FAMILY POS group than those in the NEG group (Bonferroni  $P < .05$ ). The median ANXIETY-HX and ANXIETY-YRS scores were not significantly different between the participants in the COVID-FAMILY POS and NEG groups (all Bonferroni  $P_s > .05$ ).

**Table 5.** Separation of major depression and anxiety variables by COVID-19 status.

Variable	Descriptive statistics		Mann-Whitney <i>U</i> test			Linear regression confounder analysis				
	NEG group, median (IQR)	POS group, median (IQR)	<i>U</i> statistic	<i>P</i> value	Bonferroni <i>P</i> value	$\Delta B$	B (95% CI)	<i>P</i> value	Bonferroni <i>P</i> value	
<b>COVID-DIAGNOSIS (NEG: n=340; POS: n=39)</b>										
PHQ9-SUM	6.0 (12.0)	13.0 (9.5)	4383.5	$< .001$	.001	0.25	3.26 (2.17 to 4.36)	$< .001$	$< .001$	
STAI-SUM	20.0 (21.0)	29.0 (11.0)	4538.0	.001	.004	0.08	6.64 (4.92 to 8.35)	$< .001$	$< .001$	
DEPRESSION-YRS	0.0 (4.0)	1.0 (3.5)	5612.0	.04	.21	— <sup>a</sup>	—	—	—	
ANXIETY-YRS	0.0 (2.0)	0.0 (3.0)	5765.0	.05	.32	—	—	—	—	
DEPRESSION-HX	0.0 (1.0)	0.0 (1.0)	6239.0	.23	$> .99$	—	—	—	—	
ANXIETY-HX	0.0 (1.0)	0.0 (1.0)	6102.5	.15	.90	—	—	—	—	
<b>COVID-FAMILY (NEG: n=276; POS: n=103)</b>										
PHQ9-SUM	5.0 (12.0)	9.0 (12.0)	11,444.5	.002	.01	0.55	0.96 (−0.22 to 2.15)	.11	$> .99$	
STAI-SUM	20.0 (22.0)	25.0 (18.5)	12,465.5	.03	.20	—	—	—	—	
DEPRESSION-YRS	0.0 (2.0)	0.0 (5.5)	11,800.5	.002	.01	—	—	—	—	
ANXIETY-YRS	0.0 (2.0)	0.0 (4.5)	13,033.0	.07	—	—	—	—	—	
DEPRESSION-HX	0.0 (1.0)	0.0 (1.0)	12,229.0	.004	—	—	—	—	—	
ANXIETY-HX	0.0 (1.0)	0.0 (1.0)	13,195.0	.09	—	—	—	—	—	

<sup>a</sup>Regression analysis was not performed given the nonsignificant Mann-Whitney *U* test.

### Analysis of Total Depression and Anxiety Scores Against COVID-19 Status

#### COVID-DIAGNOSIS

The median PHQ9-SUM score was significantly higher in the COVID-DIAGNOSIS POS group than in the NEG group (Bonferroni  $P < .05$ ). The median STAI-SUM score was also

significantly higher in the COVID-TEST POS group than in the NEG group (Bonferroni  $P < .05$ ; [Table 5](#)). In contrast, when DEPRESSION-HX, DEPRESSION-YRS, ANXIETY-HX, and ANXIETY-YRS were tested for differences between the COVID-DIAGNOSIS POS and NEG groups, no differences were observed (all Bonferroni  $P_s > .05$ ; [Table 5](#)).

Linear regression modeling of PHQ9-SUM from COVID-DIAGNOSIS with and without age, gender, and income as covariates suggested confounding in the relationship between PHQ9-SUM and COVID-DIAGNOSIS ( $|\Delta B| > 0.1$ ); however, the adjusted B coefficient for COVID-DIAGNOSIS was significantly greater than 0 (Bonferroni  $P < .05$ ; adjusted B reported). For the same analysis modeling STAI-SUM from COVID-DIAGNOSIS, we again found confounding ( $|\Delta B| > 0.1$ ), but with an adjusted B coefficient for COVID-DIAGNOSIS significantly greater than 0 (Bonferroni  $P < .05$ ; adjusted B reported).

### **COVID-FAMILY**

The median PHQ9-SUM score was significantly higher in the COVID-FAMILY POS group than in the NEG group (Bonferroni  $P < .05$ ; Table 5). However, the median STAI-SUM score was not significantly different between the COVID-FAMILY POS and NEG groups (Bonferroni  $P > .05$ ).

The median DEPRESSION-HX and DEPRESSION-YRS scores were significantly greater in the COVID-FAMILY POS group than in the NEG group (Bonferroni  $P < .05$ ). The median ANXIETY-HX and ANXIETY-YRS scores were not significantly different between the COVID-FAMILY POS and NEG groups (Bonferroni  $P > .05$ ).

Linear regression modeling of COVID-FAMILY from PHQ9-SUM with and without DEPRESSION-HX, DEPRESSION-YRS, and COVID-DIAGNOSIS as covariates

suggested confounding in the relationship between PHQ9-SUM and COVID-FAMILY ( $|\Delta B| > 0.1$ ), with an adjusted B coefficient not significantly more or less than 0 (Bonferroni  $P > .05$ ; adjusted B reported).

### **Analysis of Individual PHQ-9 Questions Against COVID-19 Status**

#### **COVID-DIAGNOSIS**

The median scores for PHQ9-4, PHQ9-5, PHQ9-6, PHQ9-7, PHQ9-8, and PHQ9-9 were higher in the COVID-DIAGNOSIS POS group than in the NEG group (Bonferroni  $P < .05$ ; Table 6). The median scores for PHQ9-1, PHQ9-2, and PHQ9-3 were not significantly different between the COVID-DIAGNOSIS POS and NEG groups (Bonferroni  $P > .05$ ).

Linear regressions individually modeling PHQ9-4, PHQ9-5, PHQ9-6, PHQ9-7, PHQ9-8, and PHQ9-9 from COVID-DIAGNOSIS with and without age, gender, and income as covariates suggested that the relationship of COVID-DIAGNOSIS with PHQ9-4, PHQ9-6, and PHQ9-7 was not subject to confounding ( $|\Delta B| < 0.1$ ), with crude B coefficients for COVID-DIAGNOSIS being greater than 0 (Bonferroni  $P < .05$ ; crude B reported). The models modeling PHQ9-4 and PHQ9-6 from COVID-DIAGNOSIS suggested confounding based on age, gender, or income ( $|\Delta B| > 0.1$ ). However, the adjusted B coefficients for COVID-DIAGNOSIS were significantly greater than 0 (Bonferroni  $P < .05$ ; adjusted B reported).



**Table 6.** Separation of Patient Health Questionnaire-9 (PHQ-9) item variables by COVID-19 status.

Variable	Descriptive statistics		Mann-Whitney <i>U</i> test			Linear regression confounder analysis			
	NEG group, median (IQR)	POS group, median (IQR)	<i>U</i> statistic	<i>P</i> value	Bonferroni <i>P</i> value	$\Delta$ B	B (95% CI)	<i>P</i> value	Bonferroni <i>P</i> value
<b>COVID-DIAGNOSIS (NEG: n=340; POS: n=39)</b>									
PHQ9-1	1.0 (1.0)	1.0 (2.0)	5602.5	.05	.41	__ <sup>a</sup>	—	—	—
PHQ9-2	1.0 (2.0)	1.0 (2.0)	5726.5	.07	.62	—	—	—	—
PHQ9-3	1.0 (2.0)	1.0 (2.0)	6319.5	.31	>.99	—	—	—	—
PHQ9-4	1.0 (2.0)	2.0 (1.0)	4963.5	.004	.03	0.05	0.46 (0.31 to 0.60)	<.001	<.001
PHQ9-5	0.0 (2.0)	1.0 (2.0)	5022.0	.004	.03	0.00	0.41 (0.25 to 0.58)	<.001	<.001
PHQ9-6	0.0 (2.0)	2.0 (1.0)	4049.0	<.001	<.001	0.13	0.78 (0.62 to 0.94)	<.001	<.001
PHQ9-7	0.0 (1.0)	1.0 (1.5)	4319.0	<.001	<.001	0.19	0.65 (0.50 to 0.80)	<.001	<.001
PHQ9-8	0.0 (1.0)	1.0 (2.0)	4078.0	<.001	<.001	0.17	0.68 (0.53 to 0.83)	<.001	<.001
PHQ9-9	0.0 (1.0)	1.0 (2.0)	4904.0	.001	.005	0.21	0.54 (0.39 to 0.70)	<.001	<.001
<b>COVID-FAMILY (NEG: n=276; POS: n=103)</b>									
PHQ9-1	1.0 (1.0)	1.0 (2.0)	13,012.5	.09	.80	—	—	—	—
PHQ9-2	1.0 (2.0)	1.0 (2.0)	12,291.0	.02	.14	—	—	—	—
PHQ9-3	1.0 (2.0)	1.0 (2.0)	12,577.0	.04	.32	—	—	—	—
PHQ9-4	1.0 (2.0)	1.0 (2.0)	12,579.5	.04	.33	—	—	—	—
PHQ9-5	0.0 (2.0)	1.0 (2.0)	12,315.5	.02	.14	—	—	—	—
PHQ9-6	0.0 (2.0)	1.0 (2.0)	11,674.0	.002	.02	0.90	0.03 (−0.14 to 0.20)	.75	>.99
PHQ9-7	0.0 (1.0)	1.0 (2.0)	12,635.5	.03	.31	—	—	—	—
PHQ9-8	0.0 (1.0)	0.0 (2.0)	11,903.5	.002	.02	1.18	−0.04 (−0.19 to 0.11)	.61	>.99
PHQ9-9	0.0 (1.0)	0.0 (1.5)	11,671.5	<.001	.004	0.32	0.19 (0.03 to 0.35)	.02	.20

<sup>a</sup>Regression analysis was not performed given the nonsignificant Mann-Whitney *U* test.

### COVID-FAMILY

The median scores for PHQ9-6, PHQ9-8, and PHQ9-9 were higher in the COVID-FAMILY POS group than in the NEG group (Bonferroni  $P < .05$ ; Table 6). The median scores for PHQ9-1, PHQ9-2, PHQ9-3, PHQ9-4, PHQ9-5, and PHQ9-7 were not significantly different between the COVID-FAMILY POS and NEG groups (Bonferroni  $P > .05$ ).

Linear regressions individually modeling PHQ9-6, PHQ9-8, and PHQ9-9 from COVID-FAMILY with and without DEPRESSION-HX, DEPRESSION-YRS, and COVID-DIAGNOSIS as covariates suggested that the relationships between COVID-FAMILY and all 3 PHQ-9 item scores were subject to confounding ( $|\Delta B| > 0.1$ ). Moreover, the

adjusted B coefficients for COVID-FAMILY were not significantly more or less than 0 in all the models (Bonferroni  $P > .05$ ; adjusted B reported).

### Analysis of STB Scores Against COVID-19 Status

#### COVID-DIAGNOSIS

This analysis addressed whether the 10 STB questions for the past month (short term, denoted as ST) and between 1 and 12 months ago (midterm, denoted as MT) were greater in the COVID-DIAGNOSIS POS group than in the NEG group. Scores for all STB questions except for S-SAFETY-LT (LT referring to long term) were significantly greater in the COVID-DIAGNOSIS POS group than in the NEG group (Bonferroni  $P < .05$ ; Table 7).

**Table 7.** Separation of suicidal thoughts and behavior variables by COVID-19 status.

Variable	Descriptive statistics		Mann-Whitney <i>U</i> test			Linear regression confounder analysis				
	NEG group, median (IQR)	POS group, median (IQR)	<i>U</i> statistic	<i>P</i> value	Bonferroni <i>P</i> value	$\Delta$ B	B (95% CI)	<i>P</i> value	Bonferroni <i>P</i> value	
<b>COVID-DIAGNOSIS ( NEG: n=340; POS: n=39 )</b>										
S-PASSIVE-LT	1.0 (1.0)	3.0 (3.0)	3900.0	<.001	<.001	0.28	0.73 (0.53 to 0.94)	<.001	<.001	
S-PASSIVE-MT	1.0 (1.0)	3.0 (3.0)	3976.0	<.001	<.001	0.40	0.75 (0.54 to 0.95)	<.001	<.001	
S-PASSIVE-ST	1.0 (1.0)	3.0 (3.0)	4000.0	<.001	<.001	0.16	0.77 (0.57 to 0.97)	<.001	<.001	
S-ACTIVE-LT	1.0 (0.0)	2.0 (3.0)	4466.5	<.001	<.001	0.39	0.52 (0.31 to 0.72)	<.001	<.001	
S-ACTIVE-MT	1.0 (0.0)	2.0 (3.0)	4131.5	<.001	<.001	0.22	0.69 (0.48 to 0.89)	<.001	<.001	
S-ACTIVE-ST	1.0 (0.0)	3.0 (3.0)	3719.5	<.001	<.001	0.29	0.88 (0.67 to 1.09)	<.001	<.001	
S-PLAN-LT	1.0 (0.0)	2.0 (3.0)	3809.0	<.001	<.001	0.38	0.80 (0.60 to 1.01)	<.001	<.001	
S-PLAN-MT	1.0 (0.0)	3.0 (3.0)	3950.5	<.001	<.001	0.48	0.68 (0.47 to 0.88)	<.001	<.001	
S-PLAN-ST	1.0 (0.0)	2.0 (3.0)	4128.0	<.001	<.001	0.60	0.44 (0.25 to 0.63)	<.001	<.001	
S-HISTORY-LT	1.0 (0.0)	3.0 (3.0)	3832.0	<.001	<.001	0.28	0.76 (0.58 to 0.95)	<.001	<.001	
S-HISTORY-MT	1.0 (0.0)	3.0 (3.0)	3538.0	<.001	<.001	0.32	0.76 (0.58 to 0.94)	<.001	<.001	
S-HISTORY-ST	1.0 (0.0)	2.0 (3.0)	4055.0	<.001	<.001	0.28	0.69 (0.50 to 0.87)	<.001	<.001	
S-SAFETY-LT	1.0 (2.0)	2.0	5449.5	.02	.25	__a	—	—	—	
S-SAFETY-MT	1.0 (2.0)	2.0 (3.0)	5025.0	.002	.03	0.93	0.04 (–0.17 to 0.26)	0.68	>.99	
S-SAFETY-ST	1.0 (2.0)	2.0 (3.0)	4752.0	<.001	.005	0.81	0.13 (–0.08 to 0.35)	0.22	>.99	
<b>COVID-FAMILY ( NEG: n=276; POS: n=103 )</b>										
S-PASSIVE-LT	1.0 (1.0)	1.0 (2.0)	12,915.5	.05	.77	—	—	—	—	
S-PASSIVE-MT	1.0 (1.0)	1.0 (2.0)	12,959.5	.06	.85	—	—	—	—	
S-PASSIVE-ST	1.0 (1.0)	1.0 (2.0)	12,363.0	.008	.13	—	—	—	—	
S-ACTIVE-LT	1.0 (1.0)	1.0 (2.0)	12,363.0	.008	.13	—	—	—	—	
S-ACTIVE-MT	1.0 (1.0)	1.0 (1.0)	13,318.0	.11	>.99	—	—	—	—	
S-ACTIVE-ST	1.0 (0.0)	1.0 (1.0)	12,787.0	.03	.39	—	—	—	—	
S-PLAN-LT	1.0 (0.0)	1.0 (1.0)	12,990.0	.04	.65	—	—	—	—	
S-PLAN-MT	1.0 (0.25)	1.0 (2.0)	12,316.5	.006	.09	—	—	—	—	
S-PLAN-ST	1.0 (0.0)	1.0 (2.0)	12,545.5	.01	.14	—	—	—	—	
S-HISTORY-LT	1.0 (0.0)	1.0 (1.0)	13,056.5	.05	.74	—	—	—	—	
S-HISTORY-MT	1.0 (1.0)	1.0 (2.0)	12,622.0	.02	.26	—	—	—	—	
S-HISTORY-ST	1.0 (0.0)	1.0 (1.0)	12,839.5	.03	.42	—	—	—	—	
S-SAFETY-LT	1.0 (0.0)	1.0 (0.5)	13,264.5	.08	>.99	—	—	—	—	
S-SAFETY-MT	1.0 (2.0)	1.0 (2.0)	13,360.0	.15	>.99	—	—	—	—	

Variable	Descriptive statistics		Mann-Whitney <i>U</i> test			Linear regression confounder analysis			
	NEG group, median (IQR)	POS group, median (IQR)	<i>U</i> statistic	<i>P</i> value	Bonferroni <i>P</i> value	$\Delta$ B	B (95% CI)	<i>P</i> value	Bonferroni <i>P</i> value
S-SAFETY-ST	1.0 (2.0)	1.0 (2.0)	13,409.0	.16	>.99	—	—	—	—
S-PASSIVE-LT	1.0 (2.0)	1.0 (2.0)	12,881.5	.05	.75	—	—	—	—

<sup>a</sup>Regression analysis was not performed given the nonsignificant Mann-Whitney *U* test.

Linear regressions individually modeling all STB scores, excluding S-SAFETY-LT, from the COVID-DIAGNOSIS group with and without age, gender, and income as covariates suggested that the relationships between COVID-DIAGNOSIS and all 14 scores were confounded ( $|\Delta B| > 0.1$ ). However, the adjusted B coefficients for COVID-DIAGNOSIS were significantly greater than 0 (Bonferroni  $P < .05$ ; adjusted B reported) in all models except those modeling S-SAFETY-MT and S-SAFETY-ST.

Across time frames (ie, past 1 year or more, past 1 month to 1 year, and past 1 month), significant effects (Bonferroni  $P < .05$ ) were only observed for COVID-DIAGNOSIS NEG participants for having a suicide plan. No other STB variables differed across time frames for participants with a positive or negative COVID-19 diagnosis. It should be noted that similar results were observed for COVID-TEST NEG participants, who also showed differences over time frames in desire for self-harm.

### COVID-FAMILY

This analysis addressed the same question as COVID-DIAGNOSIS but for COVID-FAMILY. The scores for all STB questions did not significantly differ between the COVID-FAMILY POS and NEG groups (Bonferroni  $P > .05$ ; Table 7). Consequently, no linear regression was performed for these scores.

Across time frames (ie, past 1 year or more, past 1 month to 1 year, and past 1 month), significant effects (Bonferroni  $P < .05$ ) were only observed for COVID-FAMILY NEG participants for having a suicide plan.

## Discussion

### Principal Findings

This study produced 4 sets of findings. One set of findings showed an increase in depression and anxiety scores for individuals with COVID-19 infection history compared to those without, and showed this despite a similar incidence of prior history of depression or anxiety in those with and without COVID-19 history. The second set of findings was that the depression-associated symptoms reported by individuals with COVID-19 history included elevated fatigue, problems with concentration, psychomotor retardation or agitation, altered appetite, feelings of guilt, and elevated suicidality. The first 4 of these symptoms overlap significantly with PCS and raise the hypothesis that the experience of PCS and guilt associated with contracting COVID-19 may drive an increase in total depression-associated symptoms and diagnosis. The third set of findings showed that elevated STB scores in individuals with COVID-19 history preceded COVID-19 diagnosis, suggesting the possibility that participants with greater pre-existing STB

were more predisposed to contracting COVID-19. Furthermore, time frame analysis indicated that differences in STB before and during the pandemic were more pronounced in those without COVID-19 history, raising the possibility that the experience of COVID-19 had a greater effect on STB in the uninfected population. Finally, the fourth set of findings showed that an increase in depression-associated symptoms in those with family or friends adversely affected by COVID-19 might correlate with an elevation in prior history of depression. Moreover, the increase in depression symptoms was no greater than expected in the general population. In what follows, each of these findings is discussed in greater detail.

This study found that participants reporting a COVID-19 clinician diagnosis (COVID-DIAGNOSIS POS) tended to produce greater cumulative depression (PHQ9-SUM) and state anxiety (STAI-SUM) scores than those denying a diagnosis of COVID-19 (NEG). Of note, the median cumulative depression score for those reporting COVID-19 fell within a range associated with mild depression (13.0), whereas the median score for those denying COVID-19 fell within a range associated with minimal depression (6.0). We found no significant difference in self-reported history of depression (DEPRESSION-HX and DEPRESSION-YRS) or anxiety disorder (ANXIETY-HX or ANXIETY-YRS) between those with and without COVID-19 history, and no evidence for demographic variables confounding the relationships between cumulative depression scores and COVID-19 infection. This suggested that experiencing COVID-19 was the primary driver of participants' increased depression- and anxiety-associated symptoms at the time of participation. Although the durations between participants' diagnoses and survey participation times were not precisely known (however, as noted in the Methods, over 95% of participants were expected to have participated in the questionnaire >2 weeks following infection), the specific symptom profiles related to increases in PHQ-9 scores were consistent with PCS, further suggesting that survey participation occurred after primary COVID-19 infection resolution. This agrees with a previous work that observed increased depression and anxiety scores up to 6 months after infection [21-24].

Five specific depression-associated symptom scores were elevated in those reporting COVID-19 history. One might hypothesize that PHQ9-4, which assesses fatigue; PHQ9-7, which assesses difficulty with concentration; and PHQ9-8, which assesses unusually slow or restless behavior, are directly capturing typical PCS symptoms including fatigue and "brain fog" (or concentration loss) [3,4,27]. PHQ9-5 assesses poor appetite and overeating and may capture anosmia and ageusia associated with PCS. PHQ9-6 assesses feelings of guilt, which may capture feelings of guilt associated with the contraction of COVID-19 [31], as well as some participants' feelings of guilt

in spreading COVID-19 to close relatives. PHQ-9 assesses suicidal ideation and needs to be interpreted in the context of the broader set of STB questions discussed in the next section. Regarding the clinical significance of the 1- or 2-point differences in the PHQ-9 item score medians between the groups, it is important to note that a 1-point increase corresponds to one-third of the entire symptom scoring interval. A movement from 0 to 1, which was frequently observed, corresponds to a shift from the absence of symptoms to symptoms appearing “several days” over the past 2 weeks. These observations were noted in the absence of significant differences between the groups for history of depression or years of depression. Given the significant statistical findings after correction for multiple comparisons, the detected changes in PHQ-9 item scores are likely to be clinically significant. Overall, the current results around specific depression-associated symptoms are consistent with the hypothesis that the experience of PCS and guilt associated with contracting COVID-19 drive an increase in cumulative depression score and a population shift toward a higher frequency of depression diagnosis among people who contract COVID-19.

A caveat of this analysis is that although PCS-associated symptoms may drive depression scores and, in turn, the expected number of depression diagnoses upward, we cannot state whether depression symptoms associated with PCS contribute to an increase in “true” cases of depression. For instance, whether fatigue associated with PCS is pathophysiologically similar enough to fatigue associated with standard depression to produce similar disease courses and treatment responses is unclear. Although this requires further investigation, we believe that depression is already considered a highly heterogeneous diagnosis [42,43]. Moreover, evidence suggests that biomarkers of PCS-associated depression align with those already linked to major depressive disorders [44].

Along with the PHQ-9 question about suicidal ideation, we assessed 5 STB questions over 3 time windows. We observed that the median scores in the COVID-19–positive group were almost universally greater than those in the COVID-19–negative group. Median scores in the COVID-19–positive group largely suggested STB symptoms were experienced “rarely” or “sometimes,” whereas median scores in the COVID-19–negative group suggested STB symptoms were experienced “never.” Given that many of the STB symptoms assessed (in particular, active ideation, suicide planning, and suicide history) are considered abnormal at any frequency, the statistically significant differences detected herein are likely to be clinically relevant. Most of these differences were robust to potential confounders, with the notable exception of scores corresponding to questions about a suicide safety plan (S-SAFETY-LT, S-SAFETY-MT, and S-SAFETY-ST). Importantly, the differences in median scores for questions inquiring about STB more than 12 months ago—prior to the approximate onset time of the COVID-19 pandemic in the United States and when we expect our participants to have contracted COVID-19—imply that elevated STB scores preceded COVID-19 diagnosis. One possible interpretation is that participants with greater pre-existing STB were more predisposed to contracting COVID-19. This interpretation is consistent with previous work

showing that observed STB is associated with a contemporaneous desire for deliberate COVID-19 exposure among the general population [13] and further suggests that elevated STB predating COVID-19 may be associated with actualized contraction of COVID-19. This is also linked to earlier reports that people with increased STB deliberately sought HIV exposure [45-47]. It did not appear that scores for STB questions differed among short-term, midterm, and long-term versions of the same question within those reporting COVID-19 infection, suggesting that the contraction of COVID-19 did not correlate with an increase in suicidality. It should be noted that scores on short-term suicidality questions (inquiring about the past month) remained higher in those reporting COVID-19 infection than in those who did not, which corroborates previous findings that median PHQ-9 scores were greater in those reporting COVID-19 infection.

This study observed an increase in depression-associated symptoms among those with family or friends adversely affected by COVID-19. Although we found that cumulative depression scores and 3 specific PHQ-9 questions were elevated in these participants (ie, PHQ-6, PHQ-8, and PHQ-9), other analyses suggested that these differences might be explained by depression history variables (eg, DEPRESSION-HX, DEPRESSION-YRS, and COVID-DIAGNOSIS control variables). Thus, no unique relationships could be established between COVID-FAMILY and any other variables studied. Although we are surprised by this finding, we suspect that participants who were emotionally closer to friends or family with COVID-19 were likely to be physically close to them and to be among the 22.3% (23/103) of participants who also became infected with COVID-19. Our results may suggest that people who witnessed more distant relations develop severe or fatal COVID-19 may experience depression, anxiety, and STB at a level equivalent to the portion of the general population that has also been psychologically affected by COVID-19.

### Limitations

The primary limitation of this study is the small sample size: 379 participants of whom 39 (10.3%) reported a COVID-19 diagnosis. As a result, our statistical power was limited, and we were unable to substantively consider confounding the main relationships by using categorical demographic variables reflecting ethnicity and employment. In addition, with respect to our target variables, we could not always distinguish between participants infected with COVID-19 and those with family or friends adversely affected by COVID-19. This resulted from a large population overlap, likely reflecting a general tendency for COVID-19 infection to be shared among close friends or family members. Therefore, it may be fundamentally difficult to separate the personal experience of having COVID-19 from experiencing a friend or close family member with it. Relatedly, it is possible that symptoms of grief may present similarly to symptoms of PHQ-9. Although we did not observe increased PHQ-9 scores among participants who experienced COVID-19 among close friends or family members after adjustment for personal COVID-19 status, we note that any PHQ-9 score elevation in this context may have reflected an increase in grief versus depression.

As a questionnaire-based study, our data are subject to forms of response bias, including acquiescence bias (likely leading to inflated PHQ-9, STAI, and STB scores), social desirability bias (likely leading to deflated PHQ-9, STAI, and STB scores, as well as potential inauthenticity about positive COVID-19 status), and extreme responding (likely leading to extreme PHQ-9, STAI, and STB scores). We expect a degree of this bias to be uncorrelated and therefore be mitigated in our core analyses, which involved internal comparisons of participants with respect to COVID-19 status. Although we warn that the existence of these biases suggests that group-wide medians for PHQ-9, STAI, and STB may be elevated, we note that the same phenomenon would be expected when these tools are deployed in a clinical setting—perhaps to an even greater extent because responses directly influence patient medical management and relationships with providers. Most concerning, we admit the possibility that correlations between the misreporting of depression, anxiety, or COVID-19 history and biased PHQ-9, STAI, and STB responses may exist that compromise the study's main finding. Although readers should be mindful of this possibility and seek to corroborate this study's findings with other studies that do not depend on questionnaires, this possibility reflects a universal weakness of questionnaire-based studies and studies done on a big data scale that cannot easily incorporate laboratory-based experiments.

Other limitations include that the survey participants may have been subject to recall bias and specifically have conflated current STB with past STB. Regarding the question order bias, questions pertaining to depression and anxiety history, COVID-19 history, and PHQ-9/STAI/STB were all separated by more than 10 items. In addition, this study would have benefited from collecting the

date of COVID-19 diagnosis or positive test and examining the duration of acute symptoms to accommodate a more precise analysis (eg, consideration of PCS symptomology with respect to illness duration and distance from diagnosis). Relatedly, it is possible that 5% of our COVID-19–positive participants may have been actively subject to acute COVID-19 symptoms. Lastly, possession of occupational information about the participants would have brought more depth to the analyses with regard to the targeted COVID-19 health outcomes.

This study examined US survey participants' current and recent mental health in association with the change in COVID-19 status between the onset of the COVID-19 pandemic in the United States (approximately March 2020) and the time of data collection (early March 2021). In the context of its limitations, it was found that cumulative depression and anxiety scores were significantly higher in those reporting COVID-19 infection despite a similar prior diagnostic history of depression or anxiety in those with and without COVID-19 infection. The majority of depression-associated symptoms overlapped with those reported for PCS, and reports of increased STB commonly preceded the onset of the pandemic. Where there were STB differences across time frames related to times before and during the pandemic, significant differences were only observed in those who had not had COVID-19 infection. Lastly, increases in depression-associated symptoms were observed in those with family or friends adversely affected by COVID-19, which appeared to be related to an increase in prior history of depression in this group. Altogether, these observations argue that the relationship of COVID-19 with depression or anxiety diagnoses and STB is not obvious and will require a more detailed study along with serial longitudinal assessments.

---

## Acknowledgments

SB and NV share second authorship. AKK and HB share corresponding authorship. Funding for this work was provided to HB by the Office of Naval Research (award N00014-21-1-2216) and by HB from the Warren Wright Adolescent Center at Northwestern University, Feinberg School of Medicine.

---

## Authors' Contributions

SFW, SB, NV, AKK, and HB provided the study concept and design. Acquisition of original data was performed by SFW, SB, NV, AKK, and HB, and coding of statistical tools was done by SFW (with input from NV, SB, AKK, and HB). Analysis of data was performed by SFW (with input from SB, NV, HB, and AK) and interpretation of data by SFW, SB, NV, AKK, and HB (with input from KS, SL, LS, BWK, and NM). Statistical assessment was performed by SFW (with input from SB, NV, AKK, and HB). The original draft was prepared by SFW (with input from HB and AKK). SFW generated the figures, and revision of the manuscript for content was done by all authors. All authors approved the final version of the paper for submission.

---

## Conflicts of Interest

None declared.

---

## Multimedia Appendix 1

This is a supplement with additional information about ethical approval, participant recruitment, and statistical methods. [[DOCX File, 230 KB - formative\\_v6i10e36656\\_app1.docx](#)]

---

## References

1. COVID data tracker. Centers for Disease Control and Prevention. URL: <https://covid.cdc.gov/covid-data-tracker/> [accessed 2022-07-20]

2. Hu B, Guo H, Zhou P, Shi Z. Characteristics of SARS-CoV-2 and COVID-19. *Nat Rev Microbiol* 2021 Mar;19(3):141-154 [FREE Full text] [doi: [10.1038/s41579-020-00459-7](https://doi.org/10.1038/s41579-020-00459-7)] [Medline: [33024307](https://pubmed.ncbi.nlm.nih.gov/33024307/)]
3. Augustin M, Schommers P, Stecher M, Dewald F, Gieselmann L, Gruell H, et al. Post-COVID syndrome in non-hospitalised patients with COVID-19: a longitudinal prospective cohort study. *Lancet Reg Health Eur* 2021 Jul;6:100122 [FREE Full text] [doi: [10.1016/j.lanepe.2021.100122](https://doi.org/10.1016/j.lanepe.2021.100122)] [Medline: [34027514](https://pubmed.ncbi.nlm.nih.gov/34027514/)]
4. 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](https://pubmed.ncbi.nlm.nih.gov/32644129/)]
5. Garrigues E, Janvier P, Kherabi Y, Le Bot A, Hamon A, Gouze H, et al. Post-discharge persistent symptoms and health-related quality of life after hospitalization for COVID-19. *J Infect* 2020 Dec;81(6):e4-e6 [FREE Full text] [doi: [10.1016/j.jinf.2020.08.029](https://doi.org/10.1016/j.jinf.2020.08.029)] [Medline: [32853602](https://pubmed.ncbi.nlm.nih.gov/32853602/)]
6. 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/)]
7. Huang Y, Zhao N. Generalized anxiety disorder, depressive symptoms and sleep quality during COVID-19 outbreak in China: a web-based cross-sectional survey. *Psychiatry Res* 2020 Jun;288:112954 [FREE Full text] [doi: [10.1016/j.psychres.2020.112954](https://doi.org/10.1016/j.psychres.2020.112954)] [Medline: [32325383](https://pubmed.ncbi.nlm.nih.gov/32325383/)]
8. Ustun G. Determining depression and related factors in a society affected by COVID-19 pandemic. *Int J Soc Psychiatry* 2021 Feb;67(1):54-63 [FREE Full text] [doi: [10.1177/0020764020938807](https://doi.org/10.1177/0020764020938807)] [Medline: [32605422](https://pubmed.ncbi.nlm.nih.gov/32605422/)]
9. Forte G, Favieri F, Tambelli R, Casagrande M. The enemy which sealed the world: effects of COVID-19 diffusion on the psychological state of the Italian population. *J Clin Med* 2020 Jun 10;9(6):1802 [FREE Full text] [doi: [10.3390/jcm9061802](https://doi.org/10.3390/jcm9061802)] [Medline: [32531884](https://pubmed.ncbi.nlm.nih.gov/32531884/)]
10. Brooks SK, Webster RK, Smith LE, Woodland L, Wessely S, Greenberg N, et al. The psychological impact of quarantine and how to reduce it: rapid review of the evidence. *Lancet* 2020 Mar 14;395(10227):912-920 [FREE Full text] [doi: [10.1016/S0140-6736\(20\)30460-8](https://doi.org/10.1016/S0140-6736(20)30460-8)] [Medline: [32112714](https://pubmed.ncbi.nlm.nih.gov/32112714/)]
11. Yip PS, Chau PH. Physical distancing and emotional closeness amidst COVID-19. *Crisis* 2020 May;41(3):153-155 [FREE Full text] [doi: [10.1027/0227-5910/a000710](https://doi.org/10.1027/0227-5910/a000710)] [Medline: [32299225](https://pubmed.ncbi.nlm.nih.gov/32299225/)]
12. Zhang F, Shang Z, Ma H, Jia Y, Sun L, Guo X, et al. Epidemic area contact history and sleep quality associated with posttraumatic stress symptoms in the first phase of COVID-19 outbreak in China. *Sci Rep* 2020 Dec 31;10(1):22463 [FREE Full text] [doi: [10.1038/s41598-020-80649-8](https://doi.org/10.1038/s41598-020-80649-8)] [Medline: [33384438](https://pubmed.ncbi.nlm.nih.gov/33384438/)]
13. Ammerman BA, Burke TA, Jacobucci R, McClure K. Preliminary investigation of the association between COVID-19 and suicidal thoughts and behaviors in the U.S. *J Psychiatr Res* 2021 Feb;134:32-38. [doi: [10.1016/j.jpsychires.2020.12.037](https://doi.org/10.1016/j.jpsychires.2020.12.037)] [Medline: [33360222](https://pubmed.ncbi.nlm.nih.gov/33360222/)]
14. O'Connor RC, Wetherall K, Cleare S, McClelland H, Melson AJ, Niedzwiedz CL, et al. *Br J Psychiatry* 2021 Jun;218(6):326-333 [FREE Full text] [doi: [10.1192/bjp.2020.212](https://doi.org/10.1192/bjp.2020.212)] [Medline: [33081860](https://pubmed.ncbi.nlm.nih.gov/33081860/)]
15. Czeisler M, Lane RI, Petrosky E, Wiley JF, Christensen A, Njai R, et al. Mental health, substance use, and suicidal ideation during the COVID-19 pandemic - United States, June 24-30, 2020. *MMWR Morb Mortal Wkly Rep* 2020 Aug 14;69(32):1049-1057 [FREE Full text] [doi: [10.15585/mmwr.mm6932a1](https://doi.org/10.15585/mmwr.mm6932a1)] [Medline: [32790653](https://pubmed.ncbi.nlm.nih.gov/32790653/)]
16. Thakur V, Jain A. COVID 2019-suicides: a global psychological pandemic. *Brain Behav Immun* 2020 Aug;88:952-953 [FREE Full text] [doi: [10.1016/j.bbi.2020.04.062](https://doi.org/10.1016/j.bbi.2020.04.062)] [Medline: [32335196](https://pubmed.ncbi.nlm.nih.gov/32335196/)]
17. Tandon R. COVID-19 and suicide: just the facts. Key learnings and guidance for action. *Asian J Psychiatr* 2021 Jun;60:102695 [FREE Full text] [doi: [10.1016/j.ajp.2021.102695](https://doi.org/10.1016/j.ajp.2021.102695)] [Medline: [33992557](https://pubmed.ncbi.nlm.nih.gov/33992557/)]
18. Hoekstra PJ. Suicidality in children and adolescents: lessons to be learned from the COVID-19 crisis. *Eur Child Adolesc Psychiatry* 2020 Jun;29(6):737-738 [FREE Full text] [doi: [10.1007/s00787-020-01570-z](https://doi.org/10.1007/s00787-020-01570-z)] [Medline: [32488455](https://pubmed.ncbi.nlm.nih.gov/32488455/)]
19. Raj S, Ghosh D, Singh T, Verma SK, Arya YK. Theoretical mapping of suicidal risk factors during the COVID-19 pandemic: a mini-review. *Front Psychiatry* 2020;11:589614 [FREE Full text] [doi: [10.3389/fpsy.2020.589614](https://doi.org/10.3389/fpsy.2020.589614)] [Medline: [33551864](https://pubmed.ncbi.nlm.nih.gov/33551864/)]
20. Guo Q, Zheng Y, Shi J, Wang J, Li G, Li C, et al. Immediate psychological distress in quarantined patients with COVID-19 and its association with peripheral inflammation: a mixed-method study. *Brain Behav Immun* 2020 Aug;88:17-27 [FREE Full text] [doi: [10.1016/j.bbi.2020.05.038](https://doi.org/10.1016/j.bbi.2020.05.038)] [Medline: [32416290](https://pubmed.ncbi.nlm.nih.gov/32416290/)]
21. Hu Y, Chen Y, Zheng Y, You C, Tan J, Hu L, et al. Factors related to mental health of inpatients with COVID-19 in Wuhan, China. *Brain Behav Immun* 2020 Oct;89:587-593 [FREE Full text] [doi: [10.1016/j.bbi.2020.07.016](https://doi.org/10.1016/j.bbi.2020.07.016)] [Medline: [32681866](https://pubmed.ncbi.nlm.nih.gov/32681866/)]
22. Ismael F, Bizario JC, Battagin T, Zaramella B, Leal FE, Torales J, et al. Post-infection depressive, anxiety and post-traumatic stress symptoms: a prospective cohort study in patients with mild COVID-19. *Prog Neuropsychopharmacol Biol Psychiatry* 2021 Dec 20;111:110341 [FREE Full text] [doi: [10.1016/j.pnpbp.2021.110341](https://doi.org/10.1016/j.pnpbp.2021.110341)] [Medline: [33940097](https://pubmed.ncbi.nlm.nih.gov/33940097/)]
23. Raman B, Cassar MP, Tunnicliffe EM, Filippini N, Griffanti L, Alfaro-Almagro F, et al. Medium-term effects of SARS-CoV-2 infection on multiple vital organs, exercise capacity, cognition, quality of life and mental health, post-hospital discharge. *EClinicalMedicine* 2021 Jan;31:100683 [FREE Full text] [doi: [10.1016/j.eclinm.2020.100683](https://doi.org/10.1016/j.eclinm.2020.100683)] [Medline: [33490928](https://pubmed.ncbi.nlm.nih.gov/33490928/)]

24. Taquet M, Geddes JR, Husain M, Luciano S, Harrison PJ. 6-month neurological and psychiatric outcomes in 236379 survivors of COVID-19: a retrospective cohort study using electronic health records. *Lancet Psychiatry* 2021 May;8(5):416-427 [FREE Full text] [doi: [10.1016/S2215-0366\(21\)00084-5](https://doi.org/10.1016/S2215-0366(21)00084-5)] [Medline: [33836148](https://pubmed.ncbi.nlm.nih.gov/33836148/)]
25. Lam MH, Wing Y, Yu MW, Leung C, Ma RC, Kong AP, et al. Mental morbidities and chronic fatigue in severe acute respiratory syndrome survivors: long-term follow-up. *Arch Intern Med* 2009 Dec 14;169(22):2142-2147. [doi: [10.1001/archinternmed.2009.384](https://doi.org/10.1001/archinternmed.2009.384)] [Medline: [20008700](https://pubmed.ncbi.nlm.nih.gov/20008700/)]
26. Lee AM, Wong JG, McAlonan GM, Cheung V, Cheung C, Sham PC, et al. Stress and psychological distress among SARS survivors 1 year after the outbreak. *Can J Psychiatry* 2007 Apr;52(4):233-240. [doi: [10.1177/070674370705200405](https://doi.org/10.1177/070674370705200405)] [Medline: [17500304](https://pubmed.ncbi.nlm.nih.gov/17500304/)]
27. Vanichkachorn G, Newcomb R, Cowl CT, Murad MH, Breeher L, Miller S, et al. Post-COVID-19 syndrome (long haul syndrome): description of a multidisciplinary clinic at Mayo clinic and characteristics of the initial patient cohort. *Mayo Clin Proc* 2021 Jul;96(7):1782-1791 [FREE Full text] [doi: [10.1016/j.mayocp.2021.04.024](https://doi.org/10.1016/j.mayocp.2021.04.024)] [Medline: [34218857](https://pubmed.ncbi.nlm.nih.gov/34218857/)]
28. Sher L. Are COVID-19 survivors at increased risk for suicide? *Acta Neuropsychiatr* 2020 Oct;32(5):270 [FREE Full text] [doi: [10.1017/neu.2020.21](https://doi.org/10.1017/neu.2020.21)] [Medline: [32364491](https://pubmed.ncbi.nlm.nih.gov/32364491/)]
29. Sher L. Post-COVID syndrome and suicide risk. *QJM* 2021 Apr 27;114(2):95-98 [FREE Full text] [doi: [10.1093/qjmed/hcab007](https://doi.org/10.1093/qjmed/hcab007)] [Medline: [33486531](https://pubmed.ncbi.nlm.nih.gov/33486531/)]
30. Epstein D, Andrawis W, Lipsky AM, Ziad HA, Matan M. Anxiety and suicidality in a hospitalized patient with COVID-19 infection. *Eur J Case Rep Intern Med* 2020;7(5):001651 [FREE Full text] [doi: [10.12890/2020\\_001651](https://doi.org/10.12890/2020_001651)] [Medline: [32399450](https://pubmed.ncbi.nlm.nih.gov/32399450/)]
31. Liu Y, Cao L, Li X, Jia Y, Xia H. Awareness of mental health problems in patients with coronavirus disease 19 (COVID-19): a lesson from an adult man attempting suicide. *Asian J Psychiatr* 2020 Jun;51:102106 [FREE Full text] [doi: [10.1016/j.ajp.2020.102106](https://doi.org/10.1016/j.ajp.2020.102106)] [Medline: [32334411](https://pubmed.ncbi.nlm.nih.gov/32334411/)]
32. Mazza MG, De Lorenzo R, Conte C, Poletti S, Vai B, Bollettini I, COVID-19 BioB Outpatient Clinic Study group, et al. Anxiety and depression in COVID-19 survivors: role of inflammatory and clinical predictors. *Brain Behav Immun* 2020 Oct;89:594-600 [FREE Full text] [doi: [10.1016/j.bbi.2020.07.037](https://doi.org/10.1016/j.bbi.2020.07.037)] [Medline: [32738287](https://pubmed.ncbi.nlm.nih.gov/32738287/)]
33. Sripad MN, Pantoji M, Gowda GS, Ganjekar S, Reddi VS, Math SB. Suicide in the context of COVID-19 diagnosis in India: insights and implications from online print media reports. *Psychiatry Res* 2021 Apr;298:113799 [FREE Full text] [doi: [10.1016/j.psychres.2021.113799](https://doi.org/10.1016/j.psychres.2021.113799)] [Medline: [33611165](https://pubmed.ncbi.nlm.nih.gov/33611165/)]
34. 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/)]
35. Spielberger C, Gorsuch R, Lushene R, Vagg P, Jacobs G. Manual for the State-Trait Anxiety Inventory. Palto Alto, CA: Consulting Psychologists Press; 1983.
36. WHO Coronavirus (COVID-19) Dashboard. World Health Organization. URL: <https://covid19.who.int/> [accessed 2021-07-01]
37. QuickFacts. United States Census Bureau. URL: <https://www.census.gov/quickfacts/fact/table/US#> [accessed 2021-10-23]
38. Viswanathan V, Sheppard JP, Kim BW, Plantz CL, Ying H, Lee MJ, et al. A quantitative relationship between signal detection in attention and approach/avoidance behavior. *Front Psychol* 2017;8:122 [FREE Full text] [doi: [10.3389/fpsyg.2017.00122](https://doi.org/10.3389/fpsyg.2017.00122)] [Medline: [28270776](https://pubmed.ncbi.nlm.nih.gov/28270776/)]
39. Hyman S. Manual of Psychiatric Emergencies. Boston, Massachusetts, United States: Little Brown and Company; 1994.
40. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: synthetic minority over-sampling technique. *J Artif Intell Res* 2002 Jun 01;16:321-357. [doi: [10.1613/jair.953](https://doi.org/10.1613/jair.953)]
41. Budtz-Jørgensen E, Keiding N, Grandjean P, Weihe P. Confounder selection in environmental epidemiology: assessment of health effects of prenatal mercury exposure. *Ann Epidemiol* 2007 Jan;17(1):27-35. [doi: [10.1016/j.annepidem.2006.05.007](https://doi.org/10.1016/j.annepidem.2006.05.007)] [Medline: [17027287](https://pubmed.ncbi.nlm.nih.gov/17027287/)]
42. Goldberg D. The heterogeneity of "major depression". *World Psychiatry* 2011 Oct;10(3):226-228 [FREE Full text] [doi: [10.1002/j.2051-5545.2011.tb00061.x](https://doi.org/10.1002/j.2051-5545.2011.tb00061.x)] [Medline: [21991283](https://pubmed.ncbi.nlm.nih.gov/21991283/)]
43. Jesulola E, Micalos P, Baguley IJ. Understanding the pathophysiology of depression: from monoamines to the neurogenesis hypothesis model - are we there yet? *Behav Brain Res* 2018 Apr 02;341:79-90. [doi: [10.1016/j.bbr.2017.12.025](https://doi.org/10.1016/j.bbr.2017.12.025)] [Medline: [29284108](https://pubmed.ncbi.nlm.nih.gov/29284108/)]
44. Lorkiewicz P, Waszkiewicz N. Biomarkers of post-COVID depression. *J Clin Med* 2021 Sep 14;10(18):4142 [FREE Full text] [doi: [10.3390/jcm10184142](https://doi.org/10.3390/jcm10184142)] [Medline: [34575258](https://pubmed.ncbi.nlm.nih.gov/34575258/)]
45. Frances RJ, Wikstrom T, Alcena V. Contracting AIDS as a means of committing suicide. *Am J Psychiatry* 1985 May;142(5):656. [doi: [10.1176/ajp.142.5.656b](https://doi.org/10.1176/ajp.142.5.656b)] [Medline: [3985206](https://pubmed.ncbi.nlm.nih.gov/3985206/)]
46. Papatomopoulos E. Intentional infection with the AIDS virus as a means of suicide. *Couns Psychol Q* 2007 Sep 27;2(1):79-81. [doi: [10.1080/09515078908256665](https://doi.org/10.1080/09515078908256665)]
47. Tompkins-Rosenblatt P. Intentional HIV contraction: implications for direct child and youth care. *Resident Treatment Children Youth* 1997 Apr 24;14(4):15-30. [doi: [10.1300/j007v14n04\\_02](https://doi.org/10.1300/j007v14n04_02)]

## Abbreviations

**PCS:** post-COVID-19 syndrome

**PHQ-9:** Patient Health Questionnaire-9  
**SMOTE:** synthetic minority oversampling technique  
**STAI:** State Trait Anxiety Index  
**STB:** suicidal thoughts and behavior

*Edited by A Mavragani; submitted 19.01.22; peer-reviewed by K Falkenhain, H Akram; comments to author 19.02.22; revised version received 30.04.22; accepted 01.05.22; published 25.10.22.*

*Please cite as:*

*Woodward SF, Bari S, Vike N, Lalvani S, Stetsiv K, Kim BW, Stefanopoulos L, Maglaveras N, Breiter H, Katsaggelos AK  
Anxiety, Post-COVID-19 Syndrome-Related Depression, and Suicidal Thoughts and Behaviors in COVID-19 Survivors: Cross-sectional Study*

*JMIR Form Res 2022;6(10):e36656*

*URL: <https://formative.jmir.org/2022/10/e36656>*

*doi: [10.2196/36656](https://doi.org/10.2196/36656)*

*PMID: [35763757](https://pubmed.ncbi.nlm.nih.gov/35763757/)*

©Sean F Woodward, Sumra Bari, Nicole Vike, Shamal Lalvani, Khrystyna Stetsiv, Byoung Woo Kim, Leandros Stefanopoulos, Nicos Maglaveras, Hans Breiter, Aggelos K Katsaggelos. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 25.10.2022. 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

# COVID-Bot, an Intelligent System for COVID-19 Vaccination Screening: Design and Development

Chinedu Wilfred Okonkwo<sup>1</sup>, PhD; Lateef Babatunde Amusa<sup>1</sup>, PhD; Hossana Twinomurizi<sup>1</sup>, PhD

Centre for Applied Data Sciences, College of Business and Economics, University of Johannesburg, Johannesburg, South Africa

**Corresponding Author:**

Chinedu Wilfred Okonkwo, PhD  
Centre for Applied Data Sciences  
College of Business and Economics  
University of Johannesburg  
Bunting Road Campus  
Auckland Park  
Johannesburg, 2006  
South Africa  
Phone: 27 0835796557  
Email: [chineduo@uj.ac.za](mailto:chineduo@uj.ac.za)

## Abstract

**Background:** Coronavirus continues to spread worldwide, causing various health and economic disruptions. One of the most important approaches to controlling the spread of this disease is to use an artificial intelligence (AI)-based technological intervention, such as a chatbot system. Chatbots can aid in the fight against the spread of COVID-19.

**Objective:** This paper introduces COVID-Bot, an intelligent interactive system that can help screen students and confirm their COVID-19 vaccination status.

**Methods:** The design and development of COVID-Bot followed the principles of the design science research (DSR) process, which is a research method for creating a new scientific artifact. COVID-Bot was developed and implemented using the SnatchBot chatbot application programming interface (API) and its predefined tools, which are driven by various natural language processing algorithms.

**Results:** An evaluation was carried out through a survey that involved 106 university students in determining the functionality, compatibility, reliability, and usability of COVID-Bot. The findings indicated that 92 (86.8%) of the participants agreed that the chatbot functions well, 85 (80.2%) agreed that it fits well with their mobile devices and their lifestyle, 86 (81.1%) agreed that it has the potential to produce accurate and consistent responses, and 85 (80.2%) agreed that it is easy to use. The average obtained  $\alpha$  was .87, indicating satisfactory reliability.

**Conclusions:** This study demonstrates that incorporating chatbot technology into the educational system can combat the spread of COVID-19 among university students. The intelligent system does this by interacting with students to determine their vaccination status.

(*JMIR Form Res* 2022;6(10):e39157) doi:[10.2196/39157](https://doi.org/10.2196/39157)

**KEYWORDS**

chatbot; COVID-Bot; COVID-19; students; vaccine; exemption letter; vaccination; artificial intelligence

## Introduction

**Background**

COVID-19 has spread worldwide, affecting many aspects of daily life, including education. This pandemic has resulted in the deaths of millions of people, job losses, illnesses, physical communication restrictions, and changes in organizational operating methods [1-3]. Various response strategies, including

lockdowns, social distancing, wearing of face masks, and vaccination, have been introduced to control the spread of the virus. Scientists have developed several vaccine candidates to help combat the disease, including Pfizer/BioNTech, Moderna mRNA vaccines, the Johnson and Johnson viral vector vaccine, and the AstraZeneca viral vector vaccine [4]. Although these vaccines aid in the fight against disease spread, other strategies are required to reduce coronavirus infection [5,6].

Artificial intelligence (AI) approaches can aid in the dissemination of critical information worldwide, as well as reduce incorrect information concerning COVID-19 [7,8]. AI-powered chatbot systems are the latest technology innovations used to combat the COVID-19 pandemic [9,10]. Chatbots are intelligent conversational agents interacting with users to respond to their questions [11]. Chatbots can be deployed on websites or social networking platforms, such as Facebook, Skype, and WhatsApp [12,13]. For example, the World Health Organization (WHO) technology initiative developed a chatbot to combat COVID-19; this system is accessible via Facebook and WhatsApp [14]. Individuals or users can use the chatbot to get answers to their questions about how to protect themselves from coronavirus, learn more about the disease, and help prevent its spread. A chatbot is highly beneficial because it can offer a single, reliable answer to most user inquiries and delivers brief information from trusted sources, which can be easier than an extensive list of results from online search engines or social media [15].

The existence of coronavirus is no longer in doubt, and people are returning to their everyday lives, while taking necessary precautions to avoid the disease. Students in higher education are returning to schools to begin classroom or face-to-face teaching and learning. Before a student is allowed to enter the school grounds, most institutions conduct COVID-19 screening at their various entrances. This screening is done manually, which causes some inconveniences, such as long lines, time waste, fatigue, and human errors. Furthermore, this process may pose a significant challenge to the authorities or administration, necessitating new technologies and automation. AI technology, such as chatbots, can aid in the screening process.

This study presents COVID-Bot, an AI-powered chatbot that can screen higher education students, determining their receipt of the COVID-19 vaccine. COVID-Bot was built using SnatchBot, a chatbot development platform [11,16]. This research will contribute significantly to the literature on technology acceptability in health care and education systems. First, we developed a chatbot that can verify whether higher education students have received the COVID-19 vaccine. Second, an evaluation of the COVID-Bot showed that the tool is helpful to students and other stakeholders in the process of determining the COVID-19 vaccination status in educational institutions.

The remainder of this paper is structured as follows. The next section presents the background information and related works, followed by the research methodology, design and development of COVID-Bot, implementation, results, and evaluation and application of the system. The final section of the study concludes the study.

### **Theoretical Aspects**

This section discusses the theoretical aspects of the research, including chatbot technology, the educational impact of the COVID-19 pandemic, and other related works.

### **Chatbot Technology**

A chatbot system is an AI-powered technology that can interact with people and provide accurate and immediate answers

[17,18]. It is an intelligent agent that can communicate with a user, answer a series of questions, and offer the correct response [19]. Some other names for chatbots are intelligent agents, conversational agents, digital assistants, and clever bots. A chatbot system can be built using a variety of development platforms, including SnatchBot, IBM Watson, Microsoft Azure, and Google [11,20]. Developers can use these frameworks to create chatbot interactions that address a wide range of educational challenges. Chatbots are designed to enhance various working experiences and are a game changer in the inventive era of the Fourth Industrial Revolution (4IR). The introduction of chatbot technology has brought in a lot of new opportunities for a variety of industries [21], including education, health care, and agriculture.

In the education domain, chatbots are not only used to improve students' learning and interaction abilities but also help the teaching staff by introducing automation [21]. Chatbots in education improve connectedness, enhance efficiency, and decrease ambiguity in interactions. They can easily deliver a targeted, individualized, and result-oriented online learning platform [22], which is precisely 1 of the major expectations of current educational institutions. The increased use of technology in daily life impacts how students learn and absorb information. The use of chatbots in education allows for the democratization of education because they do not consider the student's location, resources, or language. Chatbot systems, in general, provide quick and rapid replies, individualized learning experiences, automated tasks, and centralized learning.

However, in the light of the growing difficulties in attending to patients during a health crisis or pandemic, such as COVID-19, the health care sector has shifted its focus to strengthening digital health care services. Health care professionals are using chatbots to assist patients 24 hours a day, 7 days a week, which is a major changer for the business [23,24]. Health care chatbots can deliver accurate, up-to-date information, while improving the patient experience. One of the best ways to manage the COVID-19 crisis is to stay safe and be informed. Accurate information about the pandemic and individuals' state of health is essential in combating the spread of the disease. Chatbots can help in the war against misinformation and provide guidance to treatment, direction to basic health care facilities, safety precautions, and self-evaluation or screening. As students return to the classroom learning system, it is essential to provide an automatic and interactive agent that can help screen their state of health anytime and anywhere, while observing safety precautions.

### ***The Educational Impact of the COVID-19 Pandemic***

Millions of people worldwide have died due to the novel coronavirus disease [2]. As a result, some human activities have been forced to cease. Many individuals work under severe constraints, and many others work from home. Higher education was no exception, with most institutions worldwide closing their campuses and replacing conventional learning techniques with online learning using various software tools and applications.

Naturally, this abrupt change generated uncertainty for both administrative personnel and students at most institutions, where

many policies had to be revised to reflect the new reality [25,26]. Furthermore, the administrative staff could not respond to the massive rush of inquiries and requests made by students on a variety of themes [26]. According to Igoe and Chadwick [27], the COVID-19 pandemic has posed immediate and long-term difficulties to higher education, notably in administration, financial, academic, technological, and learning possibilities.

- **Administration:** Since the emergence of COVID-19, higher education administration has struggled to develop appropriate actions and response activities in response to the pandemic. During the pandemic, higher education institutions performed their teaching and learning online [28], which created various administrative issues, such as monitoring/supervision, faculty training for e-learning technologies, and student opposition [29]. Addressing this challenge requires proper leadership and management with clear and defined responsibilities [30].
- **Funding:** The pandemic has adversely impacted global economic growth, which has affected the funding of educational activities. Because many economies are still recovering from the effects of the COVID-19 pandemic, higher education funding is projected to reduce for a while, especially for public institutions [31].
- **Remote learning:** Remote learning occurs outside of traditional classroom settings and is frequently aided by digital platforms, such as online classrooms and learning management systems [32]. Although various institutions provide their students with data bundles to connect to the online learning system, network connectivity remains a problem, particularly in rural and less developed areas. Students may face difficulties due to remote learning, such as connection troubles, a lack of information and communication technology (ICT) efficacy, overwhelming online assignments and projects, and an inability to adjust to remote learning [33].
- **Use of technology:** Online learning approach requires the use of modern technologies, such as AI. Unfortunately, some academics and students are reluctant to accept and use these technological tools. This hesitancy may be due to unfamiliarity, which influences technology acceptance behaviors. In addition, they are not ready for radical changes in the architecture of their educational system [34].

To ensure that teaching and learning activities are not halted, the United Nations Educational, Scientific, and Cultural Organisation (UNESCO) launched the Global Education Coalition (GEC), with an emphasis on the need for free support, tools, or services that help continue educational services during the COVID-19 pandemic [35]. Consequently, various automated solutions that swiftly and accurately answered various requests became necessary during and after the pandemic.

### Related Works

Although we could not find a chatbot created specifically to screen students for their coronavirus vaccination status, there are some related studies.

Sweidan et al [26] introduced a Student Interactive Assistant Android Application with Chatbot (SIAAAC). This app guides students to obtain various excellent academic services during

the COVID-19 pandemic. The chatbot includes a campus map and several alerts, can reply to both Arabic and English queries, and covers a wide variety of academic topics. In Brazil, Roque, Cavalcanti [10] created a chatbot called BotCovid, which provides accurate information about COVID-19 and can answer not less than 600 questions. The University of California, San Francisco, created and deployed a digital chatbot to screen health care workers for COVID-19 [36]. Within the first 2 months of operation, the system performed 270,000 screens. Another study created a conversational bot based on natural language processing (NLP) that acts as a personal virtual doctor for chronic patients, providing free primary health care education, information, and advice [36]. In April 2020, Standard Bank South Africa introduced a chatbot on WhatsApp to inform its clients about its services. The chatbot provides factual information about Standard Bank's reaction to COVID-19, as well as answers to some inquiries and connections to official pandemic information sources [28]. This study is unique that it is designed to determine student vaccination status.

### Gap

Although these chatbots are created for COVID-19 pandemic response purposes, none is for screening students in higher education to evaluate their vaccination status. Because the higher education setting contains many people who may be in close proximity to one another, providing an automated system that can determine COVID-19 vaccination status is deemed necessary. COVID-Bot is designed to address this gap.

### Motivation

The COVID-19 pandemic caused unprecedented disruptions in education, resulting in extended school closures and abrupt changes to routine school operations that impacted education systems worldwide [25]. The development of various COVID-19 vaccine candidates and other safety precautions have aided in controlling the spread of the disease and reducing the consequences. Still, there is no cure for the disease yet. As students return to school to resume classroom teaching and learning, there will be an increase in human-human contact and interactions. Controlling and reducing disease transmission in higher education students require knowledge of their vaccination status. Studies have suggested that a technological resource, such as an AI-based chatbot, may help achieve this need [10,26].

## Methods

### Study Design

This study adheres to the principles of the design science research (DSR) process, which is a research approach with the goal of creating a new experience (artifact) for problem solving [37,38]. The DSR process consists of 6 sequential stages: identification, definition, design and development, implementation, evaluation, and communication. In the Gap section, the research problem was identified and described. This study introduced COVID-Bot, an intelligent interactive system that can be used to screen higher education students to determine whether they have received the COVID-19 vaccination. COVID-Bot was built on the SnatchBot platform's principles, which are mostly drag-and-drop without coding. The SnatchBot

framework was chosen because it offers an appealing set of functions and a simple user interface. It can also be used to deploy a chatbot on websites and social networking platforms. The system was implemented on a website and demonstrated to a group of selected students, who were given a thorough explanation of the system's objectives and operation. COVID-Bot was evaluated following ISO/IEC 25023:2016 standards, which establish quality metrics for quantitatively evaluating the quality of system and software products [39]. A questionnaire with 2 sections, including personal profiles and system qualities, was developed and used for data collection. The techniques of purposive and convenience sampling were used. These techniques were used because the system was designed specifically for student use, and during the evaluation, all possible available students were contacted. SPSS 25 statistical software was used to analyze the collected data. The research report and practical application of the chatbot system were used to communicate the study findings to the public.

### Ethical Considerations

The study was approved by the College of Business and Economics Research Ethics Committee (approval no. 2022-069). All participants were duly informed, with a detailed description of the system.

### Design and Development of COVID-Bot

This section describes the design and development of COVID-Bot, as well as a working algorithm sample. COVID-Bot was built using the SnatchBot development platform.

#### The SnatchBot Development Platform

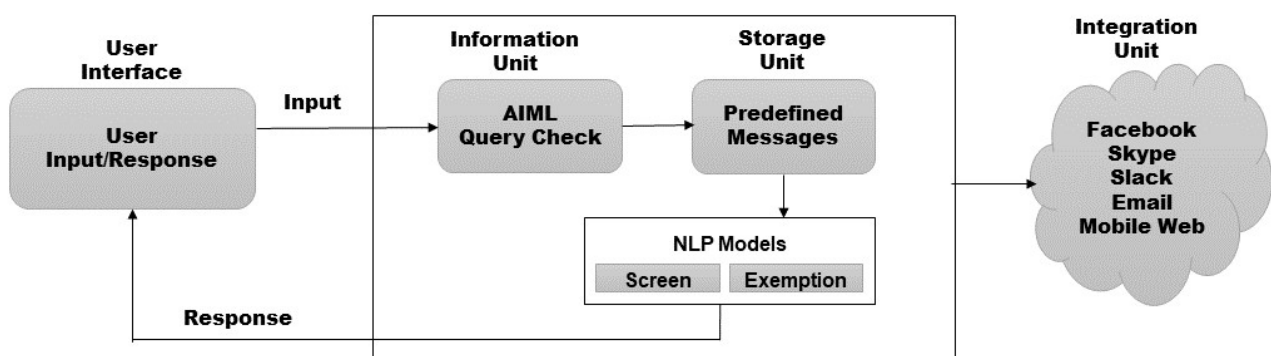
SnatchBot is a chatbot framework with an attractive range of functions and an easy-to-use interface, allowing quick deployment of a chatbot to web applications and social networking applications, such as Facebook Messenger, Skype,

SnatchApp, Viber, Slack, and Twilio. It enables developers to create visually attractive bots without coding and with a high level of expertise. SnatchBot offers the user an interface unit, a knowledge unit, a message bank, and an integration unit. The knowledge unit includes the query and artificial intelligence markup language (AIML) interfaces. AIML enables the chatbot to compare the user's input to the predefined messages in the message bank and provide the user with a matched answer. The integration unit can distribute this interaction on the website and social media network.

The channel is another distinguishing feature of SnatchBot. It is a built-in application programming interface (API) that allows chatbots to be deployed on websites and social media platforms. COVID-Bot is implemented on a website through the SnatchBot channel. Students can use smartphones and computers to access COVID-Bot via the website. SnatchBot provides a built-in editor for generating simple or sophisticated dialogues with action buttons and translations. It also allows one to establish a range of *interactions* with the chatbot's activity. Interactions are used to characterize the *subjects* in the SnatchBot platform. Interactions define a chatbot's unique activity, while subjects determine the activity's predefined content, such as messages, audios, videos, and graphs. For instance, if the chatbot is required to introduce itself, an interaction called introduction could be created, with the associated subject being the message/text that defines the chatbot's purposes. There are 2 types of connections in SnatchBot: local and global. Local connections connect to a specific interaction, while global connections connect to all interactions.

Moreover, the SnatchBot platform supports NLP model capabilities. This enables using predefined models and the creation of custom models that can support the chatbot operations. Figure 1 (adapted from Ref. [11]) depicts the SnatchBot operating sequence.

**Figure 1.** SnatchBot flow diagram. AIML: artificial intelligence markup language; NLP: natural language processing.



### Design of COVID-Bot

COVID-Bot was created using the SnatchBot API and its predefined tools, which are driven by various NLP algorithms. The design of COVID-Bot included a variety of interactions and subjects. Each interaction is given a name, a distinct ID number, and a purpose. The design involves 30 logically and locally connected interactions, including Welcome, Contents,

Screen, Exemption, Allowed, and Disallowed. The development scheme of the chatbot is shown in Figure 2.

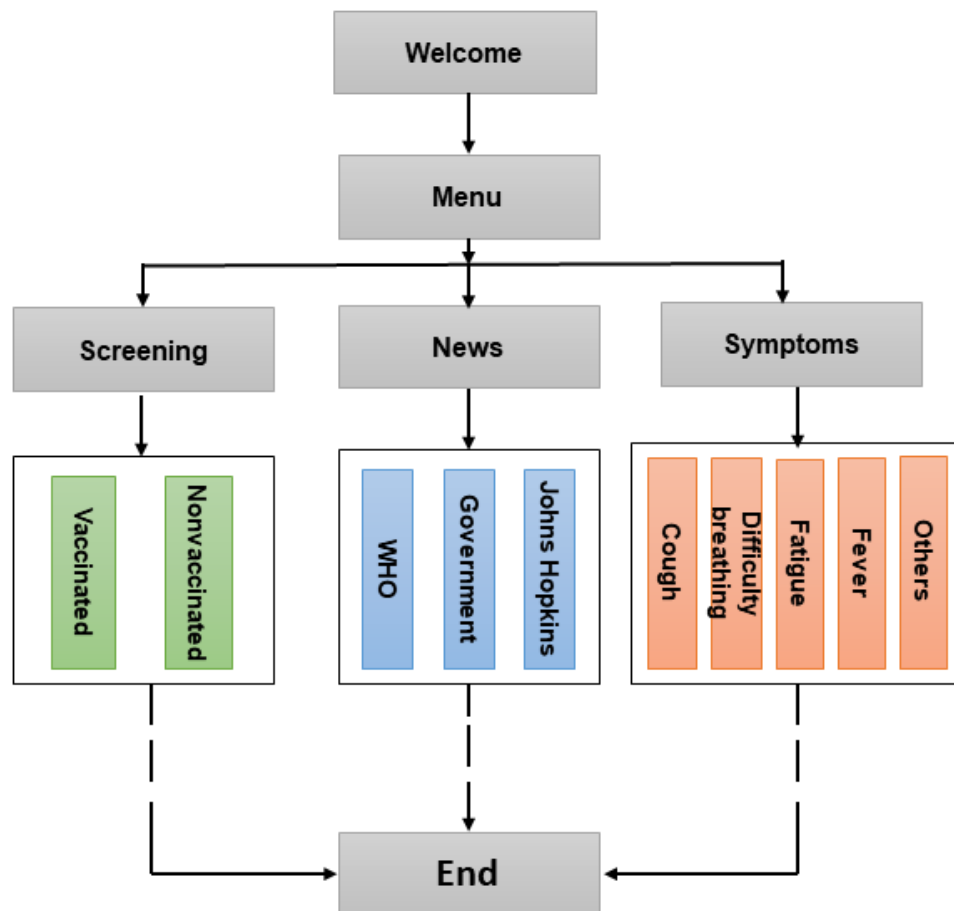
All connections are expressed as a logical statement in the following form:

*If a then b*  
*Else c*

The decision flow of COVID-Bot operations is depicted by the following algorithm, adapted from Ref. [8]. The algorithm demonstrates that the user sends a text input, as indicated in

line 1, which is compared to the sample in the storage bank or NLP models in line 2 and returns a corresponding response of access granted in line 3 or access denied in line 5 (Textbox 1).

**Figure 2.** Development scheme of COVID-Bot. WHO: World Health Organization.



**Textbox 1.** Decision flow of COVID-Bot operations.

**Algorithm 1: COVID-Bot Information-Decision**

```

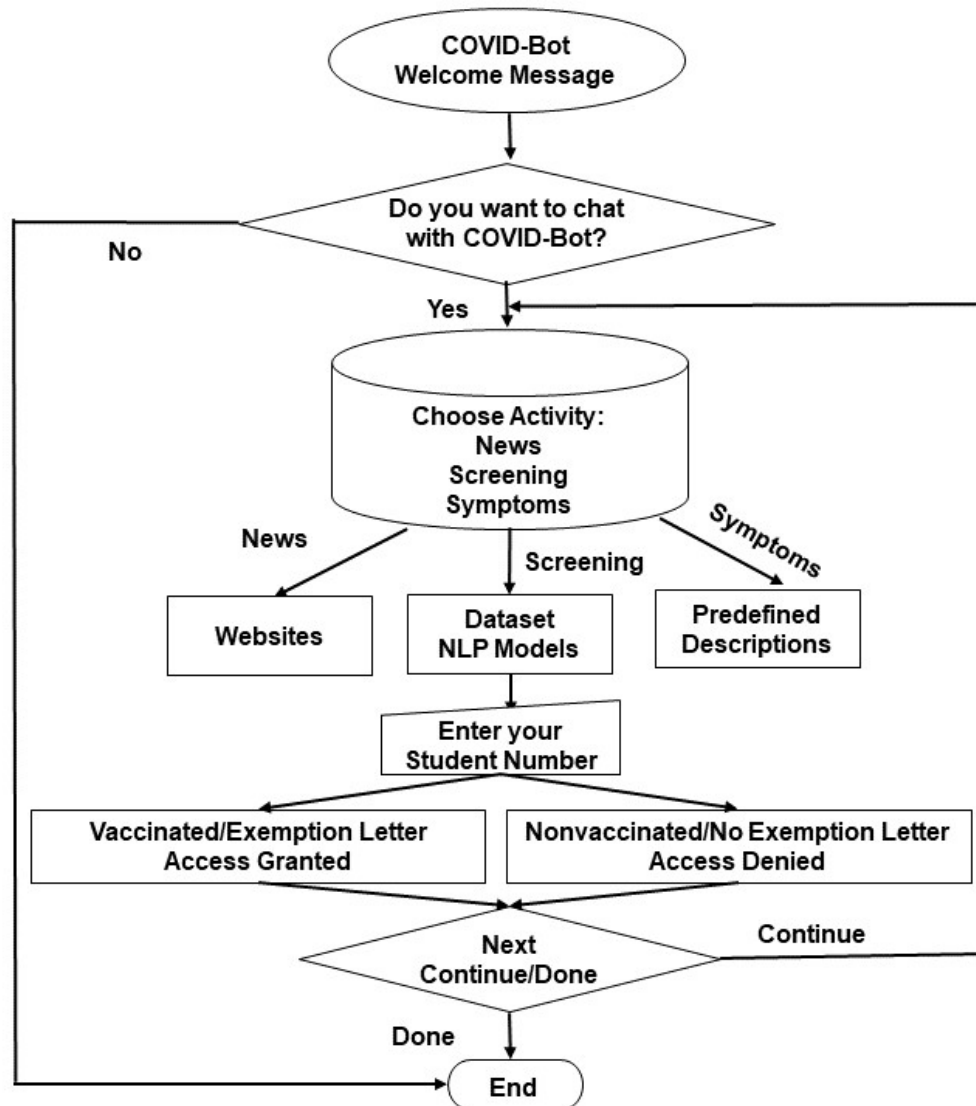
1. Input: Student Number (user text input);
2. if Student Number Matches Storage Bank then
3. return predefined response (ACCESS GRANTED);
4. else
5. return default answer (ACCESS DENIED);
end
  
```

The architecture of COVID-Bot consists of 4 main sections:

- **Welcome:** This section introduces the chatbot and confirms whether the user is ready to converse with the bot.
- **Screening:** This section checks and confirms whether the user is vaccinated or whether they have been given an exemption letter.
- **News:** This section provides the latest coronavirus news from the WHO website, the South Africa coronavirus matters website, and the Johns Hopkins University coronavirus information webpage.
- **Symptoms:** This section presents descriptions of common COVID-19 symptoms.

**Figure 3** displays an operational decision made by COVID-Bot in response to user inquiries. It demonstrates that when a user (student) connects to COVID-Bot, it delivers a welcome message asking the user whether they wish to interact with the chatbot. If the user responds with yes, COVID-Bot directs them to menu contents, including screening, news, and symptoms. The user chooses the intended activity. During the screening activity, the chatbot requests the student number to verify whether the student has been vaccinated or not vaccinated or has been given an exemption letter to grant or deny them access. If an incorrect student number is entered, the system allows 1 step back to correct it; otherwise, the process ends.

Figure 3. COVID-Bot flow diagram. NLP: natural language processing.



## Results

### Implementation Findings

In this study, we created COVID-Bot, an AI-powered intelligent system that screens higher education students and confirms their COVID-19 vaccination status. COVID-Bot begins its operation by displaying an introduction message to the user and determining whether the user wishes to interact with it. This system serves 3 purposes: (1) It verifies and confirms a user's COVID-19 vaccination status, (2) it provides the user with access to the most recent coronavirus news from recognized and trusted sources, and (3) it explains COVID-19 symptoms. COVID-Bot requests entry of the student's information, including name and number. The system takes the student's number and checks to see whether it is in the storage bank. The storage bank contains all the student numbers of those students who have been vaccinated and those who have been granted an exemption. If the student number is found in the storage bank, the student is granted access to the institution's premises and other activities; otherwise, the student is denied access. When a user mistakenly enters an incorrect student number, the system provides them with another chance to correct their mistake.

Accordingly, the user must return to the previous option by clicking the back button.

In addition, the COVID-Bot design incorporates 2 NLP models: screen and exemption. These models use machine learning techniques to make decisions based on every input from the user. The models are trained with the samples obtained from data sets of the vaccinated students' list and the nonvaccinated students' list with exemption letters. In other words, the vaccinated data set contains the names and student numbers of all students who have received the COVID-19 vaccine and have a vaccination certificate as evidence. In contrast, the nonvaccinated data set includes the names and student numbers of all students who have not received the COVID-19 vaccine for reasons approved by the institution and have an exemption letter as evidence. The screen model determines and validates the confirmed vaccinated students in the data set, whereas the exemption model determines and validates the confirmed students with exemption letters. COVID-Bot was demonstrated to a group of selected students, who were given an adequate explanation of its objectives and operation. Figures 4a and 4b show the welcome and input activities, and Figures 5a and 5b show the confirmation responses from COVID-Bot.

Figure 4. (a) Welcome and (b) activity menus of COVID-Bot.

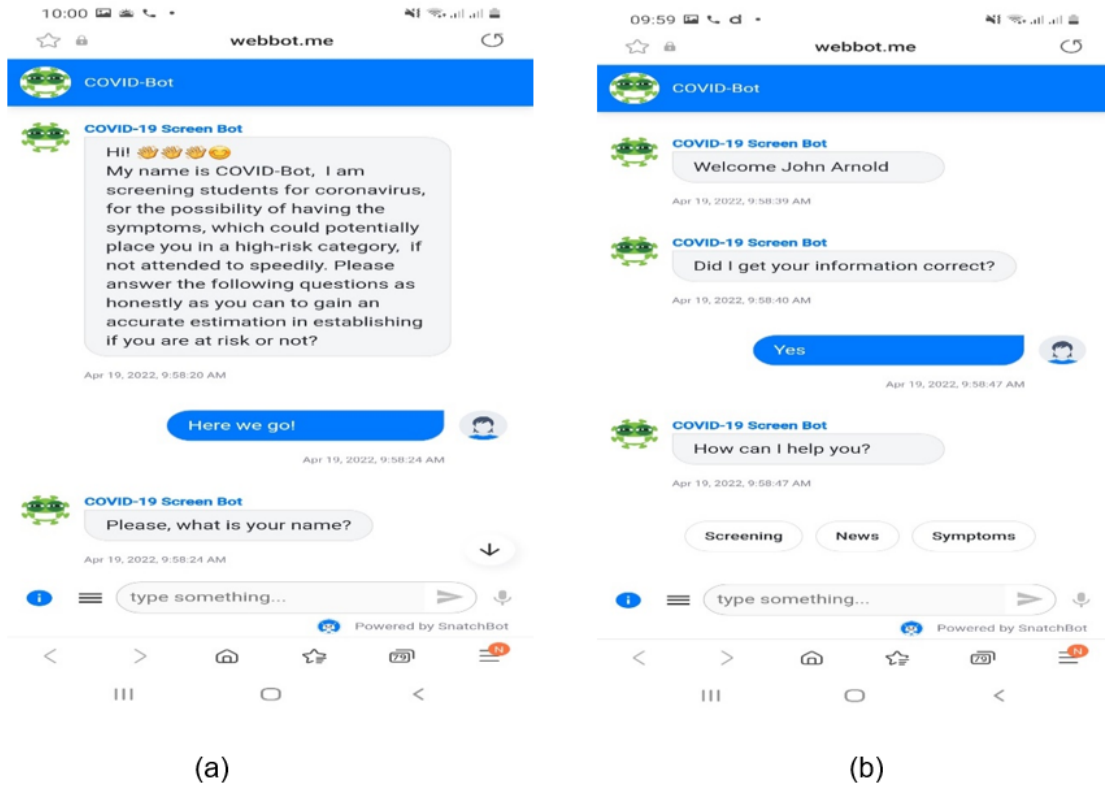
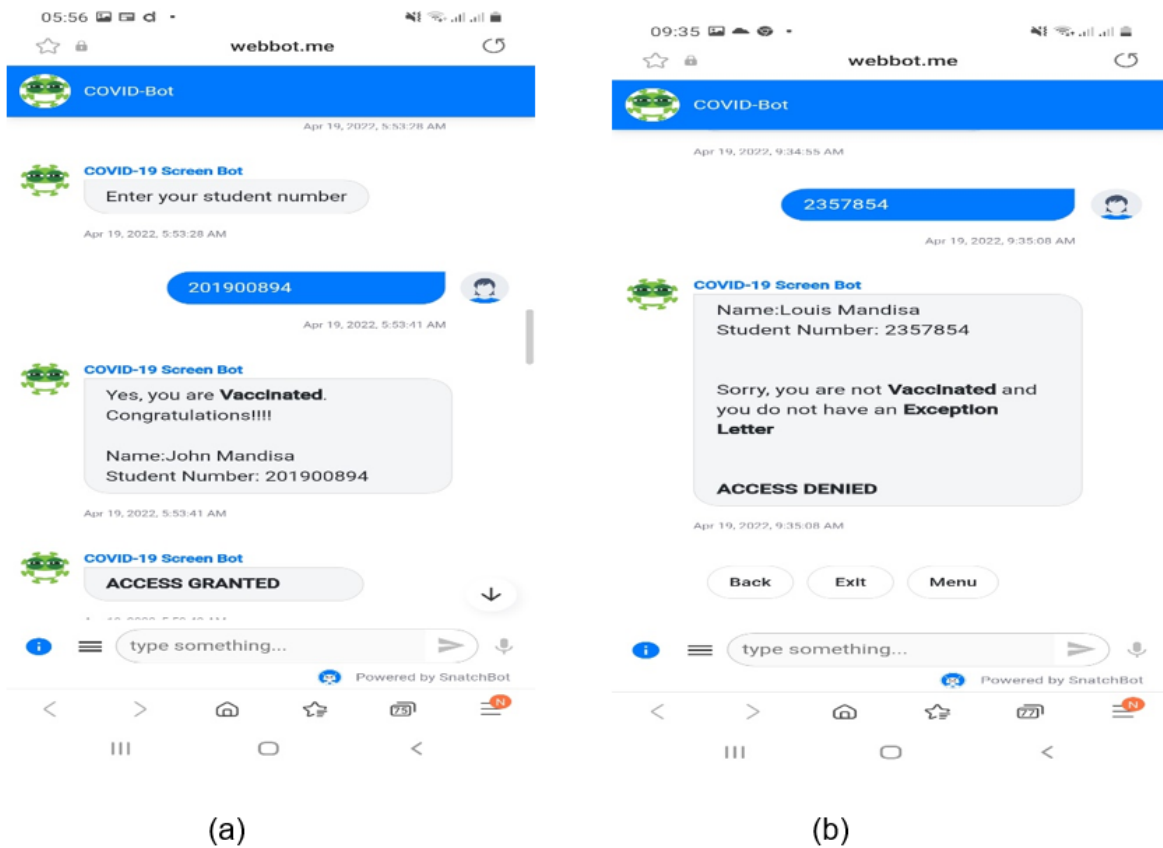


Figure 5. COVID-Bot confirmation message: (a) granting access and (b) denying access to students.



## Evaluation and Application of COVID-Bot

### Evaluation

The evaluation of COVID-Bot adhered to the ISO/IEC 25023:2016 standard, which establishes quality metrics for quantitatively evaluating the quality of system and software products [39]. It examines the following user characteristics: functionality, efficiency, compatibility, usability, reliability, maintainability, portability, and security. In this evaluation, students were asked about their perceptions of COVID-Bot's functionality, compatibility, reliability, and usability. A total of 106 students from the University of Johannesburg were chosen to participate in the study, including those who own smartphones and excluding those who do not. This audience was selected because they are potential users of the system and are actively registered students. Their student numbers were used to train the NLP models. The participants were shown a brief demonstration of the system. An online survey was conducted

**Table 1.** Reliability analysis.

Measuring items	Cronbach $\alpha$
Functionality	.88
Compatibility	.86
Reliability	.85
Usability	.85

### Participant Profile

The sample population consisted of male (n=61, 57.5%) and female (n=45, 42.5%) students at different levels of studies, including undergraduates and postgraduates within the age range of 15-35 years. The predominant age group of undergraduates was 15-25 years (n=83, 78.3%), while postgraduates were 25-35 years old (n=23, 21.7%). In addition, 95 (89.6%) participants agreed they have good knowledge of chatbot technology, while 11 (10.4%) reported having limited knowledge.

Students were asked about their opinions on the following qualities of COVID-Bot: functionality, compatibility, reliability, and usability.

#### Functionality

This test was performed to determine whether the system is functioning correctly to achieve the desired results. Of the 106 participants, 92 (86.8%) agreed or strongly agreed that the system works well and serves its purpose, 10 (9.4%) disagreed, maybe due to their limited knowledge of emerging technologies such as the Chatbot system, and 4 (3.8%) were neutral.

#### Compatibility

This was evaluated to see whether the system will be compatible with the students' existing lifestyle, social norms, and mobile devices. Of the 106 participants, 85 (80.2%) agreed or strongly agreed that the system works well with their smartphones and fits well with their lifestyle, 10 (9.4%) disagreed, possibly due to their ability to use modern technological innovations, and 11 (10.4%) were indifferent.

for quantitative analysis; participants were sent links to the system and a Google form questionnaire to collect their personal information (age, gender, and education) as well as their opinions on the chatbot's qualities. A 5-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree) was used for measurement.

#### Reliability Analysis

To determine the reliability of the measuring items, a reliability analysis was performed to obtain Cronbach  $\alpha$  coefficients of the measuring constructs. The average obtained  $\alpha$  was .87, reflecting satisfactory reliability. Furthermore, a similar correlation value ( $r=0.99$ ) was observed, indicating that the items had a strong relationship. The obtained values of  $r$  and  $\alpha$  were possible because the evaluation was conducted in a particular context (a university institution) where the participants have similar perceptions. Table 1 shows the results of the reliability analysis.

#### Reliability

This was considered to determine the consistency and accuracy of the system's responses to the user's input or queries. Of the 106 participants, 86 (81.1%) agreed that the system is able to generate appropriate information about vaccinated and nonvaccinated students, as well as those with exemption, 10 (9.4%) disagreed, and 10 (9.4%) were neutral.

#### Usability

To assess the ease with which COVID-Bot can be used to achieve the expected task, usability was evaluated. Of the 106 participants, 85 (80.2%) agreed or strongly agreed that they found the system easy to use, 8 (7.5%) disagreed that they had to repeat the process several times before getting an accurate response due to typos, and 13 (12.3%) were indifferent.

## Discussion

### Principal Findings

COVID-Bot is a conversational agent created to interact with students to confirm whether they are vaccinated or have obtained an exemption letter before being granted access to the institution's premises. This system also provides updated information about the COVID-19 pandemic from recognized sources and descriptions of the symptoms. Chatbots are already being used to battle COVID-19. They have helped eliminate disinformation, aided symptom diagnosis, motivated actions that restrict infection, and reduced mental health costs of the pandemic response [40]. The COVID-19 pandemic is a global issue that can affect anyone regardless of profession or social status. As a result, it is critical to develop a solution that can



help reduce the spread of the disease and lessen its societal impact. COVID-Bot was created with this viewpoint in mind.

The evaluation began with a demonstration of the chatbot system to the participants, followed by requesting them to fill out an online questionnaire. According to the questionnaire responses, 86.8% of the participants agreed that the system works well. This suggests the ability of COVID-Bot to determine students' vaccination status and detect those who have received an exemption letter based on the predefined data set. In terms of compatibility, the students were able to use the chatbot with their smartphones, laptops, and desktop computers at the appropriate location and time. As a result, 80.2% of the participants agreed that it is compatible with their lifestyle and activities. The system accuracy and consistency (reliability) test revealed that COVID-Bot is reliable, with 81.1% of the participants agreeing on this factor. This implies that the students' interaction with COVID-Bot produced accurate results. They inquired about their vaccination status, and the system responded with appropriate information. Regarding usability, 80.2% of the participants agreed that the system is easy to use. This means that the design of COVID-Bot is simple enough to use without any prior training or supervision.

Overall, COVID-Bot is designed in a simple manner to accommodate students' use of mobile apps, and it can provide accurate responses to questions about their vaccination status.

### Applications of COVID-Bot

The development of COVID-19 vaccine candidates aids in the fight against the spread of coronavirus. Everyone is expected to get the vaccine, especially those who work in organizations with frequent physical contact with different people, such as academic institutions. As a result, COVID-19 vaccination is required in some institutions. In theory, this study presented the

design and development of an intelligent system for COVID-19-related issues, and there are few existing studies in this area. In practice, COVID-Bot can help detect students' COVID-19 vaccination status, thereby helping in the fight against the spread of the deadly coronavirus. Furthermore, it will aid in the safety of students during teaching and learning activities while they are in the institution's environment.

### Comparison With Prior Works

COVID-Bot operations differ from previous related works. It interacts with students intelligently, using their student numbers as input to determine and validate their vaccination status or exemption letter to grant or deny them access to the institution.

### Limitations and Future Works

During the system's implementation, we encountered some limitations in gaining access to students' data sets regarding COVID-19 vaccination status and user privacy concerns. The system is only for students. The entire staff of the institution can be added to future works. It can also be expanded to perform additional functions, such as screening registered and nonregistered students and providing students with financial clearance.

### Conclusion

In this study, COVID-Bot, an intelligent interactive system that can converse with students to determine their COVID-19 vaccination status and confirm whether they can be granted access to the institution and its various activities, was created. COVID-Bot was developed using the SnatchBot chatbot API and can be installed on websites and social networking platforms. An online survey was conducted to evaluate the system. The results indicated that COVID-Bot has good qualities and can be useful in fighting against the spread and impact of COVID-19.

### Conflicts of Interest

None declared.

### References

1. Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. *Lancet Infect Dis* 2020 May;20(5):533-534. [doi: [10.1016/s1473-3099\(20\)30120-1](https://doi.org/10.1016/s1473-3099(20)30120-1)]
2. COVID-19 Dashboard by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU). URL: <https://coronavirus.jhu.edu/map.html> [accessed 2022-10-17]
3. Aruleba K, Jere N, Matarirano O. Technology adoption readiness in disadvantaged universities during covid-19 pandemic in South Africa. *Int J High Educ* 2022 Jan 23;11(2):172. [doi: [10.5430/ijhe.v11n2p172](https://doi.org/10.5430/ijhe.v11n2p172)]
4. Nagy A, Alhatlani B. An overview of current COVID-19 vaccine platforms. *Comput Struct Biotechnol J* 2021;19:2508-2517 [FREE Full text] [doi: [10.1016/j.csbj.2021.04.061](https://doi.org/10.1016/j.csbj.2021.04.061)] [Medline: [33936564](https://pubmed.ncbi.nlm.nih.gov/33936564/)]
5. Peng HT, Rhind SG, Beckett A. Convalescent plasma for the prevention and treatment of covid-19: a systematic review and quantitative analysis. *JMIR Public Health Surveill* 2021 Apr 7;7(4):e25500. [doi: [10.2196/25500](https://doi.org/10.2196/25500)]
6. Masngut N, Mohamad E. Association between public opinion and Malaysian government communication strategies about the covid-19 crisis: content analysis of image repair strategies in social media. *J Med Internet Res* 2021 Aug 04;23(8):e28074 [FREE Full text] [doi: [10.2196/28074](https://doi.org/10.2196/28074)] [Medline: [34156967](https://pubmed.ncbi.nlm.nih.gov/34156967/)]
7. Ahuja AS, Reddy VP, Marques O. Artificial intelligence and COVID-19: a multidisciplinary approach. *Integr Med Res* 2020 Sep;9(3):100434 [FREE Full text] [doi: [10.1016/j.imr.2020.100434](https://doi.org/10.1016/j.imr.2020.100434)] [Medline: [32632356](https://pubmed.ncbi.nlm.nih.gov/32632356/)]
8. Aruleba RT, Adekiya TA, Ayawei N, Obaido G, Aruleba K, Mienye ID, et al. COVID-19 diagnosis: a review of rapid antigen, RT-PCR and artificial intelligence methods. *Bioengineering (Basel)* 2022 Apr 03;9(4):153 [FREE Full text] [doi: [10.3390/bioengineering9040153](https://doi.org/10.3390/bioengineering9040153)] [Medline: [35447713](https://pubmed.ncbi.nlm.nih.gov/35447713/)]

9. Almalki M, Azeez F. Health chatbots for fighting covid-19: a scoping review. *Acta Inform Med* 2020 Dec;28(4):241-247 [FREE Full text] [doi: [10.5455/aim.2020.28.241-247](https://doi.org/10.5455/aim.2020.28.241-247)] [Medline: [33627924](https://pubmed.ncbi.nlm.nih.gov/33627924/)]
10. Roque G. BotCovid: development and evaluation of a chatbot to combat misinformation about covid-19 in Brazil. 2021 Presented at: 2021 IEEE International Conference on Systems, Man, and Cybernetics (SMC); October 17-20, 2021; Melbourne, VIC, Australia. [doi: [10.1109/smc52423.2021.9658693](https://doi.org/10.1109/smc52423.2021.9658693)]
11. Okonkwo C, Ade-Ibijola A. Python-Bot: a chatbot for teaching Python programming. *Eng Lett* 2021 Feb;29(1):25-34.
12. Walwema J. The WHO health alert: communicating a global pandemic with WhatsApp. *J Bus Tech Commun* 2020 Sep 16;35(1):35-40. [doi: [10.1177/1050651920958507](https://doi.org/10.1177/1050651920958507)]
13. Smutny P, Schreiberova P. Chatbots for learning: a review of educational chatbots for the Facebook Messenger. *Comput Educ* 2020 Jul;151:103862. [doi: [10.1016/j.compedu.2020.103862](https://doi.org/10.1016/j.compedu.2020.103862)]
14. World Health Organization. WHO Launches a Chatbot on Facebook Messenger to Combat COVID-19 Misinformation. URL: <https://www.who.int/news-room/feature-stories/detail/who-launches-a-chatbot-powered-facebook-messenger-to-combat-covid-19-misinformation> [accessed 2022-03-08]
15. Laranjo L, Dunn AG, Tong HL, Kocaballi AB, Chen J, Bashir R, et al. Conversational agents in healthcare: a systematic review. *J Am Med Inform Assoc* 2018 Sep 01;25(9):1248-1258 [FREE Full text] [doi: [10.1093/jamia/ocy072](https://doi.org/10.1093/jamia/ocy072)] [Medline: [30010941](https://pubmed.ncbi.nlm.nih.gov/30010941/)]
16. Erazo W. Chatbot implementation to collect data on possible COVID-19 cases and release the pressure on the primary health care system. 2020 Presented at: IEMCON 2020: 11th IEEE Annual Information Technology, Electronics and Mobile Communication Conference; November 4-7, 2020; Vancouver, BC, Canada. [doi: [10.1109/iemcon51383.2020.9284846](https://doi.org/10.1109/iemcon51383.2020.9284846)]
17. Okonkwo C, Ade-Ibijola A. Evaluating the ethical implications of using chatbot systems in higher education. *digitAL* 2021;2021:68 [FREE Full text]
18. Okonkwo CW, Ade-Ibijola A. Chatbots applications in education: a systematic review. *Comput Educ: AI* 2021;2:100033. [doi: [10.1016/j.caeai.2021.100033](https://doi.org/10.1016/j.caeai.2021.100033)]
19. Clarizia F. Chatbot: an education support system for student. 2018 Presented at: CSS 2018: 10th International Symposium on Cyberspace Safety and Security; October 29-31, 2018; Amalfi, Italy. [doi: [10.1007/978-3-030-01689-0\\_23](https://doi.org/10.1007/978-3-030-01689-0_23)]
20. Pham XL, Pham T, Nguyen QM, Nguyen TH, Cao TTH. Chatbot as an intelligent personal assistant for mobile language learning. 2018 Presented at: ICEEL 2018: 2nd International Conference on Education and E-Learning; November 5-7, 2018; Bali, Indonesia. [doi: [10.1145/3291078.3291115](https://doi.org/10.1145/3291078.3291115)]
21. Dsouza R, Sahu S, Patil R, Kalbande DR. 2019 Presented at: 2019 International Conference on Advances in Computing, Communication and Control (ICAC3); December 20-21, 2019; Mumbai, India. [doi: [10.1109/icac347590.2019.9036844](https://doi.org/10.1109/icac347590.2019.9036844)]
22. Cunningham-Nelson S, Boles W, Trouton L, Margerison E. A review of chatbots in education: practical steps forward. 2019 Presented at: AAEE2019: 30th Annual Conference for the Australasian Association for Engineering Education; December 8-11, 2019; Brisbane, Australia p. E.
23. Xu L, Sanders L, Li K, Chow JCL. Chatbot for health care and oncology applications using artificial intelligence and machine learning: systematic review. *JMIR Cancer* 2021 Nov 29;7(4):e27850 [FREE Full text] [doi: [10.2196/27850](https://doi.org/10.2196/27850)] [Medline: [34847056](https://pubmed.ncbi.nlm.nih.gov/34847056/)]
24. Klos MC, Escoredo M, Joerin A, Lemos VN, Rauws M, Bunge EL. Artificial intelligence-based chatbot for anxiety and depression in university students: pilot randomized controlled trial. *JMIR Form Res* 2021 Aug 12;5(8):e20678 [FREE Full text] [doi: [10.2196/20678](https://doi.org/10.2196/20678)] [Medline: [34092548](https://pubmed.ncbi.nlm.nih.gov/34092548/)]
25. Meinck S, Fraillon J, Strietholt R. The Impact of the COVID-19 Pandemic on Education: International Evidence from the Responses to Educational Disruption Survey (REDS). URL: <https://unesdoc.unesco.org/ark:/48223/pf0000380398> [accessed 2022-03-11]
26. Sweidan SZ, Abu Laban SS, Alnaimat NA, Darabkh KA. SIAAA - C: a student interactive assistant android application with chatbot during COVID - 19 pandemic. *Comput Appl Eng Educ* 2021 Apr 16;29(6):1718-1742. [doi: [10.1002/cae.22419](https://doi.org/10.1002/cae.22419)]
27. Igoe M, Chadwick V. After the Pandemic: How Will COVID-19 Transform Global Health and Development?. URL: <https://www.devex.com/news/after-the-pandemic-how-will-covid-19-transform-global-health-and-development-96936> [accessed 2022-03-11]
28. Standard Bank. Standard Bank Launches Covid 19 Chatbot. URL: <https://www.standardbank.com/sbg/standard-bank-group/whats-happening/press-releases/standard-bank-launches-Covid-19-Chatbot> [accessed 2022-03-15]
29. Nganga G, Waruru M, Nakweya G. Universities Face Multiple Challenges in Wake of COVID-19 Closures. URL: <https://www.universityworldnews.com/post.php?story=20200407162549396> [accessed 2022-03-11]
30. Salihu MJ. Agencies governing the university education system in Nigeria: an examination of their administrative powers, conflicts and advocacies. *ATBU J Sci Technol Educ* 2020;7(4):231-239.
31. Ferguson HT. Cuts Expected for State Higher Education Funding in the Wake of COVID-19. URL: [https://www.nasfaa.org/news-item/21469/Cuts\\_Expected\\_for\\_State\\_Higher\\_Education\\_Funding\\_in\\_the\\_Wake\\_of\\_COVID-19](https://www.nasfaa.org/news-item/21469/Cuts_Expected_for_State_Higher_Education_Funding_in_the_Wake_of_COVID-19) [accessed 2022-03-11]
32. Yunusa AA. The impact of the covid-19 pandemic on higher education in Nigeria: university lecturers' perspectives. *Int J Educ Dev Inf Commun Technol* 2021;17(4):43-66 [FREE Full text]
33. Rao K, Giuli C. Reaching remote learners: successes and challenges for students in an online graduate degree program in the Pacific Islands. *Int Rev Res Open Distance Learn* 2010 Mar 05;11(1):141. [doi: [10.19173/irrodl.v11i1.785](https://doi.org/10.19173/irrodl.v11i1.785)]

34. Flavell H, Harris C, Price C, Logan E, Peterson S. Empowering academics to be adaptive with eLearning technologies: an exploratory case study. *Australas J Educ Technol* 2019 Mar 21;35(1):1-15. [doi: [10.14742/ajet.2990](https://doi.org/10.14742/ajet.2990)]
35. UNESCO. Global Education Coalition. URL: <https://en.unesco.org/covid19/educationresponse/globalcoalition> [accessed 2022-03-11]
36. Judson T, Odisho AY, Young JJ, Bigazzi O, Steuer D, Gonzales R, et al. Implementation of a digital chatbot to screen health system employees during the COVID-19 pandemic. *J Am Med Inform Assoc* 2020 Jul 01;27(9):1450-1455 [FREE Full text] [doi: [10.1093/jamia/ocaa130](https://doi.org/10.1093/jamia/ocaa130)] [Medline: [32531066](https://pubmed.ncbi.nlm.nih.gov/32531066/)]
37. Peffers K, Tuunanen T, Rothenberger MA, Chatterjee S. A design science research methodology for information systems research. *J Manage Inf Syst* 2014 Dec 08;24(3):45-77. [doi: [10.2753/mis0742-1222240302](https://doi.org/10.2753/mis0742-1222240302)]
38. Iivari J, Venable JR. Action research and design science research - seemingly similar but decisively dissimilar. 2009 Presented at: ECIS 2009: 17th European Conference on Information Systems; 2009; Verona, Italy.
39. ISO. ISO/IEC 25023:2016 - Systems and Software Engineering — Systems and Software Quality Requirements and Evaluation (SQuaRE) — Measurement of System and Software Product Quality. 2016. URL: <https://www.iso.org/standard/35747.html> [accessed 2022-10-17]
40. Kocaballi AB, Quiroz JC, Rezazadegan D, Berkovsky S, Magrabi F, Coiera E, et al. Responses of conversational agents to health and lifestyle prompts: investigation of appropriateness and presentation structures. *J Med Internet Res* 2020 Feb 09;22(2):e15823 [FREE Full text] [doi: [10.2196/15823](https://doi.org/10.2196/15823)] [Medline: [32039810](https://pubmed.ncbi.nlm.nih.gov/32039810/)]

## Abbreviations

**AI:** artificial intelligence  
**AIML:** artificial intelligence markup language  
**API:** application programming interface  
**DSR:** design science research  
**NLP:** natural language processing  
**WHO:** World Health Organization

*Edited by A Mavragani; submitted 30.04.22; peer-reviewed by G Obaido, K Aruleba; comments to author 24.05.22; revised version received 11.07.22; accepted 12.07.22; published 27.10.22.*

*Please cite as:*

Okonkwo CW, Amusa LB, Twinomurinzi H  
*COVID-Bot, an Intelligent System for COVID-19 Vaccination Screening: Design and Development*  
*JMIR Form Res* 2022;6(10):e39157  
URL: <https://formative.jmir.org/2022/10/e39157>  
doi: [10.2196/39157](https://doi.org/10.2196/39157)  
PMID: [36301616](https://pubmed.ncbi.nlm.nih.gov/36301616/)

©Chinedu Wilfred Okonkwo, Lateef Babatunde Amusa, Hossana Twinomurinzi. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 27.10.2022. 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.

Corrigenda and Addenda

# Correction: Using Wake-Up Tasks for Morning Behavior Change: Development and Usability Study

Kyue Taek Oh<sup>1</sup>, MS; Jisu Ko<sup>2</sup>, BS; Jaemyung Shin<sup>3</sup>, MS; Minsam Ko<sup>1</sup>, PhD

<sup>1</sup>Department of Human-Computer Interaction, Hanyang University, Ansan, Republic of Korea

<sup>2</sup>Department of Applied Artificial Intelligence, Hanyang University, Ansan, Republic of Korea

<sup>3</sup>Delightroom, Seoul, Republic of Korea

**Corresponding Author:**

Minsam Ko, PhD

Department of Human-Computer Interaction

Hanyang University

#609, Hakyeonsan cluster center 55, Hanyangdaehak-ro, Sangnok-gu, Ansan-si, Gyeonggi-do

Ansan, 15588

Republic of Korea

Phone: 82 314001071

Email: [minsam@hanyang.ac.kr](mailto:minsam@hanyang.ac.kr)

**Related Article:**

Correction of: <https://formative.jmir.org/2022/9/e39497>

(*JMIR Form Res* 2022;6(10):e42926) doi:[10.2196/42926](https://doi.org/10.2196/42926)

In “Using Wake-Up Tasks for Morning Behavior Change: Development and Usability Study” (*JMIR Form Res* 2022;6(9):e39497) the authors noted two errors.

The affiliation of the author Jisu Ko was incorrectly mentioned as the following:

*Department of Human-Computer Interaction.*

This has been corrected to:

*Department of Applied Artificial Intelligence.*

Under “Acknowledgments” the original text read:

*This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (Ministry of Science and ICT) (grant RS-2022-00155885, Artificial Intelligence*

*Convergence Innovation Human Resources Development) and Hanyang University ERICA.*

It has been replaced by the following:

*This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (Ministry of Science and ICT) (No. RS-2022-00155885, Artificial Intelligence Convergence Innovation Human Resources Development (Hanyang University ERICA)).*

The correction will appear in the online version of the paper on the JMIR Publications website on October 3, 2022, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.

Submitted 24.09.22; this is a non-peer-reviewed article; accepted 25.09.22; published 03.10.22.

Please cite as:

Oh KT, Ko J, Shin J, Ko M

Correction: Using Wake-Up Tasks for Morning Behavior Change: Development and Usability Study

*JMIR Form Res* 2022;6(10):e42926

URL: <https://formative.jmir.org/2022/10/e42926>

doi: [10.2196/42926](https://doi.org/10.2196/42926)

PMID: [36191300](https://pubmed.ncbi.nlm.nih.gov/36191300/)

©Kyue Taek Oh, Jisu Ko, Jaemyung Shin, Minsam Ko. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 03.10.2022. 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

# Sociodemographics and Transdiagnostic Mental Health Symptoms in SOCIAL (Studies of Online Cohorts for Internalizing Symptoms and Language) I and II: Cross-sectional Survey and Botometer Analysis

Lorenzo Lorenzo-Luaces<sup>1</sup>, PhD; Jacqueline Howard<sup>1</sup>, BA; Andy Edinger<sup>2</sup>, BS; Harry Yaojun Yan<sup>2,3</sup>, BA; Lauren A Rutter<sup>1</sup>, PhD; Danny Valdez<sup>4</sup>, PhD; Johan Bollen<sup>2,5</sup>, PhD

<sup>1</sup>Department of Psychological and Brain Sciences, Indiana University-Bloomington, Bloomington, IN, United States

<sup>2</sup>Luddy School of Informatics, Computing, and Engineering, Indiana University-Bloomington, Bloomington, IN, United States

<sup>3</sup>Media School, Indiana University-Bloomington, Bloomington, IN, United States

<sup>4</sup>School of Public Health, Indiana University-Bloomington, Bloomington, IN, United States

<sup>5</sup>Cognitive Science Program, Indiana University-Bloomington, Bloomington, IN, United States

**Corresponding Author:**

Lorenzo Lorenzo-Luaces, PhD

Department of Psychological and Brain Sciences

Indiana University-Bloomington

1101 E 10th Street

Bloomington, IN, 47401

United States

Phone: 1 812 856 0866

Email: [lolorenz@indiana.edu](mailto:lolorenz@indiana.edu)

## Abstract

**Background:** Internalizing, externalizing, and somatoform disorders are the most common and disabling forms of psychopathology. Our understanding of these clinical problems is limited by a reliance on self-report along with research using small samples. Social media has emerged as an exciting channel for collecting a large sample of longitudinal data from individuals to study psychopathology.

**Objective:** This study reported the results of 2 large ongoing studies in which we collected data from Twitter and self-reported clinical screening scales, the Studies of Online Cohorts for Internalizing Symptoms and Language (SOCIAL) I and II.

**Methods:** The participants were a sample of Twitter-using adults (SOCIAL I: N=1123) targeted to be nationally representative in terms of age, sex assigned at birth, race, and ethnicity, as well as a sample of college students in the Midwest (SOCIAL II: N=1988), of which 61.78% (1228/1988) were Twitter users. For all participants who were Twitter users, we asked for access to their Twitter handle, which we analyzed using Botometer, which rates the likelihood of an account belonging to a bot. We divided participants into 4 groups: Twitter users who did not give us their handle or gave us invalid handles (*invalid*), those who denied being Twitter users (*no Twitter*; only available for SOCIAL II), Twitter users who gave their handles but whose accounts had high bot scores (*bot-like*), and Twitter users who provided their handles and had low bot scores (*valid*). We explored whether there were significant differences among these groups in terms of their sociodemographic features, clinical symptoms, and aspects of social media use (ie, platforms used and time).

**Results:** In SOCIAL I, most individuals were classified as valid (580/1123, 51.65%), and a few were deemed bot-like (190/1123, 16.91%). A total of 31.43% (353/1123) gave no handle or gave an invalid handle (eg, entered "N/A"). In SOCIAL II, many individuals were not Twitter users (760/1988, 38.23%). Of the Twitter users in SOCIAL II (1228/1988, 61.78%), most were classified as either invalid (515/1228, 41.94%) or valid (484/1228, 39.41%), with a smaller fraction deemed bot-like (229/1228, 18.65%). Participants reported high rates of mental health diagnoses as well as high levels of symptoms, especially in SOCIAL II. In general, the differences between individuals who provided or did not provide their social media handles were small and not statistically significant.

**Conclusions:** Triangulating passively acquired social media data and self-reported questionnaires offers new possibilities for large-scale assessment and evaluation of vulnerability to mental disorders. The propensity of participants to share social media handles is likely not a source of sample bias in subsequent social media analytics.

(*JMIR Form Res* 2022;6(10):e39324) doi:[10.2196/39324](https://doi.org/10.2196/39324)

## KEYWORDS

depression; anxiety; pain; alcohol; social media

## Introduction

### Background

So-called mental disorders, including depression, anxiety, substance use, and pain-related conditions, account for a substantial proportion of disabilities attributed to illness worldwide [1]. According to hierarchical models of psychopathology [2], most of these clinical problems can be grouped into dimensions that include an *internalizing* dimension, involving emotional dysfunction, and an *externalizing* dimension, involving disinhibition or antagonism. Research implicates various mechanisms in the etiology and maintenance of mental disorder symptoms, including sustained negative affect, disturbances in positive affect, disrupted social processes, disturbances in arousal and regulatory processes, sensorimotor problems, and cognitive dysfunction [3]. However, it has been extremely difficult to determine reliable mechanisms of psychopathology. Although mental disorders are very common [4], they are also highly heterogeneous in their presenting characteristics [2]. In addition, the longitudinal course of mental health symptoms is also heterogeneous, with some individuals having brief courses and others having highly chronic or relapsing-recovering courses [5].

### Social Media

Characterizing heterogeneity in psychopathology requires large samples, which, as a result, have become a staple of modern clinical research, that is, clinical trials such as STAR\*D (Sequenced Treatment Alternatives to Relieve Depression) [6], epidemiological studies [7], neuroimaging cohorts [8], and neurocognitive assessment studies [9]. More recently, analyses of naturalistic social media samples have also facilitated the collection of large samples. Social media is well-suited for collecting research data because it is ubiquitous in modern life; 72% of adults in the United States report belonging to at least one social media platform [10]. Twitter, specifically, is used by 23% of the population in the United States [10]. Although the use of Twitter has a Pareto distribution, wherein a few individuals account for most of the active Twitter activity; approximately three-fourths of Twitter users use the platform at least once a week (46% use it daily and 27% use it at least weekly). As a social media platform, Twitter is geared toward sharing frequent, brief, and introspective posts that are suitable for longitudinal, within-subject text analysis at high temporal resolutions.

We had used Twitter previously to study vulnerability to mental health symptoms. For example, in a study, we reported that individuals who had disclosed that they were diagnosed with depression in their tweets (eg, “I was diagnosed with depression

a couple of months ago...”) had different circadian patterns of Twitter activity than a random sample of Twitter users [11]. Specifically, individuals who disclosed a depression diagnosis used Twitter more frequently later into the night and used Twitter less frequently earlier in the day, possibly indicating circadian differences between the depressed users and the random sample. In another study, we measured lexical proxies of cognitive distortions, words like “should,” “must,” “have to,” “nobody,” or “always,” a concept from the literature on cognitive behavioral therapy which points to rigid or inflexible thinking [12,13]. As suggested by the generic cognitive model underlying cognitive behavioral therapy [14], individuals with depression make more use of cognitive distortions than a random sample of individuals [15]. Al-Mosaiwi and Johnstone [16] reported a similar finding with language that they deemed “absolutist.” Others have also found associations between features of written text and depressive symptoms. For example, greater use of personal pronouns (eg, “I”) in social media and other contexts appears to be correlated to symptoms of depression [17], a finding that connects with research using cognitive tasks linking depression to increased self-referential processing [18]. Similarly, greater use of negative emotional words, including those expressing depressive symptoms, appears to be related to depressive symptoms [19].

In spite of the potential offered by social media data for research into the mechanisms involved in the development and maintenance of mental disorders, there are limitations to passively acquired social media data. A limitation is that social media users are not representative of the general population [10,20]. There are data on sociodemographic differences between individuals who use specific social media sites and those who do not. Relative to the broader population, Twitter users are more likely to be male, younger, more educated, and more liberal leaning in their political orientation [10].

It has also been hypothesized that differences in variables such as need for self-disclosure [20] may bias samples of individuals who are on Twitter versus those who are not. Likewise, individuals who volunteer to give researchers access to their social media accounts may provide a biased subsample of individuals with a stronger disposition to self-disclose. Another limitation to using social media data for research is that researchers lack information to support inferences about participants’ health from their web-based activity (eg, Is someone actually depressed even if they explicitly said so?).

### This Study

To address these limitations of social media research, namely the lack of sample representativeness and inability to verify health status, we conducted the Studies of Online Cohorts for

Internalizing Symptoms and Language (SOCIAL). SOCIAL are cohort-based studies in which we triangulated self-reported disorder screening questionnaires with data acquired from social media. Participants in SOCIAL I and II completed a series of disorder screening questionnaires focused on internalizing symptoms that are meant to capture psychopathology more broadly. They were also asked to provide their Twitter handles which we subsequently verified for validity, including how closely they resembled the behavior of bots. SOCIAL I is a sample of Twitter users (Methods section) targeted to be nationally representative in terms of age, sex assigned at birth, race, and ethnicity, and SOCIAL II is a large sample of college students.

Here, we describe the baseline sociodemographic characteristics, social media use data, and mental health characteristics of individuals in SOCIAL I and II. Because we asked individuals to self-report whether they used Twitter and to give us access to their Twitter accounts, we could compare sociodemographic characteristics, social media use data, and mental health differences between groups of individuals depending on their

willingness to share their social media data. We distinguished approximately 2 groups of participants: those who provided *valid* Twitter handles pointing to their own social media content and those who did not or refused. The latter group can be separated into three subgroups: (1) users who refused to provide a valid Twitter handle (*invalid handle*), (2) users who denied being Twitter users (*not a Twitter user*), and (3) users who did provide an existing Twitter handle, but the accounts were deemed to be *bot-like* as defined by a machine learning classifier [21].

## Methods

### Overview

Both SOCIAL samples answered self-reported questionnaires probing internalizing, externalizing, somatoform, and thought disorder symptoms (Table 1). We also collected demographic information and aspects of social media use, including whether the individual was a Twitter user, whether they were willing to let us access their Twitter time line, and which other social media platforms they used.

**Table 1.** Assessment of psychopathology for the Studies of Online Cohorts for Internalizing Symptoms and Language (SOCIAL I, N=1123 and SOCIAL II, N=1988).

Construct	Measure	Item	Response options	Original range	Cronbach $\alpha$
<b>Internalizing</b>					
Depression	PHQ <sup>a</sup> -9	9	0-3 (not at all to nearly every day)	0-27	.90
Stress	MIDUS <sup>b</sup>	9	0-10 (no stress to severe stress)	0-90	.84
Social anxiety	DSM severity <sup>c</sup>	10	0-4 (never to all of the time)	0-40	.94
Panic	DSM severity	10	0-4 (never to all of the time)	0-40	.95
Agoraphobia	DSM severity	10	0-4 (never to all of the time)	0-40	.96
Worry	DSM severity	10	0-4 (never to all of the time)	0-40	.93
<b>Somatoform</b>					
Pain	PHQ-15	15	0-2 (not bothered a lot to bothered a lot)	0-30	.86
Insomnia	ISI <sup>d</sup>	7	Varies with each question	0-28	.87
<b>Externalizing</b>					
Alcohol use	AUDIT <sup>e</sup>	10	Varies with each question	0-40	.90
Substance use	DSM severity	10	0-4 (not at all to nearly every day)	0-40	.89
<b>Thought disorder</b>					
Hypomania	ASRM <sup>f</sup>	5	Varies with each question	0-25	.82

<sup>a</sup>PHQ: Patient Health Questionnaire.

<sup>b</sup>MIDUS: Midlife in the United States self-reported measure of perceived stress.

<sup>c</sup>DSM severity: Diagnostic and Statistical Manual of Mental Disorders severity measure for each symptom.

<sup>d</sup>ISI: Insomnia Severity Index.

<sup>e</sup>AUDIT: Alcohol Use Disorders Identification Test.

<sup>f</sup>ASRM: Altman Self-Rating Mania.

### Participant Recruitment

SOCIAL I purposefully sampled individuals via Qualtrics panels. We aimed to recruit approximately 1000 Twitter users,

given the budgetary constraints for this study. Individuals were recruited from July 2020 to March 2021 for a study on “social media and mental health.” The sample was selected to represent



the United States at the intersections of age, gender, race and ethnicity.

All individuals in SOCIAL I were Twitter users. Accordingly, we could not ascertain the role that being a Twitter user in itself has on potential differences between individuals in baseline sociodemographic characteristics, social media use, and mental health symptoms. To have a sample of individuals who did not use Twitter as well as to have an additional sample with which to assess the transportability of results from SOCIAL I, we began the SOCIAL II study. SOCIAL II recruited college students from a predominantly White and Asian university in the Midwest. Individuals were compensated with credits in an introductory psychology course. Individuals were recruited from September 2020 to the present date.

## Measures

For individuals in SOCIAL I and SOCIAL II, we collected information on characteristics described in the following sections.

### Demographic Characteristics

Specifically, we collected age, political orientation on a 10-point Likert scale (1=extremely liberal, 10=extremely conservative), race, ethnicity, sex assigned at birth (male, female, other or inconclusive, or prefer not to say), gender identity (male, female, nonbinary, genderqueer, agender, other, or prefer not to say), and sexual orientation (heterosexual or straight, homosexual or gay, bisexual or pansexual, other, or prefer not to say). In SOCIAL I, we asked participants for their annual household income. In SOCIAL II, we asked participants to estimate their *parents'* annual household income. We present both of these as the same variable (ie, estimated household income). In addition, in SOCIAL I, we asked participants to indicate their race by using a single category from a list (White, Black or African American, American Indian or Alaska Native, Native Hawaiian or Pacific Islander, Hispanic, or other). In SOCIAL II, we allowed participants to select multiple racial and ethnic identities, including all the possibilities in SOCIAL I along with Middle Eastern or North African. We recoded the categories in SOCIAL II to fit a version of the race variable in SOCIAL I that identified whether individuals were non-Hispanic White, non-Hispanic Black, Hispanic, Asian, or other (eg, Native Hawaiian or Pacific Islander, Middle Eastern or North African, or multiracial but not Hispanic).

### Social Media

Individuals who were Twitter users (ie, all individuals in SOCIAL I and some in SOCIAL II) were queried about how much time they spent on Twitter (less than once every few weeks; every few weeks; a few days (more like 1-2) a week; a few days (3-5) a week; about once a day; or several times a day). This item was selected on the basis that it was used by the Pew Research Center in a previous study on social media use in the United States. In addition, all individuals were queried about their use of Twitter as well as other social media platforms on a binary scale (ie, user vs nonuser of that platform). We asked for a Twitter handle for all individuals in the study who identified that they were Twitter users. Individuals could choose

to enter a valid Twitter handle or to enter text to bypass the question (eg "I don't want to give my Twitter handle").

### Mental Health

We compiled a battery of self-report disorder screening questionnaires for psychopathology (Table 1). These measures were chosen because (1) they measure symptoms that are relatively common (eg, depression) or relatively uncommon but highly impairing (eg, drug use), (2) they are indicators of some of the major domains of psychopathology as per contemporary nosologies (eg, the study by Kotov et al [2]), (3) they were freely available, and (4) they are widely used. Most of the measures we used were the Diagnostic and Statistical Manual of Mental Disorders (DSM) severity measures recommended by the American Psychiatric Association (eg, social anxiety, panic, worry, and substance use) or were measures that were eventually adapted into the *DSM* severity measures (ie, Patient Health Questionnaire (PHQ)-9 and PHQ-15 for depression and somatic symptoms, respectively). Given that all these measures have different response types and number of items, and accordingly different ranges, we standardized them all as percentage of maximum point (POMP) scores [22]. The POMP scores are defined as follows:  $POMP = ((\text{observed score} - \text{minimum possible}) / (\text{maximum possible} - \text{minimum possible})) \times 100$ . This represents the percentage of a measure's total that a specific score represents. For example, for the PHQ-9, with its score range of 0 to 27, a score of 0 is 0% of the POMP, 14 is 51.85%, and 27 is 100%. In addition to characterizing the symptoms of psychopathology that individuals currently experienced, we also asked them about whether they were aware of having received a medical diagnosis of the following mental disorders: depression, social anxiety, generalized anxiety, specific phobia, panic disorder, agoraphobia, posttraumatic stress disorder, somatic symptom disorder (or "chronic pain"), insomnia, alcohol use, drug use, or bipolar disorder (I or II). Individuals were allowed to answer "yes," "no," "no, but I should be," or "I don't know." In this study, we differentiated between individuals who were sure they had a diagnosis (ie, those answering "yes") and all others.

We conducted preliminary analyses to describe the samples, including the ranges represented in the different variables. The results of these analyses suggested that individuals gave relatively high ratings of self-reported manic symptoms, a problem that has been previously reported in the literature assessing hypomanic symptoms via self-report. Zimmerman [23] suggested that screening for bipolar disorder should be accompanied by a subsequent evaluation by a clinician. Similarly, individuals endorsed relatively few agoraphobic symptoms that were highly correlated with other internalizing symptoms. Considering these factors, we removed the mania rating scale, the Altman Self-Rating Mania Scale as well as the DSM Severity Scale For Agoraphobia from SOCIAL II leaving only a subsample of individuals with ratings on these scales (n=665).

### Twitter Status

All individuals who reported that they were Twitter users were asked to provide their Twitter handles, which identify the individual on Twitter. The Twitter application programming

interface, a free and public interface provided by Twitter, provides access to an individual's past tweets (timelines) via their individual handle (provided the tweets were public). Hence, for individuals who provided a Twitter handle, we retrieved individual timelines (a time-sorted record of their past tweets). We assessed whether the corresponding Twitter accounts were valid and belonged to real users using the Botometer application programming interface, an algorithm that uses machine learning to predict whether a given account belongs to a *bot* from its web-based behavior and content (eg, frequency of posting, specific content features, evidence that they have purchased followers, whether the account self-declares as being a bot, or whether the account has been declared a bot by others). As per recommendations of the Botometer developers, we explored the distribution of bot scores and created a cutoff of 0.42 to classify individuals as *bot-like* or *valid* users.

Individuals were classified as providing *invalid* handles if they refused to provide their handle, answered the question about handles with a response that was not a syntactically valid Twitter account (eg, "I don't want to give you this information"), or if Botometer failed to access the Twitter account. In addition to these 3 groups (ie, *invalid*, *bot-like*, and *valid*), in SOCIAL II, we included individuals who denied being Twitter users (*not a Twitter user*). We focused on the differences between these 3-4 subgroups using Twitter users who did not provide handles or who provided handles that were not syntactically valid account names (ie, the *invalid* handle group).

### Analytic Plan

All analyses were conducted using the R programming language (version 4.1.2) [24] in R Studio [25]. Given that we have collected samples that differ substantially in demographic characteristics, we report all analyses according to the study cohort (ie, first in SOCIAL I and then in SOCIAL II). For continuous variables, we provide descriptive statistics in the form of means, SDs, medians, and IQR values. For categorical variables, we present frequencies and percentages.

To assess statistically significant differences between the sociodemographic factors, social media use, and mental health variables, we tested the association of each of these variables (eg, age, frequency of Twitter use, and depression) with group membership (ie, no handle, bot-like, valid, and no Twitter [on SOCIAL II]). For continuous variables, we reported the *P* values from a Kruskal-Wallis rank-sum test. For categorical variables, we reported the *P* values from a chi-square test to assess whether Twitter group membership is significantly related to specific baseline characteristics (eg, race and gender identity) or the *P* values from Fisher exact test when a cell size is <5. To characterize the magnitude of these associations (ie, the strength

of the effect beyond its statistical significance), for binary variables, we report odds ratios (ORs) with 95% CIs when using individuals who provided invalid user names as the reference group. For nominal variables (eg, gender as male, female, or nonbinary), we report a Cramer V. For continuous variables, we report the standardized  $\beta$  values and 95% CIs, representing the differences in SD units of each variable in question.

### Ethics Approval

Both studies were approved by the Indiana University Institutional Review Board (2002549202 and 2005948214).

## Results

### SOCIAL I

#### Demographic Characteristics

In SOCIAL I (N=1123), the average participant was in their mid-30s, although there was variability in the ages represented (Table 2). Approximately half of the individuals (580/1123, 51.65%) provided valid Twitter handles. For the remainder (ie, the 543/1123, 48.35% who did not provide valid Twitter handles), most were individuals who provided invalid Twitter handles (353/1123, 31.43%) with only 16.92% (190/1123) of people providing Twitter handles that were deemed to be *bot-like*. Most of those who used Twitter reported using the platform at least "several times a day." Individuals were approximately split along the political spectrum and there appeared to be variability in sexual orientation, gender identity, and socioeconomic status. Hispanic and Asian individuals appeared to be underrepresented relative to the population from the United States.

There were various statistically significant demographic differences among individuals in SOCIAL I based on Twitter status (Table 2). In general, we focused on differences relevant to the individuals who provided valid Twitter handles versus those who refused to provide a handle or provided an invalid one (eg, we ignored differences between people who provided invalid handles vs bot-like handles). Compared with Twitter users who provided invalid handles, Twitter users who provided valid handles were more liberal ( $\beta=-0.14$ , 95% CI  $-0.20$  to  $-0.07$ ), used Twitter less (OR 0.73, 95% CI 0.56-0.95), and reported lower incomes (OR 0.58, 95% CI 0.46-0.74). In addition, compared with Twitter users who provided invalid handles, Twitter users who provided valid handles were relatively more likely to identify as genderqueer, nonbinary, or otherwise unwilling to use male or female designation than to identify as male (Cramer V=0.14, 95% CI 0.11-0.19) and were relatively more likely to identify as gay, lesbian, or bisexual than as heterosexual (OR 1.59, 95% CI 1.12-2.28).

**Table 2.** Sociodemographic characteristics of web-based panel respondents to the Studies of Online Cohorts for Internalizing Symptoms and Language (SOCIAL I), overall and by Twitter status (N=1123).

Characteristic	Twitter user, bot-like handle (n=190 <sup>a</sup> )	Twitter user, invalid handle (n=353 <sup>b</sup> )	Twitter user, valid handle (n=580 <sup>c</sup> )	P value <sup>d</sup>
<b>Age (years)</b>				<.001
Mean (SD)	38.41 (12.86)	34.20 (11.80)	33.75 (13.17)	
Median (IQR)	36.00 (29.00-45.00)	35.00 (24.00-40.00)	31.00 (22.00-41.00)	
Unknown, n (%)	3 (1.58)	18 (5.10)	29 (5.00)	
<b>Political orientation, rating (1-10; 1=extremely liberal, 10=extremely conservative)</b>				<.001
Mean (SD)	5.12 (2.58)	5.31 (2.57)	4.62 (2.52)	
Median (IQR)	5.00 (3.00-7.00)	5.00 (3.00-7.00)	5.00 (2.00-7.00)	
Unknown, n (%)	2 (1.05)	3 (0.89)	7 (1.21)	
<b>Time spent on Twitter, n (%)</b>				.17
Less than every few weeks	5 (2.63)	6 (1.70)	13 (2.2)	
Every few weeks	10 (5.26)	14 (3.97)	21 (3.6)	
A few days (eg, 1-2 days) a week	5 (2.63)	15 (4.25)	33 (5.7)	
A few days (eg, 3-5 days) a week	16 (8.42)	31 (8.78)	77 (13)	
About once a day	27 (14.21)	56 (15.86)	101 (17)	
Several times a day	127 (66.84)	231 (65.44)	335 (58)	
<b>Race and ethnicity, n (%)</b>				.13
Non-Hispanic White	148 (77.89)	262 (74.43)	416 (72)	
Non-Hispanic Black	22 (11.58)	39 (11.08)	67 (12)	
Hispanic	6 (3.16)	26 (7.39)	55 (9.5)	
Asian	13 (6.84)	17 (4.83)	32 (5.5)	
Others	1 (0.53)	8 (2.27)	10 (1.7)	
Unknown	0 (0)	1 (0.28)	0 (0)	
<b>Gender, n (%)</b>				<.001
Woman	80 (42.11)	163 (46.18)	80 (42.11)	
Genderqueer or nonbinary	2 (1.05)	3 (0.85)	2 (1.05)	
Man	108 (56.84)	187 (52.97)	108 (56.84)	
<b>Sexual orientation, n (%)</b>				.005
Heterosexual	165 (86.84)	301 (85.27)	165 (86.84)	
LGB <sup>e</sup>	25 (13.16)	52 (14.73)	25 (13.16)	
<b>Yearly income (US \$), n (%)</b>				<.001
<10,000	8 (4.21)	29 (8.22)	8 (4.21)	
10,000-19,999	15 (7.89)	21 (5.95)	15 (7.89)	
20,000-29,999	23 (12.11)	35 (9.92)	23 (12.11)	
30,000-39,999	21 (11.05)	35 (9.92)	21 (11.05)	
40,000-49,999	9 (4.74)	19 (5.38)	9 (4.74)	
50,000-59,999	7 (3.68)	25 (7.08)	7 (3.68)	
60,000-69,999	9 (4.74)	12 (3.40)	9 (4.74)	
70,000-79,999	15 (7.89)	22 (6.23)	15 (7.89)	
80,000-89,999	6 (3.16)	9 (2.55)	6 (3.16)	
90,000-99,999	12 (6.32)	18 (5.10)	12 (6.32)	

Characteristic	Twitter user, bot-like handle (n=190 <sup>a</sup> )	Twitter user, invalid handle (n=353 <sup>b</sup> )	Twitter user, valid handle (n=580 <sup>c</sup> )	P value <sup>d</sup>
100,000-149,999	34 (17.89)	71 (20.11)	34 (17.89)	
≥150,000	31 (16.32)	57 (16.15)	31 (16.32)	

<sup>a</sup>Bot-like handle: individual reported being a Twitter user and provided a Twitter handle, but it had a Botometer score >0.42.

<sup>b</sup>Invalid handle: individual reported being a Twitter user but did not provide their Twitter handle or provided a handle that was invalid.

<sup>c</sup>Valid handle: individual reported being a Twitter user and provided a Twitter handle, and it had a Botometer score ≤0.42.

<sup>d</sup>Kruskal-Wallis rank-sum test; Fisher exact test for count data with simulated *P* value (based on 2000 replicates); Pearson chi-square test.

<sup>e</sup>LGB: lesbian, gay, bisexual (or other nonheterosexual sexual orientation).

## Social Media Use

In SOCIAL I, all individuals were recruited to be Twitter users. Other social media platforms used by most of the sample were Facebook, Instagram, and YouTube (Table 3). There appeared to be several statistically significant differences in social media use by Twitter status, but these effects were mostly attributable

to bot-like users (eg, bot-like users were more likely to report being on LINE than valid users). Individuals who provided valid Twitter handles were more likely to report that they used Tumblr (OR 1.48, 95% CI 1.07-2.05) and Pinterest than individuals who provided invalid Twitter handles (OR 1.79, 95% CI 1.37-2.34).

**Table 3.** Social media platforms used by respondents to the Studies of Online Cohorts for Internalizing Symptoms and Language (SOCIAL), by cohort (SOCIAL I and SOCIAL II) and Twitter status.

Social media platforms	SOCIAL I (N=1123)				SOCIAL II (N=1988)				
	Twitter user, bot-like handle (n=190) <sup>a</sup> , n (%)	Twitter user, invalid handle (n=353) <sup>b</sup> , n (%)	Twitter user, valid handle (n=580) <sup>c</sup> , n (%)	<i>P</i> value <sup>d</sup>	Twitter user, bot-like handle (n=229), n (%)	Not a Twitter user (n=760), n (%)	Twitter user, invalid handle (n=515), n (%)	Twitter user, valid handle (n=484), n (%)	<i>P</i> value <sup>e</sup>
Twitter	190 (100)	353 (100)	580 (100)	— <sup>f</sup>	229 (100)	0 (0)	515 (100)	484 (100)	<.001
Facebook	170 (89.47)	311 (88.1)	523 (90.17)	.61	198 (86.46)	427 (56.18)	421 (81.75)	401 (82.85)	<.001
Instagram	158 (83.16)	321 (90.93)	522 (90.00)	.01	227 (99.13)	693 (91.18)	503 (97.67)	478 (98.76)	<.001
Snapchat	103 (54.21)	220 (62.32)	361 (62.24)	.12	224 (97.82)	684 (90.00)	499 (96.89)	476 (98.35)	<.001
Tumblr	47 (24.74)	68 (19.26)	151 (26.03)	.06	26 (11.35)	40 (5.26)	64 (12.43)	60 (12.4)	<.001
YouTube	164 (86.32)	291 (82.44)	486 (83.79)	.5	198 (86.46)	534 (70.26)	418 (81.17)	380 (78.51)	<.001
TikTok	75 (39.47)	178 (50.42)	267 (46.03)	.05	200 (87.34)	516 (67.89)	425 (82.52)	424 (87.6)	<.001
Reddit	61 (32.11)	98 (27.76)	176 (30.34)	.53	48 (20.96)	98 (12.89)	111 (21.55)	91 (18.8)	<.001
4Chan	12 (6.32)	11 (3.12)	9 (1.55)	.003	1 (0.44)	1 (0.13)	0 (0)	2 (0.41)	.32
Pinterest	98 (51.58)	154 (43.63)	337 (58.1)	<.001	153 (66.81)	389 (51.18)	332 (64.47)	344 (71.07)	<.001
Twitch	51 (26.84)	69 (19.55)	119 (20.52)	.11	40 (17.47)	65 (8.55)	81 (15.73)	84 (17.36)	<.001
LINE	22 (11.58)	24 (6.8)	34 (5.86)	.03	6 (2.62)	14 (1.84)	5 (0.97)	6 (1.24)	.30
Other	8 (4.21)	9 (2.55)	18 (3.1)	.57	10 (4.37)	27 (3.55)	22 (4.27)	15 (3.1)	.73

<sup>a</sup>Bot-like: individual reported being a Twitter user and provided a Twitter handle, but it had a Botometer score >0.42.

<sup>b</sup>Invalid handle: individual reported being a Twitter user but did not provide their Twitter handle or provided a handle that was invalid.

<sup>c</sup>Valid handle: Individual reported being a Twitter user and provided a Twitter handle with a Botometer score ≤0.42.

<sup>d</sup>Pearson chi-square test.

<sup>e</sup>Pearson chi-square test; Fisher exact test for count data with simulated *P* value (based on 2000 replicates).

<sup>f</sup>Not available.

## Mental Health

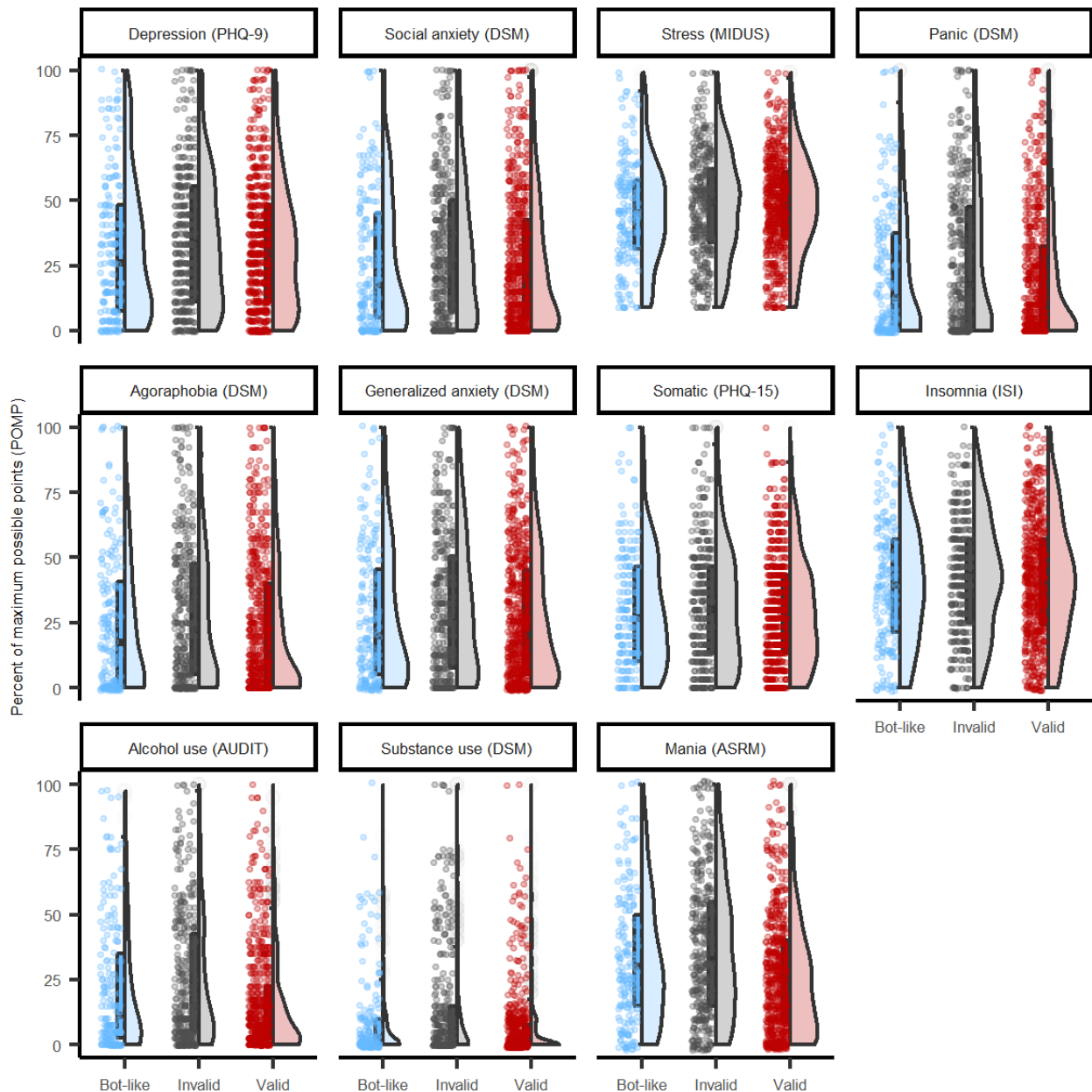
The POMP scores for the various measures of psychopathology as well as the reported diagnoses are presented in Figure 1 and Table 4. Stress and insomnia were the most commonly endorsed

symptoms. Major depression and generalized and social anxiety were the most commonly reported clinical diagnoses. There were a few statistically significant differences between the groups in clinical symptoms or diagnoses. The differences we did find were very small. For example, the largest difference

between the groups was in self-reported manic symptoms and suggested individuals who provided valid Twitter handles had lower symptoms of hypomania than individuals who did not provide Twitter handles, although this difference was small by conventional standards ( $\beta=-0.22$ , 95% CI  $-0.28$  to  $-0.15$ ). The next highest difference between the groups was in self-reported issues with alcohol and suggested that individuals who provided valid Twitter handles had lower alcohol use symptoms than

individuals who did not provide Twitter handles, although this difference was small ( $\beta=-0.19$ , 95% CI  $-0.25$  to  $-0.12$ ). Compared with individuals who provided invalid handles, individuals who provided valid handles were less likely to report relatively rare diagnoses, such as somatic symptom disorder (OR 0.38, 95% CI 0.21-0.68) and drug use disorder (OR 0.60, 95% CI 0.39-0.92)

**Figure 1.** Differences in self-reported symptoms of psychopathology for 1123 individuals in the Studies of Online Cohorts for Internalizing Symptoms and Language (SOCIAL) I. ASRM: Altman Self-Rating Mania; AUDIT: Alcohol Use Disorders Identification Test; DSM: Diagnostic and Statistical Manual of Mental Disorders; ISI: Insomnia Severity Index; MIDUS: Midlife in the United States self-reported measure of perceived stress; PHQ: Patient Health Questionnaire.



**Table 4.** Median (IQR) of symptom severity, as percentage of maximum points (0-100) and self-reported diagnoses of psychopathology in web-based respondents to the Studies of Online Cohorts for Internalizing Symptoms and Language (SOCIAL I), by Twitter status (N=1123).

Variable	Twitter user, bot-like handle (n=190 <sup>a</sup> )	Twitter user, invalid handle (n=353 <sup>b</sup> )	Twitter user, valid handle (n=580 <sup>c</sup> )	P value <sup>d</sup>
<b>Social anxiety (DSM<sup>e</sup>)</b>				.02
Median (IQR)	17.50 (5.00-45.00)	25.00 (7.50-50.00)	20.00 (5.00-42.50)	
Unknown, n (%)	2 (1.05)	1 (0.28)	0 (0)	
<b>Stress (MIDUS<sup>f</sup>)</b>				.17
Median (IQR)	45.00 (31.50-57.50)	49.00 (34.00-62.00)	48.00 (36.00-61.00)	
Unknown (n)	3 (1.58)	0 (0)	0 (0)	
<b>Depression (PHQ<sup>g-9</sup>)</b>				.21
Median (IQR)	25.93 (7.41-48.15)	33.33 (11.11-55.56)	29.63 (11.11-48.15)	
Unknown, n (%)	2 (1.05)	1 (0.28)	0 (0)	
<b>Panic (DSM)</b>				<.001
Median (IQR)	12.50 (0.00-37.50)	17.50 (0.00-47.50)	7.50 (0.00-32.50)	
Unknown, n (%)	2 (1.05)	2 (0.57)	0 (0)	
<b>Agoraphobia (DSM)</b>				<.001
Median (IQR)	17.50 (0.00-40.62)	22.50 (5.00-47.50)	12.50 (0.00-40.00)	
Unknown, n (%)	2 (1.05)	1 (0.28)	1 (0.17)	
<b>Generalized anxiety (DSM)</b>				.04
Median (IQR)	20.00 (5.00-45.62)	25.00 (7.50-50.62)	20.00 (5.00-45.00)	
Unknown, n (%)	2 (1.05)	1 (0.28)	1 (0.17)	
<b>Somatic (PHQ-15)</b>				.75
Median (IQR)	26.67 (10.00-46.67)	30.00 (13.33-46.67)	26.67 (13.33-43.33)	
Unknown, n (%)	5 (2.63)	7 (1.98)	2 (0.34)	
<b>Insomnia (ISI<sup>h</sup>)</b>				.38
Median (IQR)	39.29 (21.43-57.14)	42.86 (25.00-57.14)	39.29 (24.11-57.14)	
Unknown, n (%)	39.29 (21.43-57.14)	42.86 (25.00-57.14)	39.29 (24.11-57.14)	
Alcohol use (AUDIT <sup>i</sup> ), median (IQR)	11.25 (2.50-35.00)	15.00 (5.00-42.50)	7.50 (2.50-22.50)	<.001
<b>Substance use (DSM)</b>				<.001
Median (IQR)	2.50 (0.00-10.00)	5.00 (0.00-15.00)	0.00 (0.00-7.50)	
Unknown, n (%)	1 (0.52)	1 (0.28)	4 (0.68)	
<b>Hypomania (ASRM<sup>j</sup> scale)</b>				<.001
Median (IQR)	30.00 (15.00-50.00)	32.50 (15.00-55.00)	25.00 (8.75-40.00)	
Unknown, n (%)	3 (1.58)	1 (0.28)	0 (0)	
<b>Insomnia Dx<sup>k</sup>, n (%)</b>				.45
Yes	48 (25.53)	89 (25.36)	127 (22.2)	
Unknown	2 (1.05)	2 (0.57)	8 (1.38)	
<b>Somatic symptom Dx, n (%)</b>				.003
Yes	8 (4.26)	30 (8.62)	20 (3.50)	
Unknown	2 (1.05)	5 (1.41)	8 (1.38)	
<b>Major depression Dx, n (%)</b>				.57
Yes	73 (38.62)	147 (42.12)	247 (43.03)	

Variable	Twitter user, bot-like handle (n=190 <sup>a</sup> )	Twitter user, invalid handle (n=353 <sup>b</sup> )	Twitter user, valid handle (n=580 <sup>c</sup> )	<i>P</i> value <sup>d</sup>
Unknown	1 (0.52)	4 (1.13)	6 (1.03)	
<b>Specific phobia Dx, n (%)</b>				.10
Yes	28 (14.89)	62 (17.77)	72 (12.59)	
Unknown	2 (1.05)	4 (1.13)	8 (1.38)	
<b>Social anxiety Dx, n (%)</b>				.10
Yes	37 (19.68)	96 (27.67)	153 (26.7)	
Unknown	2 (1.05)	6 (1.70)	7 (1.21)	
<b>Panic Dx, n (%)</b>				.79
Yes	35 (18.82)	68 (19.60)	102 (17.8)	
Unknown	4 (2.10)	6 (1.70)	7 (1.21)	
<b>Posttraumatic stress Dx, n (%)</b>				.32
Yes	33 (17.74)	46 (13.26)	93 (16.29)	
Unknown	4 (2.10)	6 (1.70)	9 (1.55)	
<b>Generalized anxiety Dx, n (%)</b>				.08
Yes	50 (26.74)	92 (26.44)	188 (32.75)	
Unknown	3 (1.58)	5 (1.41)	6 (1.03)	
<b>Agoraphobia Dx, n (%)</b>				.25
Yes	9 (4.81)	26 (7.51)	29 (5.07)	
Unknown	3 (1.58)	7 (1.98)	8 (1.38)	
<b>Alcohol use Dx, n (%)</b>				.11
Yes	23 (12.30)	42 (12.07)	48 (8.35)	
Unknown	3 (1.58)	5 (1.41)	5 (0.86)	
<b>Substance use Dx, n (%)</b>				.048
Yes	23 (12.37)	46 (13.22)	48 (8.39)	
Unknown	4 (2.10)	5 (1.41)	8 (1.38)	
<b>Bipolar Dx, n (%)</b>				.68
Yes	23 (12.23)	40 (11.49)	58 (10.18)	
Unknown	2 (1.05)	5 (1.41)	10 (1.72)	

<sup>a</sup>Bot-like: individual reported being a Twitter user and provided a Twitter handle, but it had a Botometer score >0.42.

<sup>b</sup>Invalid handle: individual reported being a Twitter user but did not provide their Twitter handle or provided a handle that was invalid.

<sup>c</sup>Valid handle: individual reported being a Twitter user and provided a Twitter handle, and it had a Botometer score ≤0.42.

<sup>d</sup>Kruskal-Wallis rank-sum test; Pearson chi-square test.

<sup>e</sup>DSM: Diagnostic and Statistical Manual of Mental Disorders.

<sup>f</sup>MIDUS: Midlife in the United States self-reported measure of perceived stress.

<sup>g</sup>PHQ: Patient Health Questionnaire.

<sup>h</sup>ISI: Insomnia Severity Index.

<sup>i</sup>AUDIT: Alcohol Use Disorders Identification Test.

<sup>j</sup>ASRM: Altman Self-Rating Mania.

<sup>k</sup>Dx: diagnosis.

## Studies of Online Cohorts for Internalizing Symptoms and Language II

### *Demographic Characteristics*

In SOCIAL II (N=1988), age was more restricted to the range 18 to 22 years, as would be expected of undergraduate students (mean 19.07, SD 2.91 years; [Table 5](#)). The sample was primarily female ([Table 5](#)), as is typical of psychology students, and primarily White and Asian, which is consistent with the demographic characteristics of the institution. A total of 32.22% (760/1128) of participants denied being Twitter users. Of the Twitter users (1228/1988, 61.77% of the whole sample), most either refused to give handles (515/1228, 41.94%) or gave valid handles (484/1228, 39.41%), with a minority providing handles that were deemed to be bot-like (229/1228, 18.65%). Of those who reported being Twitter users, approximately half of the individuals used Twitter “about once a day.” Approximately

half of all the students reported that their parents made  $\geq$ US \$100,000 and the others were distributed relatively uniformly across the income categories. Politically, they reported aligning with the politics center of the liberal-conservative continuum.

There were 3 statistically significant differences between individuals based on their Twitter user status, of which 2 involved the individuals with valid Twitter user names. First, individuals who provided valid Twitter handles used Twitter more frequently than individuals who were Twitter users but did not provide their handles or provided invalid handles (OR 2.48, 95% CI 1.98-3.13). Second, there were differences in reported race and ethnicity by Twitter user status (Cramer V=0.03, 95% CI 0.02-0.07). Specifically, individuals who provided valid Twitter handles were less likely to be Hispanic than individuals who were Twitter users but did not provide their handles (OR 0.40, 95% CI 0.21-0.77).



**Table 5.** Sociodemographic characteristics of young adult respondents to the Studies of Online Cohorts for Internalizing Symptoms and Language (SOCIAL II), overall and by Twitter status (N=1988).

Characteristic	Twitter user, bot-like handle (n=229) <sup>a</sup>	Not a Twitter user (n=760)	Twitter user, invalid handle (n=515) <sup>b</sup>	Twitter user, valid handle (n=484) <sup>c</sup>	P value <sup>d</sup>
<b>Age (years)</b>					.02
Mean (SD)	18.91 (1.55)	19.11 (4.09)	19.13 (2.44)	19.03 (1.12)	
Median (IQR)	19.00 (18.00-19.00)	19.00 (18.00-19.00)	19.00 (18.00-19.00)	19.00 (18.00-20.00)	
Unknown, n (%)	15 (6.55)	50 (6.58)	41 (7.96)	25 (5.17)	
<b>Conservatives, rating (1-10; 1=extremely liberal, 10=extremely conservative)</b>					.22
Mean (SD)	4.12 (2.07)	4.18 (1.94)	4.03 (2.07)	3.96 (2.09)	
Median (IQR)	4.00 (2.00-5.25)	4.00 (3.00-5.00)	4.00 (2.00-5.00)	4.00 (2.00-6.00)	
Unknown, n (%)	5 (2.18)	27 (3.55)	30 (5.82)	8 (1.65)	
<b>Time spent on Twitter, n (%)</b>					<.001
Less than every few weeks	64 (27.95)	N/A <sup>e</sup>	115 (23.23)	29 (5.99)	
Every few weeks	46 (20.09)	N/A	85 (17.17)	43 (8.88)	
A few days (eg, 1-2 days) a week	33 (14.41)	N/A	48 (9.70)	67 (13.84)	
A few days (eg, 3-5 days) a week	19 (8.30)	N/A	30 (6.06)	39 (8.06)	
About once a day	38 (16.59)	N/A	86 (17.37)	112 (23.14)	
Several times a day	29 (12.66)	N/A	131 (26.46)	194 (40.08)	
Unknown	0 (0)	760 (100)	20 (3.88)	0 (0.00)	
<b>Race and ethnicity, n (%)</b>					<.001
Non-Hispanic White	167 (73.57)	519 (68.92)	357 (72.71)	376 (78.50)	
Asian	27 (11.89)	116 (15.41)	44 (8.96)	34 (7.10)	
Non-Hispanic Black	18 (7.93)	50 (6.64)	44 (8.96)	45 (9.39)	
Hispanic	13 (5.73)	44 (5.84)	33 (6.72)	14 (2.92)	
Others	2 (0.88)	24 (3.19)	13 (2.65)	10 (2.09)	
Unknown	2 (0.87)	7 (0.92)	24 (4.66)	5 (1.03)	
<b>Gender, n (%)</b>					.75
Woman	179 (78.85)	573 (76.10)	373 (75.97)	368 (76.83)	
Genderqueer or nonbinary	1 (0.44)	8 (1.06)	9 (1.83)	4 (0.84)	
Man	47 (20.70)	172 (22.84)	109 (22.20)	107 (22.34)	
Unknown	2 (0.87)	7 (0.92)	24 (4.66)	5 (1.03)	
<b>Sexual orientation, n (%)</b>					.16
Heterosexual	200 (88.11)	638 (84.73)	401 (81.67)	401 (83.72)	
LGB <sup>f</sup>	27 (11.89)	115 (15.27)	90 (18.33)	78 (16.28)	
Unknown	2 (0.87)	7 (0.92)	24 (4.66)	5 (1.03)	
<b>Yearly income (US \$), n (%)</b>					.99
<10,000	13 (5.75)	37 (5.00)	34 (6.98)	20 (4.22)	
10,000-19,999	5 (2.21)	23 (3.11)	16 (3.29)	18 (3.80)	
20,000-29,999	12 (5.31)	32 (4.32)	23 (4.72)	19 (4.01)	
30,000-39,999	9 (3.98)	39 (5.27)	20 (4.11)	25 (5.27)	
40,000-49,999	9 (3.98)	37 (5.00)	20 (4.11)	22 (4.64)	
50,000-59,999	13 (5.75)	37 (5.00)	21 (4.31)	22 (4.64)	

Characteristic	Twitter user, bot-like handle (n=229) <sup>a</sup>	Not a Twitter user (n=760)	Twitter user, invalid handle (n=515) <sup>b</sup>	Twitter user, valid handle (n=484) <sup>c</sup>	<i>P</i> value <sup>d</sup>
60,000-69,999	9 (3.98)	36 (4.86)	26 (5.34)	26 (5.49)	
70,000-79,999	13 (5.75)	38 (5.14)	21 (4.31)	24 (5.06)	
80,000-89,999	15 (6.64)	39 (5.27)	23 (4.72)	29 (6.12)	
90,000-99,999	13 (5.75)	51 (6.89)	40 (8.21)	37 (7.81)	
100,000-149,999	41 (18.14)	140 (18.92)	104 (21.36)	103 (21.73)	
≥150,000	74 (32.74)	231 (31.22)	139 (28.54)	129 (27.22)	
Unknown	3 (1.31)	20 (2.63)	28 (5.44)	10 (2.07)	

<sup>a</sup>Bot-like: individual reported being a Twitter user and provided a Twitter handle, but it had a Botometer score >0.42.

<sup>b</sup>Invalid handle: individual reported being a Twitter user but did not provide their Twitter handle or provided a handle that was invalid.

<sup>c</sup>Valid handle: individual reported being a Twitter user and provided a Twitter handle and it had a Botometer score ≤0.42.

<sup>d</sup>Kruskal-Wallis rank-sum test; Pearson chi-square test; Fisher exact test for count data with simulated *P* value (based on 2000 replicates).

<sup>e</sup>N/A: not applicable.

<sup>f</sup>LGB: lesbian, gay, bisexual (or other nonheterosexual sexual orientation).

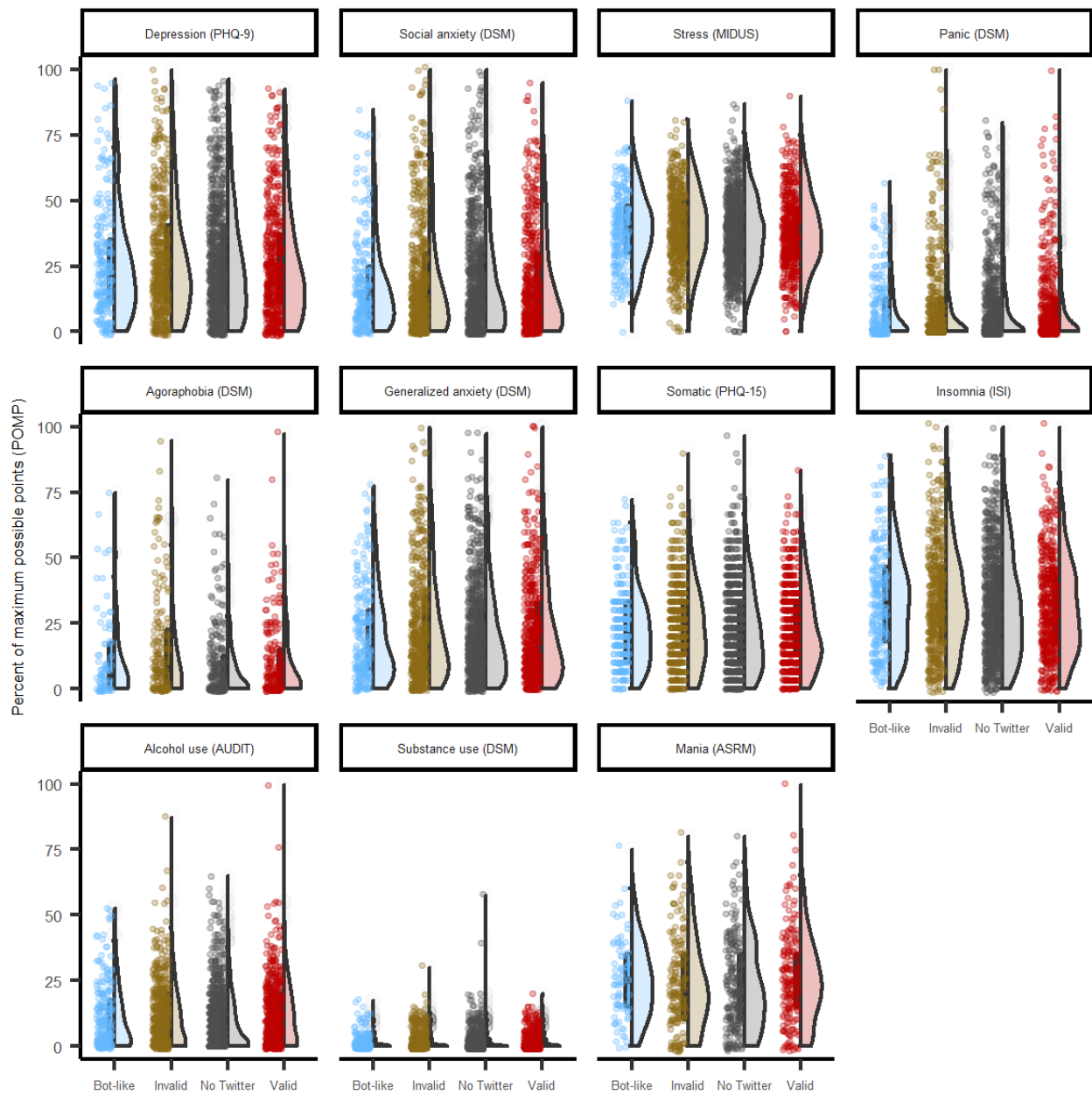
### Social Media

In SOCIAL II, Instagram and Snapchat were the most popular platforms and were used by almost all individuals (Table 3). Twitter, Facebook, YouTube, and TikTok were also quite popular, being used by 61.77% (1228/1988) to 78.72% (1530/1988) of the sample. A number of differences emerged with respect to which social media platforms the participants used. However, most of these differences indicated that individuals who denied being Twitter users were also less likely to use other platforms. The 2 exceptions were that the individuals in SOCIAL II who used Twitter and provided valid handles were more likely to also use TikTok (OR 1.50, 95% CI 1.05-2.14) and Pinterest (OR 1.35, 95% CI 1.04-1.77) than the individuals who refused to provide handles or provided invalid handles.

### Mental Health

The POMP scores for the various measures of psychopathology as well as the reported diagnoses are presented in Figure 2 and Table 6. Similar to SOCIAL I, in SOCIAL II, stress and insomnia were the most commonly endorsed symptoms, and major depression and generalized and social anxiety were the most commonly reported clinical diagnoses. Relative to SOCIAL I (Table 4), there were even fewer statistically significant differences between the groups in clinical symptoms or diagnoses. The largest difference between individuals who provided valid (vs invalid) handles was in agoraphobic symptoms, and it was relatively small in magnitude ( $\beta=-0.13$ , 95% CI 0.03-0.15).

**Figure 2.** Symptoms and self-reported diagnoses of psychopathology in 1988 college students responding to the Studies of Online Cohorts for Internalizing Symptoms and Language (SOCIAL II). ASRM: Altman Self-Rating Mania; AUDIT: Alcohol Use Disorders Identification Test; DSM: Diagnostic and Statistical Manual of Mental Disorders; ISI: Insomnia Severity Index; MIDUS: Midlife in the United States self-reported measure of perceived stress; PHQ: Patient Health Questionnaire.



**Table 6.** Median (IQR) of symptom severity, as percentage of maximum points (0-100) and self-reported diagnoses of psychopathology in web-based respondents in the Studies of Online Cohorts for Internalizing Symptoms and Language (SOCIAL I), by Twitter status (N=1123).

Variable	Twitter user, bot-like (n=229) <sup>a</sup>	Not a Twitter user (n=760)	Twitter user, invalid handle (n=515) <sup>b</sup>	Twitter user, valid handle (n=484) <sup>c</sup>	P value <sup>d</sup>
<b>Social anxiety (DSM<sup>e</sup>)</b>					.63
Median (IQR)	15.00 (5.00-25.00)	12.50 (5.00-30.00)	12.50 (5.00-30.00)	12.50 (2.50-30.00)	
Unknown, n (%)	2 (0.87)	5 (0.66)	22 (4.27)	2 (0.41)	
<b>Stress (MIDUS<sup>f</sup>)</b>					.07
Median (IQR)	39.00 (30.00-48.00)	38.00 (27.00-48.00)	39.00 (29.00-50.00)	37.00 (28.00-47.00)	
Unknown, n (%)	0 (0)	4 (0.53)	21 (4.08)	0 (0)	
<b>Depression (PHQ<sup>g-9</sup>)</b>					.33
Median (IQR)	22.22 (11.11-35.19)	22.22 (11.11-40.74)	22.22 (11.11-40.74)	22.22 (11.11-37.04)	
Unknown, n (%)	2 (0.87)	4 (0.53)	22 (4.27)	1 (0.21)	
<b>Panic (DSM)</b>					.91
Median (IQR)	2.50 (0.00-12.50)	2.50 (0.00-12.50)	2.50 (0.00-12.50)	2.50 (0.00-12.50)	
Unknown, n (%)	2 (0.87)	5 (0.66)	22 (4.27)	2 (0.41)	
<b>Agoraphobia (DSM)</b>					.006
Median (IQR)	5.00 (0.00-17.50)	5.00 (0.00-12.50)	7.50 (0.00-22.50)	5.00 (0.00-15.00)	
Unknown, n (%)	149 (65.06)	533 (70.13)	343 (66.60)	298 (61.57)	
<b>Generalized anxiety (DSM)</b>					.40
Median (IQR)	15.00 (7.50-30.00)	17.50 (7.50-30.00)	17.50 (7.50-35.00)	15.00 (7.50-32.50)	
Unknown, n (%)	2 (0.87)	7 (0.92)	23 (4.47)	2 (0.41)	
<b>Somatic (PHQ-15)</b>					.70
Median (IQR)	20.00 (10.00-33.33)	20.00 (10.00-33.33)	23.33 (13.33-33.33)	20.00 (13.33-33.33)	
Unknown, n (%)	2 (0.87)	8 (1.05)	25 (4.85)	2 (0.41)	
<b>Insomnia (ISI<sup>h</sup>)</b>					.22
Median (IQR)	32.14 (17.86-46.43)	28.57 (17.86-46.43)	32.14 (17.86-46.43)	28.57 (17.86-42.86)	
Unknown, n (%)	1 (0.44)	4 (0.53)	21 (4.08)	0 (0)	
<b>Alcohol use (AUDIT<sup>i</sup>)</b>					<.001
Median (IQR)	7.50 (2.50-17.50)	5.00 (0.00-15.00)	10.00 (2.50-20.00)	10.00 (2.50-20.00)	
Unknown, n (%)	2 (0.87)	8 (1.05)	25 (4.85)	3 (0.62)	
<b>Substance use (DSM)</b>					.02
Median (IQR)	0.00 (0.00-2.50)	0.00 (0.00-2.50)	0.00 (0.00-2.50)	0.00 (0.00-2.50)	
Unknown, n (%)	2 (0.87)	5 (0.66)	21 (4.08)	0 (0)	
<b>Mania or Hypomania (ASRM<sup>j</sup>)</b>					.19
Median (IQR)	25.00 (15.00-35.00)	20.00 (10.00-35.00)	20.00 (10.00-35.00)	25.00 (15.00-35.00)	
Unknown, n (%)	149 (65.06)	533 (70.13)	343 (66.60)	298 (61.57)	
<b>Insomnia Dx<sup>k</sup>, n (%)</b>					.17
Yes	8 (3.49)	52 (6.90)	33 (6.68)	24 (4.97)	
Unknown	0 (0)	6 (0.79)	21 (4.08)	1 (0.21)	
<b>Somatic symptom Dx, n (%)</b>					.03
Yes	2 (0.88)	11 (1.46)	1 (0.20)	1 (0.21)	
Unknown	1 (0.44)	6 (0.79)	21 (4.08)	1 (0.21)	

Variable	Twitter user, bot-like (n=229) <sup>a</sup>	Not a Twitter user (n=760)	Twitter user, invalid handle (n=515) <sup>b</sup>	Twitter user, valid handle (n=484) <sup>c</sup>	<i>P</i> value <sup>d</sup>
<b>Major depression Dx, n (%)</b>					.30
Yes	50 (21.83)	183 (24.27)	136 (27.53)	129 (26.71)	
Unknown	0 (0)	6 (0.79)	21 (4.08)	1 (0.21)	
<b>Specific phobia Dx, n (%)</b>					.58
Yes	10 (4.39)	21 (2.79)	17 (3.44)	13 (2.69)	
Unknown	1 (0.44)	6 (0.79)	21 (4.08)	1 (0.21)	
<b>Social anxiety Dx, n (%)</b>					.62
Yes	32 (14.04)	87 (11.54)	68 (13.77)	62 (12.84)	
Unknown	1 (0.44)	6 (0.79)	21 (4.08)	1 (0.21)	
<b>Panic Dx, n (%)</b>					.43
Yes	13 (5.7)	41 (5.44)	37 (7.49)	26 (5.38)	
Unknown	1 (0.44)	6 (0.79)	21 (4.08)	1 (0.21)	
<b>Posttraumatic stress Dx, n (%)</b>					.50
Yes	6 (2.63)	36 (4.77)	22 (4.45)	18 (3.73)	
Unknown	1 (0.44)	6 (0.79)	21 (4.08)	1 (0.21)	
<b>Generalized anxiety Dx, n (%)</b>					.11
Yes	69 (30.13)	193 (25.60)	155 (31.38)	146 (30.23)	
Unknown	0 (0)	6 (0.79)	21 (4.08)	1 (0.21)	
<b>Agoraphobia Dx, n (%)</b>					.88
Yes	0 (0)	4 (0.53)	2 (0.40)	3 (0.62)	
Unknown	1 (0.44)	6 (0.79)	21 (4.08)	1 (0.21)	
<b>Alcohol use Dx, n (%)</b>					.34
Yes	0 (0)	6 (0.80)	4 (0.81)	1 (0.21)	
Unknown	1 (0.44)	6 (0.79)	21 (4.08)	1 (0.21)	
<b>Substance use Dx, n (%)</b>					.23
Yes	1 (0.44)	16 (2.12)	9 (1.82)	5 (1.04)	
Unknown	1 (0.44)	6 (0.79)	21 (4.08)	1 (0.21)	
<b>Bipolar Dx, n (%)</b>					.005
Yes	0 (0)	23 (3.05)	9 (1.82)	6 (1.24)	
Unknown	1 (0.44)	6 (0.79)	21 (4.08)	1 (0.21)	

<sup>a</sup>Bot-like: individual reported being a Twitter user and provided a Twitter handle, but it had a Botometer score >0.42.

<sup>b</sup>Invalid handle: individual reported being a Twitter user but did not provide their Twitter handle or provided a handle that was invalid.

<sup>c</sup>Valid handle: individual reported being a Twitter user and provided a Twitter handle, and it had a Botometer score ≤0.42.

<sup>d</sup>Kruskal-Wallis rank-sum test; Pearson chi-square test; Fisher exact test for count data with simulated *P* value (based on 2000 replicates).

<sup>e</sup>DSM: Diagnostic and Statistical Manual of Mental Disorders.

<sup>f</sup>MIDUS: Midlife in the United States self-reported measure of perceived stress.

<sup>g</sup>PHQ: Patient Health Questionnaire.

<sup>h</sup>ISI: Insomnia Severity Index.

<sup>i</sup>AUDIT: Alcohol Use Disorders Identification Test.

<sup>j</sup>ASRM: Altman Self-Rating Mania.

<sup>k</sup>Dx: diagnosis.

## Discussion

### Principal Findings

Our results suggest that it is feasible to collect social media data from individuals who also provide information about a breadth of mental health symptoms. We found no evidence that individuals who provide valid Twitter accounts are a biased sample when compared with individuals who provide invalid handles, do not provide their handles, do not use Twitter, or are classified as bot-like. The widespread availability of social media [10] has facilitated research on large samples with longitudinal observations [26-28]. Although the nature of social media activity can be fairly simple (eg, posting short bits of text and sharing audiovisual content), researchers have made well-supported inferences from this activity about the way mood [26-28], sleep patterns [11], social relations [26], and personality [28] manifest in real-world contexts. Most of this work lacks the measurement of clinically relevant variables, such as the validated assessments of depression, anxiety, and other mental disorder symptoms that we used. For example, we conducted a study characterizing the language of individuals who self-identified as having received a clinical diagnosis of depression [15], finding that they use language that is more negative and rigid than that of a random sample [15]. Although the findings obtained using individual self-identification are interesting, they are subject to a variety of possible sample and observation biases and bear replication against validated clinical screening scales such as the ones we used in this study.

We conducted the SOCIAL I and SOCIAL II studies to triangulate data and meta-data obtained from social media with a range of validated clinical self-reports of symptoms of distress (ie, depression, stress, and generalized anxiety), fear (ie, panic and social anxiety), substance use (ie, alcohol and other drugs), somatoform problems (ie, insomnia and chronic pain), and potential thought disorder symptoms (ie, symptoms consistent with hypomania). However, a concern about studies triangulating clinical data and social media data remains that individuals who volunteer their social media accounts in such studies are not representative of individuals on social media in general [20]. In this report, we compared the baseline sociodemographic, clinical, and social media variables of individuals who were Twitter users who provided valid Twitter handles to Twitter users who provided handles associated with accounts with high bot scores, Twitter users who provided invalid account names, and, in SOCIAL II, non-Twitter users. In both cohorts, individuals who provided valid Twitter handles tended to use Twitter less than individuals who did not provide handles or who provided invalid handles, although these differences were small, and most individuals reported using Twitter "several times a day." By and large, the differences between the groups were not statistically significant, and when they were statistically significant, they were small in magnitude. This suggests that prior work that focuses on individuals who self-disclose valid Twitter handles is generalizable, at least with regard to the demographic, clinical, and social media features measured here. We observed other demographic differences between the 2 cohorts. For example, in SOCIAL I, cisgender women were more likely to provide their handles, as were

lesbian, gay, bisexual, and queer individuals (vs heterosexual individuals) and those who reported lower (vs higher) incomes. Nonetheless, in all cases in which we did detect differences, there was complete overlap in the distributions of continuous and ordinal variables, and the differences in effect sizes were relatively small in magnitude. Again, these results are encouraging regarding the generalizability of research on people who volunteer their handles to social media users more broadly, and therefore, do not support the critique that relying on a sample of users who are willing to provide their Twitter handles will lead to significant sample bias.

### Limitations and Strengths

Some limitations inherent in our data are worth considering. First, social media use, especially frequent social media use, is not a random and normally distributed variable. Evidence suggests, for example, that a small portion of users are responsible for a large number of tweets. Thus, future analyses of the SOCIAL I and II data sets and related data sets should consider the frequency of social media activity as well as the nature of that activity. In addition, the decision to enter a study focused on social media may in itself introduce a selection bias that we cannot guard against. Although our samples allow us to study mental health and social media across units of analysis (ie, self-report, text data, and meta-data), we lack more objective data including biomarkers or even observer reports of mental health symptoms. Importantly, although we did not conduct semistructured interviews about mental disorder diagnoses, the diagnosis of mental disorders is largely influenced by the severity of symptoms [29]. For many clinical problems such as depression [30], anxiety, and alcohol use [31], scores on disorder screening scales such as the ones we used are excellent predictors of diagnoses in clinical interviews.

### Future Directions

Despite the fact that social media samples are not representative of the entire population, social media users represent 20%-70% of all individuals in the United States [10], thereby providing a sample that constitutes a plurality of the entire population in the United States. In SOCIAL I, we collected a relatively heterogeneous sample of Twitter users. SOCIAL II was a more homogeneous sample, but it had the advantage of containing a subsample of individuals who did not use Twitter or were unwilling to share these data. We collected an assortment of transdiagnostic features of psychopathology representing the most common symptoms of poor mental health. These data will allow us to assess how the spectrum and range of psychopathology manifests itself in natural language and social networks.

With these data sets, we can triangulate self-reported clinical data and data collected from social media. In both samples, mental health symptoms were relatively well represented, making them good, large-scale samples for studying psychopathology. Our current analyses suggest that individuals in these data sets who volunteered to give their Twitter handle, and provided a valid handle, were not different from other individuals in terms of their demographic characteristics, social media use, and mental health. A future direction for this line of work is to use self-reported mental health to replicate findings

in which mental health is inferred through social media activity. Another direction is to extend the data collection to include ecological momentary assessments to triangulate to what extent social media behavior is a valid window into individuals' mental health.

## Conflicts of Interest

None declared.

## References

1. Murray CJ, Vos T, Lozano R, Naghavi M, Flaxman AD, Michaud C, et al. Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* 2012 Dec 15;380(9859):2197-2223. [doi: [10.1016/S0140-6736\(12\)61689-4](https://doi.org/10.1016/S0140-6736(12)61689-4)] [Medline: [23245608](https://pubmed.ncbi.nlm.nih.gov/23245608/)]
2. Kotov R, Krueger RF, Watson D, Achenbach TM, Althoff RR, Bagby RM, et al. The Hierarchical Taxonomy of Psychopathology (HiTOP): a dimensional alternative to traditional nosologies. *J Abnorm Psychol* 2017 May;126(4):454-477. [doi: [10.1037/abn0000258](https://doi.org/10.1037/abn0000258)] [Medline: [28333488](https://pubmed.ncbi.nlm.nih.gov/28333488/)]
3. Insel T, Cuthbert B, Garvey M, Heinssen R, Pine DS, Quinn K, et al. Research domain criteria (RDoC): toward a new classification framework for research on mental disorders. *Am J Psychiatry* 2010 Jul;167(7):748-751. [doi: [10.1176/appi.ajp.2010.09091379](https://doi.org/10.1176/appi.ajp.2010.09091379)] [Medline: [20595427](https://pubmed.ncbi.nlm.nih.gov/20595427/)]
4. Moffitt TE, Harrington H, Caspi A, Kim-Cohen J, Goldberg D, Gregory AM, et al. Depression and generalized anxiety disorder: cumulative and sequential comorbidity in a birth cohort followed prospectively to age 32 years. *Arch Gen Psychiatry* 2007 Jun;64(6):651-660. [doi: [10.1001/archpsyc.64.6.651](https://doi.org/10.1001/archpsyc.64.6.651)] [Medline: [17548747](https://pubmed.ncbi.nlm.nih.gov/17548747/)]
5. Lorenzo-Luaces L. Heterogeneity in the prognosis of major depression: from the common cold to a highly debilitating and recurrent illness. *Epidemiol Psychiatr Sci* 2015 Dec;24(6):466-472 [FREE Full text] [doi: [10.1017/S2045796015000542](https://doi.org/10.1017/S2045796015000542)] [Medline: [26081748](https://pubmed.ncbi.nlm.nih.gov/26081748/)]
6. Rush AJ, Trivedi MH, Wisniewski SR, Nierenberg AA, Stewart JW, Warden D, et al. Acute and longer-term outcomes in depressed outpatients requiring one or several treatment steps: a STAR\*D report. *Am J Psychiatry* 2006 Nov;163(11):1905-1917. [doi: [10.1176/ajp.2006.163.11.1905](https://doi.org/10.1176/ajp.2006.163.11.1905)] [Medline: [17074942](https://pubmed.ncbi.nlm.nih.gov/17074942/)]
7. Kessler RC, Angermeyer M, Anthony JC, DE Graaf R, Demyttenaere K, Gasquet I, et al. Lifetime prevalence and age-of-onset distributions of mental disorders in the World Health Organization's World Mental Health Survey Initiative. *World Psychiatry* 2007 Oct;6(3):168-176 [FREE Full text] [Medline: [18188442](https://pubmed.ncbi.nlm.nih.gov/18188442/)]
8. Casey BJ, Cannonier T, Conley MI, Cohen AO, Barch DM, Heitzeg MM, ABCD Imaging Acquisition Workgroup. The adolescent brain cognitive development (ABCD) study: imaging acquisition across 21 sites. *Dev Cogn Neurosci* 2018 Aug;32:43-54 [FREE Full text] [doi: [10.1016/j.dcn.2018.03.001](https://doi.org/10.1016/j.dcn.2018.03.001)] [Medline: [29567376](https://pubmed.ncbi.nlm.nih.gov/29567376/)]
9. Passell E, Dillon DG, Baker JT, Vogel SC, Scheuer LS, Mirin NL, et al. Digital cognitive assessment: results from the TestMyBrain NIMH Research Domain Criteria (RDoC) field test battery report. *PsyArXiv* 2019 Jan 26:1-123. [doi: [10.31234/osf.io/dcszr](https://doi.org/10.31234/osf.io/dcszr)]
10. Perrin A. Social media usage: 2005-2015. Pew Research Center. 2015 Oct 8. URL: <https://www.pewresearch.org/internet/2015/10/08/social-networking-usage-2005-2015/> [accessed 2022-06-10]
11. Ten Thij M, Bathina K, Rutter LA, Lorenzo-Luaces L, van de Leemput IA, Scheffer M, et al. Depression alters the circadian pattern of online activity. *Sci Rep* 2020 Oct 14;10(1):17272 [FREE Full text] [doi: [10.1038/s41598-020-74314-3](https://doi.org/10.1038/s41598-020-74314-3)] [Medline: [33057099](https://pubmed.ncbi.nlm.nih.gov/33057099/)]
12. Lorenzo-Luaces L, German RE, DeRubeis RJ. It's complicated: the relation between cognitive change procedures, cognitive change, and symptom change in cognitive therapy for depression. *Clin Psychol Rev* 2015 Nov;41:3-15. [doi: [10.1016/j.cpr.2014.12.003](https://doi.org/10.1016/j.cpr.2014.12.003)] [Medline: [25595660](https://pubmed.ncbi.nlm.nih.gov/25595660/)]
13. Lorenzo-Luaces L, Keefe JR, DeRubeis RJ. Cognitive-behavioral therapy: nature and relation to non-cognitive behavioral therapy. *Behav Ther* 2016 Nov;47(6):785-803. [doi: [10.1016/j.beth.2016.02.012](https://doi.org/10.1016/j.beth.2016.02.012)] [Medline: [27993333](https://pubmed.ncbi.nlm.nih.gov/27993333/)]
14. Beck AT, Haigh EA. Advances in cognitive theory and therapy: the generic cognitive model. *Annu Rev Clin Psychol* 2014;10:1-24. [doi: [10.1146/annurev-clinpsy-032813-153734](https://doi.org/10.1146/annurev-clinpsy-032813-153734)] [Medline: [24387236](https://pubmed.ncbi.nlm.nih.gov/24387236/)]
15. Bathina KC, Ten Thij M, Lorenzo-Luaces L, Rutter LA, Bollen J. Individuals with depression express more distorted thinking on social media. *Nat Hum Behav* 2021 Apr;5(4):458-466. [doi: [10.1038/s41562-021-01050-7](https://doi.org/10.1038/s41562-021-01050-7)] [Medline: [33574604](https://pubmed.ncbi.nlm.nih.gov/33574604/)]
16. Al-Mosaiwi M, Johnstone T. In an absolute state: elevated use of absolutist words is a marker specific to anxiety, depression, and suicidal ideation. *Clin Psychol Sci* 2018 Jul;6(4):529-542 [FREE Full text] [doi: [10.1177/2167702617747074](https://doi.org/10.1177/2167702617747074)] [Medline: [30886766](https://pubmed.ncbi.nlm.nih.gov/30886766/)]
17. Edwards T, Holtzman NS. A meta-analysis of correlations between depression and first person singular pronoun use. *J Res Pers* 2017 Jun;68:63-68. [doi: [10.1016/j.jrp.2017.02.005](https://doi.org/10.1016/j.jrp.2017.02.005)]
18. LeMoult J, Kircanski K, Prasad G, Gotlib IH. Negative self-referential processing predicts the recurrence of major depressive episodes. *Clin Psychol Sci* 2017 Jan;5(1):174-181 [FREE Full text] [doi: [10.1177/2167702616654898](https://doi.org/10.1177/2167702616654898)] [Medline: [28286705](https://pubmed.ncbi.nlm.nih.gov/28286705/)]

19. Balsters MJ, Kraemer EJ, Swerts MG, Vingerhoets AJ. Verbal and nonverbal correlates for depression: a review. *Curr Psych Rev* 2012 Jun 01;8(3):227-234. [doi: [10.2174/157340012800792966](https://doi.org/10.2174/157340012800792966)]
20. Dozois DJ. Depressive cognition on Twitter. *Nat Hum Behav* 2021 Apr;5(4):414-415. [doi: [10.1038/s41562-021-01054-3](https://doi.org/10.1038/s41562-021-01054-3)] [Medline: [33574603](https://pubmed.ncbi.nlm.nih.gov/33574603/)]
21. Yang KC, Ferrara E, Menczer F. Botometer 101: social bot practicum for computational social scientists. *J Comput Soc Sci* (forthcoming) 2022 Aug 20:1-18 [FREE Full text] [doi: [10.1007/s42001-022-00177-5](https://doi.org/10.1007/s42001-022-00177-5)] [Medline: [36035522](https://pubmed.ncbi.nlm.nih.gov/36035522/)]
22. Cohen P, Cohen J, Aiken LS, West SG. The problem of units and the circumstance for POMP. *Multivariate Behav Res* 1999 Jul;34(3):315-346. [doi: [10.1207/s15327906mbr3403\\_2](https://doi.org/10.1207/s15327906mbr3403_2)]
23. Zimmerman M. Misuse of the Mood Disorders Questionnaire as a case-finding measure and a critique of the concept of using a screening scale for bipolar disorder in psychiatric practice. *Bipolar Disord* 2012 Mar;14(2):127-134. [doi: [10.1111/j.1399-5618.2012.00994.x](https://doi.org/10.1111/j.1399-5618.2012.00994.x)] [Medline: [22420588](https://pubmed.ncbi.nlm.nih.gov/22420588/)]
24. R Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing. 2021. URL: <https://www.R-project.org/> [accessed 2022-07-21]
25. RStudio Team. RStudio: Integrated development environment for R. RStudio. 2020. URL: <http://www.rstudio.com/> [accessed 2022-07-21]
26. Bollen J, Gonçalves B, van de Leemput I, Ruan G. The happiness paradox: your friends are happier than you. *EPJ Data Sci* 2017 May 18;6(1):4. [doi: [10.1140/epjds/s13688-017-0100-1](https://doi.org/10.1140/epjds/s13688-017-0100-1)]
27. Bollen J, Gonçalves B, Ruan G, Mao H. Happiness is assortative in online social networks. *Artif Life* 2011;17(3):237-251. [doi: [10.1162/artl\\_a\\_00034](https://doi.org/10.1162/artl_a_00034)] [Medline: [21554117](https://pubmed.ncbi.nlm.nih.gov/21554117/)]
28. Eichstaedt JC, Smith RJ, Merchant RM, Ungar LH, Crutchley P, Preotiu-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/)]
29. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders (DSM-5). 5th edition. Washington, DC, USA: American Psychiatric Association; 2013.
30. Zimmerman M, Coryell W. Screening for major depressive disorder in the community: a comparison of measures. *Psychol Assess* 1994 Mar;6(1):71-74. [doi: [10.1037/1040-3590.6.1.71](https://doi.org/10.1037/1040-3590.6.1.71)]
31. Dawson DA, Grant BF, Stinson FS, Zhou Y. Effectiveness of the derived Alcohol Use Disorders Identification Test (AUDIT-C) in screening for alcohol use disorders and risk drinking in the US general population. *Alcohol Clin Exp Res* 2005 May;29(5):844-854. [doi: [10.1097/01.alc.0000164374.32229.a2](https://doi.org/10.1097/01.alc.0000164374.32229.a2)] [Medline: [15897730](https://pubmed.ncbi.nlm.nih.gov/15897730/)]

## Abbreviations

**DSM:** Diagnostic and Statistical Manual of Mental Disorders

**OR:** odds ratio

**PHQ:** Patient Health Questionnaire

**POMP:** percentage of maximum point

**SOCIAL:** Studies of Online Cohorts for Internalizing Symptoms and Language

*Edited by A Mavragani; submitted 09.05.22; peer-reviewed by A Ghosh, Z Xie; comments to author 04.07.22; revised version received 12.08.22; accepted 26.08.22; published 20.10.22.*

*Please cite as:*

Lorenzo-Luaces L, Howard J, Edinger A, Yan HY, Rutter LA, Valdez D, Bollen J  
*Sociodemographics and Transdiagnostic Mental Health Symptoms in SOCIAL (Studies of Online Cohorts for Internalizing Symptoms and Language) I and II: Cross-sectional Survey and Botometer Analysis*  
*JMIR Form Res* 2022;6(10):e39324

URL: <https://formative.jmir.org/2022/10/e39324>

doi: [10.2196/39324](https://doi.org/10.2196/39324)

PMID: [36264616](https://pubmed.ncbi.nlm.nih.gov/36264616/)

©Lorenzo Lorenzo-Luaces, Jacqueline Howard, Andy Edinger, Harry Yaojun Yan, Lauren A Rutter, Danny Valdez, Johan Bollen. Originally published in *JMIR Formative Research* (<https://formative.jmir.org/>), 20.10.2022. 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

# Examining the Twitter Discourse on Dementia During Alzheimer's Awareness Month in Canada: Infodemiology Study

Juanita-Dawne Bacsu<sup>1</sup>, PhD; Allison Cammer<sup>2</sup>, PhD; Soheila Ahmadi<sup>3</sup>, BSc; Mehrnoosh Azizi<sup>3</sup>, MSc; Karl S Grewal<sup>1</sup>, BSc; Shoshana Green<sup>4</sup>, MEd; Rory Gowda-Sookchoff<sup>4</sup>, BComm; Corinne Berger<sup>3</sup>, PhD; Sheida Knight<sup>3</sup>, PhD; Raymond J Spiteri<sup>3</sup>, PhD; Megan E O'Connell<sup>1</sup>, PhD

<sup>1</sup>Department of Psychology, Canadian Centre for Health and Safety in Agriculture, University of Saskatchewan, Saskatoon, SK, Canada

<sup>2</sup>College of Pharmacy and Nutrition, University of Saskatchewan, Saskatoon, SK, Canada

<sup>3</sup>Department of Computer Science, University of Saskatchewan, Saskatoon, SK, Canada

<sup>4</sup>Department of Psychology, University of Saskatchewan, Saskatoon, SK, Canada

**Corresponding Author:**

Juanita-Dawne Bacsu, PhD

Department of Psychology

Canadian Centre for Health and Safety in Agriculture

University of Saskatchewan

Arts 182, 9 Campus Drive

Saskatoon, SK, S7N 5A5

Canada

Phone: 1 306 250 4399

Email: [juanita.bacsu@usask.ca](mailto:juanita.bacsu@usask.ca)

## Abstract

**Background:** Twitter has become a primary platform for public health campaigns, ranging from mental health awareness week to diabetes awareness month. However, there is a paucity of knowledge about how Twitter is being used during health campaigns, especially for Alzheimer's Awareness Month.

**Objective:** The purpose of our study was to examine dementia discourse during Canada's Alzheimer's Awareness Month in January to inform future awareness campaigns.

**Methods:** We collected 1289 relevant tweets using the Twint application in Python from January 1 to January 31, 2022. Thematic analysis was used to analyze the data.

**Results:** Guided by our analysis, 4 primary themes were identified: dementia education and advocacy, fundraising and promotion, experiences of dementia, and opportunities for future actions.

**Conclusions:** Although our study identified many educational, promotional, and fundraising tweets to support dementia awareness, we also found numerous tweets with cursory messaging (ie, simply referencing January as Alzheimer's Awareness Month in Canada). While these tweets promoted general awareness, they also highlight an opportunity for targeted educational content to counter stigmatizing messages and misinformation about dementia. In addition, awareness strategies partnering with diverse stakeholders (such as celebrities, social media influencers, and people living with dementia and their care partners) may play a pivotal role in fostering dementia dialogue and education. Further research is needed to develop, implement, and evaluate dementia awareness strategies on Twitter. Increased knowledge, partnerships, and research are essential to enhancing dementia awareness during Canada's Alzheimer's Awareness Month and beyond.

(*JMIR Form Res* 2022;6(10):e40049) doi:[10.2196/40049](https://doi.org/10.2196/40049)

**KEYWORDS**

Twitter; social media; dementia; Alzheimer disease; awareness; public health campaigns

## Introduction

Dementia is a rapidly growing public health challenge. In Canada, over 402,000 people live with dementia, and this number is projected to increase as the population ages [1]. Dementia is a general term used to describe a group of diseases and conditions affecting the brain, ranging from Alzheimer disease to Parkinson disease [2]. Alzheimer disease is the most common cause of dementia and accounts for approximately 64% of all dementias in Canada [3].

Stigma (eg, negative beliefs, labels, and discriminatory actions) is a significant issue experienced by people living with dementia and their care partners [4-6]. Almost 50% of Canadians would not want others to know if they had dementia [7]. Inadequate knowledge about dementia and lack of awareness contribute to stereotypes and discrimination against people living with dementia [8-10].

Stigma of dementia can lead to feelings of depression, anxiety, social isolation, and a decreased quality of life for people with dementia and their care partners [10-12]. Moreover, stigma of dementia can deter people from accessing educational information, health services, and supports, thus creating a barrier to timely diagnosis of dementia [7]. A timely diagnosis enables people with dementia to access support services, plan for the future, and access interventions that may improve their quality of life [10]. Consequently, dementia awareness campaigns are vital to reducing stigmatic beliefs, supporting cognitive health promotion, and optimizing timely diagnosis.

Globally, there are increasing efforts to increase awareness of dementia through various Alzheimer's Awareness Months [12]. For example, the United States has its Alzheimer's Awareness Month in November, the World's Alzheimer's Awareness month is in September, and Canada has its Alzheimer's Awareness Month in January. In the past, these awareness campaigns have targeted traditional media such as the radio, television, and print media (eg, magazines, newspapers, and newsletters). More recently, social media websites (such as Facebook, Instagram, and Twitter) are being used as primary venues for hosting awareness campaigns.

With a daily average of 500 million tweets, Twitter is becoming a powerful platform for public health campaigns ranging from mental health awareness [13] to breast cancer awareness month [14]. However, no studies have examined Twitter campaigns for Alzheimer's Awareness Month. Accordingly, the purpose of our study was to examine the Twitter discourse on dementia during Alzheimer's Awareness Month in Canada.

## Methods

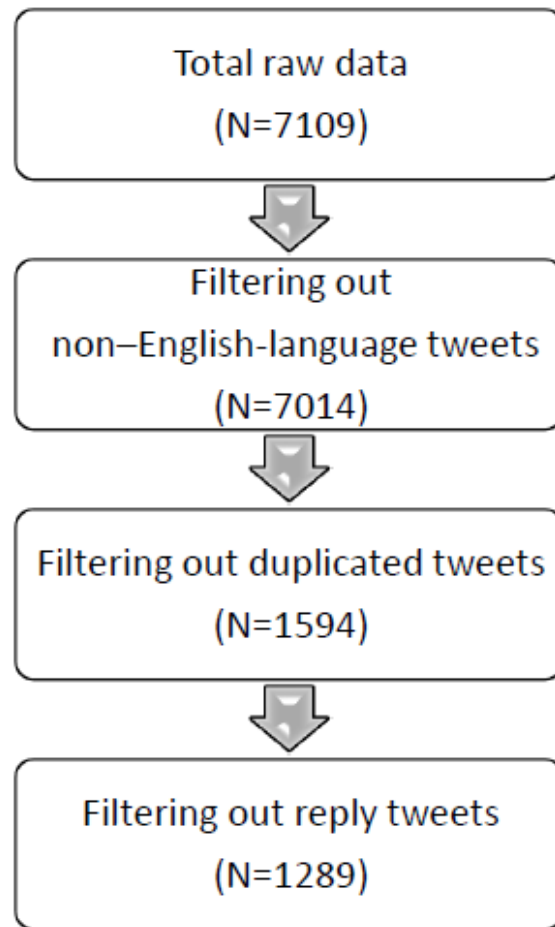
### Ethical Considerations

Compared to traditional research with human participants, tweets posted publicly on Twitter can be used for research without additional consent or ethics approval [13,15,16]. Consequently, institutional ethics approval was not obtained as our data collection focused on tweets shared within the public domain. Nevertheless, any personal identifying information (such as @handles, usernames, and URLs) was removed to ensure anonymity and protect the identity of the Twitter users.

### Recruitment

Tweets were collected using Twint, an advanced scraping tool written in Python. Twint allows users to scrape tweets without the use of Twitter's application programming interface, thus enabling it to avoid certain restrictions such as the number of tweets scraped, the frequency and time period of scrapes, and the requirement of a Twitter account [17]. Our study focused on Canada's Alzheimer's Awareness Month; consequently, we used Twint to asynchronously scrape tweets for the time frame of January 1 to January 30, 2022. Our search terms consisted of various phrases of Alzheimer's Awareness Month (ie, "#AlzheimersAwarenessMonth," "#AlzheimersAwareness," "#AlzAwareness," "#dementiawareness," "dementia month," "dementia awareness month," "dementia awareness," "Alzheimer's awareness month," "Alzheimer's awareness," "Alzheimer's month," and "January is Alzheimer's Awareness month") or tweets using a combination of either "Canada" and "dementia" or "Alzheimer's." We also searched for tweets posted from Canada's national and provincial Alzheimer organizations (ie, @alzCanada, @AlzheimerOnt, @AlzheimerSK, @DementiaAB\_NT @AlzheimerNS, @AlzheimerNB, @AlzheimerPEI, @alzheimersMB, @AlzheimerBC, @asnl2, and @FqsaAlzh). Given that our search focused on Canada's Alzheimer's Awareness Month, we only included tweets from Canada. Our initial search resulted in a total of 7109 tweets. We used filters to exclude non-English-language tweets, reply tweets, duplicate tweets, and spam tweets, resulting in 1289 tweets (see Figure 1). The remaining 1289 tweets were extracted to an Excel (Microsoft Inc) spreadsheet for thematic analysis.

It is important to note that our total number of relevant tweets (n=1289) is comparable to other Twitter awareness studies. For example, in Makita et al's [13] study on the mental health discourse on Twitter during a mental health awareness campaign, the authors analyzed 1200 tweets. In addition, Chung [14] analyzed 1018 tweets to examine the discourse during breast cancer awareness month.

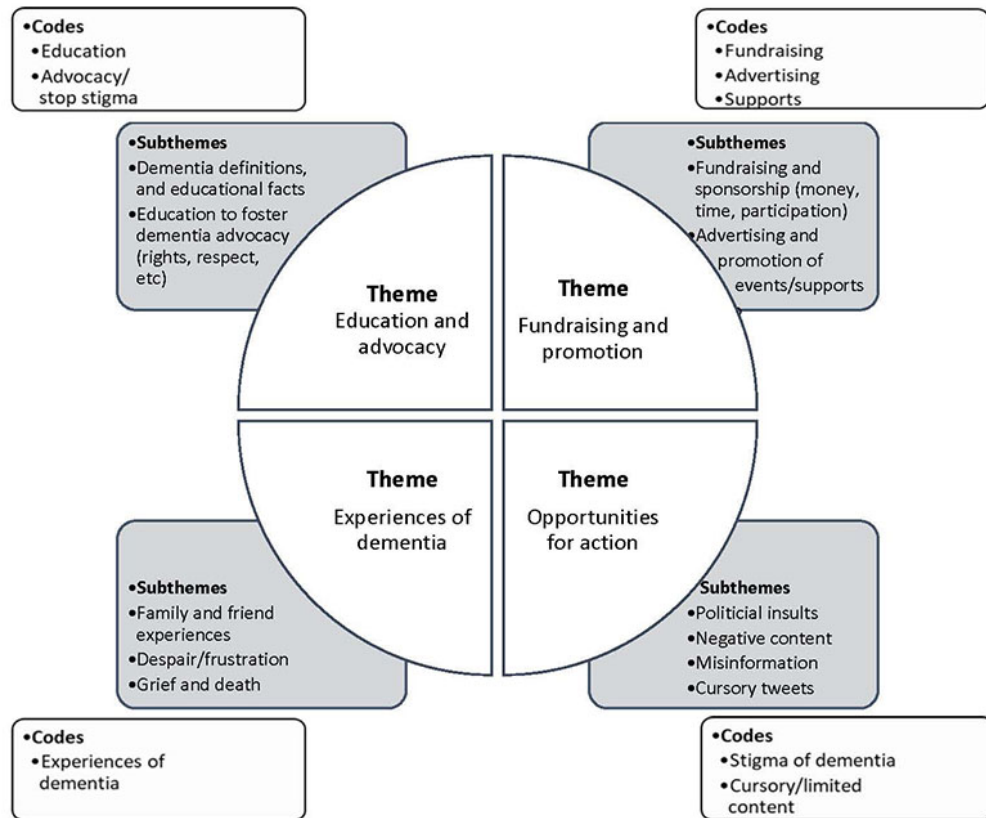
**Figure 1.** Tweet filtering process.

### Data Analysis

Tweets were analyzed using a 5-stage inductive thematic analysis process [18]. First, the researchers (JB, MEO, and RJS) immersed themselves in the by-reading and rereading of 150 tweets to support data familiarization of the Tweets. We used memos to capture any preliminary ideas about the data (eg, similarities, differences, and interesting points) to assist in the development of our initial codes [19]. Second, an open coding approach was used by rereading the initial 150 tweets and developing a code for each tweet, which led to a list of codes. Third, we created a codebook to identify each code name, code definition, and specific examples of relevant tweets for each code. We completed pilot tests with our codebook to help support consistency and intercoder reliability during the coding process. For example, our codebook was pilot tested by having our full team independently code the same set of data (eg, 75 tweets), and then compare coding to discuss coding similarities, differences, or uncertainties. Following our pilot test, tweaks were made to ensure a robust codebook (eg, adding inclusion and exclusion criteria and providing additional examples for each code) to help improve the clarity of the codes. The final codebook consisted of 8 codes, including the following: fundraising (eg, “please donate to Alzheimer’s research”), advocacy or stopping the stigma on dementia (eg, “let’s break

the stigma of dementia...”), cursory or limited content (eg, “January is Alzheimer’s Awareness Month”), advertising (eg, “attend our conference on...”), experiences with dementia (eg, “Mom died of Alzheimer’s...”), advocacy or stopping the stigma on dementia (eg, “...the only truly deranged are the ones having their Alzheimer’s medications filled...”), supports (eg, “each month we offer a caregiver support group...”), and education (eg, “dementia describes a group of symptoms, such as memory loss...”). Each tweet was coded independently by 2 coders to support intercoder reliability. All of our coders were experienced in conducting thematic analysis and have had previous experience in coding tweets [20]. Any coding uncertainties or disputes were resolved through discussion and consensus with 2 coders. However, if consensus could not be reached, the first author acted as a third reviewer to resolve any coding discrepancies. Fourth, once the coding was completed, team meetings were held to discuss relationships and theme development among the codes. More specifically, we used Braun and Clark’s [18] proposed theme generation questions to help support our theme development process. For example, we discussed the coherence of our themes, whether our data supported our themes, and whether we were missing any themes. Fifth, we used a thematic map to document our process for grouping our codes into subthemes that supported the development of our 4 main themes (see Figure 2).

Figure 2. Thematic Map.



### Trustworthiness and Rigor

We used 4 measures to support trustworthiness and rigor in our research. First, multiple coders (eg, each tweet was coded independently by 2 researchers) were used to help ensure a rich data analysis by reducing the potential for individual bias and subjectivity during our coding process [21]. Second, we calculated intercoder reliability scores to assess coding agreement between our different pairs of researchers who were coding the same sets of data. More specifically, our intercoder reliability rate was determined by calculating the percentage of coding agreement among our 6 pairs of coders to establish our overall group average, which was 88.83% (see Table 1 for a breakdown of our intercoder reliability scores). Third, memo writing was integrated into the data analysis process to enable iterative descriptions to support theme development to be discussed and reviewed by the full research team [19]. For example, researchers used analytic memos to document interesting aspects of the data and emerging impressions of

relationships and patterns (eg, similarities and differences) to help inform the development of overarching themes [22]. Illustrative tweets of each theme were also identified collaboratively by the team during the memo writing process. Fourth, interdisciplinary reflexivity (eg, reflecting on each researchers’ identity, assumptions, roles, theoretical underpinnings, and positioning) was used to help inform and avoid any potential areas for bias [23]. Team members were representative of different disciplines and academic backgrounds such as psychology, nutrition, community health and epidemiology, and computer science. Moreover, our coding pairs (ie, 2 coders analyzing each tweet) were purposefully selected to ensure researchers from different disciplines were partnered together to provide a more nuanced analysis. Drawing on this interdisciplinary lens (researchers with different disciplines, training, and expertise) enabled a more comprehensive and in-depth interpretation of the data than could have been achieved by a single discipline [23].

Table 1. Intercoder reliability.

Coder pairs	Intercoder agreement, %
JDB and CB	84.00
KSG and MA	80.00
JDB and SA	86.00
AC and RGS	91.00
MEO and RS	98.00
SG and SK	94.00
Intercoder reliability average	88.83

## Results

Based on our analysis, 4 themes were identified on Twitter during Alzheimer's Awareness Month in Canada: (1) dementia education and advocacy, (2) fundraising and promotion, (3) experiences of dementia, and (4) opportunities for future actions.

### Dementia Education and Advocacy

A predominant theme identified in the tweets was dementia education and advocacy. Several tweets shared facts and statistics about dementia in Canada. For example, tweets often provided definitions and descriptions of dementia. Numerous tweets provided dementia-related statistics, such as the number of Canadians living with dementia. Some tweets discussed statistics of dementia care partners in addition to highlighting brief facts (eg, early diagnosis benefits, risk factors, etc). These educational tweets are illustrated in the examples below.

*Dementia is used to describe a group of symptoms, such as memory loss, confusion, mood changes and communication difficulties. Alzheimers disease is the most common cause of dementia...*

*January is Alzheimer Awareness Month. #DYK that more than 500,000 people are currently living with dementia today in Canada? That number is expected to rise to 912,000 by 2030...*

*January is #AlzheimersAwarenessMonth, an opportunity to learn more about dementia. 1 in 5 Canadians have experienced caring for someone with dementia. Early diagnosis can lead to a better quality of life.*

Many tweets focused on advocacy related to stopping stigma and recognizing the rights of people living with dementia. For example, tweets on stopping stigma described the need to combat or end stigma against people living with dementia. These tweets often had a strong emphasis on the need for more education and dialogue on dementia to address misconceptions and false information. Tweets on stopping the stigma on dementia are captured in the following examples.

*January is #AlzheimersAwarenessMonth. Lets end the stigma surrounding dementia by talking about it, educating ourselves, and supporting those who live with dementia and care partners. Learn from our variety of resources...*

*... We encourage you to learn more about Alzheimer's disease throughout the month and help us smash the stigma surrounding it... get started by learning about these common myths surrounding the disease...*

Several tweets discussed respecting the rights of people with dementia. For example, family care partners often emphasized the need to recognize the rights and dignity of people living with dementia. Advocacy tweets highlighting the need to respect people with dementia are illustrated in the examples below.

*January marks #AlzheimersAwarenessMonth. People living with dementia deserve to have their rights respected and protected, including the right to*

*manage their own lives. Learn more about the rights for people living with dementia...*

*January is #AlzheimersAwarenessMonth I advocate to ensure my beloved Mum receives the care, dignity, and respect she deserves while living with dementia so much more to do!!*

### Fundraising and Promotion

Many tweets identified fundraising and promotional efforts to support dementia awareness during Alzheimer's Awareness Month in Canada. More specifically, fundraising frequently involved requests for money, donation of time, or participation in an event (eg, walk, run, bike, skate, etc), often with a focus of supporting people with dementia and fighting for better outcomes. The following examples illustrate fundraising tweets.

*We need your help to find a cure! Please donate to Alzheimer's today! #Alzheimers #AlzheimersResearch #dementia #alzheimersawareness #caregiver #dementiacare #dementiaawareness #seniorcare #aging #alzheimersdisease #mentalhealth*

*On Saturday, February 5, 2022, starting at 12:00am (midnight), Steve McNeil, a mailman and recreational hockey referee from Etobicoke, ON is coming to Windsor to skate for 19 hours and 26 minutes to raise awareness and funds for Alzheimers disease...*

*Day 12, 1317 press ups completed, 705 to go. Raising funds for Alzheimers Society and Alzheimer's awareness. Donate via the link...*

Promotional efforts did not involve requests for any type of fundraising but focused on drawing the public's attention toward an array of dementia-related efforts. These promotions took 2 forms. The first form of tweets promoted a range of opportunities to connect people to knowledge and education about dementia, including research, conferences, webinars, books, and videos. These promotional tweets focused on educational events are highlighted in the examples below.

*SAVE THE DATE! Our 10th Annual Alzheimers Awareness Conference will take place virtually on Thursday, January 27th from 9am-12:30pm. Stay tuned for further registration announcements...*

*It's #AlzheimersAwarenessMonth! Register for the @ccna\_ccnv's event... to learn more about supporting the well-being of care partners for persons living with #dementia. Sign up here...*

The second type of promotional tweets amplified social supports for people living with dementia and their care partners. Tweets highlighting social supports (eg, support groups and telephone helplines) are demonstrated in the following examples.

*It's Alzheimer's Awareness Month. We know being a caregiver for a loved one can be challenging. Call the 2-1-1 helpline to find supports in your community from home assessments to respite supports, Alzheimer's programs and more. Dial 2-1-1...*

*Are you a caregiver for someone with dementia? We have a support group for you. Join us the first Thursday of each month...*

## Experiences of Dementia

A third theme within the data described experiences related to knowing someone with dementia. Notably absent from our findings were tweets sharing the lived experiences and the direct perspectives of people living with dementia. Rather, tweets tended to focus on the experiences of family members and friends of people with dementia. Within this theme, experiences of dementia were shared in concert with the awareness month campaign to bolster the campaign, but other experience-based tweets appeared disconnected from the campaign. Appreciation for the lived experience of dementia was also heralded as an important component of Alzheimer's Awareness Month.

*Today marks one year without my uncle [sic] & I've been raving about getting a purple butterfly for him for Alzheimers awareness...*

*January is Alzheimer's Awareness Month! Let's understand the experiences of people living with #dementia to impact their lives positively...*

Although some tweets highlighted the awareness month campaign by sharing a personal story, many tweets appeared to be distinct and not connected to the advocacy campaign. Instead, these appeared to represent the regular and constant flow of personal dementia experiences shared via Twitter [24].

*Just over a month ago I moved back home to be with mum as shed received a diagnosis of Alzheimers. Im sharing this with you now as I know a few of you have been in my shoes with your own parents and Ive realised Id appreciate support & advice on this journey.*

*My dad has dementia and cant be alone. Every month my mom goes to dinner w/friends. Comes home and asks why he hasnt gone to bed. Every month I say because he is waiting for you. I cant decide if she thinks shes being cute. Because I already know shes got a selfish streak.*

*My grandpa has dementia and Im the only one in the family who visits him weekly and helps him. I asked my family to visit him at least once a month in our group text. One person said theyd visit this week. I hope he haunts them when he dies.*

Tweets expressing grief and loss were notable among these personal experiences and frustration with inadequate care at the end of life for people with dementia. Although some positive experiences with dementia were shared, many tweets featured negative experiences including despair, frustration, and issues of grief. Among these tweets, there was a predominant focus on death and the end-of-life stage among people with dementia.

*And there's the call from mom. Grandma passed away peacefully this morning after over a decade fighting alzheimer's. It was very expected, she's been in steep decline over the last month. I lost my grandma years and years ago, but it still hurts to lose her again.*

*Mom died of Alzheimers, yesterday. Last month, owing to a fall, she was hospitalized. They were overwhelmed with Covid patients, so there were no*

*beds. She was strapped into a gurney, in a hallway, confused and alone...*

## Opportunities for Action

Although numerous tweets offered educational content, several tweets lacked informative content. For example, these tweets usually made cursory references to January as Alzheimer's Awareness Month without providing any educational information or additional context. Examples of cursory tweets are illustrated below.

*Did you know that January is #Alzheimer Awareness Month in Canada?*

*January is Alzheimers Awareness Month!*

*January is Canadas Alzheimer Awareness Month #Alzheimers #alzheimersawarenessmonth #Canada #dementia #findacure #cure #memory #memoryloss...*

While these tweets promoted the campaign and increased general awareness, they also indicate an opportunity for increasing informative content to counter stigma (eg, negative attitudes, beliefs, and discriminatory behavior) of dementia [25,26]. For example, many tweets used stigmatizing language such as "suffer" and "endure" to describe people living with dementia. Moreover, several tweets made derogatory comments and political ridicule about dementia. Some tweets also shared false information about the COVID-19 vaccine causing dementia. These tweets are documented in the examples below.

*Checking out at the liquor store. Guy at register: how old are you? Me: \*genuinely trying to remember my age\* dahhhh.....twentyyyy-nine? \*Realizing how stupid I am\* I turn 30 next month. The dementia's settling in way earlier than I thought.*

*You know how some politicians are documented as saying one thing then contradicting it, within a month to a year... Shouldn't people check if they have dementia, instead of accusing of lies/being 'evil puppets' ect ect.*

*These people aren't dumb, they know exactly what they're doing. They're counting on you just writing them off as stupid. The only truly deranged are the ones having their Alzheimer's medications filled every month. But thats an entirely different conversation.*

*Alzheimers is a vax side-effect. Perhaps they are slowly introducing their advances in breakthrough treatments to what they have injected us with...*

These negative comments and misinformation only further increase the stigma against people living with dementia. Recent reports indicate that negative attitudes and misinformation can act as a barrier for people with dementia in seeking a timely dementia diagnosis [8,10]. Consequently, these examples highlight opportunities for future Twitter campaigns to provide targeted tweets to counter stigmatizing language and misinformation.

## Discussion

### Principal Findings

Our study aimed to examine the Twitter discourse on dementia during Alzheimer's Awareness Month in Canada. Understanding Twitter content during Alzheimer's Awareness Month is vital to informing targeted education and messaging strategies. A growing body of literature suggests that Twitter may provide an effective means to educate audiences on public health issues, raise awareness, and counter misinformation on stigmatized diseases such as dementia [13,27,28]. Using thematic analysis, we found 4 main themes during Alzheimer's Awareness Month ranging from fundraising to opportunities for future actions.

Our study identified many educational, advocacy, promotional, and fundraising tweets to support people living with dementia. Although several tweets provided informative content, we also identified numerous tweets that contained cursory messaging by simply referencing January as Canada's Alzheimer's Awareness Month. While these tweets promoted the campaign and increased general awareness, they also indicate an opportunity for targeting educational content to counter misinformation and correct stigmatizing language. Many tweets used a negative discourse and stigmatizing words such as "suffer" and "endure" to describe people living with dementia. However, stigmatizing language perpetuates misinformation and myths by implying that people with dementia are not able to live meaningful lives. Moreover, negative discourse against people living with dementia can have important public health implications. How society communicates about dementia influences how policy makers, practitioners, and the general public value the lives of people with dementia. Studies show that stigma can lead to inequitable access to health care services for people living with dementia [29-31]. Consequently, targeted educational campaigns are needed to address stigmatizing language and misinformation about dementia.

Existing literature on mental health and cancer show that educational awareness campaigns can have positive outcomes at both the societal and individual levels [31-33]. For example, educational campaigns can counter public stigma at the societal level and alleviate self-stigma to improve self-confidence at the individual level [31]. However, a paucity of knowledge exists on educational campaign strategies to address the stigma on dementia, especially on social media platforms such as Twitter [34,35]. Future research is necessary to develop and evaluate educational campaign strategies on social media to counter stigma and correct stigmatizing language used against dementia.

Similar to existing studies [36,37], our study on Twitter discourse found that many tweets shared experiences of dementia. More specifically, people shared personal narratives related to issues of grief, bereavement, and death of family members with dementia. Among these tweets, there was a notable focus on death and the end-of-life stage of dementia. Although dementia due to all etiologies is incurable (ie, no disease-modifying intervention currently exists; rather, symptom management and risk reduction remain the primary interventions), the public is often unaware that dementia does not progress in a linear fashion and varies from person to person

[38,39]. Tweets focused on negative aspects of the end-of-life stage may enhance the awareness of support needs during this stage, but they may also exacerbate existing stereotypes against dementia focused on frailty, suffering, and one's inability to live well with dementia [20]. Conversely, tweets sharing real-life stories of people living well with dementia may help to counteract negative stereotypes of people with dementia.

Our study found that the lived experiences and perspectives of people living with dementia were notably absent from the tweets. Research suggests that sharing diverse lived experiences of dementia can help reduce the stigma on dementia by fostering knowledge, understanding, and empowerment [38]. Accordingly, future Alzheimer awareness campaigns on Twitter could be enhanced by highlighting diverse voices and insights of people living with dementia. For example, existing Twitter studies on mental health [13], flu vaccines [40], and breast cancer [14] highlight the beneficial role of engaging social media influencers (eg, celebrities) and champions (eg, advocates, care partners, and people with lived experience) to promote awareness and counter stigmatizing discourse on social media. Accordingly, partnerships with diverse stakeholders (eg, people living with dementia, care partners, celebrities, social media influencers, etc) could play an important role in improving dementia dialogue and awareness on Twitter. Further research is necessary to examine and evaluate awareness campaigns that work in direct partnership with different stakeholders.

### Limitations

Although a rigorous process was undertaken to perform this study, our research has limitations. For example, our research focused only on tweets related to Alzheimer's Awareness Month in Canada. However, examining Alzheimer's Awareness Months in other countries may add value to understanding different social media strategies and awareness techniques. Moreover, it would be of interest to address cultural differences in perceptions of the disease and factors that may positively influence public discourse. Further research is necessary to investigate Twitter data based on Alzheimer's Awareness Months in different countries.

Although we aimed to use inclusive search criteria, it is possible that some relevant tweets related to Alzheimer's Awareness Month may have been missed. More specifically, because our inclusion criteria focused specifically on English-language tweets, it is possible that relevant tweets in other languages were overlooked. Indeed, Canada is a bilingual country, and tweets in the second official language of French were not analyzed in this study. Additional research examining Twitter content in other languages (such as French) may provide a more nuanced and in-depth understanding of the discourse on dementia during Alzheimer's Awareness Month in Canada.

Given the nature of Twitter, our study did not collect any sociodemographic information based on the Twitter users. Without this information, it is difficult to make any specific inferences or assumptions based on sociodemographic characteristics such as age, ethnicity, or sex and gender. Interviews, focus groups, or surveys with sociodemographic questionnaires may provide insight on sociodemographic

characteristics related to effective messaging campaigns tailored to specific audiences during Alzheimer's Awareness Months.

### Conclusions

Our study examined the Twitter discourse on dementia during Alzheimer's Awareness Month in Canada. Understanding Twitter content during public health campaigns is critical to informing targeted messaging and educational strategies. Drawing on thematic analysis, our study identified 4 main themes during Alzheimer's Awareness Month in Canada: dementia education and advocacy, fundraising and promotion, experiences of dementia, and opportunities for future actions.

While several of the tweets provided informative content, we also identified opportunities for action to enhance future awareness month campaigns. For example, we found numerous

tweets that contained cursory messaging by simply referencing January as the Alzheimer's Awareness Month in Canada. While these tweets promoted the campaign and increased general awareness, they also highlighted an opportunity for targeted educational content to correct stigmatizing language and misinformation about dementia. Moreover, awareness strategies partnering with diverse stakeholders (such as people living with dementia, celebrities, social media influencers, advocates, and care partners) may play a pivotal role in fostering dementia education and awareness on Twitter. Further research is needed to develop and evaluate dementia awareness campaigns on social media platforms such as Twitter. Increased knowledge, partnerships, and research are essential to enhancing dementia awareness during Canada's Alzheimer's Awareness Month and beyond.

### Acknowledgments

This work was supported by Team 15 in the Canadian Consortium on Neurodegeneration in Aging, which is supported by a grant from the Canadian Institutes of Health Research with funding from several partners including the Saskatchewan Health Research Foundation, the Centre for Aging and Brain Health, and the Alzheimer Society of Canada.

### Authors' Contributions

J-DB conceptualized the study with MEO and RJS. J-DB, MEO, SA, MA, and RJS devised the Twitter scraping approach, and MA and SA scraped for tweets. J-DB, MEO, and RJS devised the codebook. J-DB, MEO, AC, MA, SA, CB, KSG, SK, SG, RG-S, and RJS coded the tweets. All contributed example tweets and participated in the thematic analysis. J-DB wrote the first draft of the manuscript, with theme writing contributions from AC, KSG, SG, RG-S, SA, and CB. All authors reviewed the final manuscript.

### Conflicts of Interest

None declared.

### References

1. Dementia in Canada, including Alzheimer's disease. Government of Canada. URL: <https://www.canada.ca/en/public-health/services/publications/diseases-conditions/dementia-highlights-canadian-chronic-disease-surveillance.html> [accessed 2022-05-30]
2. Dementia. Mayo Foundation for Medical Education and Research. URL: <https://www.mayoclinic.org/diseases-conditions/dementia/symptoms-causes/syc-20352013> [accessed 2022-05-29]
3. A new way of looking at the impact of dementia in Canada. Alzheimer Society. URL: <https://tinyurl.com/2236deyh> [accessed 2022-05-31]
4. Bacsu J, Mateen FJ, Johnson S, Viger MD, Hackett P. Improving dementia care among family physicians: from stigma to evidence-informed knowledge. *Can Geriatr J* 2020 Dec;23(4):340-343 [FREE Full text] [doi: [10.5770/cgj.23.426](https://doi.org/10.5770/cgj.23.426)] [Medline: [33282053](https://pubmed.ncbi.nlm.nih.gov/33282053/)]
5. Bethell J, Pringle D, Chambers LW, Cohen C, Comisso E, Cowan K, et al. Patient and public involvement in identifying dementia research priorities. *J Am Geriatr Soc* 2018 Aug 07;66(8):1608-1612. [doi: [10.1111/jgs.15453](https://doi.org/10.1111/jgs.15453)] [Medline: [30084194](https://pubmed.ncbi.nlm.nih.gov/30084194/)]
6. Suárez-González A, Rajagopalan J, Livingston G, Alladi S. The effect of COVID-19 isolation measures on the cognition and mental health of people living with dementia: A rapid systematic review of one year of quantitative evidence. *EClinicalMedicine* 2021 Sep;39:101047 [FREE Full text] [doi: [10.1016/j.eclinm.2021.101047](https://doi.org/10.1016/j.eclinm.2021.101047)] [Medline: [34386758](https://pubmed.ncbi.nlm.nih.gov/34386758/)]
7. Alzheimer Society 2017 awareness survey executive summary. Alzheimer Society of Canada. URL: [https://ilivewithdementia.ca/wp-content/uploads/2018/01/2017\\_AWARENESS-SURVEY\\_EXECUTIVE\\_SUMMARY.pdf](https://ilivewithdementia.ca/wp-content/uploads/2018/01/2017_AWARENESS-SURVEY_EXECUTIVE_SUMMARY.pdf) [accessed 2022-05-23]
8. World Alzheimer Report 2019. Alzheimer's Disease International. URL: <https://www.alzint.org/resource/world-alzheimer-report-2019/> [accessed 2022-05-06]
9. A dementia strategy for Canada: together we aspire. Government of Canada. 2019. URL: <https://www.canada.ca/en/public-health/services/publications/diseases-conditions/dementia-strategy.html%20%20> [accessed 2022-05-18]
10. World Alzheimer Report 2012. Overcoming the stigma of dementia. Alzheimer's Disease International. URL: <https://www.alzint.org/u/WorldAlzheimerReport2012.pdf> [accessed 2022-05-19]



11. Lopez RP, Rose KM, Kenney L, Sanborn V, Davis JD. Managing shame: a grounded theory of how stigma manifests in families living with dementia. *J Am Psychiatr Nurses Assoc* 2020;26(2):181-188. [doi: [10.1177/1078390319832965](https://doi.org/10.1177/1078390319832965)] [Medline: [30866693](https://pubmed.ncbi.nlm.nih.gov/30866693/)]
12. Towards a dementia inclusive society. World Health Organization. 2021. URL: <https://www.who.int/publications/i/item/9789240031531> [accessed 2022-05-18]
13. Makita M, Mas-Bleda A, Morris S, Thelwall M. Mental health discourses on Twitter during mental health awareness week. *Issues Ment Health Nurs* 2021 May;42(5):437-450. [doi: [10.1080/01612840.2020.1814914](https://doi.org/10.1080/01612840.2020.1814914)] [Medline: [32926796](https://pubmed.ncbi.nlm.nih.gov/32926796/)]
14. Chung JE. Retweeting in health promotion: analysis of tweets about breast cancer awareness month. *Comput Hum Behav* 2017 Sep;74:112-119. [doi: [10.1016/j.chb.2017.04.025](https://doi.org/10.1016/j.chb.2017.04.025)]
15. Westmaas JL, Masters M, Bandi P, Majmundar A, Asare S, Diver WR. COVID-19 and tweets about quitting cigarette smoking: topic model analysis of Twitter posts 2018-2020. *JMIR Infodemiology* 2022;2(1):e36215 [FREE Full text] [doi: [10.2196/36215](https://doi.org/10.2196/36215)] [Medline: [35611092](https://pubmed.ncbi.nlm.nih.gov/35611092/)]
16. Williams ML, Burnap P, Sloan L. Towards an ethical framework for publishing Twitter data in social research: taking into account users' views, online context and algorithmic estimation. *Sociology* 2017 Dec;51(6):1149-1168 [FREE Full text] [doi: [10.1177/0038038517708140](https://doi.org/10.1177/0038038517708140)] [Medline: [29276313](https://pubmed.ncbi.nlm.nih.gov/29276313/)]
17. Twint Project. GitHub, Inc. URL: <https://github.com/twintproject/twint> [accessed 2022-05-24]
18. Braun V, Clarke V. Thematic analysis. In: Cooper H, Camic PM, Long DL, Panter AT, Rindskopf D, Sher KJ, editors. *APA handbook of research methods in psychology, Vol. 2. Research designs: Quantitative, qualitative, neuropsychological, and biological*. Washington, DC: American Psychological Association; 2012:51-71.
19. Coffey A, Atkinson P. *Making Sense of Qualitative Data: Complementary Research Strategies*. Thousand Oaks, CA: Sage Publications; 1996.
20. Bacsu J, Fraser S, Chasteen AL, Cammer A, Grewal KS, Bechard LE, et al. Using Twitter to examine stigma against people with dementia during COVID-19: infodemiology study. *JMIR Aging* 2022 Mar 31;5(1):e35677 [FREE Full text] [doi: [10.2196/35677](https://doi.org/10.2196/35677)] [Medline: [35290197](https://pubmed.ncbi.nlm.nih.gov/35290197/)]
21. Barry CA, Britten N, Barber N, Bradley C, Stevenson F. Using reflexivity to optimize teamwork in qualitative research. *Qual Health Res* 1999 Jan;9(1):26-44. [doi: [10.1177/104973299129121677](https://doi.org/10.1177/104973299129121677)] [Medline: [10558357](https://pubmed.ncbi.nlm.nih.gov/10558357/)]
22. Nowell LS, Norris JM, White DE, Moules NJ. Thematic analysis. *Int J Qual Methods* 2017 Oct 02;16(1):160940691773384. [doi: [10.1177/1609406917733847](https://doi.org/10.1177/1609406917733847)]
23. Bieler P, Bister MD, Hauer J, Klausner M, Niewöhner J, Schmid C, et al. Distributing reflexivity through co-laborative ethnography. *J Contemp Ethnogr* 2020 Nov 03;50(1):77-98. [doi: [10.1177/0891241620968271](https://doi.org/10.1177/0891241620968271)]
24. Bacsu J, O'Connell ME, Cammer A, Azizi M, Grewal K, Poole L, et al. Using Twitter to understand the COVID-19 experiences of people with dementia: infodemiology study. *J Med Internet Res* 2021 Feb 03;23(2):e26254 [FREE Full text] [doi: [10.2196/26254](https://doi.org/10.2196/26254)] [Medline: [33468449](https://pubmed.ncbi.nlm.nih.gov/33468449/)]
25. Goffman E. *Stigma: notes on the management of spoiled identity*. Englewood Cliffs, NJ: Prentice-Hall; 1963.
26. Thornicroft G, Rose D, Kassam A, Sartorius N. Stigma: ignorance, prejudice or discrimination? *Br J Psychiatry* 2007 Mar 02;190(3):192-193. [doi: [10.1192/bjp.bp.106.025791](https://doi.org/10.1192/bjp.bp.106.025791)] [Medline: [17329736](https://pubmed.ncbi.nlm.nih.gov/17329736/)]
27. Oscar N, Fox PA, Croucher R, Wernick R, Keune J, Hooker K. Machine learning, sentiment analysis, and Tweets: an examination of Alzheimer's disease stigma on Twitter. *J Gerontol B Psychol Sci Soc Sci* 2017 Sep 01;72(5):742-751. [doi: [10.1093/geronb/gbx014](https://doi.org/10.1093/geronb/gbx014)] [Medline: [28329835](https://pubmed.ncbi.nlm.nih.gov/28329835/)]
28. Apriceno M, Lytle A, Monahan C, Macdonald J, Levy SR. Prioritizing health care and employment resources during COVID-19: roles of benevolent and hostile ageism. *Gerontologist* 2021 Jan 21;61(1):98-102 [FREE Full text] [doi: [10.1093/geront/gnaa165](https://doi.org/10.1093/geront/gnaa165)] [Medline: [33119089](https://pubmed.ncbi.nlm.nih.gov/33119089/)]
29. Italy: violations of the human rights of older residents of care homes during COVID-19 pandemic. Amnesty International. URL: <https://www.amnesty.org/en/latest/news/2020/12/italyviolations-of-the-human-rights-of-older-residents-of-care-homes-during-covid-19-pandemic/> [accessed 2021-11-27]
30. Suárez-González A, Livingston G, Low LF, Cahill S, Hennelly N, Dawson WD, et al. Impact and mortality of COVID-19 on people living with dementia: cross country report. *LTC Covid*. 2020. URL: <https://ltccovid.org/wp-content/uploads/2020/08/> [accessed 2022-09-28]
31. Cook JE, Purdie-Vaughns V, Meyer IH, Busch JTA. Intervening within and across levels: a multilevel approach to stigma and public health. *Soc Sci Med* 2014 Feb;103:101-109. [doi: [10.1016/j.socscimed.2013.09.023](https://doi.org/10.1016/j.socscimed.2013.09.023)] [Medline: [24513229](https://pubmed.ncbi.nlm.nih.gov/24513229/)]
32. Griffiths KM, Carron-Arthur B, Parsons A, Reid R. Effectiveness of programs for reducing the stigma associated with mental disorders. A meta-analysis of randomized controlled trials. *World Psychiatry* 2014 Jun;13(2):161-175 [FREE Full text] [doi: [10.1002/wps.20129](https://doi.org/10.1002/wps.20129)] [Medline: [24890069](https://pubmed.ncbi.nlm.nih.gov/24890069/)]
33. Xu S, Markson C, Costello KL, Xing CY, Demissie K, Llanos AA. Leveraging social media to promote public health knowledge: example of cancer awareness via Twitter. *JMIR Public Health Surveill* 2016;2(1):e17 [FREE Full text] [doi: [10.2196/publichealth.5205](https://doi.org/10.2196/publichealth.5205)] [Medline: [27227152](https://pubmed.ncbi.nlm.nih.gov/27227152/)]
34. Bacsu J, Johnson S, O'Connell ME, Viger M, Muhajarine N, Hackett P, et al. Stigma reduction interventions of dementia: a scoping review. *Can J Aging* 2022 Jun;41(2):203-213. [doi: [10.1017/S0714980821000192](https://doi.org/10.1017/S0714980821000192)] [Medline: [34253273](https://pubmed.ncbi.nlm.nih.gov/34253273/)]

35. Herrmann LK, Welter E, Leverenz J, Lerner AJ, Udelson N, Kanetsky C, et al. A systematic review of dementia-related stigma research: can we move the stigma dial? *Am J Geriatr Psychiatry* 2018 Mar;26(3):316-331. [doi: [10.1016/j.jagp.2017.09.006](https://doi.org/10.1016/j.jagp.2017.09.006)] [Medline: [29426607](https://pubmed.ncbi.nlm.nih.gov/29426607/)]
36. Talbot CV, O'Dwyer ST, Clare L, Heaton J. The use of Twitter by people with young-onset dementia: a qualitative analysis of narratives and identity formation in the age of social media. *Dementia (London)* 2021 Oct;20(7):2542-2557 [[FREE Full text](#)] [doi: [10.1177/14713012211002410](https://doi.org/10.1177/14713012211002410)] [Medline: [33765848](https://pubmed.ncbi.nlm.nih.gov/33765848/)]
37. Bartmess M, Talbot C, O'Dwyer ST, Lopez RP, Rose KM, Anderson JG. Using Twitter to understand perspectives and experiences of dementia and caregiving at the beginning of the COVID-19 pandemic. *Dementia (London)* 2022 Jul 13;21(5):1734-1752 [[FREE Full text](#)] [doi: [10.1177/14713012221096982](https://doi.org/10.1177/14713012221096982)] [Medline: [35549466](https://pubmed.ncbi.nlm.nih.gov/35549466/)]
38. Kitwood T. Toward a theory of dementia care: ethics and interaction. *J Clin Ethics* 1998;9(1):23-34. [Medline: [9598430](https://pubmed.ncbi.nlm.nih.gov/9598430/)]
39. Gallagher R. The lived experience of people with dementia. *BCM J* 2019;61(1):39-40 [[FREE Full text](#)]
40. Bonnevie E, Smith SM, Kummeth C, Goldbarg J, Smyser J. Social media influencers can be used to deliver positive information about the flu vaccine: findings from a multi-year study. *Health Educ Res* 2021 Jul 12;36(3):286-294 [[FREE Full text](#)] [doi: [10.1093/her/cyab018](https://doi.org/10.1093/her/cyab018)] [Medline: [34252187](https://pubmed.ncbi.nlm.nih.gov/34252187/)]

*Edited by T Leung; submitted 03.06.22; peer-reviewed by S Erturk, S Mehrabi; comments to author 24.08.22; revised version received 09.09.22; accepted 13.09.22; published 26.10.22.*

*Please cite as:*

Bacsu JD, Cammer A, Ahmadi S, Azizi M, Grewal KS, Green S, Gowda-Sookochoff R, Berger C, Knight S, Spiteri RJ, O'Connell ME  
*Examining the Twitter Discourse on Dementia During Alzheimer's Awareness Month in Canada: Infodemiology Study*  
*JMIR Form Res* 2022;6(10):e40049

URL: <https://formative.jmir.org/2022/10/e40049>

doi: [10.2196/40049](https://doi.org/10.2196/40049)

PMID: [36287605](https://pubmed.ncbi.nlm.nih.gov/36287605/)

©Juanita-Dawne Bacsu, Allison Cammer, Soheila Ahmadi, Mehrnoosh Azizi, Karl S Grewal, Shoshana Green, Rory Gowda-Sookochoff, Corinne Berger, Sheida Knight, Raymond J Spiteri, Megan E O'Connell. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 26.10.2022. 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](mailto:support@jmir.org)

---

<https://www.jmirpublications.com/>