JMIR Mental Health

Journal Impact Factor (JIF) (2023): 4.8
Volume 7 (2020), Issue 11   ISSN 2368-7959   Editor in Chief: John Torous, MD, MBI

Contents

Viewpoint

Original Papers

Industry Perspective

Viewpoint

Direct-to-Consumer Genetic Ancestry Testing in Clinical Encounters: Perspectives From Psychotherapy Cases

Caryn Kseniya Rubanovich[1,2], MS; Wendy Zhang[1], MS; Cinnamon S Bloss[2,3], PhD

[1]San Diego State University/University of California San Diego Joint Doctoral Program in Clinical Psychology, San Diego, CA, United States

[2]Center for Empathy and Technology, T Denny Sanford Institute for Empathy and Compassion, University of California San Diego, La Jolla, CA, United States

[3]Division of Health Policy, Department of Family Medicine and Public Health, University of California San Diego, La Jolla, CA, United States

Corresponding Author:
Cinnamon S Bloss, PhD
Division of Health Policy
Department of Family Medicine and Public Health
University of California San Diego
9500 Gilman Drive
La Jolla, CA, 92093
United States
Phone: 1 (858) 534 9595
Email: cbloss@ucsd.edu

Abstract

Despite the fact that direct-to-consumer (DTC) genetic ancestry testing (GAT) has been available for two decades, there is a lack of evidence-based guidance for clinicians who may work with patients who raise the topic of DTC-GAT. Although DTC-GAT accounts for the majority of the DTC genetic testing marketplace, it has received less attention than health-related testing from scientific and clinical communities. Importantly, however, from our personal experience, patients have been raising the topic of DTC-GAT in clinical encounters, including psychotherapy sessions. In this viewpoint, we present two cases of patients seen by two of the authors to raise awareness of this issue. We describe the implications of DTC-GAT for patients and clinicians, offer recommendations, and suggest future directions.

(JMIR Ment Health 2020;7(11):e23596) doi:10.2196/23596

KEYWORDS

direct-to-consumer; genetic ancestry testing; therapeutic alliance; psychotherapy

Background

At a time when approximately 1 in 13 Americans have utilized direct-to-consumer (DTC) genetic testing [1], we write to raise awareness of the lack of evidence-based clinical guidance related to, and empirical studies within the clinical literature of DTC genetic ancestry testing (GAT). Although 1 in 13 may be an overestimate as it does not likely account for customers who purchase more than one GAT kit from different companies, it is clear that GAT has attracted interest from a significant proportion of the population. GAT accounts for the majority of the DTC genetic testing marketplace but has received less attention from clinical and scientific communities than health-related testing. This is concerning because patients bring up DTC-GAT in clinical encounters, as we demonstrate in two clinical case examples herein.

Two Case Examples

Patient 1 is a Hispanic-American male veteran in his mid-60s who was attending psychotherapy to manage unwanted anger. Sessions consisted of teaching mindfulness and emotion regulation skills. Upon discussion of his values and future goals, the patient identified a desire to have a better understanding of his family of origin and reported interest in undergoing DTC-GAT. The clinician probed the patient’s meaning-making of receiving ancestry results; he reported interest in making sense of “why I behave in certain ways,” and feeling closer to his father. He later sought guidance from the clinician regarding whether to utilize DTC-GAT. The clinician reviewed the available literature on the clinical impacts of DTC-GAT, which was minimal, as well as the risks and benefits reported in popular news media. Ultimately, the veteran and clinician engaged in a collaborative discussion in which the clinician worked to ensure consideration of both risks and potential benefits, while also focusing on the veteran's own personal values and motivations for considering DTC-GAT.

Patient 2 is a White, non-Hispanic male veteran in his mid-50s with a history of severe mental illness who, during his cognitive behavioral therapy, disclosed having recently found his biological mother through DTC-GAT. After years of searching for her and growing up as the only adopted child in a family that "couldn't be more different" than him, he described his mood as "over the moon." After reconnecting, he and his biological mother began having daily phone calls; he reported having always felt lonely, but that he never understood what these feelings were related to until he reconnected with her. He noted considering the implications of DTC-GAT for his identity, autobiographical narrative, and mental health. He chose not to connect with his biological father for these very reasons, as "he may have been the one who gave me up." Without relevant clinical guidance available, the clinician worked on helping the patient see links between thoughts, emotions, and behaviors with respect to this situation.

These case examples reflect clinical interactions showcasing pre- and post-DTC-GAT timepoints. Patient 1 requested clinician guidance on whether to utilize DTC-GAT, whereas Patient 2's DTC-GAT experience was integrated into treatment. Together, these cases call attention to the need for clinician preparedness to engage in conversation with patients regarding DTC-GAT.

Implications for Patients and Clinicians

It is perhaps unsurprising that DTC-GAT would come up in the clinical setting such as within the scope of psychotherapy. Indeed, it is natural for people to want to search for and understand their origins [2,3], and in today's world of technology and connectivity, individuals have turned to available information-gathering resources such as DTC-GAT and companies' online social network platforms to do so [3]. Unfortunately, for clinicians who seek to provide guidance to patients on the topic of DTC-GAT, including risks and benefits, little evidence-based guidance is available. The therapist in the first case example above raised the following topics with the patient, which have been discussed in the scientific literature and popular media: (1) privacy concerns and use of data by for-profit companies or researchers; (2) unexpected findings (eg, misattributed paternity); and (3) lack of consistency in ancestry results due to differences in DTC-GAT company approaches to sequencing and reference pools used for ancestry estimates. The patient's own underlying motivations and attitudes toward testing were explored and were a focus of the conversation. Ultimately, the therapist utilized a motivational interviewing framework, which is commonly used for a variety of health behaviors and decision-making in clinical settings. This resulted in the creation of a decisional balance [4], which involved constructing a list of the risks and benefits, tailored to the patient, which he could use to ultimately make the best decision for himself.

In the second case example, the therapist incorporated the DTC-GAT results the patient had shared into cognitive behavioral therapy, creating space for him to explore his reactions to the DTC-GAT emotionally, cognitively, and behaviorally. The therapist also took this opportunity to educate herself on the mental health implications of DTC-GAT by discussing it openly with the patient and consulting with colleagues.

Ethical Guidance for Clinicians

Applicable ethics codes for psychologists and physicians include the five general principles of the American Psychological Association's Ethical Principles of Psychologists and Code of Conduct ("Ethics Code") [5] or the nine principles of medical ethics from the American Medical Association's Code of Medical Ethics [6]. Both sets of principles are relevant for working with patients who raise the topic of DTC-GAT and have substantial overlap in their messages.

First, clinicians are to make relevant information available to the patient for decision-making. In the first case, the clinician took steps to raise critical topics related to DTC-GAT such as privacy, learning new and possibly unexpected information, and the fallibility of results. These points of conversation are not unlike those a clinician might be expected to raise in the scope of other clinical testing or assessments with patients. Second, clinicians are to recognize the boundaries of their competence, be transparent about any deficits, and seek out additional consultation and resources as needed. Clinicians are to strive for accuracy, honesty, and truthfulness, providing information grounded in scientific research whenever possible. For instance, for many DTC-GAT takers, results may be purely informational, but for certain individuals or subgroups of people, the results may be "genealogically disorienting" [7] (eg, a patient learns of new or unexpected ancestral information) or emotionally distressing [8,9] (eg, new information calls one to question presumed ethnic/racial identities and is negatively internalized).

Next, clinicians have a duty to consider the welfare of other potentially affected individuals, communities, and the public more generally. Ultimately, clinicians have a duty to their patients and should respect a patient's autonomy and decision-making. Clinicians should strive to help patients make the best decision for themselves and their unique situations. If a patient has decided to pursue or has completed DTC-GAT, as in the second case example, clinicians should be prepared to work with the patient to process any new or unexpected information brought to light by the testing results, and their potential impacts on the test-taker and their biological or chosen families, for example. This would also require clinicians to bring to awareness and actively work on any of their own assumptions or biases that may become relevant as the patient discovers new aspects of their ancestry and possibly renegotiates parts of their identity.

Recommendations and Future Directions

First, we recommend surveying clinicians to determine the prevalence of DTC-GAT–related discussions in clinical encounters. The cases described above suggest patient-initiated interest and discussion of DTC-GAT; however, it is unclear at this time how common these occurrences might be. We also recommend empirical study of the short- and long-term psychological and behavioral impacts of DTC-GAT, including health care utilization and health-related decision-making. Little research has been conducted in this area, despite the fact that DTC-GAT and related topics are being raised in health care settings with expectations that clinicians are in a position to help patients navigate them, as highlighted by the two cases described herein. Genetic testing might disproportionately impact some patients such as those who are adoptees, egg or sperm donor–conceived, or who have learned of misattributed paternity. For these subgroups in particular, the stakes are high as testing could reveal and open access to biological family members or provide new genetic-relative family health history [10]. These discoveries could prompt a patient to initiate a new interpersonal relationship, as was the case with the patient in the second case example, who reconnected with his biological mother. It is also not hard to imagine how such results could impact a patient's future medical encounters (eg, testing for *BRCA* mutations or Lynch syndrome) if a patient discovers that their family history warrants this.

Second, we recommend training and education related to DTC-GAT and similar technologies for clinical psychologists, psychiatrists, and other psychotherapists, in addition to physicians, generally. One topic to focus on, for instance, might be test limitations such as the potential for variability in results from different companies [11,12], inaccurate or unreliable results [12,13], and suboptimal company practices [14]. Research has been performed to identify gaps in physician knowledge and preparedness for discussing DTC genetic testing generally [15], and as a result, has spurred movement toward integrating communication of genetic- and genomic-related information and data within medical genetics education and training. The same cannot be said for mental health providers such as clinical psychologists. This represents a critical gap as patients in therapy often see their therapists more frequently than their physicians (eg, on a weekly or biweekly basis); typically have longer sessions (ie, 45-90 minutes) than physicians with room for lengthy discussions; seek out therapy to discuss topics related to identity, family, and health-related matters that often perpetuate or coexist alongside mental health concerns; and are oriented to examine their thoughts, emotions, and behaviors. In sum, the therapy setting is rife with opportunities to explore a patient's beliefs and desires related to DTC-GAT, and the information they might seek to learn from it.

Third, contingent on empirical findings from the first recommendation, we might suggest the American Psychiatric Association consider integrating questions about DTC-GAT into one of the supplementary modules of the Cultural Formulation Interview (CFI) [16]. The CFI is a tool used by clinicians to ensure they are considering the intersections of culture and identity to elicit the patient's own attitudes of cause, context, support, coping, help-seeking, and barriers related to their treatment-seeking and outlook on clinician-patient relationships. The CFI's supplementary modules are optional resources for assessors who might elect to probe additional aspects of identity, perspectives, or beliefs. Specifically, the CFI may benefit from inclusion of an optional question and probe in one of the supplementary modules (eg, Cultural Identity; Coping and Help-seeking) such as: "Have you ever engaged in or thought about doing genetic testing like an ancestry or health testing kit (eg, Ancestry.com, 23andMe) to better understand your identity, family, or health? If so, how has that experience impacted you?" As a result, clinicians may be able to get a better sense of patients' attitudes toward and potential experiences with DTC-GAT and other genetic testing. Importantly, however, given the potentially high-stakes issues, some of which have been described in this viewpoint, clinicians should also use clinical judgment in weighing the possible risks and benefits of introducing the topic of DTC testing to a patient. A disclaimer indicating some of the high-stakes issues surrounding DTC-GAT could be helpful to include for clinicians who may never have considered them, and could help guide clinician-patient conversations.

Conclusion

As DTC-GAT has drastically grown in popularity, has become more affordable, and is increasingly more integrated into mainstream society (ie, kits are now sold at local pharmacies in the United States), test-takers have expanded from the "early adopters" or "worried well" to individuals from all walks of life who sometimes have important and sensitive reasons for engaging in DTC-GAT (ie, finding biological parents). Patients will continue to bring up DTC-related topics (eg, GAT) and clinicians must be prepared to engage in these types of conversations [12]. The review of professional ethics and the three recommendations outlined above are possible first steps in better preparing clinicians who are at the frontlines in offering support to these patients.

Acknowledgments

The authors would like to thank the two US military veterans whose personal experiences served as inspiration for this article. Both patients verbally consented to sharing their stories in hopes of benefiting others. Both have reviewed or discussed the contents of this article with one of the authors (CR or WZ). This work was supported in part by a grant from the National Human Genome Research Institute (R01 HG008753, principal investigator CB), a philanthropic gift to UC San Diego as part of the T. Denny Sanford Institute for Empathy and Compassion, a UC San Diego Chancellor's Interdisciplinary Collaboratory Fellowship, and a San Diego State University Graduate Fellowship.

Conflicts of Interest

None declared.

References

1. Regalado A. More than 26 million people have taken an at-home ancestry test. MIT Technology Review. 2019. URL: https://www.technologyreview.com/s/612880/more-than-26-million-people-have-taken-an-at-home-ancestry-test/ [accessed 2019-11-01]
2. Palacios J, Brodzinsky D. Review: Adoption research: Trends, topics, outcomes. Int J Behav Dev 2010 Apr 19;34(3):270-284. [doi: 10.1177/0165025410362837]
3. Cai J, Kim A, Lee R. Psychological correlates of interest in genetic testing among Korean American adoptees and their parents. J Genet Couns 2020 Jun;29(3):460-470. [doi: 10.1002/jgc4.1237] [Medline: 32125051]
4. Prochaska JO, Velicer WF, Rossi JS, Goldstein MG, Marcus BH, Rakowski W, et al. Stages of change and decisional balance for 12 problem behaviors. Health Psychol 1994;13(1):39-46. [doi: 10.1037/0278-6133.13.1.39]
5. American Psychological Association. Ethical Principles of Psychologists and Code of Conduct. Am Psychol 2002;57(12):1060-1073. [doi: 10.1037/0003-066x.57.12.1060]
6. American Medical Association. AMA Code of Medical Ethics. Council on Ethical and Judicial Affairs. 2001. URL: https://www.ama-assn.org/sites/ama-assn.org/files/corp/media-browser/principles-of-medical-ethics.pdf [accessed 2019-11-01]
7. Nelson A. Bio science: genetic genealogy testing and the pursuit of African ancestry. Soc Stud Sci 2008 Oct;38(5):759-783. [doi: 10.1177/0306312708091929] [Medline: 19227820]
8. Panofsky A, Donovan J. Genetic ancestry testing among white nationalists: From identity repair to citizen science. Soc Stud Sci 2019 Oct 02;49(5):653-681 [FREE Full text] [doi: 10.1177/0306312719861434] [Medline: 31264517]
9. Roth WD, Ivemark B. Genetic Options: The Impact of Genetic Ancestry Testing on Consumers' Racial and Ethnic Identities. Am J Sociol 2018 Jul;124(1):150-184. [doi: 10.1086/697487]
10. May T, Strong KA, Zusevics KL, Jeruzal J, Farrell MH, LaPean Kirschner A, et al. Does Lack of "Genetic-Relative Family Health History" Represent a Potentially Avoidable Health Disparity for Adoptees? Am J Bioeth 2016 Dec 30;16(12):33-38. [doi: 10.1080/15265161.2016.1240255] [Medline: 27901440]
11. Bolnick DA, Fullwiley D, Duster T, Cooper RS, Fujimura JH, Kahn J, et al. Genetics. The science and business of genetic ancestry testing. Science 2007 Oct 19;318(5849):399-400. [doi: 10.1126/science.1150098] [Medline: 17947567]
12. Royal CD, Novembre J, Fullerton SM, Goldstein DB, Long JC, Bamshad MJ, et al. Inferring genetic ancestry: opportunities, challenges, and implications. Am J Hum Genet 2010 May 14;86(5):661-673 [FREE Full text] [doi: 10.1016/j.ajhg.2010.03.011] [Medline: 20466090]
13. Tandy-Connor S, Guiltinan J, Krempely K, LaDuca H, Reineke P, Gutierrez S, et al. False-positive results released by direct-to-consumer genetic tests highlight the importance of clinical confirmation testing for appropriate patient care. Genet Med 2018 Dec 22;20(12):1515-1521. [doi: 10.1038/gim.2018.38] [Medline: 29565420]
14. Barrera J, Foxcroft T. Heredity or hoax?: How dog DNA helped uncover a suspected Indian status scam. CBC News.: CBC/Radio-Canada; 2018. URL: https://newsinteractives.cbc.ca/longform/dna-ancestry-test [accessed 2020-10-25]
15. Rubanovich CK, Cheung C, Mandel J, Bloss CS. Physician preparedness for big genomic data: a review of genomic medicine education initiatives in the United States. Hum Mol Genet 2018 Aug 01;27(R2):R250-R258 [FREE Full text] [doi: 10.1093/hmg/ddy170] [Medline: 29750248]
16. Lewis-Fernández R, Aggarwal N, Hinton L, Hinton D, Kirmayer L. DSM-5® handbook on the cultural formulation interview. Washington, DC: American Psychiatric Association Publishing; 2015.

Abbreviations

CFI: Cultural Formulation Interview
DTC: direct-to-consumer
GAT: genetic ancestry testing

Edited by J Torous; submitted 17.08.20; peer-reviewed by W Roth, A Phillips; comments to author 08.10.20; revised version received 20.10.20; accepted 21.10.20; published 26.11.20.

Please cite as:
Rubanovich CK, Zhang W, Bloss CS
Direct-to-Consumer Genetic Ancestry Testing in Clinical Encounters: Perspectives From Psychotherapy Cases
JMIR Ment Health 2020;7(11):e23596
URL: http://mental.jmir.org/2020/11/e23596/
doi:10.2196/23596
PMID:33242016

Original Paper

Health Personnel's Perceived Usefulness of Internet-Based Interventions for Parents of Children Younger Than 5 Years: Cross-Sectional Web-Based Survey Study

Hege Therese Størksen[1], PhD; Silje Marie Haga[1], PhD; Kari Slinning[1], PhD; Filip Drozd[1], PhD

[1] Department for Infant Mental Health, Regional Centre for Child and Adolescent Mental Health, Eastern and Southern Norway, Oslo, Norway

Corresponding Author:
Hege Therese Størksen, PhD
Department for Infant Mental Health
Regional Centre for Child and Adolescent Mental Health, Eastern and Southern Norway
Gullhaugveien 1-3
Oslo, 0484
Norway
Phone: 47 95972822
Email: hege.storksen@r-bup.no

Abstract

Background: Approximately 10%-15% of children struggle with different socioemotional and psychological difficulties in infancy and early childhood. Thus, health service providers should have access to mental health interventions that can reach more parents than traditional face-to-face interventions. However, despite increasing evidence on the efficacy of internet-based mental health interventions, the pace in transferring such interventions to health care has been slow. One of the major suggested barriers to this may be the health personnel's attitudes to perceived usefulness of internet-based interventions.

Objective: The purpose of this study was to examine health professionals' perceived usefulness of internet-based mental health interventions and to identify the key areas that they consider new internet-based services to be useful.

Methods: Between May and September 2018, 2884 leaders and practitioners of infant and child health services were recruited to a cross-sectional web-based survey through the following channels: (1) existing email addresses from the Regional Centre for Child and Adolescent Mental Health, Eastern and Southern Norway, course database, (2) an official mailing list to infant and child health services, (3) social media, or (4) other recruitment channels. Respondents filled in background information and were asked to rate the usefulness of internet-based interventions for 12 different infant and child mental health problem areas based on the broad categories from the Diagnostic Classification of Mental Health and Developmental Disorders of Infancy and Early Childhood (DC:0-5). Perceived usefulness was assessed with 1 global item: "*How often do you think internet-based self-help programs can be useful for following infant and child mental health problems in your line of work?*" The answers were scored on a 4-point scale ranging from 0 (*never*) to 3 (*often*).

Results: The participants reported that they sometimes or often perceived internet-based interventions as useful for different infant and child mental health problems (scale of 0-3, all means>1.61). Usefulness of internet-based interventions was rated acceptable for sleep problems (mean 2.22), anxiety (mean 2.09), and social withdrawal and shyness (mean 2.07), whereas internet-based interventions were rated as less useful for psychiatric problems such as obsessive behaviors (mean 1.89), developmental disorders (mean 1.91), or trauma (mean 1.61). Further, there were a few but small differences in perceived usefulness between service leaders and practitioners (all effect sizes<0.32, all *P*<.02) and small-to-moderate differences among daycare centers, well-baby clinics, municipal child welfare services, and child and adolescent mental health clinics (all effect sizes<0.69, all *P*<.006).

Conclusions: Internet-based interventions for different infant and child mental health problems within services such as daycare centers, well-baby clinics, municipal child welfare services, and child and adolescent mental health clinics are sometimes or often perceived as useful. These encouraging findings can support the continued exploration of internet-based mental health interventions as a way to improve parental support.

(JMIR Ment Health 2020;7(11):e15149) doi:10.2196/15149

KEYWORDS

internet; parent support; children; mental health; acceptability; health care services

Introduction

Background

Children aged 0-5 years depend on their relationship with their primary caregiver for survival and development. The main task of the parents during these years is to be socially and emotionally available, identify and understand their child's needs, and respond to his or her needs appropriately [1]. For most parents, these parenting skills are intuitive. However, during the many and frequent developmental shifts in the child's early years of life, most parents will find some to be more challenging than others. Children may struggle with different socioemotional and psychological difficulties such as sleep problems [2], excessive crying, or aggressive behavior, which present opportunities for unhealthy developmental pathways in children, which in turn can affect the parent-child relationship, including increased risk of neglect and violence. These are problems that the child may carry with him or her into adolescence and young adulthood. Therefore, primary prevention and early intervention in terms of parental support or training programs, wherein the purpose is to strengthen parenting and parent-child relationship and ensure children's right to care and protection may be highly beneficial for preventing early life difficulties turning into longer lasting problems such as mental illness and their potential consequences [3].

Epidemiological research with children younger than 3 years is limited, but a few studies have indicated prevalence rates of 10%-15% of subclinical/clinical symptoms of mental health disorders [4,5]. These prevalence rates seem to increase from the age of 3 years, as it is estimated that 15%-20% of older children have a reduced functioning due to symptoms most commonly associated with anxiety, affective, and behavioral disorders [6]. Approximately 7%-8% of preschool and school-aged children have symptoms that are compatible with a psychiatric diagnosis at the time of examination [7,8].

Systematic reviews have shown that face-to-face parenting interventions can be effective for children with severe attachment problems [9] and internalizing [10] and externalizing problems [11]. However, to effectively reach many parents and potentially target a broader range of problems, it is necessary to exploit the inherent characteristics of modern technology such as the internet that has a high reach at low cost. Internet-based interventions can also help reduce geographical and social inequalities in health care by providing web-based access for patients to health treatments in remote and rural areas that may lack treatment options and trained health care providers. Internet-based interventions can reach many parents, are in line with their preferences for self-help, have few side effects and high response rates, and can increase the capacity of health care professionals and cover some of the lack of trained personnel [12].

There have been relatively few studies on internet-based parenting programs among families with children aged 0-3 years. Meta-analytic studies suggest that digital cognitive-behavioral programs can be effective for children from the age of 3 years and upwards for targeting depression and childhood anxiety [13,14], disruptive behaviors [15], and somatic conditions [16]. It is therefore reasonable to think that digital interventions may be useful for children younger than 3 years as well. However, there are several barriers to the adoption of new treatments by professionals into their daily practice and few available and accessible e-mental health solutions [17]. Increased availability of e-mental health solutions itself can be positive for health personnel's knowledge and acceptability of digital solutions [18]. However, this is unlikely to be sufficient for uptake and sustained use. In addition to the availability of new digital services, one of the major barriers to their uptake is clinicians' attitudes [17,19].

Attitudes and acceptability toward information technology, conceptualized as perceived usefulness, has, over the years, been one of the major determinants of usage intentions [20,21]. Researchers have proposed and tested several models to explain and predict user acceptance and use of information technology. In 2003, Venkatesh et al [21] integrated elements from these acceptance models to create the Unified Theory of Acceptance and Use of Technology (UTAUT), which identifies 4 key factors (ie, performance expectancy, effort expectancy, social influence, and facilitating conditions) and 4 moderators (ie, age, gender, experience, and voluntariness) related to describe the processes underlying the development and change of attitudes toward digitally delivered interventions primarily in organizational contexts. There are several similar constructs pertaining to user acceptance identified in other models; hence, the term "performance expectancy" is used interchangeably with "perceived usefulness." These similar concepts confirm from different angles that perceived usefulness plays an important part in forming users' attitudes or behavioral intentions.

Despite limited research on child health practitioners (and parental) user acceptance and opinions of e-mental health services, there are still a few studies available; one of these studies showed that obstetricians may be skeptical toward the use of eHealth solutions [22]. However, child and youth mental health workers seem cautiously positive [23,24], particularly in the prevention and treatment of mild-to-moderate mental health problems [24,25]. Furthermore, previous research has shown that leadership and different leadership styles may also have an impact on the adoption and perceived usefulness of technology [26,27]. Thus, it is necessary to consider perceived usefulness from the perspective of different professions/health care services for young children and professional roles (eg, practitioners vs leaders), which, to the best of our knowledge, has not been previously studied directly. In summary, it is therefore important to examine health professionals' perceived usefulness of e-mental health programs and identify the key areas in which they consider new e-mental health services to be useful.

Aims of This Study

The aim of this study was to examine how often health personnel in prenatal, infant, and child health care services for children younger than 5 years would find internet-based parent support interventions useful for different target groups and problem areas. More specifically, we wanted to examine if there are differences in the perception of the usefulness of e-mental health intervention between leaders, middle managers, and practitioners and among daycare centers, well-baby clinics, municipal child welfare services (CWS), and child and adolescent mental health clinics (CAMHS).

Methods

Study Design and Participants

This study was conducted as a cross-sectional web-based survey and approved by the Norwegian Centre for Research Data [28]. We aimed to assess the views of all types of infant and child health practitioners. Participants who were 16 years or older were recruited either through any of the following channels: (1) existing email addresses from the Regional Centre for Child and Adolescent Mental Health (RBUP), Eastern and Southern Norway, course database, (2) an official mailing list to infant and child health services, (3) social media or (4) other recruitment channels (ie, national labor unions). Administrative staff was excluded from this study.

We extracted 5050 unique email addresses from the course database at the time of recruitment based on participation in supplementary education and courses for leaders and clinical staff working with parents and children younger than 5 years. Owing to ethical considerations, participants who registered in the course database before January 2013 and with a private email address were not invited to participate in the study (1448/5050, 28.7%). This was in line with Norwegian ethical codes. Thus, 3602 (71.3%) unique and eligible participants received a study invitation.

The official mailing list [29] contained 7345 unique official email addresses to relevant infant and child mental health services. These included the leaders and clinical staff in the following primary care services: (1) municipal services (n=590), (2) state CWS (n=22), (3) daycare centers (n=5627), (4) family counselling services (n=59), (5) district medical officers (n=190), (6) leading public health nurses (n=247), (7) municipality psychologists (n=108), and (8) educational and psychological counselling services (n=89). The following secondary care services were also included: leaders and clinical staff in CAMHS (n=178) and maternity wards in hospitals (n=20). Finally, 215 email addresses were to other services (eg, private practices and low-threshold services).

Participants were also recruited by promoting the study in 14 relevant Facebook groups with a total of 16,555 group members. These were, however, not unique members (ie, a participant may be a member of more than one group). In addition, 3 national labor unions, that is, the Norwegian Midwife Association, Norwegian Association for Clinical Pedagogy, and the Norwegian Psychologist Association were contacted by mail or telephone and encouraged to send study information and consent to its member list per email or promote the study on their website.

Data Collection and Measures

The web-based questionnaire was developed by 3 experienced researchers following literature studies and an interdisciplinary discussion with other staff members working at RBUP. The questions were self-developed and subsequently pilot tested independently by 6 health care professionals sampled from our intended study population prior to data collection to ensure face validity. Their feedback did not lead to any changes to the included questions, but the length of the questionnaire was on the verge of being unacceptable. We, therefore, tailored the survey such that participants did not have to complete questions about infant and mental health problems they rarely worked with, as this was considered meaningless and annoying. Consequently, we could only afford to include brief measures and we tailored the survey according to 12 different mental health conditions.

Data were collected between May and September 2018 using the survey platform Confirmit. All infant and child health practitioners working with children younger than 5 years who were able to complete the survey in Norwegian were eligible for the study. Practitioners were provided with written information explaining the purpose of the study and that participation was voluntary. Informed web-based consent was obtained from all participants. After providing informed consent, participants completed the web-based questionnaire, which took, on average, 20 minutes. If no web-based consent was provided, the potential participants could not access the web-based questions. Participants who did not respond to the study invitation after consenting to participate received up to 2 reminders (ie, 1 per week).

First, participants were requested to fill in background information (ie, age, sex, and education). Then, based on the broad categories from the Diagnostic Classification of Mental Health and Developmental Disorders of Infancy and Early Childhood (DC:0-5; [30]), respondents were asked to indicate which of the following 12 different infant and child mental health problems they worked with: (1) parent-child relationship and attachment problems, (2) developmental delays, (3) dysregulation, (4) behavioral problems, (5) social withdrawal and shyness, (6) sleep problems, (7) developmental disorders, (8) breastfeeding and eating problems, (9) anxiety, (10) trauma (developmental and physical), (11) depression, and (12) obsessive behaviors and disorders. Sensory processing disorder, which is a category in DC:0-5, was not included in our list because the disorder is little known and widespread in services and requires specialized services. In addition, we pulled out behavioral difficulties and social withdrawal/shyness: behavioral difficulties because it is subsumed under “Mood disorders” in DC:0-5 and social withdrawal/shyness because it is considered a gateway disorder to other disorders and something many in the Norwegian health care system are trained to look for.

Participants who indicated that they “never” or “rarely” worked with one or several of these mental health problems or that a mental health problem was “not relevant” for their work did not receive any further questions about the perceived usefulness of internet-based parenting interventions for that particular problem. Thereafter, respondents were presented with a brief written introduction, defining and explaining the practical use of self-directed internet-based interventions (see Textbox 1).

Textbox 1. Brief textual information about internet-based interventions adapted and translated from Norwegian.

Experience and knowledge of internet-based self-help programs

- Internet-based interventions are often developed by researchers and clinicians and usually consist of 6-12 weekly consultations. In web-based guidance, parents learn about the challenges they face, do exercises, and receive weekly homework. Interactive content is used and standard internet technology is used.
- The content is based on recognized theories and methods in psychology that are often used by professionals in their work with pregnant women, parents, and children. One of the most common approaches is cognitive behavioral therapy, but psychodynamic and other approaches are also used.
- Internet-based self-help programs do not require any prior knowledge or expertise. You will be trained for 2-4 days and receive guidance from qualified personnel with expertise within the relevant problem area. The guidance takes place over time and as needed.
- Your task as a professional is to support the parents in carrying out the web-based program and solving challenges along the way (eg, low motivation and adaptation of tasks and exercises to the family's own situation). Everything takes place either on the internet or in combination with consultations and is included as part of the ordinary service offering. With web-based guidance or treatment, you will use an average of 10-20 minutes per family per consultation.
- Please answer the following questions based on the information you have now received about internet-based self-help programs.

Perceived usefulness of internet-based parenting interventions was assessed with 1 global item: "*How often do you think internet-based self-help programs can be useful for following infant and child mental health problems in your line of work?*" derived from the global usefulness items in the usefulness scale in the Technology Acceptance Model (ie, "*I would find WriteOne useful in the MBA program*" [20]) and Performance Expectancy scale in the UTAUT (ie, "*I find mobile internet useful in my daily life*" [21]). The answers were scored on a 4-point scale and coded as (0) "*never,*" (1) "*rarely,*" (2) "*sometimes,*" and (3) "*often.*" There was also a "*Not sure/I don't know*" category. We defined a score ≥2 on any of the mental health conditions as being perceived as "*useful*" by practitioners and leaders herein. The job level was assessed with the question: "*Check the description that best fits your current position.*" The answers were coded as (1) "*top manager (eg, service leader),*" (2) "*middle manager (eg, team leader or project leader),*" (3) "*practitioner (eg, caseworker or therapist),*" and (4) "*other*" (ie, self-employed).

Information on service was assessed from the question: "*Where do you work?*" Responses were coded into 5 health care services: (1) "*daycare centers,*" (2) "*well-baby clinics,*" (3) "*municipal CWS,*" (4) "*CAMHS,*" and (5) "*other services*" (ie, private practice or neonatal intensive care unit).

Statistical Methods

Descriptive analyses were applied to summarize participant characteristics, missing data, and scores of perceived usefulness for different infant and child mental health problems, including means and standard deviations for continuous variables and frequency counts and percentages for categorical variables. Participants who did not complete the survey, that is, those who had one or more missing data were counted as missing and were compared against those who provided complete data. Participants with missing data were analyzed using chi-square and independent sample two-sided *t* tests for categorical and continuous data, respectively.

Participants in the "*Other*" category were excluded from analyses at job level. The number of participants identifying neither as a top-level, mid-level manager, or practitioner was too low for any comparisons and not of main interest for our research purposes (13/2884, 0.5%). Participants categorized as working in "*other services*" were not included in comparisons of services on perceived usefulness (481/2884, 16.7%). First, other services consisted of a wide range of primary and secondary care services, thereby making any meaningful comparisons practically impossible. Second, we were primarily interested in differences between daycare centers, well-baby clinics, municipal CWS, and CAMHS, as stated in the aims of this study. A one-way between-groups analysis of variance was conducted to explore the impact of leaders and practitioners, and daycare centers, well-baby clinics, municipal CWS, and CAMHS on perceived usefulness of internet-based parenting interventions for different infant and child mental health problems. The data set was assessed for skewness and kurtosis, and a histogram was plotted for outcome variables to check whether they had a normal distribution. All variables were within acceptable range. Posthoc comparisons using Bonferroni correction were used to investigate differences between job level and services. Effect sizes were calculated for systematic differences and expressed as Cohen *d*, which were interpreted as small (0.2), medium (0.5), and large (0.8). The statistical package SPSS version 23 (SPSS Inc) was used for all statistical analyses.

Results

Participant Characteristics

A total of 2884 infant and child health leaders and practitioners provided their consent to participate in this study and responded to the web-based questionnaire. Norway has 18 counties as of 2018. We received responses from each county, with the fewest responses from Finnmark (36/2884, 1.3%) and the most responses from Oslo (379/2884, 13.2%). This proportion reflects the population in these counties according to Statistics Norway (2019; ie, 1.4% and 12.7% of the population lived in Finnmark and Oslo during the study period, respectively) [31]. The characteristics of the study population are presented in Table 1. Participants were primarily middle-aged women who had attended college or university for 1-3 years. In addition, most participants reported working in daycare centers, while the least number of participants reported working in CAMHS. Most respondents were practitioners; however, 29.3% (844/2884) were top-level managers.

Table 1. Characteristics of the infant and child health leader and practitioner population in this study (N=2884).

Characteristic of the health care professionals	Values	Missing data[a], values
Sex, n (%)		2 (0.1)
Male	187 (6.5)	N/A[b]
Female	2695 (93.4)	N/A
Age (years), mean (SD)	46.1 (10.0)	5 (0.2)
College/University education (years), n (%)		1 (0.0)
≤1-3 years	1969 (68.3)	N/A
≥4-5 years	914 (31.7)	N/A
Services, n (%)		57 (2.0)
Daycare center	1215 (42.1)	N/A
Well-baby clinic	701 (24.3)	N/A
Municipal child welfare service	321 (11.1)	N/A
Child and adolescent mental health clinic	109 (3.8)	N/A
Other services	481 (16.7)	N/A
Job level, n (%)		94 (3.3)
Top manager (eg, service leader)	844 (29.3)	N/A
Middle manager (eg, team leader or project leader)	653 (22.6)	N/A
Practitioner (eg, caseworker or therapist)	1280 (44.4)	N/A
Other	13 (0.5)	N/A

[a] Participants who did not complete the survey, that is, those who had one or more missing data were counted as missing.

[b] N/A: Not applicable.

Missing Data

Missing data for participants' background characteristics are reported in Table 1 above. All respondents were required to fill in demographical data; however, not all respondents received questions about the perceived usefulness of internet-based interventions for all child problems, as explained above. Responses to usefulness were thus conditional in "sometimes" or "often" working with the respective infant or child mental health problems. This also means that each respondent received a varying number of mental health problems for the assessment of the usefulness of internet-based parenting interventions, which partly explains the lesser number of participants used for the analyses in the subsequent tables shown below. Importantly, these were not defined as missing. Only respondents who did not complete all questions were defined as study dropouts. The analysis showed an association between missingness and job level (χ_3^2=13.27, *P*=.004). Fewer middle managers (169/653, 25.9%) dropped out from the study than practitioners (387/1280, 30.2%) and top managers (271/844, 32.1%). There were no systematic differences in missingness, neither for sex, age, education, or health service (all *P*>.08).

Overall Perceived Usefulness of Internet-Based Interventions

Participants were asked to indicate how often internet-based parent support interventions could be useful for 12 different infant and child mental health problems. Our results showed that the majority of practitioners and leaders reported that they "*sometimes*" or "*often*" perceived internet-based parenting interventions for the different mental health problems as useful (on a scale from 0 to 3, >76% scored 2=*sometimes* or 3=*often* for 8 out of the 12 conditions, all means>1.61, Table 2). Usefulness of internet-based parenting interventions was rated acceptable for problem areas such as sleep problems, anxiety, and social withdrawal and shyness. More caution was reported toward usefulness for psychiatric problems such as trauma. However, more than half of the participants reported that internet-based interventions for trauma "*sometimes*" or "*often*" could be useful.

Table 2. Overall ranking of the perceived usefulness of internet-based programs for infant and child mental health problems.

Rank	Infant and child mental health problem	Participants (n)	Mean (SD)[a]	PU[b] ≥2, n (%)	Don't know/not sure (N=2884), n (%)
1	Sleep problems	1115	2.22 (0.67)	992 (89.0)	80 (2.8)
2	Anxiety	1104	2.09 (0.73)	915 (82.9)	92 (3.2)
3	Social withdrawal and shyness	1182	2.07 (0.69)	971 (82.1)	114 (4.0)
4	Dysregulation	1113	2.03 (0.71)	909 (81.7)	79 (2.7)
5	Behavioral problems	1535	2.09 (0.72)	1250 (81.4)	129 (4.5)
6	Breastfeeding and eating problems	938	2.06 (0.73)	761 (81.1)	74 (2.6)
7	Parent-child relationship and attachment problems	1399	2.02 (0.78)	1088 (77.8)	93 (3.2)
8	Developmental delays	1378	1.99 (0.76)	1047 (76.0)	137 (4.7)
9	Depression	847	1.91 (0.74)	621 (73.3)	75 (2.6)
10	Obsessive behaviors and disorders	507	1.89 (0.75)	370 (73.0)	49 (1.7)
11	Developmental disorders	1085	1.91 (0.77)	785 (72.4)	114 (4.0)
12	Trauma	772	1.61 (0.84)	434 (56.2)	85 (2.9)

[a] Scale from 0 to 3 where 0=never, 1=rarely, 2=sometimes, and 3=often.

[b] PU: perceived usefulness.

Job Level and Usefulness of Internet-Based Parenting Interventions

A one-way between-group analysis of variance was performed to investigate the impact of job level on the perceived usefulness of internet-based parenting interventions for different infant and child mental health problems (Table 3). Participants were divided into 3 groups according to their job level (ie, practitioner, middle manager, and top manager). Overall tests suggested significant differences at the $P<.05$ level for 7 out of 12 mental health problems: (1) behavioral problems, (2) parent-child relationship and attachment problems, (3) social withdrawal and shyness, (4) trauma, (5) dysregulation, (6) developmental delays, and (7) developmental disorders. Despite reaching statistical significance, actual mean differences between the groups were small. Posthoc comparisons using the Bonferroni adjustment indicated an overall tendency that most differences were between practitioners and top-level managers (Multimedia Appendix 1). Top-level managers found internet-based parenting interventions more useful than practitioners in 6 areas: (1) behavioral problems, (2) parent-child relationship and attachment problems, (3) social withdrawal and shyness, (4) trauma, (5) developmental delays, and (6) developmental disorders. In addition, posthoc comparisons indicated that there were differences in the mean scores between practitioners and middle managers for developmental delays in that middle managers expressed more positive attitudes to internet-based interventions than practitioners. All effect sizes, calculated using Cohen *d*, were small (all $d<0.32$, all $P<.02$).

Table 3. Comparisons of the perceived usefulness of internet-based interventions for 12 different infant and child mental health problems by practitioners at different job levels (employment status) by using one-way analysis of variance.

Mental health problems, job level of health care practitioner	Participants (n)	Mean (SD)	*F (df)*	*P* value
Breastfeeding and eating problems			0.44 (2)	.64
Practitioner	490	2.04 (0.72)		
Middle manager	231	2.09 (0.71)		
Top manager	215	2.08 (0.76)		
Anxiety			0.61 (2)	.55
Practitioner	526	2.07 (0.73)		
Middle manager	256	2.11 (0.75)		
Top manager	319	2.12 (0.70)		
Behavioral problems			5.34 (2)	.005[a]
Practitioner	675	2.03 (0.75)		
Middle manager	386	2.09 (0.73)		
Top manager	472	2.17 (0.66)		
Depression			2.88 (2)	.06
Practitioner	402	1.86 (0.77)		
Middle manager	203	1.92 (0.71)		
Top manager	242	2.00 (0.73)		
Parent-child relationship and attachment problems			8.01 (2)	<.001[b]
Practitioner	679	1.94 (0.79)		
Middle manager	329	2.02 (0.79)		
Top manager	388	2.14 (0.74)		
Social withdrawal and shyness			4.60 (2)	<.001[b]
Practitioner	526	2.00 (0.69)		
Middle manager	304	2.10 (0.72)		
Top manager	350	2.14 (0.64)		
Sleep problems			0.36 (2)	.70
Practitioner	550	2.22 (0.67)		
Middle manager	268	2.25 (0.68)		
Top manager	294	2.20 (0.66)		
Trauma			4.18 (2)	.02[c]
Practitioner	434	1.54 (0.83)		
Middle manager	164	1.65 (0.88)		
Top manager	174	1.75 (0.82)		
Obsessive behaviors and disorders			1.26 (2)	.29
Practitioner	287	1.84 (0.75)		
Middle manager	100	1.95 (0.67)		
Top manager	120	1.94 (0.78)		
Dysregulation			3.43 (2)	.03[c]
Practitioner	562	1.97 (0.73)		
Middle manager	252	2.08 (0.69)		
Top manager	296	2.08 (0.67)		

Mental health problems, job level of health care practitioner	Participants (n)	Mean (SD)	*F (df)*	*P* value
Developmental delays			14.79 (2)	<.001[b]
Practitioner	597	1.87 (0.77)		
Middle manager	355	2.06 (0.76)		
Top manager	425	2.11 (0.74)		
Developmental disorders			6.56 (2)	<.001[b]
Practitioner	495	1.81 (0.78)		
Middle manager	263	1.97 (0.78)		
Top manager	324	2.00 (0.74)		

[a] This value was significant at *P*<.01.

[b] This value was significant at *P*<.001.

[c] This value was significant at *P*<.05.

Services and Usefulness of Internet-Based Parenting Interventions

A one-way between-group analysis of variance was performed to investigate the impact of services on the perceived usefulness of internet-based parenting interventions for different infant and child mental health problems. Overall tests between health care services suggested systematic differences at the *P*<.05 level for all mental health problems, except breastfeeding and eating problems (Table 4). However, as with the job level, actual differences in the mean scores between the services were small.

Posthoc tests with Bonferroni corrections showed a few small significant differences between health services (Multimedia Appendix 2). An overall tendency was that most differences were between daycare centers and the 3 remaining services. Daycare centers considered internet-based parenting interventions as more useful for behavioral problems, parent-child relationship and attachment problems, social withdrawal and shyness, trauma, dysregulation, developmental delays, and developmental disorders than the other services. All effect sizes, calculated using Cohen *d*, were small to moderate (all *d*<0.69, all *P*<.006).

Table 4. Comparisons of the perceived usefulness of internet-based interventions by different services for 12 different infant and child mental health problems by using one-way analysis of variance.

Mental health problems, service	Participants (n)	Mean (SD)	*F (df)*	*P* value
Breastfeeding and eating problems			1.46 (3)	.23
CAMHS[a]	53	2.15 (0.57)		
Daycare centers	305	2.12 (0.73)		
Well-baby clinics	369	2.02 (0.71)		
Municipal CWS[b]	55	2.11 (0.69)		
Anxiety			3.08 (3)	.03[c]
CAMHS	54	2.24 (0.55)		
Daycare centers	465	2.15 (0.73)		
Well-baby clinics	250	2.12 (0.72)		
Municipal CWS	148	1.97 (0.71)		
Behavioral problems			5.33 (3)	.001[d]
CAMHS	59	2.03 (0.59)		
Daycare centers	720	2.18 (0.72)		
Well-baby clinics	344	2.09 (0.69)		
Municipal CWS	193	1.96 (0.76)		
Depression			3.95 (3)	.008[d]
CAMHS	53	1.92 (0.58)		
Daycare centers	330	2.04 (0.75)		
Well-baby clinics	155	1.86 (0.75)		
Municipal CWS	147	1.82 (0.73)		
Parent-child relationship and attachment problems			12.91 (3)	<.001[e]
CAMHS	65	1.68 (0.69)		
Daycare centers	541	2.18 (0.76)		
Well-baby clinics	349	2.00 (0.77)		
Municipal CWS	197	1.92 (0.77)		
Social withdrawal and shyness			7.06 (3)	<.001[e]
CAMHS	53	2.09 (0.53)		
Daycare centers	561	2.16 (0.68)		
Well-baby clinics	271	2.06 (0.68)		
Municipal CWS	125	1.86 (0.66)		
Sleep problems			4.07 (3)	.007[d]
CAMHS	56	2.27 (0.59)		
Daycare centers	426	2.21 (0.68)		
Well-baby clinics	369	2.31 (0.63)		
Municipal CWS	105	2.08 (0.66)		
Trauma			5.51 (3)	.001[d]
CAMHS	62	1.60 (0.78)		
Daycare centers	194	1.81 (0.87)		
Well-baby clinics	150	1.47 (0.86)		
Municipal CWS	190	1.55 (0.76)		

Mental health problems, service	Participants (n)	Mean (SD)	*F (df)*	*P* value
Obsessive behaviors and disorders			2.65 (3)	.049[c]
CAMHS	45	2.04 (0.64)		
Daycare centers	149	1.96 (0.77)		
Well-baby clinics	97	1.77 (0.68)		
Municipal CWS	102	1.78 (0.74)		
Dysregulation			8.15 (3)	<.001[e]
CAMHS	52	1.88 (0.76)		
Daycare centers	418	2.11 (0.68)		
Well-baby clinics	321	2.11 (0.64)		
Municipal CWS	133	1.82 (0.69)		
Developmental delays			17.95 (3)	<.001[e]
CAMHS	57	1.88 (0.68)		
Daycare centers	652	2.16 (0.76)		
Well-baby clinics	316	1.84 (0.74)		
Municipal CWS	167	1.84 (0.67)		
Developmental disorders			10.79 (3)	<.001[e]
CAMHS	61	2.07 (0.70)		
Daycare centers	466	2.04 (0.78)		
Well-baby clinics	228	1.76 (0.74)		
Municipal CWS	157	1.75 (0.70)		

[a] CAMHS: child and adolescent mental health clinics.

[b] CWS: child welfare services.

[c] This value was significant at *P*<.05.

[d] This value was significant at *P*<.01.

[e] This value was significant at *P*<.001.

Discussion

Overview of the Findings

Despite the rapid development of e-mental health services and the promising evidence for their utility [32], less attention has been paid to whether practitioners are positive toward internet-based interventions and would find these useful in their practice. Therefore, in this cross-sectional study of 2884 infant and child health leaders and practitioners, we aimed to investigate health professionals' perceived usefulness of e-mental health programs and identify the key areas in which they consider new e-mental health services to be useful, as well as differences between service practitioners and leaders, and different prenatal, infant, and child health care services for children aged 0-5 years. The results showed that a majority reported that they would sometimes or often find internet-based parenting interventions for different infant and child mental health problems as useful. Usefulness of internet-based interventions was rated acceptable for sleep problems, anxiety, and social withdrawal and shyness, whereas fewer reported that it would be useful for psychiatric problems such as obsessive disorders or trauma (eg, child maltreatment). Moreover, there were a few but small differences in the perceived usefulness between service leaders and practitioners (all effect sizes<0.32, all *P*<.02) and small-to-moderate differences between daycare centers, well-baby clinics, CWS, and CAMHS (all effect sizes<0.69, all *P*<.006).

Our findings are in line with results of previous studies that show that practitioners generally consider internet-based interventions useful but that attitudes may range from skepticism to positivity for prevention and treatment of mild-to-moderate problems [24,25,33]. These studies have also found that health personnel hold more negative views toward their usefulness for severe disorders. In our study, we compared a broad range of different infant and child mental health problems, which, to our knowledge, has not been done before. Participants had a positive attitude toward the use of internet-based interventions for infant and child health problems such as sleep problems, anxiety, social withdrawal, and dysregulation. The results revealed a more ambivalent attitude towards the use of e-mental health services for problem areas such as developmental delays and trauma. However, even for such problem areas, most practitioners recognized the potential of e-mental health interventions and seemed to be aware of its usefulness. With respect to the more ambivalent attitude of eHealth for some infant and child health problems, a possible explanation for this is that practitioners may perceive these problems as more severe and in need of different kinds of treatments. For example, Stallard et al [24] have shown that concerns about using eHealth with children and adolescents has 4 themes: limited potential, risk management, support and understanding, and lack of therapeutic relationship. Issues relating to the importance of therapeutic alliance are also found in adult literature [34]. Such concerns may apply more to disorders that are considered as clinically more severe and where there is a need for more interdisciplinary and frequent follow-up than mild-to-moderate problems. Furthermore, lack of knowledge about eHealth among practitioners may also be a possible explanation for the more ambivalent attitude to using e-mental health services for some problems.

To our knowledge, this study is the first to compare differences in the perceived usefulness of internet-based interventions for different infant and child mental health problems between leaders and practitioners and between different infant and child health services. Our results showed that top-level managers found internet-based interventions more useful than practitioners for several infant and child mental health problems. As this study is the first to compare differences in perceived usefulness between leaders and practitioners, the causes of the difference in opinion between the 2 occupational levels are not clearly identified. However, the UTAUT model hypothesizes that individual-level contextual factors such as gender, age, actual experience, and voluntariness of use would moderate the effect of behavioral intention [21]. Hence, potential differences in these variables among leaders and practitioners (ie, leaders are often older, more experienced, and exhibit more voluntariness of use) may contribute to explain why top-level managers found internet-based interventions more useful than practitioners on several infant and child mental health problems.

Previous research has shown that implementation leadership is a critical factor for organizational changes [35]. Hence, leadership may either promote or inhibit the adoption of e-mental health in services [36]. Even if practitioners are open to e-mental health, effective leadership may help to support implementation climate and efforts [35]. Some leaders may develop plans, anticipate and address implementation challenges, and have clear priorities and expectations (ie, proactive leadership), while others may give up when they face obstacles or fail to address challenges effectively (ie, nonperseverant leadership, [35]). A transformational leadership (ie, proactive) style seems to be important for the adoption of information technology [27], and previous studies have reported similar experiences with implementation of an internet-based postpartum depression intervention that show the importance of leadership [26].

According to the UTAUT model, higher-level contextual factors such as organization attributes would also influence technology acceptance and use [21]. Differences in perceived usefulness of internet-based interventions could therefore be expected for people working in different fields of the health care system. Contrary to this, we found few differences in perceived usefulness between infant and child health services. Although there are many similarities between infant and child health services [37], there may still be differences in how acceptable, feasible, and suitable e-mental health services can be. Hence, our findings make it important to further examine what opportunities exist within services, as there are also challenges and opportunities unique to each [35]. Some services may perceive that e-mental health services for a specific condition can be useful, but organizational conditions can still make them unworkable or inappropriate. This can, for example, in line with the UTAUT model be more structural and organizational conditions such as climate, organizational culture, leadership or grant schemes for parental support interventions, and other (policy) guidelines from official health authorities. As a future direction, such things may be important to consider from the outset before developing e-mental health services to increase the likelihood of new interventions being taken up and used in services.

Overall, infant and child health practitioners' attitudes toward internet-based interventions were positive, suggesting that many practitioners may be open to taking advantage of internet-based interventions. Only a minor proportion of those who could benefit from evidence-based parenting programs seem to receive these. Hence, our findings are encouraging, considering the potential of technology to expand the delivery portfolio to overcome barriers associated with face-to-face delivery and increase the availability and accessibility of e-mental health interventions to bridge the gaps in the provision of care. However, despite the general positivity toward such interventions, few web-based interventions are available for families with children aged 0-5 years. Programs for this age group are also mostly for disruptive child behaviors [15], and there appears to be only few programs that have been studied and made available in non-English languages. Hence, additional efforts are needed to develop, implement, and disseminate interventions for families with infants and young children.

Strengths and Limitations

This study has both strengths and limitations that should be recognized. To our knowledge, this study was the first to address health professionals' perceived usefulness of e-mental health programs for different infant and child problems and to compare different services and job levels. The advantage is that this study provides new insights. However, the disadvantage is that there are no comparable studies to rely on. For that reason, we may have omitted relevant aspects. Another notable strength is the large number of participants as well as a national sample consisting of practitioners from all major infant and child health services in Norway, likely to be targeted in dissemination efforts. However, as the study was web-based and practitioners were only contacted via email or social media, a selection bias may have been introduced. Practitioners facing greater practical barriers to the use of computers and internet (ie, lower computer fluency and reduced access to technology at work) may have been more reluctant to participate in the study. This has also been shown to have a negative impact on people's perceived usefulness of information technology [38]. However, 1 study suggested that prevalence estimates of exposure and outcome but not estimates of exposure-outcome associations are biased due to self-selection. Hence, it is important to bear in mind that selection bias does not necessarily influence results much when associations between variables are investigated [39]. Other drawbacks associated with web-based surveys are also important to keep in mind such as researchers cannot check whether participants have understood and interpreted the questions in the same way as the researchers meant. Further, participants cannot elaborate on their answers; therefore, we may have potentially missed nuances that would have yielded valuable insights. However, the web-based questionnaire was pilot tested independently by 6 health care professionals sampled from our intended study population, prior to data collection so as to minimize the risk for misinterpretation. Another limitation was that perceived usefulness of internet-based interventions was only assessed with 1 global item. However, to our knowledge, no general instruments exist for measuring perceived usefulness across different types of information technology, as these always must be modified to accommodate the specifics of the attitude object that is being evaluated. Thus, conducting studies with more extensive questions may have allowed for further insight into the topic, although our overall results are also consistent with those of previous studies. This supports our findings and indicates that a global single-item assessment of perceived usefulness is fully possible. However, using more questions may have added more variation, thereby making the results between different job levels and health and social services more distinguishable. This was not practically possible for our study, without making the survey potentially unacceptable and too time-consuming to complete. Other limitations include the cross-sectional nature of this study as well as multiple testing. Even when adjusting for multiple testing, it is possible that the significant results of some items may be due to chance. Thus, any significant finding must be interpreted with caution. Furthermore, the study was conducted in Norway. Thus, the results may not be applicable to countries with widely different health care systems, which may potentially limit the generalizability of our study findings. A final limitation may be that even the provision of brief textual information about e-mental health services can influence people's attitudes toward e-mental health [40,41], which may have affected participants' evaluations in our study.

Future Research

Future research should include more detailed questions about factors that could influence perceived usefulness, for example, general openness to new treatments, organizational support, and practical problems and barriers that limit successful implementation. For e-mental health to have the large public health impact that it is often praised for, there is a need for improving the translation of e-mental health research into clinical practice. Therefore, there is also a need for more research on cocreation of interventions adapted to both services and parents' needs, as well as conditions in clinical practice such as lack of time, resources, and low visibility [17]. At last, considering the COVID-19 pandemic and the consequent increase in the use of digital solutions, it can also be useful to conduct a new study as this crisis may have affected health professionals' perceived usefulness of e-mental health programs for different infant and child problems.

Conclusion

This study shows that internet-based interventions for different infant and child mental health problems within services such as daycare centers, well-baby clinics, municipal CWS, and CAMHS are sometimes or often perceived as useful. These are encouraging findings and support the continued exploration of internet-based mental health interventions as way to improve parental support. In turn, these insights may inform processes of technological development, clinical use, and organizational implementation of internet-based interventions.

Acknowledgments

We thank Monica Werner and Unni Rosenkilde, RBUP, for their help and support in data collection. We are indebted to the infant and child health practitioners who took part in this study.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Results from multiple comparisons utilizing the posthoc Bonferroni test_job level.
[DOCX File , 19 KB - mental_v7i11e15149_app1.docx]

Multimedia Appendix 2
Results from multiple comparisons utilizing the posthoc Bonferroni test_service.
[DOCX File , 26 KB - mental_v7i11e15149_app2.docx]

References

1. Saunders H, Kraus A, Barone L, Biringen Z. Emotional availability: theory, research, and intervention. Front Psychol 2015;6:1069 [FREE Full text] [doi: 10.3389/fpsyg.2015.01069] [Medline: 26283996]
2. Kempler L, Sharpe L, Miller C, Bartlett DJ. Do psychosocial sleep interventions improve infant sleep or maternal mood in the postnatal period? A systematic review and meta-analysis of randomised controlled trials. Sleep Med Rev 2016 Oct;29:15-22. [doi: 10.1016/j.smrv.2015.08.002] [Medline: 26555938]

3. Perrin E, Leslie L, Boat T. Parenting as Primary Prevention. JAMA Pediatr 2016 Jul 01;170(7):637-638. [doi: 10.1001/jamapediatrics.2016.0225] [Medline: 27182902]
4. Briggs-Gowan M, Carter A, Skuban E, Horwitz SM. Prevalence of social-emotional and behavioral problems in a community sample of 1- and 2-year-old children. J Am Acad Child Adolesc Psychiatry 2001 Jul;40(7):811-819. [doi: 10.1097/00004583-200107000-00016] [Medline: 11437020]
5. Skovgaard A, Houmann T, Christiansen E, Landorph S, Jørgensen T, Olsen EM, et al. The prevalence of mental health problems in children 1(1/2) years of age - the Copenhagen Child Cohort 2000. J Child Psychol Psychiatry 2007 Jan;48(1):62-70 [FREE Full text] [doi: 10.1111/j.1469-7610.2006.01659.x] [Medline: 17244271]
6. Accessed April 29, 2019. Norwegian Institute of Public Health (2018) Mental disorders among children and adolescents in Norway. URL: https://www.fhi.no/en/op/hin/groups/mental-health-children-adolescents/ [accessed 2019-04-29] [WebCite Cache ID 77zLb8xU4]
7. Wichstrøm L, Berg-Nielsen T, Angold A, Egger H, Solheim E, Sveen TH. Prevalence of psychiatric disorders in preschoolers. J Child Psychol Psychiatry 2012 Jun;53(6):695-705. [doi: 10.1111/j.1469-7610.2011.02514.x] [Medline: 22211517]
8. Heiervang E, Stormark K, Lundervold A, Heimann M, Goodman R, Posserud M, et al. Psychiatric disorders in Norwegian 8- to 10-year-olds: an epidemiological survey of prevalence, risk factors, and service use. J Am Acad Child Adolesc Psychiatry 2007 Apr;46(4):438-447. [doi: 10.1097/chi.0b013e31803062bf] [Medline: 17420678]
9. Wright B, Barry M, Hughes E, Trépel D, Ali S, Allgar D, et al. Clinical effectiveness and cost-effectiveness of parenting interventions for children with severe attachment problems: a systematic review and meta-analysis. Health Technol Assess 2015 Jul;19(52):vii-xxviii, 1 [FREE Full text] [doi: 10.3310/hta19520] [Medline: 26177494]
10. Yap M, Morgan A, Cairns K, Jorm A, Hetrick S, Merry S. Parents in prevention: A meta-analysis of randomized controlled trials of parenting interventions to prevent internalizing problems in children from birth to age 18. Clin Psychol Rev 2016 Dec;50:138-158. [doi: 10.1016/j.cpr.2016.10.003] [Medline: 27969003]
11. Tarver J, Daley D, Lockwood J, Sayal K. Are self-directed parenting interventions sufficient for externalising behaviour problems in childhood? A systematic review and meta-analysis. Eur Child Adolesc Psychiatry 2014 Dec;23(12):1123-1137. [doi: 10.1007/s00787-014-0556-5] [Medline: 24842197]
12. McGoron L, Ondersma S. Reviewing the need for technological and other expansions of evidence-based parent training for young children. Children and Youth Services Review 2015 Dec;59:71-83 [FREE Full text] [doi: 10.1016/j.childyouth.2015.10.012]
13. Rooksby M, Elouafkaoui P, Humphris G, Clarkson J, Freeman R. Internet-assisted delivery of cognitive behavioural therapy (CBT) for childhood anxiety: systematic review and meta-analysis. J Anxiety Disord 2015 Jan;29:83-92 [FREE Full text] [doi: 10.1016/j.janxdis.2014.11.006] [Medline: 25527900]
14. Pennant M, Loucas C, Whittington C, Creswell C, Fonagy P, Fuggle P, et al. Computerised therapies for anxiety and depression in children and young people: a systematic review and meta-analysis. Behav Res Ther 2015 Apr;67:1-18. [doi: 10.1016/j.brat.2015.01.009] [Medline: 25727678]
15. Baumel A, Pawar A, Kane J, Correll CU. Digital Parent Training for Children with Disruptive Behaviors: Systematic Review and Meta-Analysis of Randomized Trials. J Child Adolesc Psychopharmacol 2016 Oct;26(8):740-749. [doi: 10.1089/cap.2016.0048] [Medline: 27286325]
16. Vigerland S, Lenhard F, Bonnert M, Lalouni M, Hedman E, Ahlen J, et al. Internet-delivered cognitive behavior therapy for children and adolescents: A systematic review and meta-analysis. Clin Psychol Rev 2016 Dec;50:1-10 [FREE Full text] [doi: 10.1016/j.cpr.2016.09.005] [Medline: 27668988]
17. Feijt M, de Kort YA, Bongers IM, IJsselsteijn WA. Perceived Drivers and Barriers to the Adoption of eMental Health by Psychologists: The Construction of the Levels of Adoption of eMental Health Model. J Med Internet Res 2018 Apr 24;20(4):e153. [doi: 10.2196/jmir.9485] [Medline: 29691215]
18. Topooco N, Riper H, Araya R, Berking M, Brunn M, Chevreul K, et al. Attitudes towards digital treatment for depression: A European stakeholder survey. Internet Interv 2017 Jun;8:1-9 [FREE Full text] [doi: 10.1016/j.invent.2017.01.001] [Medline: 30135823]
19. Cook J, Biyanova T, Coyne JC. Barriers to adoption of new treatments: an internet study of practicing community psychotherapists. Adm Policy Ment Health 2009 Mar;36(2):83-90 [FREE Full text] [doi: 10.1007/s10488-008-0198-3] [Medline: 19104928]
20. Davis F, Bagozzi R, Warshaw P. User Acceptance of Computer Technology: A Comparison of Two Theoretical Models. Management Science 1989 Aug;35(8):982-1003. [doi: 10.1287/mnsc.35.8.982]
21. Venkatesh V, Thong JYL, Xu X. Consumer Acceptance and Use of Information Technology: Extending the Unified Theory of Acceptance and Use of Technology. MIS Quarterly 2012;36(1):157. [doi: 10.2307/41410412]
22. Grassl N, Nees J, Schramm K, Spratte J, Sohn C, Schott T, et al. A Web-Based Survey Assessing the Attitudes of Health Care Professionals in Germany Toward the Use of Telemedicine in Pregnancy Monitoring: Cross-Sectional Study. JMIR Mhealth Uhealth 2018 Aug 08;6(8):e10063 [FREE Full text] [doi: 10.2196/10063] [Medline: 30089606]
23. Fleming T, Merry S. Youth Work Service Providers' Attitudes Towards Computerized CBT for Adolescents. Behav Cogn Psychother 2012 May 17;41(3):265-279 [FREE Full text] [doi: 10.1017/s1352465812000306]

24. Stallard P, Richardson T, Velleman S. Clinicians' attitudes towards the use of computerized cognitive behaviour therapy (cCBT) with children and adolescents. Behav Cogn Psychother 2010 Oct;38(5):545-560. [doi: 10.1017/S1352465810000421] [Medline: 20615273]
25. Vigerland S, Ljótsson B, Bergdahl Gustafsson F, Hagert S, Thulin U, Andersson G, et al. Attitudes towards the use of computerized cognitive behavior therapy (cCBT) with children and adolescents: A survey among Swedish mental health professionals. Internet Interventions 2014 Jul;1(3):111-117 [FREE Full text] [doi: 10.1016/j.invent.2014.06.002]
26. Drozd F, Haga S, Lisøy C, Slinning K. Evaluation of the implementation of an internet intervention in well-baby clinics: A pilot study. Internet Interv 2018 Sep;13:1-7 [FREE Full text] [doi: 10.1016/j.invent.2018.04.003] [Medline: 30206512]
27. Schepers J, Wetzels M, de Ruyter K. Leadership styles in technology acceptance: do followers practice what leaders preach? Managing Service Quality 2005 Dec;15(6):496-508. [doi: 10.1108/09604520510633998]
28. Project no. 60041. Norwegian Centre for Research Data. URL: https://nsd.no/nsd/english/index.html [accessed 2020-03-01]
29. Pedlex. URL: https://www.pedlex.no/ [accessed 2020-03-01]
30. Zeanah C, Carter A, Cohen J, Egger H, Gleason M, Keren M, et al. Diagnostic classification of mental health and developmental disorders of infancy and early childhood. Infant Ment Health J 2016 Sep;37(5):471-475. [doi: 10.1002/imhj.21591] [Medline: 27570937]
31. Statistics Norway (2019) Population. URL: https://www.ssb.no/befolkning/faktaside/befolkningen [accessed 2019-06-03]
32. Andersson G, Cuijpers P, Carlbring P, Riper H, Hedman E. Guided Internet-based vs. face-to-face cognitive behavior therapy for psychiatric and somatic disorders: a systematic review and meta-analysis. World Psychiatry 2014 Oct;13(3):288-295 [FREE Full text] [doi: 10.1002/wps.20151] [Medline: 25273302]
33. Gun S, Titov N, Andrews G. Acceptability of Internet treatment of anxiety and depression. Australas Psychiatry 2011 Jun;19(3):259-264. [doi: 10.3109/10398562.2011.562295] [Medline: 21682626]
34. Flückiger C, Del Re AC, Wampold B, Horvath AO. The alliance in adult psychotherapy: A meta-analytic synthesis. Psychotherapy (Chic) 2018 Dec;55(4):316-340. [doi: 10.1037/pst0000172] [Medline: 29792475]
35. Finn N, Torres E, Ehrhart M, Roesch S, Aarons GA. Cross-Validation of the Implementation Leadership Scale (ILS) in Child Welfare Service Organizations. Child Maltreat 2016 Aug;21(3):250-255. [doi: 10.1177/1077559516638768] [Medline: 27002137]
36. Vis C, Mol M, Kleiboer A, Bührmann L, Finch T, Smit J, et al. Improving Implementation of eMental Health for Mood Disorders in Routine Practice: Systematic Review of Barriers and Facilitating Factors. JMIR Ment Health 2018 Mar 16;5(1):e20 [FREE Full text] [doi: 10.2196/mental.9769] [Medline: 29549072]
37. Aarons GA. Mental Health Provider Attitudes Toward Adoption of Evidence-Based Practice: The Evidence-Based Practice Attitude Scale (EBPAS). Ment Health Serv Res 2004 Jun;6(2):61-74. [doi: 10.1023/b:mhsr.0000024351.12294.65]
38. Becker E, Jensen-Doss A. Therapist attitudes towards computer-based trainings. Adm Policy Ment Health 2014 Nov;41(6):845-854. [doi: 10.1007/s10488-013-0522-4] [Medline: 24150441]
39. Nilsen R, Vollset S, Gjessing H, Skjaerven R, Melve K, Schreuder P, et al. Self-selection and bias in a large prospective pregnancy cohort in Norway. Paediatr Perinat Epidemiol 2009 Nov;23(6):597-608. [doi: 10.1111/j.1365-3016.2009.01062.x] [Medline: 19840297]
40. Casey L, Joy A, Clough BA. The impact of information on attitudes toward e-mental health services. Cyberpsychol Behav Soc Netw 2013 Aug;16(8):593-598. [doi: 10.1089/cyber.2012.0515] [Medline: 23679567]
41. Apolinário-Hagen J, Harrer M, Kählke F, Fritsche L, Salewski C, Ebert DD. Public Attitudes Toward Guided Internet-Based Therapies: Web-Based Survey Study. JMIR Ment Health 2018 May 15;5(2):e10735 [FREE Full text] [doi: 10.2196/10735] [Medline: 29764797]

Abbreviations

CAMHS: child and adolescent mental health clinics
CWS: child welfare services
DC 0-5: Diagnostic Classification of Mental Health and Developmental Disorders of Infancy and Early Childhood
RBUP: Regional Centre for Child and Adolescent Mental Health
UTAUT: Unified Theory of Acceptance and Use of Technology

Edited by G Eysenbach, R Leung; submitted 26.06.19; peer-reviewed by R Otte, SGS Shah, C Salewski; comments to author 30.03.20; revised version received 09.06.20; accepted 28.10.20; published 18.11.20.

Please cite as:
Størksen HT, Haga SM, Slinning K, Drozd F
Health Personnel's Perceived Usefulness of Internet-Based Interventions for Parents of Children Younger Than 5 Years: Cross-Sectional Web-Based Survey Study
JMIR Ment Health 2020;7(11):e15149
URL: http://mental.jmir.org/2020/11/e15149/
doi:10.2196/15149
PMID:33206058

Original Paper

Health Care Management Models for the Evolution of Hospitalization in Acute Inpatient Psychiatry Units: Comparative Quantitative Study

Susel Góngora Alonso[1], MSc; Beatriz Sainz-De-Abajo[1], PhD; Isabel De la Torre-Díez[1], PhD; Manuel Franco-Martin[2], MD

[1]Department of Signal Theory and Communications, and Telematics Engineering, University of Valladolid, Valladolid, Spain
[2]University Rio Hortega Hospital, Valladolid, Spain

Corresponding Author:
Isabel De la Torre-Díez, PhD
Department of Signal Theory and Communications, and Telematics Engineering
University of Valladolid
Paseo de Belén, 15
Valladolid, 47011
Spain
Phone: 34 983423000 ext 3703
Email: isator@tel.uva.es

Abstract

Background: Mental health disorders are a problem that affects patients, their families, and the professionals who treat them. Hospital admissions play an important role in caring for people with these diseases due to their effect on quality of life and the high associated costs. In Spain, at the Healthcare Complex of Zamora, a new disease management model is being implemented, consisting of not admitting patients with mental diseases to the hospital. Instead, they are supervised in sheltered apartments or centers for patients with these types of disorders.

Objective: The main goal of this research is to evaluate the evolution of hospital days of stay of patients with mental disorders in different hospitals in a region of Spain, to analyze the impact of the new hospital management model.

Methods: For the development of this study, a database of patients with mental disorders was used, taking into account the acute inpatient psychiatry unit of 11 hospitals in a region of Spain. SPSS Statistics for Windows, version 23.0 (IBM Corp), was used to calculate statistical values related to hospital days of stay of patients. The data included are from the periods of 2005-2011 and 2012-2015.

Results: After analyzing the results, regarding the days of stay in the different health care complexes for the period between 2005 and 2015, we observed that since 2012 at the Healthcare Complex of Zamora, the total number of days of stay were reduced by 64.69%. This trend is due to the implementation of a new hospital management model in this health complex.

Conclusions: With the application of a new hospital management model at the Healthcare Complex of Zamora, the number of days of stay of patients with mental diseases as well as the associated hospital costs were considerably reduced.

(JMIR Ment Health 2020;7(11):e15776) doi:10.2196/15776

KEYWORDS

acute inpatient psychiatry unit; database; hospitalizations; mental health; readmission; SPSS

Introduction

Having good mental health allows us to develop the social and intellectual skills that are needed to face new challenges in everyday life [1]. The World Health Organization has determined that mental health is a cornerstone of general health. Large-scale collection of mental health–related data is difficult and is done infrequently. It is a challenge for researchers to evaluate seasonal, weekly, or diurnal trends [2].

Mental illnesses can cause mild to severe disorders in thinking and behavior; they can incapacitate patients, preventing them from carrying out the ordinary demands and routines of life [3]. Some of the most common disorders are clinical depression, bipolar disorder, dementia, schizophrenia, and anxiety disorders.

The problem for people who suffer from these disorders is when specialists do not interpret the symptoms correctly. Symptoms may include changes in mood, personality, or personal habits and/or social withdrawal. Mental health problems may be related to excessive stress due to a particular situation or a series of events [4]; their prevalence is high worldwide. At least 1% of any population is incapacitated by a serious mental disorder at a specific time. The percentage of people affected in any period of their lives is 10% [1].

Hospital admissions are important events in the care of people with mental disorders due to the associated costs and their possible effect on quality of life [5]. Despite the application of various personalized treatments, the rate of relapse among the mentally ill is relatively high. It is estimated that the relapse rate among people with schizophrenia is between 50% and 92%. This implies high morbidity and high readmission rate. As a consequence, this relapse rate has a high cost to the health care system and community services [6].

In Fleury et al [7], the authors showed that only 17% of patients had received a follow-up appointment before hospital discharge. Best practices recommend brief hospitalizations and postdischarge follow-ups to improve social integration and recovery. Psychiatric care is still necessary for a small subgroup of patients who cannot be treated safely or effectively at home [8].

The majority of hospitalized psychiatric patients can be discharged without extensive follow-up. However, patients with a serious mental illness need long-term aftercare [9]. Early psychiatric readmission serves as a negative indicator of the quality of care in mental health services. Some studies report that days of stay of hospitalized patients under 28 days increase readmission rates [10].

In the Healthcare Complex of Zamora, Spain, one of the hospital centers analyzed in this work, a new management model has been applied since 2012. It consists of not admitting patients with mental diseases to the hospital. Instead, they are supervised in sheltered apartments or centers for patients with these types of disorders. Hence, the main objective of this research is to evaluate the evolution of hospital admissions and days of stay of patients with mental diseases in 11 hospital centers of a region of Spain. From the data, we can see the effects of the application of a new model on monitoring patients outside the hospital.

There are similar studies that show us the feasibility of our research. In Cooper et al [11], the authors described the service provision of 32 hospitals and evaluated the changes in the management and quality of the service, comparing it with the results of a previous study over a period of 10 years. Steeg et al [12] presented a study that applied methods of multiple imputation and propensity score. Four types of hospital management were related to patients who self-harm and a risk of suicide in these patients in the following 12 months. As a result, it was concluded that the propensity score adjustment seemed to mitigate only some of the greatest risks observed and that the differences between the treatment groups had little impact on the reduction of suicide.

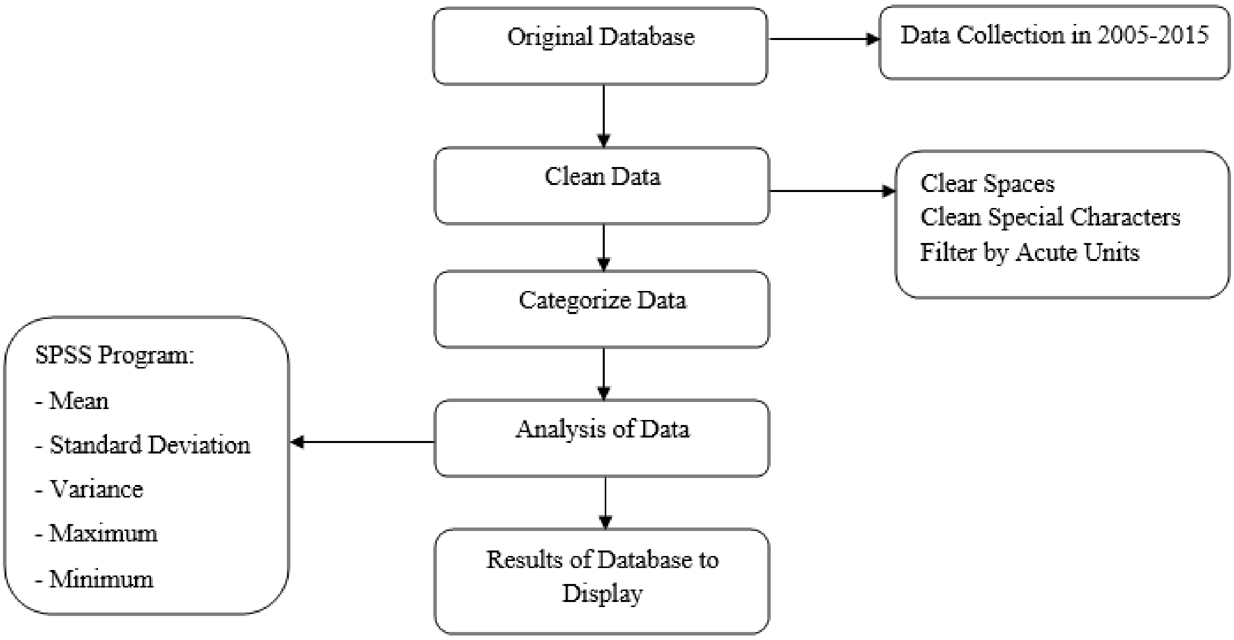
Below we show the methodology used in this study, the results achieved, and, finally, the discussion and conclusions of the investigation.

Methods

The hospital admission records for this study were extracted from an anonymized database of patients with mental diseases. The database includes a total of 53,641 records from 11 public health care complexes in Castilla and León, Spain. Once the data were processed, we included in the study the acute inpatient psychiatry unit (AIPU) of each hospital, which included a total of 49,824 admissions. The data follow the International Classification of Diseases, Ninth Revision (ICD-9), and the study period covers the years from 2005 to 2015. The data include admissions of patients with the following mental disorders: schizophrenia, bipolar disorder, Alzheimer disease, depression, autism, disorders due to drugs and alcohol, affective disorders, and other psychoses.

The database includes the name of the hospital, the gender of the patient, the year of admission, the number of days of stay, the date of admission, the date of discharge, the diagnosis, and the therapies used according to the diagnosis. For this study, the three selected variables were (1) the name of the hospital, (2) the days of stay, and (3) the year of admission. The rest of the variables were excluded. In addition, null values, double blanks, and special characters were removed. Figure 1 shows the flowchart followed in this study.

Figure 1. Study flowchart.



To obtain the descriptive and inferential statistics of the data from the 11 health care complexes during the periods 2005-2011 and 2012-2015, we used SPSS for Windows, version 23.0 (IBM Corp). We calculated the following parameters: mean, standard deviation, variance, minimum number of days of stay, and maximum number of days of stay.

Results

For the study, 49,824 out of 53,641 (92.88%) database records from anonymous hospital admissions, from 2005 to 2015, were included. All the records were part of the AIPUs of 11 health care complexes in Castilla and León, Spain. The data obtained from the total number of days of stay per year are shown in Table 1 for Zamora, Ávila, Burgos, León, and Palencia, and in Table 2 for Salamanca, Soria, Segovia, University Clinical Hospital of Valladolid, The Bierzo Hospital, and the University Hospital of Rio Hortega.

Table 1. Total number of days of stay per admission year for hospital centers in Zamora, Ávila, Burgos, León, and Palencia.

Year	Days of stay per health care complex, n (%)				
	Healthcare Complex of Zamora (N=59,789)	Healthcare Complex of Ávila (N=40,842)	Healthcare Complex of Burgos (N=131,948)	Healthcare Complex of León (N=70,371)	Healthcare Complex of Palencia (N=54,395)
2005	8336 (13.94)	2511 (6.15)	13,579 (10.29)	6060 (8.61)	4883 (8.98)
2006	8012 (13.40)	2415 (5.91)	12,920 (9.79)	6033 (8.57)	4842 (8.90)
2007	5979 (10.00)	2268 (5.55)	13,509 (10.24)	6574 (9.34)	4631 (8.51)
2008	6171 (10.32)	5401 (13.23)	13,189 (9.99)	6769 (9.62)	4845 (8.91)
2009	7166 (11.99)	2648 (6.48)	12,465 (9.45)	4815 (6.84)	4799 (8.82)
2010	6953 (11.63)	3871 (9.48)	11,406 (8.64)	6874 (9.77)	5073 (9.33)
2011	7135 (11.93)	5878 (14.39)	12,860 (9.75)	6700 (9.52)	6348 (11.67)
2012	3412 (5.71)	5261 (12.88)	10,777 (8.17)	7603 (10.81)	5056 (9.29)
2013	1217 (2.04)	3802 (9.31)	11,002 (8.34)	6812 (9.68)	4677 (8.60)
2014	2051 (3.43)	3535 (8.66)	9375 (7.10)	6451 (9.17)	4584 (8.43)
2015	3357 (5.61)	3252 (7.96)	10,866 (8.24)	5680 (8.07)	4657 (8.56)
Total	59,789 (100)	40,842 (100)	131,948 (100)	70,371 (100)	54,395 (100)

Table 2. Total number of days of stay per admission year for hospital centers in Salamanca, Soria, Segovia, Valladolid, and El Bierzo.

Year	Days of stay per health care complex, n (%)					
	Healthcare Complex of Salamanca (N=62,031)	Healthcare Complex of Soria (N=98,242)	Healthcare Complex of Segovia (N=54,783)	University Clinical Hospital of Valladolid (N=97,303)	The Bierzo Hospital (N=44,069)	University Hospital of Rio Hortega (N=34,896)
2005	7186 (11.58)	8806 (8.96)	5470 (9.98)	7713 (7.93)	4360 (9.89)	0 (0)
2006	7171 (11.56)	8102 (8.25)	5621 (10.26)	7686 (7.90)	3849 (8.73)	0 (0)
2007	5346 (8.62)	8353 (8.50)	4488 (8.19)	8762 (9.00)	4259 (9.67)	0 (0)
2008	5910 (9.53)	9081 (9.25)	3997 (7.30)	9629 (9.90)	4679 (10.62)	0 (0)
2009	5194 (8.37)	8360 (8.51)	4852 (8.86)	8726 (8.97)	4526 (10.27)	233 (0.67)
2010	4504 (7.26)	7637 (7.77)	4460 (8.14)	9053 (9.30)	4241 (9.62)	5920 (16.96)
2011	5367 (8.65)	9354 (9.52)	4948 (9.03)	8122 (8.35)	4051 (9.19)	5239 (15.01)
2012	5477 (8.83)	9659 (9.83)	4614 (8.42)	9405 (9.66)	3920 (8.90)	5817 (16.67)
2013	6597 (10.64)	10,542 (10.73)	5366 (9.80)	9318 (9.58)	3333 (7.56)	5873 (16.83)
2014	4899 (7.90)	10,482 (10.67)	5425 (9.90)	9676 (9.94)	3494 (7.93)	5784 (16.58)
2015	4380 (7.06)	7866 (8.01)	5542 (10.12)	9213 (9.47)	3357 (7.62)	6030 (17.28)
Total	62,031 (100)	98,242 (100)	54,783 (100)	97,303 (100)	44,069 (100)	34,896 (100)

The results show that the Healthcare Complex of Burgos had the highest number of days of stay during the 11 years analyzed; it represents 17.62% (131,948/748,669) of the total sample. The Healthcare Complex of Soria had the next highest total number of days of stay, with 13.12% (98,242/748,669) of the total sample. The University Hospital of Rio Hortega in Valladolid had the lowest number of days of stay of 4.66% (34,896/748,669) of the total, but we must consider that the data have only been recorded for this hospital since 2009. The Healthcare Complex of Ávila represents 5.46% (40,842/748,669) of the total sample of days of stay. In 2012, a new hospital management model was implemented at the Healthcare Complex of Zamora; under this model, patients are supervised in sheltered homes or other similar centers. Tables 1 and 2 show the decrease in the number of patient days of stay in that hospital.

Through the SPSS statistics program, we obtained the following statistical parameters: mean, standard deviation, variance, minimum, and maximum, taking into account the total number of days of stay per hospital complex in the periods 2005-2011 and 2012-2015 (see Table 3).

Table 3. Descriptive statistics of total days of stay per hospital over 11 years.

Health care complex and years	Days of stay			
	Mean (SD)	Variance	Minimum	Maximum
Healthcare Complex of Zamora				
2005-2011	7107.43 (866.60)	750,987.62	5979	8336
2012-2015	2509.25 (1066.70)	1,137,846.92	1217	3412
Healthcare Complex of Ávila				
2005-2011	3570.29 (1514.97)	2,295,136.57	2268	5878
2012-2015	3962.50 (894.32)	799,809.67	3252	5261
Healthcare Complex of Burgos				
2005-2011	12,846.86 (743.74)	553,155.81	11,406	13,579
2012-2015	10,505.00 (758.99)	576,071.33	9375	11,002
Healthcare Complex of León				
2005-2011	6260.71 (719.70)	517,970.57	4815	6874
2012-2015	6636.50 (798.80)	638,075.00	5680	7603
Healthcare Complex of Palencia				
2005-2011	5060.14 (582.57)	339,392.14	4631	6348
2012-2015	4743.50 (212.13)	45,000.33	4584	5056
Healthcare Complex of Salamanca				
2005-2011	5811.14 (1020.81)	1,042,054.14	4504	7186
2012-2015	5338.25 (951.30)	904,962.25	4380	6597
Healthcare Complex of Soria				
2005-2011	8527.57 (591.55)	349,928.95	7637	9354
2012-2015	9637.25 (1247.66)	1,556,658.25	7866	10,542
Healthcare Complex of Segovia				
2005-2011	4833.71 (576.97)	332,890.91	3997	5621
2012-2015	5236.75 (421.56)	177,712.92	4614	5542
University Clinical Hospital of Valladolid				
2005-2011	8527.29 (721.09)	519,967.91	7686	9629
2012-2015	9403.00 (198.21)	39,286.00	9213	9676
The Bierzo Hospital				
2005-2011	4280.71 (278.77)	77,712.91	3849	4679
2012-2015	3526.00 (272.07)	74,023.33	3333	3920
University Hospital of Rio Hortega				
2005-2011	3797.33 (3105.53)	9,644,294.33	233	5920
2012-2015	5876.00 (109.04)	11,890.00	5784	6030

Table 4 shows a comparison of total number of days of stay in 2012 for each hospital in relation to the mean days of stay in the period from 2005 to 2011. In the case of the University Hospital of Rio Hortega, the mean is calculated from data between 2009 and 2011, because in the previous years there were no records. The results show that since 2012, the number of days of stay in the Healthcare Complex of Zamora decreased considerably. In other health care complexes, the days of stay increase in some cases and in others the behavior is not too variable.

Table 4. Mean days of stay in the period from 2005 to 2011 compared to the total days of stay from the year 2012 for each hospital.

Health care complex	Days of stay from 2005 to 2011, mean (SD)	Total days of stay in 2012, n
Healthcare Complex of Zamora	7107.43 (866.60)	3412
Healthcare Complex of Ávila	3570.29 (1514.97)	5261
Healthcare Complex of Burgos	12,846.86 (743.74)	10,777
Healthcare Complex of León	6260.71 (719.70)	7603
Healthcare Complex of Palencia	5060.14 (582.57)	5056
Healthcare Complex of Salamanca	5811.14 (1020.81)	5477
Healthcare Complex of Soria	8527.57 (591.55)	9659
Healthcare Complex of Segovia	4833.71 (576.97)	4614
University Clinical Hospital of Valladolid	8527.29 (721.09)	9405
The Bierzo Hospital	4280.71 (278.77)	3920
University Hospital of Rio Hortega	3797.33 (3105.53)	5817

Table 5 reports the percentage increase and decrease in the number of days of stay for each hospital in 2012 with respect to the mean days of stay in the period from 2005 to 2011. The records for the University Hospital of Rio Hortega are limited to the years between 2009 and 2011. The results show that the percentage decrease in the number of days of stay is 52% higher than in the rest of the care complexes.

Table 5. Percentage increase and decrease in the number of days of stay in 2012 with respect to the mean days of stay from 2005 to 2011.

Health care complex	Increase or decrease in days of stay, %
Increase	
Healthcare Complex of Ávila	47.36
Healthcare Complex of León	21.44
Healthcare Complex of Soria	13.27
University Clinical Hospital of Valladolid	10.29
University Hospital of Rio Hortega	53.19
Decrease	
Healthcare Complex of Zamora	52.00
Healthcare Complex of Burgos	16.11
Healthcare Complex of Palencia	0.08
Healthcare Complex of Salamanca	5.75
Healthcare Complex of Segovia	4.55
The Bierzo Hospital	8.43

The results from Tables 4 and 5 show how the days of stay at the Healthcare Complex of Zamora decreased considerably in 2012, with respect to other hospital centers. In Table 6, we show the evolution of the number of days of stay in this hospital from 2012 to 2015, with respect to the mean days of stay from 2005 to 2011.

Table 6. Percentage decrease in the number of days of stay for each year from 2012 to 2015 with respect to the mean days of stay from 2005 to 2011 for the Healthcare Complex of Zamora.

Year	Days of stay
2005-2011, mean (SD)	7107.43 (866.60)
2012, % decrease	52.00
2013, % decrease	82.88
2014, % decrease	71.15
2015, % decrease	52.77

Discussion

Once a serious mental illness occurs it tends to become chronic, and a patient may need repeated hospitalizations that affect daily life and social integration. Therefore, early diagnosis, proper treatment, and follow-up of mental health disorders are crucial for disease prevention [13].

In this study, we used a total of 49,824 records of anonymous hospital admissions of patients with mental diseases. The sample corresponds to a time period of 11 years, from 2005 to 2015, and includes 11 AIPUs.

It is necessary to highlight the variation of the behavior of days of stay per year in the Healthcare Complex of Zamora since 2012; in Tables 1, 2, and 3, the percentage of days of stay of this hospital compared to other hospitals reduced the mean number of days of stay by 64.69%.

This favorable variation of the total number of days of stay per year is due to the hospital management model that was implemented; that is, patients with mental diseases do not go to the hospital for treatment. Instead, the specialists travel to the primary health care center to see their patients. This allows for the integration of levels of care in primary and hospital care. The sheltered homes are another aspect of the applied management model, where the patient feels integrated into society and their daily lives.

Since 2012, in the Healthcare Complex of Ávila and the University Hospital of Rio Hortega, the percentage of days of stay with respect to the mean days of stay from 2005 to 2011 has increased by 47.36% and 53.19%, respectively (see Tables 4 and 5). Table 5 indicates that in the Healthcare Complex of Burgos and The Bierzo Hospital, the days of stay decreased by 16.11% and 8.43%, respectively. These values are lower compared to the value of days of stay in the Hospital of Zamora in 2012 (n=3412), which represents a decreased mean of 52.00% from the previous years (mean 7107.43, SD 866.60).

Table 6 shows the decrease in the number of days of stay for patients with mental diseases since 2012. In 2013, the highest decrease of 82.88% was observed, which represents a total of 1217 days of stay registered in that year.

In relation to expenses, a stay in these health care complexes costs between €325 and €408 (US $384.71 and US $482.96) [14]. Taking into account the Healthcare Complex of Zamora, the mean number of days of stay between 2005 and 2011 represents a cost of €2,309,914.75 to €2,999,731 (US $2,734,280.74 to US $3,550,826.58). When applying the hospital management model, the cost of days of stay was reduced in 2012 by 51.99%; it corresponds to a value of €1,108,900 to €1,392,096 (US $1,312,621.56 to US $1,647,844.92). Therefore, in the 2012-2015 period, the cost of the mean days of stay was reduced by 64.69%.

These results allow us to demonstrate the efficiency of the management model applied in the Healthcare Complex in Zamora. Hence, we propose the following as future lines of study: (1) apply the hospital management model to the rest of the health care complexes to improve management efficiency, (2) analyze the results in subsequent years using the same model, comparing it with what was obtained previously, (3) analyze the trend of mental health diseases in the data set and determine the main disorders in this region, and (4) apply machine learning techniques to the database in order to obtain predictions of the most prevalent mental disorders in patients.

Acknowledgments

This research has been supported by Gerencia Regional de Salud (GRS 1801/A/18) of the Junta de Castilla y León, under the project Análisis de Eficiencia y Sostenibilidad del Método Reticular de Atención en Salud Mental y Estrategia de Mejora.

Conflicts of Interest

None declared.

References

1. Thenmozhi R, Kabali B, Kapali SC. Effect of administrative stress on mental health in a group of office workers: A cross-sectional study. In: Proceedings of the 2014 International Conference on Science Engineering and Management Research (ICSEMR). 2014 Presented at: 2014 International Conference on Science Engineering and Management Research (ICSEMR); November 27-29, 2014; Chennai, India p. 1-3. [doi: 10.1109/ICSEMR.2014.7043570]
2. Watkins NG, Lovell NH, Larsen ME. SMCT: An innovative tool for mental health analysis of Twitter data. In: Proceedings of the 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). 2018 Jul Presented at: 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC); July 18-21, 2018; Honolulu, HI p. 4114-4117. [doi: 10.1109/EMBC.2018.8513257]
3. Välimäki M, Kannisto KA, Vahlberg T, Hätönen H, Adams CE. Short text messages to encourage adherence to medication and follow-up for people with psychosis (Mobile.Net): Randomized controlled trial in Finland. J Med Internet Res 2017 Jul 12;19(7):e245 [FREE Full text] [doi: 10.2196/jmir.7028] [Medline: 28701292]
4. Jain S, Gangwar M. A mining analysis over psychiatric database for mental health classification. i-Manager's J Comput Sci 2017;5(1):7-13. [doi: 10.26634/jcom.5.1.13792]

5. Cronin RM, Hankins JS, Byrd J, Pernell BM, Kassim A, Adams-Graves P, et al. Risk factors for hospitalizations and readmissions among individuals with sickle cell disease: Results of a US survey study. Hematology 2019 Dec;24(1):189-198 [FREE Full text] [doi: 10.1080/16078454.2018.1549801] [Medline: 30479187]
6. Adebiyi MO, Mosaku SK, Irinoye OO, Oyelade OO. Socio-demographic and clinical factors associated with relapse in mental illness. Int J Afr Nurs Sci 2018;8:149-153 [FREE Full text] [doi: 10.1016/j.ijans.2018.05.007]
7. Fleury MJ, Fortin M, Rochette L, Grenier G, Huỳnh C, Pelletier É, et al. Assessing quality indicators related to mental health emergency room utilization. BMC Emerg Med 2019 Jan 15;19(1):8 [FREE Full text] [doi: 10.1186/s12873-019-0223-8] [Medline: 30646847]
8. Steeg S, Emsley R, Carr M, Cooper J, Kapur N. Routine hospital management of self-harm and risk of further self-harm: Propensity score analysis using record-based cohort data. Psychol Med 2018 Jan;48(2):315-326. [doi: 10.1017/S0033291717001702] [Medline: 28637535]
9. Higgins N, Meehan T, Dart N, Kilshaw M, Fawcett L. Implementation of the Safewards model in public mental health facilities: A qualitative evaluation of staff perceptions. Int J Nurs Stud 2018 Dec;88:114-120. [doi: 10.1016/j.ijnurstu.2018.08.008] [Medline: 30236863]
10. Roos E, Bjerkeset O, Steinsbekk A. Health care utilization and cost after discharge from a mental health hospital: An RCT comparing community residential aftercare and treatment as usual. BMC Psychiatry 2018 Nov 12;18(1):363 [FREE Full text] [doi: 10.1186/s12888-018-1941-2] [Medline: 30419894]
11. Cooper J, Steeg S, Bennewith O, Lowe M, Gunnell D, House A, et al. Are hospital services for self-harm getting better? An observational study examining management, service provision and temporal trends in England. BMJ Open 2013 Nov 19;3(11):e003444 [FREE Full text] [doi: 10.1136/bmjopen-2013-003444] [Medline: 24253029]
12. Steeg S, Carr M, Emsley R, Hawton K, Waters K, Bickley H, et al. Suicide and all-cause mortality following routine hospital management of self-harm: Propensity score analysis using multicentre cohort data. PLoS One 2018;13(9):e0204670 [FREE Full text] [doi: 10.1371/journal.pone.0204670] [Medline: 30261030]
13. Boege I, Herrmann J, Wolff JK, Hoffmann U, Koelch M, Kurepkat M, et al. CCSchool: A multicentre, prospective study on improving continuum of care in children and adolescents with mental health problems associated with school problems in Germany. BMC Health Serv Res 2018 Dec 06;18(1):947 [FREE Full text] [doi: 10.1186/s12913-018-3713-6] [Medline: 30522471]
14. Decreto 25/2010, de 17 de junio, sobre precios públicos por actos asistenciales y servicios sanitarios prestados por la Gerencia Regional de Salud. Boletín Oficial de Castilla y León. Valladolid, Spain: Junta de Castilla y Leon; 2013 Dec 30. URL: https://www.saludcastillayleon.es/transparencia/es/transparencia/informacion-datos-publicos/gestion-economica/coste-servicios/precios-publicos-actos-asistenciales-servicios-sanitarios.ficheros/1199560-01%20ANEXO%20PRECIOS%20P%C3%9ABLICOS.pdf [accessed 2020-11-19]

Abbreviations

AIPU: acute inpatient psychiatry unit
ICD-9: International Classification of Diseases, Ninth Revision

Edited by G Eysenbach; submitted 05.08.19; peer-reviewed by M Sharma, A Khanna; comments to author 12.09.19; revised version received 13.09.19; accepted 23.09.19; published 30.11.20.

Please cite as:
Góngora Alonso S, Sainz-De-Abajo B, De la Torre-Díez I, Franco-Martin M
Health Care Management Models for the Evolution of Hospitalization in Acute Inpatient Psychiatry Units: Comparative Quantitative Study
JMIR Ment Health 2020;7(11):e15776
URL: http://mental.jmir.org/2020/11/e15776/
doi:10.2196/15776
PMID:33252351

Original Paper

Shaping Blended Care: Adapting an Instrument to Support Therapists in Using eMental Health

Hanneke Kip[1,2], MSc; Jobke Wentzel[1,3], PhD; Saskia M Kelders[1,4], PhD

[1]Centre for eHealth and Wellbeing Research, Department of Psychology, Health & Technology, University of Twente, Enschede, Netherlands
[2]Department of Research, Transfore, Deventer, Netherlands
[3]Research Group IT Innovations in Health Care, Windesheim University of Applied Sciences, Zwolle, Netherlands
[4]Optentia Research Focus Area, North-West University, Vanderbijlpark, South Africa

Corresponding Author:
Hanneke Kip, MSc
Centre for eHealth and Wellbeing Research
Department of Psychology, Health & Technology
University of Twente
Drienerlolaan 5
Enschede, 7500 AE
Netherlands
Phone: 31 53 489 9180
Email: h.kip@utwente.nl

Abstract

Background: Although eMental health interventions, especially when delivered in a blended way, have great potential to improve the quality and efficiency of mental health care, their use in practice lags behind expectations. The Fit for Blended Care (FfBC) instrument was developed to support therapists and clients in shaping blended care in a way that optimally fits their needs. However, this existing version cannot be directly applied to specific branches of mental health care as it is too broad and generic.

Objective: The goal of this study is to adapt the existing FfBC instrument to fit a specific, complex setting—forensic mental health care—by means of participatory development with therapists.

Methods: The participatory process was divided into 4 phases and was executed by a project team consisting of 1 manager, 3-5 therapists, and 1 researcher. In phase 1, general requirements for the adaptation of the existing instrument were discussed in 2 focus groups with the project team. In phase 2, patient-related factors that influence the use of an existing web-based intervention were elicited through semistructured interviews with all 18 therapists working at an outpatient clinic. In phase 3, multiple focus groups with the project teams were held to create the first version of the adapted FfBC instrument. In phase 4, a digital prototype of the instrument was used with 8 patients, and the experiences of the 4 therapists were discussed in a focus group.

Results: In phase 1, it became clear that the therapists' main requirement was to develop a much shorter instrument with a few items, in which the content was specifically tailored to the characteristics of forensic psychiatric outpatients. The interviews showed a broad range of patient-related factors, of which 5 were used in the instrument: motivation for blended treatment; writing about thoughts, feelings, and behavior; conscientiousness; psychosocial problems; and social support. In addition, a part of the instrument was focused on the practical necessary preconditions that patients should fill by themselves before the treatment was developed. The use of the web-based prototype of the instrument in treatment resulted in overall positive experiences with the content; however, therapists indicated that the items should be formulated in a more patient-centered way to encourage their involvement in discussing the factors.

Conclusions: The participatory, iterative process of this study resulted in an adapted version of the FfBC instrument that fits the specific forensic context and supports shared decision making. In general, the adaptiveness of the instrument is important: its content and implementation should fit the type of care, the organization, and eHealth intervention. To adapt the instrument to other contexts, the guidelines described in this paper can be followed.

(JMIR Ment Health 2020;7(11):e24245) doi:10.2196/24245

KEYWORDS

eHealth; blended care; implementation science; participatory development; forensic psychiatry; mobile phone

Introduction

The Benefits of Blended Care

eMental health interventions are a potentially effective and efficient way to improve the quality of care in a mental health care system that is under pressure due to shortages in staff and money [1-3]. eMental health refers to the use of technology for the treatment or prevention of mental health disorders [4]. Although there are different types of technologies that can be used [4,5], web-based interventions are currently the most predominant form in both research and practice. The content of these types of interventions is based on existing treatment models such as cognitive behavioral therapy or mindfulness, and they offer treatment via multiple modalities such as written text, assignments, and short videos [6]. Studies have shown that these types of interventions can result in clinical outcomes that are comparable with those of standard in-person treatments [7-10]. In addition, they have the potential to increase the efficiency of care by replacing parts of the in-person treatment by web-based treatment [3,11]. Combining this *offline* in-person treatment with *web-based* technologies in mental health care is referred to as blended care [12]. By integrating both approaches, we can have the best of both worlds: offering low-threshold web-based treatment, independent of place and time, which increases the patient's sense of ownership while maintaining the advantages of a strong therapeutic alliance of in-person treatment [3,11,13]. Despite the benefits of blended care, implementation in practice is very challenging [11,14-16], partly due to the barriers experienced by therapists [17]. Among other things, they often do not think of introducing the possibility of using eMental health interventions to their clients as these are *not in their system* and thus are not on top of their mind [17,18]. Furthermore, especially therapists with little eMental health experience are unsure about the topics that they need to address when introducing or discussing the use of eMental health interventions in treatment with a patient [15,17]. In addition, therapists often decide whether to use eMental health by themselves, based on their own estimation of the patient, instead of considering its applicability together with the patient via shared decision making [18]. If eMental health interventions are used, they are often viewed as a separate addition instead of an equal, fully integrated part of the treatment [19]. Blended care is often delivered in a standardized *one-size-fits-all* way, whereas ideally, the way eHealth is integrated in treatment should be personalized based on characteristics and preferences of individual clients [14,20]. These reasons for the lack of successful implementation of blended care indicate that there is a need to support therapists in shaping their blended treatment in such a way that it can be embedded in treatment to fit the preferences and skills of the patient.

Fit for Blended Care Instrument

A tool that was designed to support therapists in mixing web-based and offline mental health care is the *Fit for Blended Care* (FfBC) instrument. This instrument aims to support shared decision making in shaping blended treatment in mental health care [12]. To achieve this, it provides topics for therapists and patients to think about and discuss as well as decide on topics related to the needs, characteristics, and skills of a patient regarding blended care. On the basis of a literature review, multiple focus groups and interviews with both therapists and clients as well as a document containing instructions for and the items of the FfBC instrument were created [12]. The instrument consists of 4 main parts, which are briefly described in Table 1. All items of the instrument are provided in the left column of the table in Multimedia Appendix 1. Although the FfBC instrument is considered to be a valuable tool to shape blended care in practice [12,20,21], its current factors seem to be too generic and broad for application in specific domains of mental health care [20]. For example, there are many differences in patient characteristics and treatment goals in the treatment of addiction, anxiety and mood disorders, or delinquent behavior. If the instrument does not optimally fit the characteristics of patients, therapists, and health care, its applicability in practice is low. This implies that there is a need for multiple versions of the FfBC instrument, each adapted to the characteristics of different types of mental health care settings.

Table 1. A brief description of the 4 parts of the Fit for Blended Care instrument.

Part of the instrument	Number of items	Examples
Part 1: Practical, necessary prerequisites that need to be met to be able to start blended treatment	A total of 4 items for the patient and 4 items for the therapists to be filled out individually before starting the treatment	Items on clients' access to a computer; their internet skills; and the presence of acute, severe psychiatric or medical problems that would hinder the use of blended care
Part 2: Possible barriers that might hinder blended treatment	A total of 10 items filled out by therapist and patient together during a treatment session	Items on topics such as a client's cognitive problems or sensitivity to a psychological crisis
Part 3: Possible facilitators that can facilitate blended treatment	A total of 5 items filled out by the therapist and patient together during a treatment session	Preference for blended care because of practical reasons and a client's discipline and social support
Part 4: An overview of the previously discussed barriers and facilitators	N/A[a]	An overview of the first 3 parts to prompt therapists and clients to discuss and decide on the composition of blended treatment

[a] N/A: not applicable.

Blended Care in Forensic Mental Health Care

An adapted version of the FfBC instrument would be especially relevant for the treatment of forensic psychiatric patients. Forensic mental health care is a complex branch of mental health care. The main difference between forensic and regular mental health care is that the main goal of forensic mental health care is to prevent delinquent behavior; therefore, treatment takes place at the intersection between law and psychiatry [22]. Forensic mental health care focuses on the treatment of patients who have committed or are on the verge of committing an aggressive or sexual offense, fully or partly caused by a psychiatric disorder [23]. The use of eMental health in this unique domain appears to be very challenging, which can partly be explained by the characteristics of the forensic psychiatric patients population. Many patients have hardly had any education and have difficulties with reading or writing. Furthermore, forensic psychiatric patients have a broad range of disorders and have committed different types of offenses [24], making the current predominant one-size-fits-all approach toward eMental health interventions not very applicable [25]. In addition, as treatment is often part of a sentence and thus obligatory, many patients are not motivated to be in therapy [24], making it even harder to engage them in eHealth interventions. By integrating eHealth interventions in treatment, the quality of forensic mental health care can be improved, for example, by tailoring eHealth interventions to patient characteristics, by adding persuasive elements that can increase engagement and adherence, or by offering new ways for patients to work on their treatment [25,26]. However, as is the case in mental health care in general, successful implementation of eHealth interventions in existing treatment is considered to be a major barrier [18,26,27].

Objective

Many branches of mental health care have much to gain from successful blended treatment; however, implementation is a main barrier. To overcome this barrier, the FfBC instrument can be a useful tool. However, to ensure that the FfBC instrument fits the characteristics of a specific form of mental health care, the existing version needs to be adapted. In line with the recommendations on eHealth development, this should be done in close collaboration with end users to ensure that it fits their needs and wishes [28-31]. Consequently, the goal of this study is to adapt the existing FfBC instrument to fit forensic mental health care by means of participatory development with therapists. This will result not only in a new, ready-to-use version of the instrument for forensic mental health care but also in a blueprint for steps that need to be taken to adapt it to other types of (mental) health care.

Methods

Setting

This study took place in a Dutch organization that offers forensic mental health care to both inpatients and outpatients. The organization has 2 main outpatient clinics where approximately 85% of all patients are treated. This study took place in one of these outpatient clinics that treats approximately 50% of the organization's entire patient population. The patient population of this clinic is characterized by a relatively low education level: 46% had only primary and/or secondary education. Furthermore, patients had a broad range of diagnoses, such as personality, attention deficit, sexual, anxiety, depression, schizophrenia, and substance use disorders.

The focus of this project was to adapt the FfBC instrument to an eMental health intervention that was already used by the organization: a web-based intervention platform that contains a collection of over 200 different modules, developed by a Dutch commercial company. The platform is suitable for all types of mental health care and consequently contains modules on, among other things, mindfulness, depression, substance abuse, aggression regulation, relaxation exercises, and social skills. Each module consists of multiple sessions that are provided in a fixed order and can be accessed via a browser or mobile app. These sessions consist of a combination of elements, for example, written information about the topic, a story from a peer (in video or text), written assignments derived from cognitive behavioral therapy, and informative videos. Within the clinic, these modules are used in a blended manner. This means that therapists must first introduce the intervention platform to a patient and select an appropriate module. During usage, the patient is asked to complete assignments in the modules by themselves, on which the therapist must then provide written feedback via the platform in between in-person sessions. Log data analysis has shown that the intervention has been used for over 5 years; however, the uptake in practice is considered disappointing: only 18% of the patients started a module, whereas the goal was to use the intervention with all patients. In addition, among the patients who started, 82% did not finish the module and thus can be characterized as nonadherent. Furthermore, only half of the organization's therapists used the intervention, of which most used it only several times [18].

Study Design

Several methods have been used to adapt the existing FfBC instrument to optimally fit the treatment of forensic psychiatric outpatients. The existing version can be found in a paper by Wentzel et al [12], and a summarized version is provided in Table 1. To create an adapted version, an agile approach was applied, in which several subproducts were created, regularly evaluated with therapists, and adapted accordingly [32]. These formative evaluation cycles are in line with current recommendations on eHealth development and support developers in ensuring that the final product fits the needs and characteristics of the end users and their contexts [33]. The phases of this study and the accompanying methods used are presented in Table 2.

Table 2. An overview of the methods used to adapt the Fit for Blended Care instrument for forensic mental health care.

Research method	Main research goal
Phase 1: Requirements for adaptations	
A total of 2 focus groups with 3 therapists and 1 manager	Identifying the preferences and ideas of the therapists, managers, and researchers to determine the general layout and structure of the to-be-adapted FfBC[a] instrument
Phase 2: Identifying factors	
Semistructured interviews with all 18 therapists of 1 outpatient clinic	Identifying specific forensic psychiatric patients–related factors that influence the use of the eMental health intervention
Phase 3: Content generation	
Focus group with 3 therapists, 1 manager, and 2 researchers	Formulating the items that should be integrated in the to-be-adapted version of the instrument, based on previously identified factors
Focus group with 6 therapists, 1 manager, and 2 researchers	Formulating tips and recommendations for therapists on how to deal with different types of patient-related factors
Prototyping	Developing a functioning, interactive prototype of the adapted version of the FfBC instrument
Phase 4: Testing	
Pilot study with 5 therapists	Gaining insight into the experiences of therapists and practical feasibility of using the instrument
Focus group with 5 therapists	Identifying points of improvement for the adapted version of the FfBC instrument
Prototyping	Developing an improved version of the FfBC instrument based on the points of improvement of previous phases

[a] FfBC: Fit for Blended Care.

Throughout the entire process, a project team was actively involved. This team consisted of the researcher who led the focus groups (HK), a minimum of 2 and a maximum of 5 therapists, and the team manager. Therapists were included by the team manager based on their motivation to improve the use of eMental health. To ensure different perspectives, not all participating therapists were very positive about the intervention. All members had at least three years of experience working in mental health care, and all had used the eMental health intervention at least once. The composition of the project team changed throughout the process for various reasons: 1 member was replaced by another due to personal circumstances, and later in the process, 3 new members were added to expand the expertise and experience of the project team. In addition, not all members could join all focus group meetings due to conflicting appointments.

Materials and Procedure

Phase 1: Requirements for Adaptations

As can be seen in Table 2, the goal of the first phase was to map the requirements, that is, the needs and wishes of the involved therapists regarding adaptations to the existing version of the FfBC instrument (Table 1).

Participants

In the first phase of the process, two 1-hour focus groups were held with 1 manager and 3 therapists—of which 1 was a social worker and 2 were psychologists; all were members of the project team.

Data Collection

In the first focus group, the participants studied the existing FfBC instrument and discussed its potential usefulness to determine whether adapting it would be of added value for the organization. After agreeing on its usefulness, the second focus group focused on the therapists' needs and wishes regarding an adapted version of the instrument via a brainstorming session about required adaptations. The main discussion points centered on the content of the items, the length of the instrument, the way of filling it out, and the way in which the questions were asked.

Analysis and Product

On the basis of the notes that were taken by the researcher (HK), a document was created with the stakeholders' requirements regarding the adapted version of the FfBC, which was checked and verified by the participating therapists.

Phase 2: Identification of Factors

To create the content of the adapted version of the FfBC instrument that was specifically tailored to the use of the web-based modules in forensic mental health care, semistructured interviews with therapists were conducted to gain insight into the patient-related factors that, according to the therapists, are related to the use of the web-based modules.

Participants

To avoid self-selection bias, all 18 therapists working at the forensic outpatient clinic were included in the interview study. All therapists were expected to use the eMental health intervention when offering therapy. The included therapists had different occupations: 8 psychologists, 6 social workers, 2 system therapists, 1 trauma therapist, and 1 forensic nurse were interviewed. At the time of interviewing, they had been working in forensic care for an average of 13.18 years (SD 8.68), with a range of 8 months to 29 years.

Data Collection

After the interview's goal and content were discussed, informed consent was obtained. The entire interview consisted of 6 main categories with open-ended questions: (1) sociodemographic questions, (2) experiences with the introduction of the intervention, (3) the way in which the intervention was used with patients and/or reasons for not using the intervention, (4) the potential added value of the intervention, (5) the ideal situation regarding implementation in practice, and (6) barriers related to the use of the intervention. Consequently, patient-related factors were discussed throughout the interview.

Analysis and Product

To identify patient-related factors, an inductive, bottom-up approach was applied to analyze the transcripts. For this study, 195 fragments on patient-related factors that are related to the use of the eMental health intervention were identified. Next, an initial coding scheme was created based on these fragments using the method of constant comparison [34]. Overall, 2 researchers used the initial coding scheme to code 20% of the fragments, resulting in a joint probability agreement of 89%. No further adaptations to the underlying structure of the code scheme were required. Owing to the high interrater reliability, 1 researcher coded the remaining fragments and discussed them with the other researcher in case of doubt.

Phase 3: Content Generation

As shown in Table 3, the goal of the third phase was to combine the requirements of phase 1 and patient-related factors identified in phase 2 into an adapted version of the FfBC instrument that fit the needs and wishes of the therapists and the characteristics of forensic outpatient care. This was achieved in 2 stages, of which the first stage focused on the creation of items for the instrument and the second stage on the tips and guidelines that the instrument should offer. This resulted in the creation of a functional prototype of the FfBC instrument.

Table 3. Main codes and the number of interviews in which they were identified (Nint=18) and the total number of times that a code was identified (Ntot=195).

Main code	Interviews in which the main codes were identified, n	Total number of times the code was identified, n
Treatment motivation	14	40
Conscientiousness	14	27
Literacy levels	14	22
Perceived benefits	14	22
Psychosocial situation	13	28
Technological skills	12	17
Availability of technological resources	11	18
Reflective skills	11	21

Participants (Focus Groups Round 1)

A total of 2 researchers, 5 therapists, and 1 manager participated in the first focus group.

Data Collection (Focus Groups Round 1)

A 1-hour focus group took place to discuss which patient-related factors identified in phase 2 should be included in the adapted FfBC instrument. One researcher led the focus group by explaining the previously identified factors and asking the participants whether these factors would be suitable for inclusion in the instrument.

Analysis and Product (Focus Groups Round 1)

On the basis of the discussion, 2 researchers created a table with (1) the factors from the original instrument, (2) comparable or similar factors from the previously conducted interviews, (3) a suggested adaptation for the adapted version of the FfBC instrument, and (4) substantiation and explanation for the adaptation. Furthermore, for each factor, 3 *multiple-choice* options to indicate the extent to which a factor was estimated to be present within a patient were added. A document with the factors, a brief explanation, and the 3 options were discussed with 5 therapists and 1 manager in a new focus group and adapted accordingly.

Participants (Focus Groups Round 2)

In the second focus group, the same 5 therapists and the manager participated, and 1 researcher was present.

Data Collection (Focus Groups Round 2)

A focus group was conducted with the project team (all 5 therapists, a researcher, and a manager) and 1 additional researcher who was actively involved in developing the existing version of the FfBC instrument. In this focus group, concrete tips and guidelines on how to deal with specific patient-related factors were generated. All therapists participating in the focus group had experience using the intervention and were asked to use their own experiences to formulate the tips and recommendations. The researcher and manager also actively participated in the brainstorming session. Each factor was discussed separately, and general tips and guidelines were discussed as well.

Analysis and Product (Focus Groups Round 2)

The researcher kept extensive notes. On the basis of these notes, a document with tips per patient-related factor was created. This document was validated by the participants of the focus group, and several minor adaptations were made accordingly. This resulted in 1 document with all tips and guidelines that had to be integrated into the adapted FfBC instrument. On the basis of the previously identified items and the tips that were generated in phase 3, a working prototype of the instrument was created in Qualtrics (SAP SE), a web-based survey system. In this prototype, therapists were able to select 1 answering option per patient-related factor, resulting in tailored advice for each factor.

Phase 4: Testing

The goal of the fourth phase was to gain insight into the experiences and identify points for the improvement of the functioning prototype of the FfBC instrument by testing it in practice.

Participants

In the pilot study, 5 therapists were asked to use the prototype of the FfBC instrument with 3 patients per participating therapist, resulting in an intended number of 15 patients. A total of 4 therapists participated in the focus group.

Data Collection

In total, the instrument was used with 8 patients. In the focus group, 4 therapists were asked about their experiences with the content, usability, and integration in treatment. To ensure that all relevant topics were discussed, a semistructured approach was used in which the following topics were discussed: the way in which the instrument was used, opinion about the instrument, and recommendations for improvement.

Analysis and Product

On the basis of the outcomes of the focus group, changes were made to the initial prototype to ensure an optimal fit with the needs and wishes of therapists, which was again evaluated by the therapists. This version of the instrument will be further developed and implemented in clinical practice.

Results

In this section, the results are discussed for each of the 3 phases (Table 2) and their accompanying research methods.

Phase 1: Requirements for Adaptations

The first focus group showed that therapists saw the potential of the FfBC instrument in addressing the current implementation problems. On the one hand, it was seen as a way to offer concrete and relevant topics to discuss to identify the most optimal way to shape blended care. On the other hand, if implemented well, the instrument could be seen as a *reminder* that could help therapists in remembering to bring up the use of technology, as therapists often forgot to introduce the possibility of blended care or decided for themselves that a patient would not benefit from eMental health. However, participants indicated that the instrument needed to be adapted to better fit the forensic context and to account for several practical limitations.

The second focus group resulted in the following broad requirements for the new version of the FfBC instrument:

- The adapted version should be shorter and contain fewer texts and fewer items. Therapists found that the existing instrument contained too many items and thus would be too time consuming.
- Each item of the adapted version should be accompanied by 3 to 4 multiple-choice options. Therapists indicated that open-ended questions would require too much time.
- Each multiple-choice option should be accompanied by tailored advice and tips and tricks specific for that option. Therapists indicated that these tips and tricks could support them in initiating and continuing the use of the modules.
- The items of the existing version need to be specified to fit the treatment of forensic psychiatric outpatients. Therapists found the items in the existing version too broad and generic for use in forensic mental health care; therefore, the adapted version of the instrument should be based on patient-related factors that specifically influence the use of eHealth interventions in forensic mental health care.
- There should be a web-based version of the instrument. Therapists believed that a web-based version would be easier to fill than a paper-based version.
- Patients have to answer several questions about the necessary preconditions for using eMental health in advance by themselves. Therapists stated that this could avoid them from discussing these practicalities in treatment, which would demand valuable time. This means that therapists wanted to keep the existing distinction between the first part and the additional items of the instrument [12], where the first part should be filled out by the patient and the second part should be discussed by the therapist and patient together.

Phase 2: Identification of Factors

The patient-related factors that, according to therapists, influence the use of the eMental health intervention that arose from the interview study are presented in Table 3.

Treatment Motivation

Motivation refers to the extent to which a patient is enthusiastic or open toward working with the eMental health intervention in treatment. Although some patients were described as motivated for blended treatment, therapists indicated that a large proportion of the patients were not eager to work with the eMental health intervention. A lack of motivation was not only observed at the beginning of the blended treatment but also when the patients were using the intervention. This lack of motivation is illustrated by the following quote:

> *But I think it will be very difficult for a patient who already is not very motivated, to also encourage him to log in again, and to read things again, because there's a lot of text sometimes. And to work on assignments. It would be better to lower the threshold a bit at first.*

Conscientiousness

Conscientiousness refers to the extent to which a patient adheres to agreements regarding the independent use of the intervention outside of treatment, which was described by 1 participant as follows:

> *But you have to actually do it, you really have to get into it. And even though they can practically do it, they still have to put their mind to it. Plan a moment*

> *for it, do things, take steps. And a lot of patients don't get to that point.* [PP 3]

Therapists indicated that it often required a lot of their time and effort to ensure that patients performed the activities that they agreed on, such as working on and completing assignments. A comparison was made with doing homework, with which a large share of the patients, who often received little education, had difficulties.

Literacy Levels

Literacy levels refer to the patient's ability to write, read, and understand treatment-related information in the intervention. One therapist described this problem in the following way:

> *But you do meet people who cannot even write. I don't want to call them illiterate, but they are very ashamed of a lot of linguistic errors and things like that. That's a barrier with which you'd have to help them first, so that it isn't about the sentence construction or errors, but what's going on in their head. Just try to write that down in your own words. And people often find that difficult.* [PP 10]

In addition, therapists indicated that patients had difficulties not only with writing but also with reading, as the intervention contained several words that were perceived as difficult.

Perceived Benefits

Perceived benefits refer to the extent to which a patient experiences or expects to experience a positive influence on his or her treatment because of the use of the technology. Therapists indicated that if patients did not directly see how a module fits their problems or could be of added value for them, the chances of them using the module were lower. A therapist said the following about this:

> *There can be a lot of reasons for that. It might be that some have heard about it from others, that it's helpful. Or that some modules fit well. And also what I've said before, that it fits the needs of the patient. So if you offer a sleeping module for someone with difficulties with sleeping, there's a greater chance that he will continue.*

Psychosocial Situation

Psychosocial situations refer to difficult circumstances or events in a patient's personal life and/or mental state that influence the use of technology for treatment. This can refer to patients who are in a crisis such as a psychosis or severe depression and to those with problems related to their daily life, such as fights with neighbors or loved ones, no current place to stay, or money problems. This is further illustrated by the following quote:

> *For two patients it wasn't possible to complete the assignments. And one of them is someone of whom I think, there's just too much going on. That person has lost his job, the emotions are all over the place, and that makes it more difficult to work on a session, even though it might be beneficial.*

Overall, therapists indicated that it is important for patients to have a relatively steady life when using the intervention, because they otherwise have no *mental space* to work on the module.

Technological Skills

This code refers to the level of skills required for successful use of information and communication technologies such as computers or smartphones. Therapists indicated that several patients, especially older ones, have difficulties with using technologies. These difficulties could be with either using the actual technology, such as a computer, or navigating through the intervention itself. One therapist said the following:

> *I can definitely imagine that with young people, who already sit behind the computer a lot, it might fit a bit better. I can really imagine that.* [PP 3]

Availability of Technological Resources

This code refers to the patient's access to a technological device, an appropriate working area, and a good internet connection that is necessary to use the technology. The importance of a suitable work space is described in the following quote:

> *I think that in their own environment, where they like doing it. They have to be able to do it privately, not that there is someone around the entire time. So privacy is important for them, I think. We can't facilitate that; they have to arrange that themselves. Or we'd have to offer them a place to work here, so they can sit behind a computer here.* [PP 7]

Reflective Skills

Reflective skills refer to the patient's ability to independently reflect on and write about emotions, cognitions, and behaviors in the technology. Often, patients are not used to talking about their problems, and writing individually about these situations is often even more difficult. Therapists also indicated that reflecting individually could also lead to intense emotions and adverse consequences because of a patient's inability to independently deal with them, as explained in the following quote:

> *And also that it elicits too much emotions that they cannot directly deal with by talking to someone. Basically, you'd have to inhibit the direct gratification of your own needs. Yes, they can chat, but they do not receive an answer immediately. And some patients keep on thinking about it, running it through their head, because they do not get support directly.* [PP 16]

Phase 3: Content Generation

In the first focus group of this phase, 3 therapists, 2 researchers, and a manager decided on items that should be integrated in the instrument based on the factors identified in phase 2. In line with the requirements identified in phase 1, the number of factors that arose from the interviews needed to be reduced. To create an overview of relevant items, a table was created by 2 researchers, in which the factors of the original instrument were combined with the suggested change to the item and a substantiation of the change by means of the results of the previous phases. In Table 4, a part of this table is provided to illustrate this process. The complete table can be found in Multimedia Appendix 1.

Table 4. Examples of the table used to create items of the adapted version based on the original version of the Fit for Blended Care instrument.

Item from original version	Item for adapted version	Rationale behind change
10. *Motivation and trust* (discuss and answer): • Do you (client) trust that a blended treatment can help you with your complaints? • Are you (client) motivated to do a blended treatment?	*Motivation forweb-basedtreatment:* to what extent is the patient motivated to work on the eMental health intervention in his or her treatment	• Motivation is an important issue in forensic mental health care according to the interviews • Lack of trust in effectiveness was not an important topic in the interviews; therefore, remove it for conciseness • Rephrase item because the instrument has to be filled out by a therapist (after discussing with the patient)
15. *Working alliance*: Is there a good working alliance or do you (therapist and client) expect that a good working alliance will be developed? Note: Here it is important that you (the client) recognize your own contribution to the therapy and are aware of what is expected of you.	N/A[a]	• Remove to make the instrument more concise • Person administering this instrument might be someone other than the therapist (eg, the *intaker*) • Hard to assess in the first meetings, especially in forensic patients who are obliged to attend treatment; they might have a different attitude than later in the treatment process • Not an important topic in interviews

[a] N/A: not applicable.

A total of 2 researchers (HK and JW) combined the findings from the interviews with the factors from the existing instrument, which resulted in 5 items. The researchers also created 3 *multiple-choice* answering options from which a therapist had to choose the most fitting option. These 5 items and multiple-choice options were combined into a document with the first version of the adapted FfBC. This document was checked by 3 therapists from the project team, and slight changes in phrasing were made accordingly. The way in which the instrument should be administered was discussed in the focus group. It was decided that the first part, focusing on the practical preconditions, should be filled out by a patient individually, ideally before beginning the treatment. The second part should be filled out by the therapist based on a discussion of all 5 factors with the patient, either at the beginning or during treatment. A summarized version of the instrument is provided in Textbox 1; the text of the entire adapted version of the instrument can be found in Multimedia Appendix 2.

Textbox 1. The main topics and summary of the content and questions of the adapted Fit for Blended Care instrument.

Content of part 1: Necessary preconditions

- Reading and writing
    - Are you able to read and write short texts?
- Workstation and devices
    - Do you have access to a device (computer, laptop, smartphone, or tablet), does it have a well-functioning internet connection, and is there a place where you can work on web-based treatment in a calm and familiar manner?
- Internet skills
    - Are you able to send emails, watch videos on the web, use the internet to read short texts, use social media, and use the internet to send messages to others?

Content of part 2: Patient-related factors that can influence blended care

- Motivation for blended treatment
    - To what extent is a patient motivated to work with the eMental health intervention in his or her treatment?
- Writing about thoughts, feelings, and behavior
    - To what extent is a patient able to independently write and reflect on his or her thoughts, feelings, and behavior?
- Conscientiousness/working with discipline
    - To what extent is a patient capable of sticking to appointments on blended care for matters such as forgetfulness, concentration, or planning skills?
- Psychosocial problems
    - To what extent are there problems in the patient’s private life and/or severe psychiatric problems that can have a negative impact on using the eMental health intervention?
- Social support
    - To what extent does a patient have a social network (partner, parents, and friends) that is able to support him or her in using the eMental health intervention?

In the second focus group, the therapists, researchers, and a manager formulated tips and recommendations for therapists on how to deal with different types of patient-related factors, resulting in a document with tips for all 5 items of the second part of the instrument. Each multiple-choice answering option was accompanied by tailored advice specific to that option. On the basis of the outcomes of the focus group, a researcher (HK) created a document with the recommendations, which was emailed to the researcher, therapists, and manager. On the basis of their input, several minor changes to phrasing were made. In Textbox 2, one tip or recommendation per item is provided for illustration purposes. To prevent therapists from using the instrument as a reason for not using eMental health interventions, the members of the focus group decided that the tips should never suggest not to use eMental health. Instead, the tips should encourage therapists to think outside the box to involve difficult-to-engage patients in blended care or to delay the use of eMental health to a later point.

Textbox 2. Examples of advice or recommendations provided per item of the instrument.

General advice

- Before you start with a module, it is important to discuss with the patient what the added value of the module should be. Make sure you set clear goals that you both agree on. On the basis of that, you can regularly evaluate how the blended treatment is going.

Motivation for blended treatment

- When a patient is not motivated at all, it is important to figure out why this is the case by means of an open discussion, instead of just accepting it. It might be that a patient foresees obstacles that are actually easy to overcome.

Writing about thoughts, feelings, and behavior

- You can take away a patient's fear for writing by clearly stating that you do not expect flawless spelling or elegant phrasing but that the goal is to think about thoughts, feelings, and behaviors. You might suggest the patient to use very short sentences or terms.

Conscientiousness/working with discipline

- If the patient did not hand in an assignment although this was agreed on, you can send a reminder to ask the patient why he or she has not completed the assignment yet and if he or she is able to still finish it.

Psychosocial problems

- If a patient is experiencing a crisis such as current psychosis or suicidality, it is often not recommended to directly start with an eMental health intervention because the crisis has to be dealt with first. However, it is possible to start the intervention at a later point. It is advised to regularly evaluate with the patient to determine if blended treatment can be initiated after a while.

Social support

- If a patient indicates that one or more loved ones can actively support him or her, you can look for possibilities to actively involve those in blended treatment. A loved one might support the patient in working on assignments.

The final activity of the third phase was the development of a working prototype of the instrument to enable therapists to test it with several patients. A digital prototype was created based on the needs and wishes of the therapists that were identified in phase 1. The prototype was made in Qualtrics, as this software provides tailored advice per chosen answering option.

Phase 4: Testing

The goal of the fourth and final phase was to identify experiences with and accompanying points of improvement of the previously developed prototype. After 2 months, 5 therapists used the instrument with a total of 8 patients. This was about half of the expected 15 patients. The most important experiences and conclusions of the final focus group in which the prototype was evaluated are as follows:

- Therapists indicated that the 5 factors were useful to discuss and that the current content sufficed: no factors should be removed or added.
- The first part of the FfBC instrument was considered as useful, but therapists indicated that it was difficult to remember to ask the patients to fill it out beforehand. It was considered important to integrate the first part in existing structures, for example, in a web-based welcoming module
.
- The second part of the FfBC instrument was used several times but not as often as expected, as therapists were asked to try the second part with at least three patients. The main explanation for this was that they did not remember to administer the instrument during their treatment routines. They indicated the importance of reminders to support them in remembering the use of the FfBC instrument in treatment. Other reasons for the lower usage were not provided; it was mostly attributed to not remembering to use the instrument, and therapists expressed the intention to use it more.
- Therapists indicated that the instrument can be of added value during multiple points in treatment. For example, it can be used at the beginning of the treatment to get an idea of the type of patient and to plan blended care; however, it can also be used throughout the course of the treatment, for example, if a patient stops using a module or if the use of the module is not going as expected.
- The prototype was designed in such a way that therapists had to fill out the instrument individually, after discussing the factor with the patient. However, therapists indicated that, in practice, they preferred to fill the instrument together with the patient and expressed a need for a patient-centered version, including easier phrasing.
- In addition to a web-based version, several therapists expressed the need for a paper-based version that they could fill out together with the patient when, for example, no laptop was available in case of home visits or if they preferred not to sit behind their computer with the patient.
- To use the instrument, therapists had to use the Qualtrics prototype, which was considered inconvenient as they often could not retrieve the link, which was e-mailed to them. They indicated that it would be easier to integrate the instrument in one of the existing systems they used, among which were the platform of the eMental health intervention and the electronic patient file. Consequently, the importance of integrating the instrument in these systems was emphasized to prevent the use of the instrument as an additional time-consuming activity.

In the focus group, the project team decided that the first part of the FfBC instrument should be integrated in a to-be-developed web-based *welcoming module*, which is expected to be followed by all patients that start the treatment. The questions of the second part of the FfBC instrument have to be integrated into an existing system that has all existing questionnaires that are used in treatment to ensure that the FfBC is used in the same way as other questionnaires that are used in forensic mental health care and to ensure that they do not require any additional work. Furthermore, based on the outcomes of this focus group, a second patient-centered version of the FfBC instrument was developed; the content and a screenshot of the instrument are provided in Multimedia Appendix 3. In this additional version, the phrasing of the 5 factors is targeted at patients, that is, shorter and simpler sentences without jargon. Participants of the focus group indicated that these items can be printed on paper, for example, in the form of a poster or as 5 separate cards that can be placed on the desks of therapists, to ensure that the items are always visible, which was expected to serve as an additional reminder. Together, the patient and therapist can discuss these items during a therapy session to determine which answering option best fits the patient, instead of the therapist deciding individually on which option fits best afterward.

Discussion

Principal Findings

This study described the development of an adapted version of the existing FfBC instrument to optimally fit forensic mental health care. In phase 1, it became clear that therapists wanted a shorter, easier-to-use version of the instrument, ideally on the web, containing factors that were more specific for forensic mental health care. In phase 2, a broad range of patient-related factors were identified in a systematic interview with all therapists working at the outpatient clinic. In phase 3, these factors were translated into a functioning prototype of the instrument, using the broad requirements from phase 1. The instrument consisted of 2 parts: one to be filled out by the patient individually, targeting practical necessary preconditions, and one with 5 items that should be discussed by the patient and therapist to shape blended treatment. These factors were motivation for blended treatment; writing about thoughts, feelings, and behavior; conscientiousness/working with discipline; psychosocial problems; and social support. In phase 4, the prototype was used in practice. The adapted version of the instrument was seen as useful and promising but was not used as often as expected in the pilot study. Therapists indicated that the main reason for this was a lack of integration in existing systems and procedures, showing that a fit between the instrument and their current practices was deemed essential for its success and added value for clinical practice. On the basis of the outcomes of this final phase, a second, more patient-centered version was developed, with items that are phrased in a shorter and simpler manner.

Adapting the Instrument

Although in this study an adapted version for the use of a web-based intervention platform in forensic mental health care was created, the FfBC instrument can be adapted to fit many different types of mental health care and even for other types of health care where eHealth interventions are used, such as physiotherapy [20] or by general practitioners. This study showed that to prevent an overload of factors resulting in an impractical and time-consuming instrument, only the most important ones should be included. Patient-related factors that are most important might differ per branch of (mental) health care. For example, the interview study and focus groups showed that conscientiousness is seen as a very important topic for forensic psychiatric patients: therapists stated that they often have difficulty working independently on modules and doing their *homework* [6,18]. However, this issue might be less relevant in other domains of mental health care. To illustrate, it is known that highly educated women are most adherent to eHealth interventions [35,36], and although these types of patients are underrepresented in forensic mental health care, they are more prevalent in the treatment of, for example, anxiety or mood disorders [37]. This might imply that conscientiousness is a less relevant factor in that domain. Consequently, the version of the instrument that was developed in this study cannot be copy-pasted to be used in other settings.

To adapt the instrument to ensure that it fits a specific form of health care, the approach used in this study can be used as a guideline. Each new project should start with the generation of general requirements regarding adaptation, either using the original elaborate version of the FfBC Instrument (Multimedia Appendix 1) or the version developed in this study (Multimedia Appendices 2 and 3). In phase 2, we conducted a semistructured interview to identify the factors. Other projects can apply the same approach; however, as it is quite time consuming to conduct an entire interview study, the factors that were identified in this or other studies on eHealth usage for a specific setting might be used, as long as they are validated with therapists and possibly patients. In phase 3, the actual instrument was developed. The findings of this study can be used as the foundation. For example, other instruments can also use the distinction between part 1, which focuses on practical necessary preconditions and must be filled out by patients themselves, and part 2, which contains approximately 5 items that must be discussed in treatment. However, although the existing and adapted version of the instrument can be used, changes should always be discussed with therapists to ensure a participatory approach. Finally, before implementing the instrument in practice, it has to be pilot tested and changes should be made accordingly, as was done in phase 4 of this study. As became clear in this study, including many patients in a pilot test with a prototype that is not yet integrated in existing systems can be challenging from a practical point of view. However, in usability testing, the general guideline is that 5 to 7 tests are often enough to identify most flaws of the prototype [38]. Possibly, this reasoning can be extended to the pilot test, which might mean that testing the instrument with this number of patients might suffice to identify the most important points of improvement. However, future research should show whether this is actually the case.

In general, this study has shown that developing an adapted version of the FfBC instrument requires multiple phases that are connected by continuous formative evaluation cycles with active end-user involvement. The main challenge for the adaptation of other versions will be to identify an approach that is thorough yet not too time consuming. The guidelines and content that were generated in this study can support other researchers in setting up an efficient yet thorough development process.

Shared Decision Making

An important finding that gradually became clearer throughout the process of adapting the FfBC instrument was the importance of shared decision making in shaping blended care. In current clinical practice, the decision on whether and how to use eMental health interventions is often made in a top-down manner, with the therapist deciding the intervention that will be used, the frequency of usage, and the mode of communication about it [17]. However, therapists participating in this study clearly expressed the need for an instrument that facilitated shared decision making as much as possible. Although the initial prototype encouraged therapists to discuss the factors elaborately with patients, the therapists had to decide on the most suitable option individually, and the text of the instrument was focused on the therapist. The pilot test showed the need for an additional version to be filled out together with the patient during treatment. In this project, cards with these patient-centered items were developed; however, there are other possibilities to further support shared decision making in shaping blended care, such as digital versions using tablets or mobile phones or gamified versions, which can be developed in further research. Using the FfBC instrument is an excellent way to prevent top-down processes and fits within the models of shared decision making such as that of Elwin et al [39]. Consequently, in line with current movements such as positive health, the use of the FfBC instrument results in a more prominent role of patients in their own health and health care, which can increase their sense of ownership and self-management [40].

Future Research

This study mostly focused on the development and formative evaluation of a therapist- and client-centered version of the FfBC instrument. Although this instrument is well-substantiated and can be used in clinical practice, more research is required. First, the instrument needs to be used in clinical practice by more therapists and in more organizations to further optimize it. In line with this, it is important to note that this version of the instrument should not be seen as fixed; it should constantly be adapted based on experiences, new insights, and changes in treatment or context. Second, a necessary precondition to further optimize the instrument is that it is actually used in practice. The results of the pilot study showed that, even though therapists saw the items as valuable, it was not used as much as expected beforehand. This touches upon a larger problem related to the implementation of new innovations in clinical practice [17,41]. On the basis of the outcomes of this study, the main reason for this seems to be that the therapists simply forget about using this new instrument during their daily routines. This conclusion is in line with the work on the implementation of eHealth: although health care professionals see the added value of an intervention, they often do not use it because it is *not in their system* and does not seamlessly fit their regular activities [18,42]. However, as there might be a broad range of other reasons for possible low acceptance of the instrument, such as a negative attitude toward using eHealth in general or a lack of skills to fill out the instrument together with the patient [17,43], future research should use the instrument in a larger sample of therapists and investigate the reasons for nonacceptance.

Third, although the main goal of the instrument is to shape blended care in a fitting way, a secondary goal is to help the therapist remember to introduce blended care, as this is often overlooked in day-to-day practice [5]. By using the instrument with all patients, the uptake of eMental health in practice might improve. Further research using log data analyses can study whether the use of this instrument actually results in increased usage and whether different scores on the 5 items can be related to different ways in which a module is used. In addition, it is expected that a better fit between the patient’s needs and a blended treatment will result in better adherence to and effectiveness of the intervention, as personalized interventions can result in improved outcomes [44,45]. However, not much is known about this topic within the domain of blended care; therefore, research that aims to determine whether increased use of the instrument indeed results in higher usage of and engagement with eHealth interventions is recommended.

The instrument has the potential to not only benefit clinical practice but also add value for research on eMental health interventions. In line with the previous recommendation, items of the FfBC instrument might serve as predictors for the effectiveness of eMental health interventions. It appears to be difficult to predict whether and why users are nonadherent to an intervention and whether it is effective for an individual [46,47]. Often, sociodemographic factors are identified as predictors [47]; however, these factors do not provide much information that is useful in clinical practice as they are often fixed, for example, we cannot change someone’s age to increase the effectiveness of the intervention. However, the items of the FfBC instrument might be potential moderators for effectiveness. For example, if the level of motivation appears to be an important predictor of effectiveness, therapists might be encouraged to increase a patient’s motivation for blended care, which might increase the chances of an intervention being effective. An accompanying advantage is that the outcomes of FfBC and log data from clinical practice can be used to identify predictors. These results will have more ecological validity as most research on the predictors of effectiveness is conducted with data from randomized controlled trials, which take place in controlled settings as opposed to data from eHealth use in the *real world*. However, to achieve this, the instrument should be adapted for use as a research tool instead of a clinical tool. Among other things, the items should be accompanied by consistent and validated scoring options, and a study on its reliability and validity as a research tool is needed.

Strengths and Limitations

The main strength of this study is its iterative nature with multiple formative evaluation cycles. Applying such a bottom-up approach where products are created based on collected data and evaluated with end users results in a final product that is ecologically valid and closely fits the requirements from practice

[32,48]. Throughout this process, the main focus of data collection was on therapists as they were the main target group of the instrument. However, the patients were not actively involved. This might have resulted in a bias, overlooking factors that were important for the use of eMental health according to patients. Nevertheless, as most involved therapists had much experience with using eMental health with patients and had discussed reasons for nonusage with their patients, the chance of missing important factors is considered fairly low. In addition, comparable factors have been identified in other studies in which patients and other stakeholders are involved [6,49]. However, it is recommended to actively involve the patient perspective in following research to verify whether the identified items are in line with their experiences as well. Furthermore, data collection took place at one outpatient clinic of one organization. Although this was a deliberate decision to ensure that the adapted version of the instrument seamlessly fitted this organization, the specific focus raises questions about the generalizability of the results. Although the involved therapists had much experience in forensic mental health care, it is still advised to pilot test the instrument and its implementation in other forensic organizations to ensure that it also fits their needs and way of delivering blended care. Finally, another limitation regarding generalizability is related to the country in which this study took place. Although many similarities exist between Dutch mental health care and that of other countries, there are also many differences. This implies that this version of the instrument cannot be copy-pasted into the forensic mental health care of other countries. Therefore, we stress that the instrument should always be adapted to fit specific settings, and this applies to health care in other countries as well. The guidelines developed for adapting the instrument can be used for this purpose.

Conclusions

This study showed that the iterative, participatory development of an FfBC instrument resulted in an adapted version that fits the context by incorporating the needs and wishes of therapists and patient-related factors that are relevant for the use of web-based interventions in forensic mental health care. This instrument can further support shared decision making in blended care, as this is an important yet often overlooked topic. The instrument's adaptability is important: its content, design, and implementation in existing care should fit the specific type of health care, organization, and eHealth intervention for which it is used; it is not a one-size-fits-all tool. To adapt this instrument to other contexts, the guidelines described in this paper can be used. By using such approaches to better integrate in-person care and eHealth interventions, we can combine the best of both worlds and increase the quality of care.

Acknowledgments

Funding for this study was provided by Stichting Vrienden van Oldenkotte. The authors would like to thank Floor Sieverink, who helped with analyzing the interviews, and all members of the project team who were involved in this study.

Conflicts of Interest

None declared.

Multimedia Appendix 1
The table used to create items of the adapted version based on the original version of the Fit for Blended Care instrument, interviews and focus groups.
[DOCX File , 38 KB - mental_v7i11e24245_app1.docx]

Multimedia Appendix 2
Fit for Blended Care instrument – Therapist-centered version.
[DOCX File , 19 KB - mental_v7i11e24245_app2.docx]

Multimedia Appendix 3
Fit for Blended Care instrument – Patient-centered version.
[DOCX File , 163 KB - mental_v7i11e24245_app3.docx]

References

1. Gilburt H. Mental Health Under Pressure. London, UK: The King's Fund; 2015.
2. Truth A. No Health Without a Workforce. Geneva, Switzerland: World Health Organisation; 2013.
3. Krausz R, Ramsey D, Wetterlin F, Tabiova K, Thapliyal A. Accessible and Cost-Effective Mental Health Care Using E-Mental Health (EMH). New York, USA: Springer; 2019.
4. Schueller S. Mental health and eHealth technology. In: Ehealth Research, Theory and Development: an Multi-disciplinary Approach. London, UK: Routledge; 2018.
5. Hilty DM, Chan S, Hwang T, Wong A, Bauer AM. Advances in mobile mental health: opportunities and implications for the spectrum of e-mental health services. Focus (Am Psychiatr Publ) 2018 Jul;16(3):314-327 [FREE Full text] [doi: 10.1176/appi.focus.16301] [Medline: 32015712]

6. Kip H, Oberschmidt K, Bierbooms J, Dijkslag D, Kelders S, Roelofsen B. Technologie in De Forensische Zorg-crossing Borders. Utrecht: KFZ; 2019.
7. Andersson G, Cuijpers P. Internet-based and other computerized psychological treatments for adult depression: a meta-analysis. Cogn Behav Ther 2009;38(4):196-205. [doi: 10.1080/16506070903318960] [Medline: 20183695]
8. Cuijpers P, Donker T, van Straten A, Li J, Andersson G. Is guided self-help as effective as face-to-face psychotherapy for depression and anxiety disorders? A systematic review and meta-analysis of comparative outcome studies. Psychol Med 2010 Dec;40(12):1943-1957. [doi: 10.1017/S0033291710000772] [Medline: 20406528]
9. Carlbring P, Andersson G, Cuijpers P, Riper H, Hedman-Lagerlöf E. Internet-based vs. face-to-face cognitive behavior therapy for psychiatric and somatic disorders: an updated systematic review and meta-analysis. Cogn Behav Ther 2018 Jan;47(1):1-18. [doi: 10.1080/16506073.2017.1401115] [Medline: 29215315]
10. Andersson G, Cuijpers P, Carlbring P, Riper H, Hedman E. Guided Internet-based vs. face-to-face cognitive behavior therapy for psychiatric and somatic disorders: a systematic review and meta-analysis. World Psychiatry 2014 Oct;13(3):288-295 [FREE Full text] [doi: 10.1002/wps.20151] [Medline: 25273302]
11. Fairburn CG, Patel V. The impact of digital technology on psychological treatments and their dissemination. Behav Res Ther 2017 Jan;88:19-25 [FREE Full text] [doi: 10.1016/j.brat.2016.08.012] [Medline: 28110672]
12. Wentzel J, van der Vaart R, Bohlmeijer ET, van Gemert-Pijnen JE. Mixing online and face-to-face therapy: how to benefit from blended care in mental health care. JMIR Ment Health 2016 Feb 9;3(1):e9 [FREE Full text] [doi: 10.2196/mental.4534] [Medline: 26860537]
13. Kenter RM, van de Ven PM, Cuijpers P, Koole G, Niamat S, Gerrits RS, et al. Costs and effects of Internet cognitive behavioral treatment blended with face-to-face treatment: Results from a naturalistic study. Internet Interv 2015 Mar;2(1):77-83. [doi: 10.1016/j.invent.2015.01.001]
14. van der Vaart R, Witting M, Riper H, Kooistra L, Bohlmeijer ET, van Gemert-Pijnen LJ. Blending online therapy into regular face-to-face therapy for depression: content, ratio and preconditions according to patients and therapists using a Delphi study. BMC Psychiatry 2014 Dec 14;14:355 [FREE Full text] [doi: 10.1186/s12888-014-0355-z] [Medline: 25496393]
15. Schueller SM, Tomasino KN, Mohr DC. Integrating human support into behavioral intervention technologies: the efficiency model of support. Clin Psychol Sci Pract 2016 Nov 17;24(1):27-45. [doi: 10.1111/cpsp.12173]
16. Wozney L, Newton AS, Gehring ND, Bennett K, Huguet A, Hartling L, et al. Implementation of eMental health care: viewpoints from key informants from organizations and agencies with eHealth mandates. BMC Med Inform Decis Mak 2017 Jun 2;17(1):78 [FREE Full text] [doi: 10.1186/s12911-017-0474-9] [Medline: 28577543]
17. Feijt MA, de Kort YA, Bongers IM, IJsselsteijn WA. Perceived drivers and barriers to the adoption of emental health by psychologists: the construction of the levels of adoption of emental health model. J Med Internet Res 2018 Apr 24;20(4):e153 [FREE Full text] [doi: 10.2196/jmir.9485] [Medline: 29691215]
18. Kip H, Sieverink F, van Gemert-Pijnen LJ, Bouman YH, Kelders SM. Integrating people, context, and technology in the implementation of a web-based intervention in forensic mental health care: mixed-methods study. J Med Internet Res 2020 May 26;22(5):e16906. [doi: 10.2196/16906] [Medline: 32348285]
19. Robertson L, Smith M, Castle D, Tannenbaum D. Using the internet to enhance the treatment of depression. Australas Psychiatry 2006 Dec;14(4):413-417. [doi: 10.1080/j.1440-1665.2006.02315.x] [Medline: 17116083]
20. Kloek CJ, Janssen J, Veenhof C. Development of a checklist to assist physiotherapists in determination of patients' suitability for a blended treatment. Telemed J E Health 2020 Aug;26(8):1051-1065. [doi: 10.1089/tmj.2019.0143] [Medline: 31804904]
21. Titzler I, Saruhanjan K, Berking M, Riper H, Ebert DD. Barriers and facilitators for the implementation of blended psychotherapy for depression: a qualitative pilot study of therapists' perspective. Internet Interv 2018 Jun;12:150-164 [FREE Full text] [doi: 10.1016/j.invent.2018.01.002] [Medline: 30135779]
22. Arboleda-Flórez J. Forensic psychiatry: contemporary scope, challenges and controversies. World Psychiatry 2006 Jun;5(2):87-91 [FREE Full text] [Medline: 16946941]
23. Mullen PE. Forensic mental health. Br J Psychiatry 2000 Apr;176(4):307-311. [doi: 10.1192/bjp.176.4.307] [Medline: 10827876]
24. Drieschner KH, Boomsma A. The treatment motivation scales for forensic outpatient treatment (TMS-F): construction and psychometric evaluation. Assessment 2008 Jun;15(2):224-241. [doi: 10.1177/1073191107311650] [Medline: 18250180]
25. Kip H, Bouman YH, Kelders SM, van Gemert-Pijnen LJ. Ehealth in treatment of offenders in forensic mental health: a review of the current state. Front Psychiatry 2018;9:42 [FREE Full text] [doi: 10.3389/fpsyt.2018.00042] [Medline: 29515468]
26. Kip H, Oberschmidt K, Bierbooms J, Dijkslag D, Kelders S, Roelofsen B. Technologie in De Forensische Zorg? Crossing Borders. Utrecht: Kwaliteit Forensische Zorg; 2019.
27. Kip H, Oberschmidt K, Bierbooms JJ. Ehealth technology in forensic mental healthcare: recommendations for achieving benefits and overcoming barriers. Int J Forensic Ment Health 2020 Aug 24:- epub ahead of print. [doi: 10.1080/14999013.2020.1808914]
28. van Gemert-Pijnen JE, Kelders S, Kip H, Sanderman R. Ehealth Research, Theory and Development: a Multi-disciplinary Approach. London, UK: Routledge; 2018.

29. van de Belt TH, Engelen LJ, Berben SA, Schoonhoven L. Definition of Health 2.0 and Medicine 2.0: a systematic review. J Med Internet Res 2010 Jun 11;12(2):e18 [FREE Full text] [doi: 10.2196/jmir.1350] [Medline: 20542857]
30. Yardley L, Morrison L, Bradbury K, Muller I. The person-based approach to intervention development: application to digital health-related behavior change interventions. J Med Internet Res 2015 Jan 30;17(1):e30 [FREE Full text] [doi: 10.2196/jmir.4055] [Medline: 25639757]
31. Mohr DC, Lyon AR, Lattie EG, Reddy M, Schueller SM. Accelerating digital mental health research from early design and creation to successful implementation and sustainment. J Med Internet Res 2017 May 10;19(5):e153 [FREE Full text] [doi: 10.2196/jmir.7725] [Medline: 28490417]
32. Hekler EB, Klasnja P, Riley WT, Buman MP, Huberty J, Rivera DE, et al. Agile science: creating useful products for behavior change in the real world. Transl Behav Med 2016 Jun;6(2):317-328. [doi: 10.1007/s13142-016-0395-7] [Medline: 27357001]
33. van Gemert-Pijnen JE, Kip H, Kelders SM, Sanderman R. Introducing ehealth. In: Ehealth Research, Theory and Development: A Multi-disciplinary Approach. London, UK: Routledge; 2018.
34. Boeije H. A purposeful approach to the constant comparative method in the analysis of qualitative interviews. Qual quant 2002;36(4):391-409. [doi: 10.4135/9781849209403.n920]
35. Kontos E, Blake KD, Chou WS, Prestin A. Predictors of ehealth usage: insights on the digital divide from the health information national trends survey 2012. J Med Internet Res 2014 Jul 16;16(7):e172. [doi: 10.2196/jmir.3117] [Medline: 25048379]
36. Hansson Vikström N, Wasteson E, Lindam A, Samuelsson E. Anxiety and depression in women with urinary incontinence using e-health. Int Urogynecol J 2020 Feb 24:- epub ahead of print. [doi: 10.1007/s00192-020-04227-2] [Medline: 32095954]
37. Ruiter GD. Inzicht in De Geestelijke Gezondheidszorg. Enschede: Vektis; 2017. URL: https://www.vektis.nl/uploads/Publicaties/Zorgthermometer/Zorgthermometer%20Inzicht%20in%20de%20ggz%202017.pdf
38. Virzi RA. Refining the test phase of usability evaluation: how many subjects is enough? Hum Factors 2016 Nov 23;34(4):457-468. [doi: 10.1177/001872089203400407]
39. 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] [Medline: 22618581]
40. Huber M, van Vliet M, Giezenberg M, Winkens B, Heerkens Y, Dagnelie PC, et al. Towards a 'patient-centred' operationalisation of the new dynamic concept of health: a mixed methods study. BMJ Open 2016 Jan 12;6(1):e010091 [FREE Full text] [doi: 10.1136/bmjopen-2015-010091] [Medline: 26758267]
41. Greenhalgh T, Wherton J, Papoutsi C, Lynch J, Hughes G, A'Court C, et al. Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. J Med Internet Res 2017 Nov 1;19(11):e367 [FREE Full text] [doi: 10.2196/jmir.8775] [Medline: 29092808]
42. Wentzel J, van Drie-Pierik R, Nijdam L, Geesing J, Sanderman R, van Gemert-Pijnen JE. Antibiotic information application offers nurses quick support. Am J Infect Control 2016 Jun 1;44(6):677-684. [doi: 10.1016/j.ajic.2015.12.038] [Medline: 26905792]
43. Kip H, Kelders SM, Weerink K, Kuiper A, Brüninghoff I, Bouman YH, et al. Identifying the added value of virtual reality for treatment in forensic mental health: a scenario-based, qualitative approach. Front Psychol 2019;10:406 [FREE Full text] [doi: 10.3389/fpsyg.2019.00406] [Medline: 30873093]
44. Andrews G, Williams AD. Internet psychotherapy and the future of personalized treatment. Depress Anxiety 2014 Nov;31(11):912-915. [doi: 10.1002/da.22302] [Medline: 25407580]
45. Musiat P, Hoffmann L, Schmidt U. Personalised computerised feedback in e-mental health. J Ment Health 2012 Aug;21(4):346-354. [doi: 10.3109/09638237.2011.648347] [Medline: 22315961]
46. Cuijpers P, Reynolds CF, Donker T, Li J, Andersson G, Beekman A. Personalized treatment of adult depression: medication, psychotherapy, or both? A systematic review. Depress Anxiety 2012 Oct;29(10):855-864. [doi: 10.1002/da.21985] [Medline: 22815247]
47. Karyotaki E, Kleiboer A, Smit F, Turner DT, Pastor AM, Andersson G, et al. Predictors of treatment dropout in self-guided web-based interventions for depression: an 'individual patient data' meta-analysis. Psychol Med 2015 Oct;45(13):2717-2726. [doi: 10.1017/S0033291715000665] [Medline: 25881626]
48. Patrick K, Hekler E, Estrin D, Mohr D, Riper H, Crane D, et al. The pace of technologic change: implications for digital health behavior intervention research. Am J Prev Med 2016 Nov;51(5):816-824. [doi: 10.1016/j.amepre.2016.05.001] [Medline: 27745681]
49. Bierbooms J, Bouman Y, Dijkslag D, Kimpen R, Muller J, Wieske R. Do's en Don'ts Van E-health in De Forensische Ggz. Utrecht: Kwaliteit Forensische Zorg (KFZ); 2015.

Abbreviations

FfBC: Fit for Blended Care

Edited by J Torous; submitted 10.09.20; peer-reviewed by M Feijt, J Torous; comments to author 22.09.20; revised version received 05.10.20; accepted 05.10.20; published 13.11.20.

Please cite as:
Kip H, Wentzel J, Kelders SM
Shaping Blended Care: Adapting an Instrument to Support Therapists in Using eMental Health
JMIR Ment Health 2020;7(11):e24245
URL: http://mental.jmir.org/2020/11/e24245/
doi:10.2196/24245
PMID:33185559

Original Paper

Virtual Reality Behavioral Activation as an Intervention for Major Depressive Disorder: Case Report

Margot Paul[1], MS; Kim Bullock[2], MD; Jeremy Bailenson[3], PhD

[1]PGSP-Stanford PsyD Consortium, Palo Alto, CA, United States

[2]Department of Psychiatry & Behavioral Sciences, Stanford School of Medicine, Stanford, CA, United States

[3]Department of Communication, Stanford University, Stanford, CA, United States

Corresponding Author:
Margot Paul, MS
PGSP-Stanford PsyD Consortium
1791 Arastradero Rd
Palo Alto, CA, 94304
United States
Phone: 1 781 572 4136
Email: mdpaul@stanford.edu

Abstract

Background: Major depressive disorder (MDD) is a global problem with an increasing incidence and prevalence. There has additionally been an increase in depression due to the COVID-19 pandemic. Behavioral activation is considered an evidence-based treatment for MDD. However, there are many barriers that could hinder one's ability to engage in behavioral activation, with COVID-19 "shelter-in-place" and social distancing orders being current and large impediments. Virtual reality (VR) has been successfully used to help treat a variety of mental health conditions, but it has not yet been used as a method of administering behavioral activation to a clinical population. Using VR to engage in behavioral activation could eliminate barriers that pandemic precautions place and help decrease symptoms of depression that are especially exacerbated in these times.

Objective: The following case report examines the feasibility, acceptability, and tolerability of VR behavioral activation for an adult with MDD during a global pandemic. This participant was part of a larger pilot study, and the case serves as a description of the VR intervention.

Methods: The participant engaged in a weekly 50-minute psychotherapy Zoom session for 4 weeks, in which a modified behavioral activation protocol was administered using a VR headset to simulate activities. Data on mood ratings, homework compliance, and headset use were obtained from the headset. Acceptability, tolerability, and depression symptoms were obtained using self-report rating scales.

Results: The intervention was feasible, acceptable, and tolerable, as reported by this participant. The participant's depressive symptoms decreased by five-points on the Patient Health Questionnaire-9 over a month, with a beginning score of 10 (moderate depression) and a final score of 5 (mild depression).

Conclusions: The implications of these findings for future research are discussed.

Trial Registration: ClinicalTrials.gov NCT04268316; http://clinicaltrials.gov/ct2/show/NCT04268316

(JMIR Ment Health 2020;7(11):e24331) doi:10.2196/24331

KEYWORDS

virtual reality; case report; major depressive disorder; behavioral activation; VR; depression; COVID-19; behavior; intervention; feasibility; acceptability; telehealth; pilot study

Introduction

Depression, classified as major depressive disorder (MDD) by the Diagnostic and Statistical Manual of Mental Disorders, 5th Edition (DSM-5), is a global problem with an increasing incidence and prevalence [1]. MDD is characterized by the experience of at least five of the following nine symptoms nearly every day during the same 2 week period (with at least one of the symptoms being either depressed mood or loss of interest or pleasure): depressed mood; diminished interest or pleasure in activities; significant weight loss when not dieting, weight gain, or a change in appetite; insomnia or hypersomnia;

psychomotor agitation or retardation; fatigue; feelings of worthlessness or excessive and inappropriate guilt; difficulty concentrating; and suicidal thoughts [2]. As such, MDD is heterogeneous, given that two people can hold the diagnosis with only one overlapping symptom. However, regardless of an individual's symptom presentation, those who meet diagnostic criteria experience significant distress or impairment in areas of functioning [2].

More than 322 million people worldwide experience symptoms of depression and about 788,000 people die each year from suicide, with suicide being the second global leading cause of death for people aged 15-29 years [1]. COVID-19 has further led to an increased risk for people developing depression worldwide, due to containment measures such as confinement to one's home with "shelter-in-place" and community shutdown orders lasting for months [3,4]. As a result of these containment measures and their subsequent negative consequences on individuals, such as social isolation and increased rates of unemployment, there have been estimates that potentially 50,000 more individuals could commit suicide worldwide [5].

Depressive disorders are worldwide the "single largest contributor to non-fatal health loss" and are among the leading drivers for years lived with disability [1]. This latter metric accounts for the number of people who are affected by the disorder as well as the "severity and disability associated with their symptoms" [6]. Specifically, MDD is estimated to be the 11th leading cause of disability and mortality worldwide and the second leading cause of disability among all disease and injury in the United States [7,8]. Data analyzed from 36,309 US adults between 2012 and 2013 found that the lifetime prevalence of MDD was 20.6%, while a 1-year prevalence was 10.4% [9]. This means that, over a person's lifetime, there is more than a one in five chance of having a major depressive episode. Among adults with MDD, about 64% had a severe impairment [10].

Given the severe impact that depressive symptoms have on individuals and society, it is imperative to identify effective treatment options. Many evidence-based treatments have been identified for MDD, behavioral activation (BA) being one of them [11]. BA is defined "as a structured, brief psychotherapeutic approach that aims to (a) increase engagement in adaptive activities (which often are those associated with the experience of pleasure or mastery), (b) decrease engagement in activities that maintain depression or increase risk for depression, and (c) solve problems that limit access to reward or that maintain or increase aversive control" [12]. This is in direct response to the behavioral theory of depression, which states that a dearth of response-contingent positive reinforcement catalyzes symptoms of depression due to less frequent engagement in pleasant activities or behavioral avoidance [12]. Thus, by helping people who have depression to become behaviorally activated through engaging in activities that they find pleasurable or that lead to a sense of accomplishment and mastery, they are able to regain the lost positive reinforcement and improve mood symptoms.

Despite BA's effectiveness and ease of dissemination and use in primary care settings, obstacles exist to its implementation. First, finding the ongoing internal motivation to become behaviorally activated is not an easy task for people who struggle with depression, due to the nature of the symptoms themselves [12]. There may also be external obstacles that prevent engagement in pleasant activities, such as finances and mobility. For example, an individual may find pleasure in travel or an adrenaline-filled activity, which may be too costly to engage in or not available due to safety concerns from physical conditions. Another person may be unable to engage in activities they previously enjoyed, such as hiking or visiting distant places, due to mobility constraints, lack of social connections, and community or pandemic restrictions. As previously mentioned, the COVID-19 outbreak led to widespread confinement to one's home with "shelter-in-place" and community shutdown orders lasting for months, preventing individuals from partaking in the activities they used to enjoy. Thus, it is vital to consider alternative treatment methods that patients may access and more easily engage in, especially for those that may be unable to receive in-person treatment.

The use of technology as an adjunct to or a method of delivering mental health treatments is becoming increasingly popular as a way to fill this treatment gap [13]. One technology medium, virtual reality (VR), is defined as a "computer-generated simulation, such as a set of images and sounds that represents a real place or situation, that can be interacted with, in a seemingly real or physical way by a person using special electronic equipment. It can transmit visual, auditory, and various sensations to users through a headset to make them feel as if they are in a virtual or imagined environment" [13]. VR has been successfully used to help treat a variety of mental health conditions, and the use of VR could help eliminate many of the aforementioned barriers to care due to a sense of presence that can match real-world activities [13-15]. Unlike engaging in real-world activities, VR is readily accessible and can consistently be used, making it a potentially beneficial therapeutic modality when other activities are barred.

Although there is minimal risk when using a VR headset, studies have indicated that the side effects may include cybersickness, often comprising three subscales: nausea (N), oculomotor (O), and disorientation (D). N includes increased salivation, sweating, nausea, upset stomach, or burping; O includes fatigue, headache, eyestrain, or difficulty focusing; D includes vertigo, dizziness, and blurred vision [16-18]. The cause of cybersickness is largely unknown, but there are many theories. One of the predominant hypotheses is that cybersickness is due to a mismatch between visual and vestibular cues [16,18]. In other words, the person using the VR headset is perceiving movement without feeling the movement or doing so themselves, causing feelings of sickness [16,18]. Research has illustrated that the rates of cybersickness increase with time wearing the headset [16]. Thus, although the exact cause of cybersickness remains unknown and can be unpleasant, there are precautions that can be taken to minimize the risk.

This case represents results from a single subject who was part of a larger pilot study currently being explored to test the feasibility, acceptability, and tolerability of using a VR headset as a way to administer BA during the COVID-19 pandemic. If pleasant activities can be successfully simulated and found effective using a VR headset, this would eliminate many obstacles to receiving care or engaging in pleasant activities such as cost-related impediments, ability-related obstacles, or other access-related difficulties such as the COVID-19 “shelter-in-place” orders. Given the plethora of VR options readily available online for free and the cheaper headset selections, VR is now more publicly accessible than in previous years [19]. If using VR to simulate pleasurable activities decreases the symptoms of depression, it could potentially provide relief for many people who would otherwise not be able to engage in such activities. This case hypothesized that VR BA would be an acceptable, feasible, and tolerable method of delivering a BA intervention for an individual diagnosed with MDD, and there would be a decrease in symptoms of depression after using the headset.

Methods

Materials and Apparatus

A VR headset supplied by Limbix, now partnered with BehaVR, was used. This headset had a 5.5-inch screen size with 2560 x 1440-pixel resolution, a screen aspect ratio of 16:9, a 92° field of view, 3 degrees of freedom, and a refresh rate of 70 Hz [20] (see Figure 1). This headset was chosen due to the fact that it was a wireless system with preprogrammed content that did not require the use of a phone or computer. The headset was easily turned on and off, and used the motion of the participant’s head to pinpoint the desired content with a visually simulated white circle that could then be clicked with a side button by one’s finger. The ease of use was imperative, due to the participant’s need to have the headset at home for engaging in activities between sessions. The Limbix devices were returned to the protocol director after study completion.

The immersive 360° videos were chosen from 360° videos already accessible on YouTube (see Multimedia Appendix 1). A total of 37 videos were selected based on activities from the Pleasant Events Schedule [21] and the subsequent quality of the available image (at least 4K resolution). These videos were uploaded onto the headset and each video was sorted into at least one of the five categories: animals; sports, dance, or arts; adrenaline; travel; and hiking or outdoors (see Multimedia Appendix 2). These categories were chosen to provide participants a diversity of options that most align with their values and interests. For example, a person experiencing symptoms of depression who previously enjoyed travel could explore the beauty of the Maldives (see Figure 2) or the majesty of the Eifel Tower in Paris (see Figure 3). Another participant may enjoy nature and could choose to experience swimming with dolphins (see Figure 4) or visiting a waterfall in Venezuela (see Figure 5). The videos ranged in length from 1 minute and 2 seconds to 10 minutes, to minimize the risk of cybersickness [16].

Figure 1. Virtual reality headset.

Figure 2. Screenshot of Visit Maldives.

Figure 3. Screenshot of Paris.

Figure 4. Screenshot of Swim with Dolphins.

Figure 5. Screenshot of Angel Falls.

Measures

Demographics

The participant was asked to confirm his name and date of birth during the intake assessment screening. During the intake visit, the participant completed a written demographic form over Zoom (Zoom Video Communications, Inc), which asked him to identify information about gender identity, racial identity, mental health treatment history, VR use history, and history of epilepsy and treatment (see Multimedia Appendix 3).

The Mini International Neuropsychiatric Interview

The Mini International Neuropsychiatric Interview (MINI) was verbally administered during the Zoom intake and was used to evaluate inclusion and exclusion criteria. It is a short diagnostic structured interview that assesses the 17 most common psychiatric disorders in the DSM-5. In a study comparing the validity and reliability of the MINI to the Structured Clinical Interview for DSM, it was found that the MINI had good reliability and validity, and took half the amount of time [22].

Mood

The participant's mood was primarily measured using the Patient Health Questionnaire-9 (PHQ-9) at four timepoints and was verbally administered by the protocol director. The PHQ-9 is a nine-item self-report that measures an individual's depression symptoms by mapping onto the nine criteria in the DSM-5 and rating them from 0 (not at all) to 3 (nearly every day), with a score range of 0-29, with 29 indicating the most severe depression and 0 indicating a lack of depression. A PHQ-9 score of 10 or higher indicates the presence of MDD with a sensitivity of 88% and a specificity of 88% [23]. The PHQ-9 has an excellent construct validity and a high internal reliability (Cronbach alpha .89) [23].

Presence

The intensity of presence felt in the virtual environment was adapted from the five-question telepresence scale outlined by Nowak and Biocca [24], which has an alpha of .88. The adapted scale used in this study incorporated three questions with five response choices ranging from "Not at all" to "Very Strongly." Question 3 was modified from Nowak and Biocca's [24] "To what extent did you feel immersed in the environment you saw/heard?" to "How much did it feel as if you visited another place?" to clarify the wording and make it more distinct from question 2. These questions were completed on a worksheet by the participant after each VR activity and securely emailed to the protocol director before each session.

Tolerability

Physical tolerability was measured using the Simulator Sickness Questionnaire (SSQ) [25]. This questionnaire includes 16 symptoms that each load differently onto the three clusters of sickness: oculomotor (O; eyestrain, difficulty focusing, blurred vision, headache; Cronbach alpha .91), disorientation (D; dizziness, vertigo; Cronbach alpha .88), and nausea (N; nausea, stomach awareness, increased salivation, burping; Cronbach alpha .84) [26]. Each symptom has a rating choice of "No more than usual" (0), "Slightly more than usual" (1), "Moderately more than usual" (2), or "Severely more than usual" (3). Emotional tolerability was measured by using the Brief Agitation Measure. This measure consists of three items with each item scored from "Strongly Disagree" (1) to "Strongly Agree" (7). This measure has a high internal consistency with a coefficient alpha of .91 [27]. These questions were completed on a worksheet by the participant after each VR activity and securely emailed to the protocol director before each session.

Acceptability

Acceptability was assessed using an adapted version of the technology acceptance model (TAM), a valid and reliable (Cronbach alpha ranging from .73 to .94) measure [28]. The TAM used in this study encompassed 13 questions with the subcategories of "Perceived Usefulness," "Perceived Ease of Use," "Attitudes Toward Use," and "Intention to Use Technology." Participants were given the option of circling one of five choices ranging from "Strongly Disagree" to "Strongly Agree." These questions were completed on a worksheet by the participant after each VR activity and securely emailed to the protocol director before each session.

Protocol

Recruitment

This case report is part of a current study that aims to continue recruitment until either 30 participants are enrolled or the timeline of January 15, 2021, whichever comes first. Participants were recruited from a study flyer posted at Stanford School of Medicine, Department of Psychiatry & Behavioral Sciences located at 401 Quarry, Palo Alto, CA. The flyer and description of the study were also listed on the Department of Psychiatry & Behavioral Sciences at Stanford University School of Medicine's currently recruiting studies website and ClinicalTrials.gov. Individuals calling into Stanford Psychiatry's intake team were also informed about the study and given the protocol director's contact information, if interested. Curify, a health-technology startup, also assisted in recruitment by advertising the study on Facebook.

Screening (Part 1)

The participant who contacted the protocol director expressing interest in the study was scheduled for an initial phone screen, where he was briefly assessed for initial eligibility and provided with the opportunity to ask questions about the study. Initial eligibility was determined by a PHQ-8 score of 10 or greater [23] as well as a brief questionnaire that was designed to be the first preliminary screener for inclusion criteria (see Multimedia Appendix 4). After the participant met initial eligibility and stated that he was still interested in participation, a formal intake was scheduled via Zoom, due to COVID-19 restrictions, and he was securely emailed the consent form to review prior to meeting.

Screening (Part 2)

After reviewing any questions and securely emailing the signed consent form back to the protocol director, the intake session occurred. During the intake session, the participant was asked to verbally complete the demographic questionnaire while the protocol director shared her screen via Zoom. The participant was subsequently administered the MINI by the protocol director. The participant was then informed of his eligibility and was scheduled for his first session via Zoom.

Randomization

Before the first session, this participant was randomized into one of the three study arms using five opaque, concealed envelopes in sets of six to preserve balance throughout the study. The participant had a one in three chance of being randomized into each group.

Intervention: VR BA Arm

The VR BA study arm followed the protocol for brief BA based on the guidance of Funderburk et al [29] and Jacobson et al [30]. That is, the treatment incorporated the four component parts that Jacobson et al [30] outlined: establishing the therapeutic relationship, developing goals for treatment, conducting a functional analysis, and treatment review with relapse prevention. Funderburk et al's [29] brief treatment protocol assisted in outlining the flow of information per session, with the first session focusing on establishing rapport, identifying activities that the participant valued or felt a sense of mastery or pleasure in from the past, introducing the activity log, and setting activity goals. The second session focused on reviewing homework and the connection between mood and activities, addressing barriers and problem-solving, and scheduling new activity goals. The third and fourth sessions similarly reviewed materials, addressed barriers to completing goals and problem-solving these barriers, and created new activity goals.

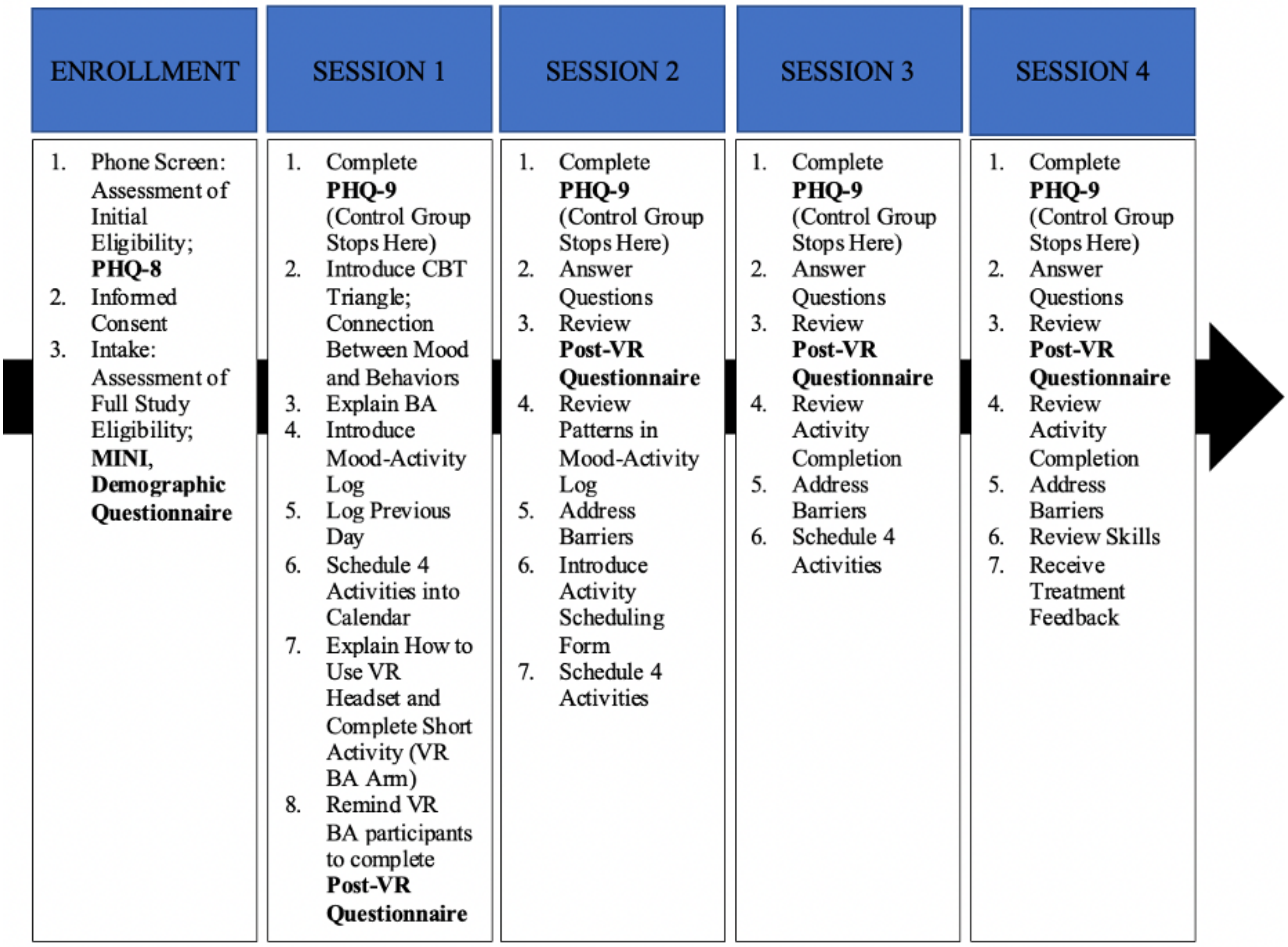
Specifically, the participant in the VR BA arm of the study met with the protocol director once per week for 4 weeks over Zoom for 50 minutes to receive BA therapy. He was securely emailed the mood activity log, the VR list of activities, and the post-VR questionnaire prior to the first session. The VR headset was shipped to his address prior to the first session as well.

During the first session, psychoeducation around the connection between thoughts, behaviors, and feelings was discussed, and the cognitive behavioral therapy triangle was shown via screen sharing. The participant was then introduced to the idea of BA and briefly explained the theory behind the evidence outlined by Lewinsohn [31]. The protocol director explained the difference between pleasure and mastery activities, and mapped out the participant's previous day, hour-by-hour, with him to determine how often he engaged in pleasurable or mastery activities. The participant was then asked to use the mood activity log and schedule, in session, at least four VR "activities" that he may enjoy into his upcoming week, as well as complete the log (see Multimedia Appendix 5).

The participant was then shown how to use the headset and asked to complete a short activity in VR during the session to ensure proper use of the headset. He was informed that he could move his head and body, but he should remain seated for his safety. The participant was asked to complete the post-VR questionnaire assessing spatial presence, simulator sickness, agitation, and acceptability every time he finished an activity in VR (see Multimedia Appendix 6). Additionally, the headset prompted him to rate his mood on a scale of 1-10 (1=worst ever felt; 10=best ever felt) before and after each activity. Barriers were anticipated and problem-solving strategies were discussed.

During session two, the protocol director reviewed the mood activity log with the participant and checked-in regarding goal attainment. Barriers to completion of activities and problem-solving strategies were discussed. New activity goals were then introduced and scheduled in session using an activity scheduling form, which was securely emailed prior to session two (see Multimedia Appendix 7). Session three followed the same structure as session two, with review of the activity scheduling form instead of the mood activity log. During session four, the treatment and skills were reviewed, and feedback was attained (see Figure 6 for study timeline).

Figure 6. Study timeline. BA: behavioral activation; CBT: cognitive behavior therapy; MINI: Mini International Neuropsychiatric Interview; PHQ: Patient Health Questionnaire; VR: virtual reality.



Results

Demographic and Background Information

This case report used the information and data gathered from the first participant who completed the VR BA arm of the study. This participant contacted the protocol director, stating that he was interested in participating in the study. After the initial phone screen and intake session were complete, he was informed that he met criteria to participate in the study. He was then randomized into the VR BA arm of the study.

The participant was a Caucasian male in his early 40s with a history of depression, who had been treated with psychotherapy and medication management. The participant stated that he had never used a VR device.

Medications

The participant was taking Fluoxetine and Mirtazapine (dosages unknown) to treat his symptoms of depression.

Psychiatric History

The participant stated that he had engaged in weekly talk therapy for years and was currently engaged in weekly psychotherapy for an hour and was being followed for medication management. The participant experienced six prior episodes of depression and met criteria for current MDD on the MINI. The participant denied current suicidal ideation or intent but stated that he had one previous suicide attempt over 10 years ago. The participant had a family history of bipolar disorder, with two paternal first cousins being diagnosed, but he denied ever experiencing any manic or hypomanic symptoms and did not meet criteria for bipolar disorder. The participant met criteria for mild cannabis use disorder, in early remission, given that he reported abstaining from cannabis for over 6 months. The participant additionally met criteria for bulimia nervosa, with about one inappropriate compensatory behavior a week after an eating binge.

Medical History

The participant denied any history of seizures. He also denied any underlying medical illness.

Mental Status Examination

During all meetings with the protocol director, the participant was alert and oriented to person, place, and time, and did not endorse any auditory or visual hallucinations. The participant did not appear to be in acute distress and denied suicidal ideation. He was appropriately dressed and well groomed. His speech, volume, and prosody were within normal limits. The participant's affect was content congruent, and he was agreeable to the protocol director's instruction and questioning.

Intake and Session 1

During the initial phone session, the participant's PHQ-8 score was a 10. The first session took place 9 days later and his PHQ-9 score was an 8. The subsequent three sessions took place exactly a week apart from each other.

Session 2

At the beginning of session two, the participant's PHQ-9 score decreased to a 7. During this session, the participant reflected that he enjoyed using the VR headset during the week because "VR can give [him] new experiences that [he] would not normally be able to do." He also shared that using the VR headset had helped him to feel better, with the hypothesis that the "novelty helps with depression." The participant further noted that he found himself "more motivated to do other things after using the headset." He provided the feedback that the videos were not too short, and he liked that the length of each activity was provided; however, he thought that the motorcycle video was too long and repetitive at 10 minutes.

Session 3

At the beginning of session three, the participant's PHQ-9 score decreased to 6. During this session, the participant remarked that it was easier for him to engage in VR activities than activities in real life because he keeps the headset nearby, knows it is a short time commitment, and that he will feel better after using the headset. He further stated that after completing a VR activity, he experienced increased motivation to partake in a real life activity, enumerating his thought of "I may as well try something else" after using the headset. The participant also remarked that he used the VR headset to help replace activities he wanted to decrease, such as social media.

Session 4

At the beginning of session four, the participant's PHQ-9 score was 5. The participant stated that due to activity monitoring he was completing more activities than usual, both in VR and in real life. He also mentioned that using the VR headset provided him with a sense of accomplishment and something to look forward to. He attributed the overall 5-point decrease in his PHQ-9 scores to the fact that he had been increasing his activities, had more motivation after using the VR headset, and was feeling a greater sense of mastery. He stated that the appeal of using VR was the ability to gain exposure to new things, which provided the impetus for him to engage in novel activities in real life, such as visiting new parks. The participant remarked that although there were only 37 VR activities to choose from, he still felt that the experiences were novel since he could look in different directions during each activity. He recalled that the key to frequently using the device was keeping it close by so that he would remember to use it instead of other, less helpful activities. He noted that 5 days prior his psychiatrist increased his dosage of Fluoxetine but stated that he did not currently feel a difference and was informed that the effects would not be felt for "a week or two."

Feasibility, Acceptability, Tolerability

The feasibility, or degree to which VR could successfully be integrated into BA treatment, was measured by commenting on qualitative barriers to use observed. Barriers were assessed by rates of dropout, adverse events, the number of times the headset was used, and the level of presence felt in the headset. The level of presence was calculated on a scale of 0 (not at all) through 4 (very strongly) for each question; and with three questions, there was a possibility of yielding a score between 0 and 12. The average total presence for the participant was then calculated as 9.53 out of 12. The participant completed the study and did not report any adverse events during the study. Although full homework completion required using the VR headset a minimum of four times per week, yielding a minimum total of 12 times, this participant used the headset a total of 21 times, while completing 15 post-VR worksheets. The participant used the headset six times during his first week, 10 times during his second week, and five times during his third week. Although the participant engaged in 15 of the 37 potential activities, he chose to participate in "Cats in Living Room," "Bungee Swinging Canyon," "Rollercoaster," "Motorcycle Race," "Swimming with Sharks," and "Elephant" twice each, while engaging in the other nine activities once.

Acceptability of the VR BA treatment was measured using the TAM, with the agreeance choice on the Likert scale represented from 0 (Strongly Disagree) through 4 (Strongly Agree). The number of questions in each category determined the outcome range, which were then averaged. The participant maintained the same agreeance choices on all 15 questionnaires, indicating "Strongly Agree" for "Perceived Ease of Use" (average score: 12 out of 12), "Attitudes Towards Use" (average score: 16 out of 16), and "Intention to Use Technology" (average score: 12 out of 12), while in the "Perceived Usefulness" category, he rated "Strongly Agree" for questions 2 and 3 but "Agree" for question 1 (average score: 11 out of 12).

Physical tolerability of the VR headset was assessed by using the SSQ, and the emotional tolerability of the VR headset was assessed by using the Brief Agitation Measure. Physical tolerability was broken into each item and scaled from 0 (No more than usual) to 3 (Severely more than usual) for each item. The participant endorsed symptoms of *nausea*, *general discomfort*, *stomach awareness*, *sweating*, *increased salivation*, *vertigo*, *burping*, and *dizzy* (*eyes open*; see Table 1). Specifically, the participant endorsed the resulting symptoms after participating in the following four activities: "Cats in Living Room" (nausea), "Motorcycle Race" (nausea, general discomfort, stomach awareness, sweating, increased salivation, vertigo, dizzy [eyes open]), "Rollercoaster" (nausea, general discomfort, stomach awareness, sweating, increased salivation, vertigo, dizzy [eyes open]), and "Skiing" (nausea).

The total scores for physical tolerability were then summed, yielding a total score of 1.8 out of a potential 48, indicating high physical tolerability. Emotional tolerability was scored from 1 (Strongly Disagree) to 7 (Strongly Agree) per question; and with three questions, there was a possibility of yielding a score between 3 and 21. The average emotional tolerability score for this participant was a 3, indicating high emotional tolerability.

Table 1. Physical tolerability.

Measure (range of score)	Score, mean
Nausea (0-3)	0.33
General discomfort (0-3)	0.2
Stomach awareness (0-3)	0.27
Sweating (0-3)	0.27
Increased salivation (0-3)	0.13
Vertigo (0-3)	0.4
Burping (0-3)	0
Difficulty concentrating (0-3)	0
Difficulty focusing (0-3)	0
Eyestrain (0-3)	0
Fatigue (0-3)	0
Headache (0-3)	0
Blurred vision (0-3)	0
Dizzy (eyes open; 0-3)	0.2
Dizzy (eyes closed; 0-3)	0
Fullness of head (0-3)	0

Discussion

Principal Findings

This case demonstrates that VR BA was a feasible, acceptable, and tolerable method of delivering BA for an individual diagnosed with MDD during the COVID-19 pandemic. It also describes in detail the intervention that is being studied in a pilot randomized controlled trial (RCT) to be completed in 2021.

The participant used the headset more than was required for homework, did not verbally report any adverse events, and experienced an average presence of nearly 80% while using the headset. Although this is a relatively high presence rating, the authors conjecture that it was not higher for two reasons. First, the Limbix headset created a subtle effect that one is looking at the image through a screen, due to the simple device technology. Second, in using a 360° video, to give the illusion of movement, the image moves while the participant remained still, rather than the participant being able to walk around the virtual environment. It is possible that with a more advanced device, the presence rating would be higher. Still, his presence ratings were not correlated with his pre-VR to post-VR mood ratings.

The largest feasibility issue was working through COVID-19 “shelter-in-place” and transitioning the study from an in-person to telehealth design. However, the issue was solved by shipping the VR headset to the participant and using Zoom for the session meetings. Thus, the participant was still able to engage in weekly face-to-face sessions with the protocol director.

The participant rated the use of the headset as highly acceptable, giving the highest ratings of acceptability to all but one question. He indicated that using the headset was emotionally tolerable, denying any symptoms of agitation. He also rated the use of the headset as largely physically tolerable, providing a rating of about 3.8% intolerability. Although the participant did endorse varying degrees of nausea (N), general discomfort, stomach awareness (N), sweating (N), increased salivation (N), vertigo (D), and dizziness (eyes open; D) during three of the adrenaline activities and one nonadrenaline activity, the participant did not discuss these symptoms during sessions and still rated his post-VR mood as the same as or higher than his pre-VR mood after each of the adrenaline activities [16]. Given that the majority of his cybersickness symptoms occurred during the adrenaline activities of riding a rollercoaster, skiing, and riding a motorcycle, the authors hypothesize that this was a result of the mismatch between his vestibular and visual cues, since the movement of the image during adrenaline activities happens more quickly than when watching a sunset or observing nature [16,18]. Interestingly, despite endorsing symptoms of cybersickness during “Cats in Living Room,” “Rollercoaster,” and “Motorcycle Race” in week 1, he chose to engage in those three activities again during week 2. The activity of “Skiing” was participated in once, during his last week (week 3).

The participant experienced a 5-point decrease in depression symptoms on the PHQ-9 over a month, providing an initial rating of 10, which indicated moderate depression, and a final rating of 5, indicating very mild depression [23]. This decrease is clinically significant and illustrated that, despite the restrictions in place due to COVID-19, an individual was able to decrease his symptoms of depression using VR BA. He attributed this decrease in depression symptoms to increasing the number of his real life and virtual activities, a hypothesis that is in accordance with the behavioral theory of depression [12]. However, without a powered RCT, this finding could be due to a placebo effect. This is because the participant’s PHQ-9 score decreased from 10 during the initial intake to 8 at the beginning of the first session, before any intervention was provided. It should also be noted that the participant increased the dosage of his antidepressant medication simultaneously, adding a confound. However, the medication adjustment was done in between sessions 3 and 4, and there was already a downward trend of his PHQ-9 scores.

The strength of this VR BA intervention is that the mood ratings, activities completed, and amount of times using the headset were captured objectively and were standardized on the Limbix headset. Consequently, accurate home practice measurements were made, and the possibility of inaccurate homework reporting was eliminated. Additionally, this study took place during COVID-19 “shelter-in-place,” when real life activities were limited. The fact that this participant experienced mood increases after using the headset provides some evidence and possible potential for using VR to increase mood when real life activities are limited or treatment is being delivered remotely through telehealth.

This case study has several limitations. First, many of the quantitative and qualitative measures were subjective and completed by the participant. Although the participant engaged with the headset 21 times, he only completed 15 post-VR questionnaires, and thus, the complete data set was not able to be analyzed after every activity. Qualitative data collection that included further context was needed and is an important consideration for future research. In addition, given that this is a case report on one individual, the results may not be generalizable or help us identify causality. Results may not be applicable to all populations struggling with symptoms of depression, due to the heterogeneity of the disorder. Specifically, the participant began treatment with moderate symptoms of depression, and although BA has been shown to be effective for those with more severe symptoms, it is unknown whether using a VR headset to perform BA would yield these findings [12]. Unlike BA in real life, the participant was limited to the 37 activities on the headset, which were chosen primarily based on quality of image. Additionally, there was no follow-up, and thus, it is unknown whether the mood gains were lasting. Last, the participant had to return the Limbix headset to the protocol director upon study completion, and although there are low-cost VR options that the participant was educated to use, it is unknown if they will yield the same outcomes or compliance.

Conclusion

To the authors’ knowledge, this was the first reported use of VR to administer BA for a person with MDD. This is a retrospective case report that used the data from the first VR BA participant in a larger, three-arm study. This case was a combination of VR1 and VR2 methodologies, as outlined by Birckhead et al [32], since it both discussed in-depth user feedback of the prototype VR intervention, and it also evaluated the feasibility, acceptability, and tolerability of the VR intervention for a participant diagnosed with MDD in a clinical setting. We believe that these findings will inspire other researchers to investigate and explore the use of VR BA as a method of treating individuals diagnosed with MDD. We also believe that these encouraging findings may inspire other researchers to pursue VR3 trials (powered RCT) to compare outcomes between using VR to administer BA and a control condition for individuals diagnosed with MDD [32].

Acknowledgments

The authors wish to thank Talia Lyric Weiss for her early assistance with the study design.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Video URLs.
[DOCX File , 17 KB - mental_v7i11e24331_app1.docx]

Multimedia Appendix 2
List of virtual reality activities.
[DOCX File , 16 KB - mental_v7i11e24331_app2.docx]

Multimedia Appendix 3
Demographic questionnaire.
[DOCX File , 14 KB - mental_v7i11e24331_app3.docx]

Multimedia Appendix 4
Telephone screen questions.
[DOCX File , 13 KB - mental_v7i11e24331_app4.docx]

Multimedia Appendix 5
Activity monitoring form.

[DOCX File , 14 KB - mental_v7i11e24331_app5.docx]

Multimedia Appendix 6
Post–virtual reality questionnaire.
[DOCX File , 19 KB - mental_v7i11e24331_app6.docx]

Multimedia Appendix 7
Activity scheduling form.
[DOCX File , 14 KB - mental_v7i11e24331_app7.docx]

References

1. Depression and other common mental disorders: global health estimates. World Health Organization. 2017. URL: https://apps.who.int/iris/bitstream/handle/10665/254610/WHO-MSD-MER-2017.2-eng.pdf
2. Diagnostic and Statistical Manual of Mental Disorders (5th edition). Arlington, VA: American Psychiatric Association; 2013.
3. 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;395(10227):912-920. [doi: 10.1016/s0140-6736(20)30460-8]
4. Hawryluck L, Gold WL, Robinson S, Pogorski S, Galea S, Styra R. SARS control and psychological effects of quarantine, Toronto, Canada. Emerg Infect Dis 2004 Jul;10(7):1206-1212 [FREE Full text] [doi: 10.3201/eid1007.030703] [Medline: 15324539]
5. Weems CF, Mccurdy B, Carrion V, Scozzafava MD. Increased risk of suicide due to economic and social impacts of social distancing measures to address the Covid-19 pandemic: a forecast. Res Prepr Preprint posted online March 2020. [doi: 10.13140/RG.2.2.21601.45926]
6. Ferrari A, 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]
7. 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] [Medline: 23245608]
8. Murray C, Atkinson C, Bhalla K, Birbeck G, Burstein R, Chou D, Foreman, Lopez, Murray, Dahodwala, Jarlais, Fahami, Murray, Jarlais, Foreman, Lopez, Murray, U.S. Burden of Disease Collaborators. The state of US health, 1990-2010: burden of diseases, injuries, and risk factors. JAMA 2013 Aug 14;310(6):591-608 [FREE Full text] [doi: 10.1001/jama.2013.13805] [Medline: 23842577]
9. Hasin D, 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] [Medline: 29450462]
10. Major depression. National Institute of Mental Health. 2019. URL: https://www.nimh.nih.gov/health/statistics/major-depression.shtml [accessed 2020-03-05]
11. Halverson J, Beevers C, Kamholz B. Clinical practice review for major depressive disorder. Anxiety and Depression Association of America. 2016 Feb 02. URL: https://adaa.org/resources-professionals/practice-guidelines-mdd [accessed 2020-04-11]
12. Dimidjian S, Barrera M, Martell C, Muñoz RF, Lewinsohn PM. The origins and current status of behavioral activation treatments for depression. Annu Rev Clin Psychol 2011;7:1-38. [doi: 10.1146/annurev-clinpsy-032210-104535] [Medline: 21275642]
13. Park M, Kim DJ, Lee U, Na EJ, Jeon HJ. A literature overview of virtual reality (VR) in treatment of psychiatric disorders: recent advances and limitations. Front Psychiatry 2019;10:505. [doi: 10.3389/fpsyt.2019.00505] [Medline: 31379623]
14. Chirico A, Gaggioli A. When virtual feels real: comparing emotional responses and presence in virtual and natural environments. Cyberpsychol Behav Soc Netw 2019 Mar;22(3):220-226. [doi: 10.1089/cyber.2018.0393] [Medline: 30730222]
15. Wang P, Lane M, Olfson M, Pincus HA, Wells KB, Kessler RC. Twelve-month use of mental health services in the United States: results from the National Comorbidity Survey Replication. Arch Gen Psychiatry 2005 Jun;62(6):629-640. [doi: 10.1001/archpsyc.62.6.629] [Medline: 15939840]
16. Davis S, Nesbitt K, Nalivaiko E. A systematic review of cybersickness. 2014 Presented at: 2014 Conference on Interactive Entertainment; December 2014; New York, NY. [doi: 10.1145/2677758.2677780]
17. Srivastava K, Das RC, Chaudhury S. Virtual reality applications in mental health: challenges and perspectives. Ind Psychiatry J 2014;23(2):83-85 [FREE Full text] [doi: 10.4103/0972-6748.151666] [Medline: 25788795]
18. Wilson JR. Effects of participating in virtual environmentsa review of current knowledge. Saf Sci 1996 Jun;23(1):39-51. [doi: 10.1016/0925-7535(96)00026-4]

19. Jerdan S, Grindle M, van Woerden HC, Kamel Boulos MN. Head-mounted virtual reality and mental health: critical review of current research. JMIR Serious Games 2018 Jul 06;6(3):e14 [FREE Full text] [doi: 10.2196/games.9226] [Medline: 29980500]
20. Kyoto K. Pico Goblin specifications. SizeScreens.com. 2017. URL: https://www.sizescreens.com/pico-goblin-specifications/ [accessed 2020-09-10]
21. MacPhillamy D, Lewinsohn PM. The pleasant events schedule: studies on reliability, validity, and scale intercorrelation. J Consult Clin Psychol 1982 Jun;50(3):363-380. [doi: 10.1037/0022-006x.50.3.363]
22. Sheehan D, Lecrubier Y, Harnett Sheehan K, Janavs J, Weiller E, Keskiner A, et al. The validity of the Mini International Neuropsychiatric Interview (MINI) according to the SCID-P and its reliability. Eur Psychiatry 1997;12(5):232-241. [doi: 10.1016/s0924-9338(97)83297-x]
23. 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. [doi: 10.1046/j.1525-1497.2001.016009606.x] [Medline: 11556941]
24. Nowak K, Biocca F. The effect of the agency and anthropomorphism on users' sense of telepresence, copresence, and social presence in virtual environments. Presence Teleoperators Virtual Environments 2003 Oct;12(5):481-494. [doi: 10.1162/105474603322761289]
25. Kennedy R, Lane NE, Berbaum KS, Lilienthal MG. Simulator sickness questionnaire: an enhanced method for quantifying simulator sickness. Int J Aviat Psychol 1993 Jul;3(3):203-220. [doi: 10.1207/s15327108ijap0303_3]
26. Sevinc V, Berkman MI. Psychometric evaluation of Simulator Sickness Questionnaire and its variants as a measure of cybersickness in consumer virtual environments. Appl Ergon 2020 Jan;82:102958. [doi: 10.1016/j.apergo.2019.102958] [Medline: 31563798]
27. Ribeiro J, Bender TW, Selby EA, Hames JL, Joiner TE. Development and validation of a brief self-report measure of agitation: the Brief Agitation Measure. J Pers Assess 2011 Nov;93(6):597-604. [doi: 10.1080/00223891.2011.608758] [Medline: 21999383]
28. Manis K, Choi D. The virtual reality hardware acceptance model (VR-HAM): extending and individuating the technology acceptance model (TAM) for virtual reality hardware. J Business Res 2019 Jul;100:503-513. [doi: 10.1016/j.jbusres.2018.10.021]
29. Funderburk J, Pigeon WR, Shepardson RL, Maisto SA. Brief behavioral activation intervention for depressive symptoms: patient satisfaction, acceptability, engagement, and treatment response. Psychol Serv 2019 Feb 04. [doi: 10.1037/ser0000328] [Medline: 30714752]
30. Jacobson NS, Martell CR, Dimidjian S. Behavioral activation treatment for depression: returning to contextual roots. Clin Psychol Sci Pract 2006. [doi: 10.1093/clipsy.8.3.255]
31. Lewinsohn PM. A behavioral approach to depression. In: Friedman RJ, Katz MM, editors. The Psychology of Depression: Contemporary Theory and Research. New York, NY: Wiley; 1974:157-185.
32. Birckhead B, Khalil C, Liu X, Conovitz S, Rizzo A, Danovitch I, et al. Recommendations for methodology of virtual reality clinical trials in health care by an international working group: iterative study. JMIR Ment Health 2019 Jan 31;6(1):e11973 [FREE Full text] [doi: 10.2196/11973] [Medline: 30702436]

Abbreviations

BA: behavioral activation
D: disorientation
DSM-5: Diagnostic and Statistical Manual of Mental Disorders, 5th Edition
MDD: major depressive disorder
MINI: Mini International Neuropsychiatric Interview
N: nausea
O: oculomotor
PHQ-9: Patient Health Questionnaire-9
RCT: randomized controlled trial
SSQ: Simulator Sickness Questionnaire
TAM: technology acceptance model
VR: virtual reality

Edited by J Torous; submitted 15.09.20; peer-reviewed by DA Rohani, J Torous; comments to author 21.09.20; revised version received 04.10.20; accepted 04.10.20; published 03.11.20.

Please cite as:
Paul M, Bullock K, Bailenson J
Virtual Reality Behavioral Activation as an Intervention for Major Depressive Disorder: Case Report
JMIR Ment Health 2020;7(11):e24331
URL: https://mental.jmir.org/2020/11/e24331
doi:10.2196/24331
PMID:33031046

Original Paper

A Mobile Health Mindfulness Intervention for Women With Moderate to Moderately Severe Postpartum Depressive Symptoms: Feasibility Study

Lyndsay A Avalos[1], PhD, MPH; Sara Aghaee[1], MPH; Elaine Kurtovich[1], MPH; Charles Quesenberry Jr[1], PhD; Linda Nkemere[1], BA; MegAnn K McGinnis[1], MPH; Ai Kubo[1], MPH, PhD

Kaiser Permanente Northern California, Division of Research, Oakland, CA, United States

Corresponding Author:
Lyndsay A Avalos, PhD, MPH
Kaiser Permanente Northern California
Division of Research
2000 Broadway
Oakland, CA, 94612
United States
Phone: 1 510 891 3426
Email: Lyndsay.A.Avalos@kp.org

Abstract

Background: Approximately 20% of women suffer from postpartum depression (PPD). Due to barriers such as limited access to care, half of the women with PPD do not receive treatment. Therefore, it is critical to identify effective and scalable interventions. Traditional mindfulness programs have been effective in reducing depressive symptoms, however access remains a barrier. A self-paced mobile health (mHealth) mindfulness program may fit the lifestyle of busy mothers who are unable to attend in-person classes. However, little is known regarding the feasibility or efficacy of mHealth mindfulness interventions in postpartum women with depressive symptoms.

Objective: This study aims to assess the feasibility, acceptability, and preliminary efficacy of an mHealth mindfulness intervention for postpartum women with moderate to moderately severe depressive symptoms.

Methods: We conducted a single-arm feasibility trial of an mHealth mindfulness intervention within Kaiser Permanente Northern California (KPNC), a large integrated health care system. Participants were identified through clinician referral and electronic health records via KPNC's universal perinatal depression screening program and recruited by the study team. Inclusion criteria included the following: English-speaking, up to 6 months postpartum with a Patient Health Questionnaire (PHQ-8) score of 10 to 19, and no regular mindfulness/meditation practice. Participants were asked to use a mindfulness app, Headspace, 10 to 20 min/day for 6 weeks. Baseline and postintervention surveys captured data on patient-reported outcomes (depression and stress symptoms, sleep quality, and mindfulness). Semistructured interviews captured acceptability. Retention and adherence were used to assess feasibility.

Results: Of the 115 women who were contacted and met the eligibility criteria or declined participation before eligibility assessment, 27 (23%) were enrolled. In addition, 70% (19/27) completed the study. The mean age of participants was 31 years (SD 5.2), 30% (8/27) were non-Hispanic White, and, on average, participants were 12.3 weeks postpartum (SD 5.7). Of the women who completed the study, 100% (19/19) used the Headspace app at least once, and nearly half (9/19, 47%) used the app on ≥50% of the days during the 6-week intervention period. Of the 16 participants who completed the postintervention interview, 69% (11/16) reported that they were *very* or *extremely* satisfied with the app. Interviews indicated that women appreciated the variety of meditations and felt that the program led to reduced anxiety and improved sleep. Significant improvements in pre- and postintervention scores were observed for depressive symptoms (PHQ-8: −3.8, *P*=.004), perceived stress (10-item Perceived Stress Scale: −6.0, *P*=.005), and sleep quality (Pittsburgh Sleep Quality Index: −2.1, *P*=.02, indicating less sleep disturbance). Improvements in mindfulness were also significant (Five Facet Mindfulness Questionnaire-Short Form: 10.9, *P*=.01).

Conclusions: An mHealth mindfulness intervention for postpartum women with moderate to moderately severe depressive symptoms is feasible and acceptable. An efficacy trial is warranted.

(JMIR Ment Health 2020;7(11):e17405) doi:10.2196/17405

KEYWORDS

depression; postpartum; health services, mental; eHealth; mental health; internet-based intervention; mindfulness; behavioral intervention; mobile phone

Introduction

Postpartum Depression

Postpartum depression (PPD) is the number 1 complication of childbirth [1,2], affecting up to 20% of postpartum women. It is a life-threatening, debilitating, and costly mood disorder that emerges within a year of delivery [1,3,4]. Symptoms of PPD include loss of interest or energy, depressed mood, fluctuations in sleep or eating patterns, reduced ability to think or concentrate, feelings of worthlessness, and recurrent suicidal ideation; PPD can also result in infanticide [5,6]. PPD can have multigenerational consequences, substantially affecting the health of the mother and the child. For example, women with PPD are more likely to demonstrate hostile and/or coercive behaviors and disengagement from their infants [7], resulting in negative mother-infant interactions [8]. Women with PPD are less likely to breastfeed and are at increased risk of early cessation of breastfeeding [9], and their infants receive fewer preventive services, such as recommended immunizations [10,11]. Children of women with PPD have poorer cognitive function [12-14], are at increased risk of behavioral and developmental disorders, such as attention deficit hyperactivity disorder [15] and psychiatric disorders [16,17], such as depression, anxiety, and conduct disorders. The societal costs of untreated perinatal mood disorders for all US births in 2017 were estimated at US $14.2 billion [18], and unfortunately, half of the women with a perinatal mood disorder (which includes PPD) do not receive the treatment they need.

Recent guidelines by the United States Preventive Services Task Force [19] and several specialized medical societies, including the American Academy of Pediatrics and the American College of Obstetricians and Gynecologists [20,21], have established perinatal depression screening and treatment as essential components of postpartum care. Despite these recommendations, several patient- and system-level barriers to the receipt of current treatment options (eg, psychotherapy and antidepressant medications) remain. For example, although psychotherapy is an effective nonpharmacological treatment option, numerous barriers to receiving care exist, including the shortage of mental health care providers [22], limited access to care, financial constraints, and lack of time, transportation, and childcare [23]. In addition, most pregnant and postpartum women (83-95%) prefer nonpharmaceutical treatments [24]. Further, the COVID-19 pandemic has had a significant impact on the mental health of pregnant and postpartum women, with more than one-third of women reporting significant depression symptoms [25]. Although the rates of perinatal depression have increased, access to in-person delivered health care has drastically diminished [26,27]. Therefore, it is critical to identify safe, effective, patient-centered, and scalable intervention options for postpartum women with heavy depression symptom burden.

Mindfulness Interventions for Depression

Mindfulness, a psychological process of bringing attention to the present moment [28-30], has demonstrated its effectiveness as an intervention for reducing symptoms of depression in many populations [31-35]. However, gold standard mindfulness training often requires 30 or more hours of *in-person* instruction with 45 min of daily homework [28,36]. Thus, despite their known efficacy in reducing depression symptoms, traditional mindfulness programs pose similar accessibility barriers to those associated with counseling services, reducing its potential to help busy women with PPD. Technology is becoming an increasingly popular method for delivering lifestyle and behavioral interventions, and there has been a steady rise in mobile health (mHealth) interventions, particularly as they fit the lifestyles of individuals who are unable to attend regular in-person classes. Thus, a self-paced, mHealth mindfulness-based intervention has potential as a scalable behavioral intervention that addresses barriers to traditional mindfulness programs.

Research is needed to ascertain the effectiveness of mHealth mindfulness interventions in postpartum women with depressive symptoms. As a first step, we conducted a feasibility study of an mHealth mindfulness-based intervention for women with moderate to moderately severe PPD symptoms within a large integrated health care delivery system. This study investigates the feasibility and acceptability of the mHealth mindfulness intervention while also reporting on the preliminary efficacy of patient-reported outcomes to determine whether conducting a randomized control trial of the intervention is warranted.

Methods

Study Setting

The study was conducted within Kaiser Permanente Northern California (KPNC), an integrated health care delivery system serving over 4.4 million racially and socioeconomically diverse members representative of the Northern California population [37,38]. Standard postpartum care includes screening for depression at the fourth to eighth week postpartum visit using the 9-item Patient Health Questionnaire (PHQ-9) [39,40].

Study Design and Population

A mixed-methods single-arm trial of a 6-week mHealth mindfulness intervention was conducted between March 2018 and June 2019. Women seeking postpartum care were recruited from 7 of the 44 KPNC obstetrics and gynecology clinics. Women aged at least 18 years, within 6 months of giving birth, with a PHQ-9 score of 10 to 19 (indicating moderate to moderately severe depressive symptoms), English-speaking, with access to a smartphone, tablet, or computer with internet access were eligible for the study. Women who engaged in regular mindfulness, meditation, or yoga practice 3 or more times per week or enrolled in a mindfulness program were excluded. Participants were asked to complete a web survey at baseline and immediately after the intervention to assess patient-reported outcomes of depression, stress, sleep quality, and mindfulness. Semistructured interviews were conducted within 3 weeks of completion of the intervention to assess the acceptability of the mHealth mindfulness intervention. This study was approved by the KPNC institutional review board.

Participant Identification and Recruitment

Potential participants were identified via 2 strategies: (1) a postpartum PHQ-9 score of 10 to 19, identified through the KPNC electronic health records (EHRs), and (2) self- or clinician-referral from KPNC obstetrics and gynecology clinical staff or study brochures. Potential eligible participants were contacted about the study via email and phone by a research assistant and rescreened for depression symptoms using the 8-item Patient Health Questionnaire (PHQ-8; see the *Measures* subsection for more information). Women who met all eligibility criteria and had a PHQ-8 score of 10 to 19 were enrolled. Participants who completed both baseline and 6-week follow-up surveys received a US $25 gift card and an additional year-long subscription to the mindfulness app.

Intervention

On signing the informed consent and completing the baseline survey, participants were provided access to a commercially available mindfulness app, Headspace. Headspace was chosen because it was identified as the best commercially available mindfulness mobile app in a review published in a peer-reviewed journal [41], and most of the previous studies, including ours, have reported that Headspace is an accessible and effective tool for delivering training to increase mindfulness in various populations [42-48]. Headspace provides self-paced, guided mindfulness meditations through a website or mobile app (iOS and Android). The home screen displays the next meditation in the series. Much of the program follows a linear pathway of daily, progressive meditations (ie, each day builds upon previous content) designed to deepen the understanding of mindfulness and encourage its integration into daily life.

The women were asked to use the app for 10 to 20 min a day during the 6-week study period. Each participant was given a study-specific log-in ID and encouraged to complete the 30-day *Basics* course first and then choose from the other themed sessions (eg, anxiety, relationships) for the remainder of the 6-week study period. When the study staff noted that a participant had completed fewer than 3 sessions in the past week, they called the participant to remind her to use the app.

Measures

Feasibility

We assessed 2 feasibility measures: adherence and retention.

- Retention: retention was calculated as the proportion of enrolled participants who completed both the baseline and postintervention surveys;
- Adherence to the intervention: the date, time, duration, and type of each meditation session that participants completed were collected by Headspace using the study-specific log-in ID. Adherence was assessed for all enrolled women and for women who completed the study.

Acceptability

Acceptability was assessed through responses in a semistructured interview. Participants were asked open-ended questions about their experience with the study and Headspace, recommended changes to the study procedures, perceived effects or benefits of practicing mindfulness, and perceived need for additional health system support for pregnant and postpartum women. Participants were also asked to respond to the question, “What was your overall experience with the Headspace program?” with 1 of the 4 responses: extremely useful/satisfied, very useful/satisfied, somewhat useful/satisfied, and not at all useful/satisfied. Participants' responses were written down as close to verbatim as possible by the interviewer.

Preliminary Efficacy of Participant-Reported Outcomes

The 4 participant-reported outcomes assessed were depression, stress, sleep quality, and mindfulness.

- Depression: the PHQ-8 [49] depression screener is a validated instrument adapted from the PHQ-9, which was used to assess current depression symptoms at recruitment and follow-up. The PHQ-8 excludes the question regarding suicidal thoughts. The PHQ-8 scores ranged from 0 to 24. Scores of 1 to 4 suggest minimal depression, 5 to 9 mild depression, 10 to 14 moderate depression, 15 to 19 moderately severe depression, and 20 to 24 severe depression;
- Stress: the 10-item Perceived Stress Scale [50] assesses the degree to which a respondent perceives situations in his or her life in the previous month as stressful through a 5-point Likert scale (0=*never* to 4=*very often*). The scores are summed to give a total score ranging from 0 to 40;
- Sleep quality: the 19-item Pittsburgh Sleep Quality Index [51] asks about sleep quality during the previous month, including questions on sleep duration, sleep disturbance, and use of sleep-inducing medications. A global score ranging from 0 to 21 is calculated using 7 components of sleep. Higher scores indicate poorer sleep quality;
- Mindfulness: the 24-item Five Facet Mindfulness Questionnaire-Short Form [52] uses a 5-point Likert scale to measure mindfulness and includes subscales to assess 5 elements of mindfulness—observing, describing, acting with awareness, nonjudging of inner experience, and nonreactivity to inner experience. Responses vary between 1 (never or very rarely true) to 5 (very often or always true). Scores for overall mindfulness range from 24 to 120. The scores for observing range from 4 to 20, whereas the scores for describing, acting with awareness, nonjudging of inner experience, and nonreactivity to inner experience range from 5 to 25. Higher scores indicate greater mindfulness.

Analytic Methods

Quantitative Data Analyses

Baseline characteristics differences between completers and noncompleters were assessed using the analysis of variance (ANOVA) for continuous variables and the Fisher exact test for categorical variables. Baseline mean and SD were calculated for all participant-reported outcome measures. Pre-post changes in scores and P values were derived using paired t tests, and a repeated measures ANOVA analysis was conducted to compare pre-post changes in scores by adherence to the intervention (meditated <50% of the days vs ≥ 50% of the days). All analyses were conducted using SAS software version 9.4 (Cary, NC).

Qualitative Assessment and Analysis

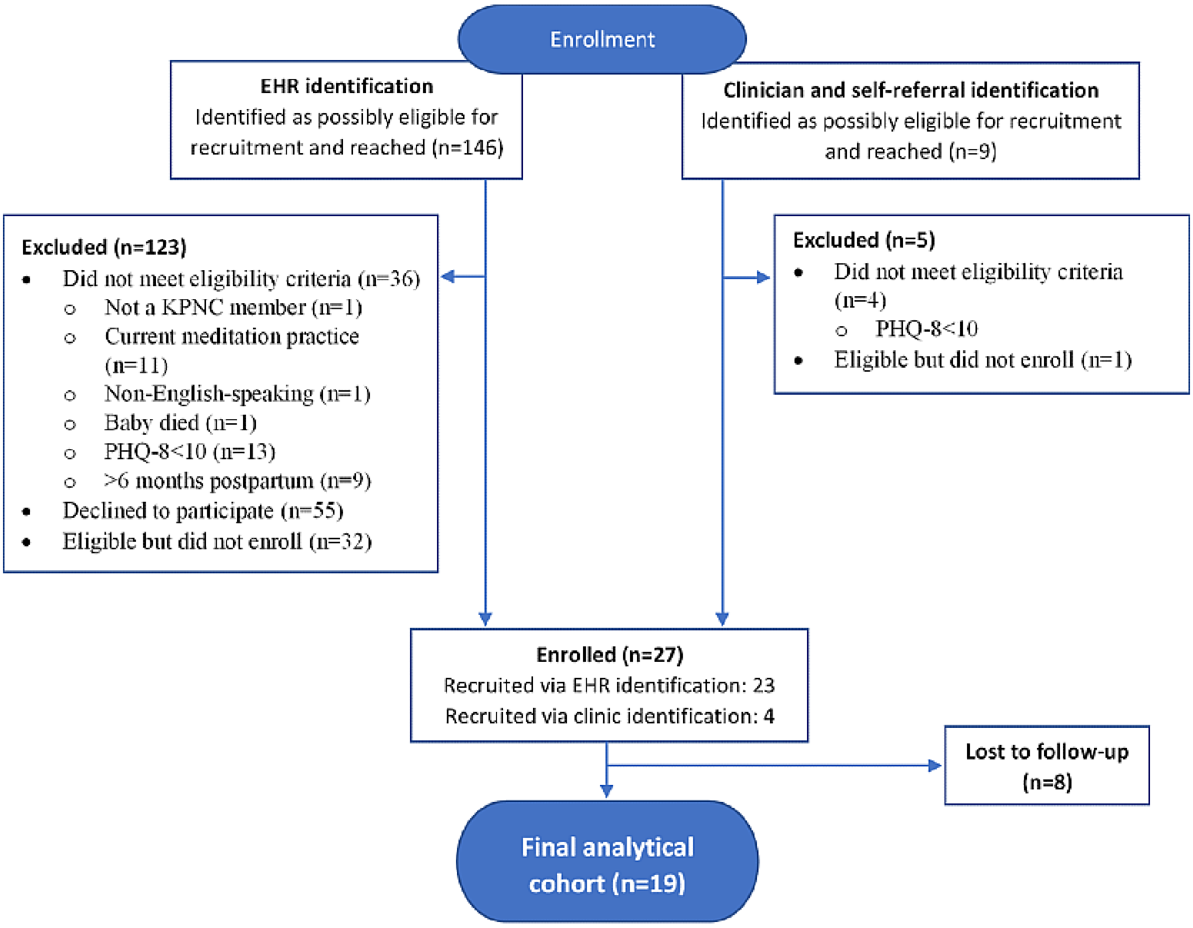
Qualitative data were uploaded into the NVivo qualitative data analysis software (QSR International Pty Ltd version 12, 2018). Inductive thematic analysis was used to identify and develop codes on themes related to mindfulness benefits, interface experience, experience with the study, and suggested changes. Each interview was coded by 2 primary coders (MM and LN); a third coder (EK) reviewed all coded transcripts to ensure the accuracy of codes.

Results

Recruitment

We contacted and reached 155 potentially eligible women, identified through the EHR (n=146) or clinician referral (n=9), by phone and assessed their eligibility; some women declined to participate (Figure 1). Of the women who were contacted, 26% (40/155) did not meet the eligibility criteria, 36% (55/155) declined to participate, and 21% (33/155) were eligible and interested but did not complete the enrollment process and thus were excluded from the study. The most common reasons for ineligibility were PHQ-8 scores outside of the eligible range and existing meditation practice. The most common reasons for declining participation were being too busy, too tired, already sought other treatments for depression, and lack of interest. A total of 27 of the 115 women who were either eligible or who declined to participate before eligibility was assessed, enrolled in the study, corresponding to a conservative 24% (27/115) recruitment rate.

Figure 1. Study recruitment flowchart for a mobile health mindfulness feasibility study for postpartum women with moderate to moderately severe symptoms of depression in Kaiser Permanente Northern California between 2018 and 2019.



Baseline Characteristics

Of the 27 women recruited, 9 (33%) were Hispanic, 8 (30%) were White, and 5 (19%) were Black. More than half (15/27, 56%) of the participants did not have a college degree, nearly a quarter (6/27, 22%) received Medicaid benefits, and most (20/27, 74%) had household income <US $100,000. The mean baseline PHQ-8 score was 14.3 (range 10-23, SD 3.8), and about half of the women had a current depression diagnosis (13/27, 48%; 6/27, 22%; postpartum only; and 7/27, 26%, postpartum and history of depression; Table 1). There were no significant differences between women who completed the study and those who did not complete it.

Table 1. Baseline characteristics of participants in a feasibility mobile health mindfulness study of postpartum women with moderate to moderately severe symptoms of depression in Kaiser Permanente Northern California between 2018 to 2019, overall and by study completion status.

Characteristics	Total (n=27)	Completed (n=19)	Not completed (n=8)
Baseline PHQ-8[a] score, mean (SD)	14.3 (3.8)	15.2 (4.0)	12.1 (2.0)
Number of weeks postpartum, mean (SD)	12.3 (5.7)	11.5 (5.0)	14.3 (7.2)
Age (years), mean (SD)	30.9 (5.2)	31.4 (5.3)	29.6 (5.1)
Depression diagnosis, n (%)			
Postpartum only	6 (22)	4 (21)	2 (25)
History of depression and postpartum	7 (26)	7 (37)	0 (0)
None	14 (52)	8 (42)	6 (75)
Race and ethnicity, n (%)			
Non-Hispanic White	8 (30)	7 (37)	1 (13)
Non-Hispanic Black	5 (19)	3 (16)	2 (25)
Asian	1 (4)	1 (5)	0 (0)
Hispanic	9 (33)	4 (21)	5 (63)
Multiracial	4 (15)	4 (21)	0 (0)
Education, n (%)			
Less than college graduate	15 (56)	9 (47)	6 (75)
College graduate	12 (44)	10 (53)	2 (25)
Household income, n (%)			
<US $100,000	20 (74)	13 (68)	7 (88)
≥US $100,000	7 (26)	6 (32)	1 (13)
Medicaid status, n (%)			
Yes	6 (22)	2 (11)	4 (50)
No	18 (67)	15 (79)	3 (38)
Unknown	3 (11)	2 (11)	1 (13)
Marital status, n (%)			
Married/living with partner	19 (70)	15 (79)	4 (50)
Single	7 (26)	3 (16)	4 (50)
Unknown	1 (4)	1 (5)	0 (0)
Parity, n (%)			
0	13 (48)	10 (53)	3 (38)
1+	14 (52)	9 (47)	5 (63)
Primary device used for the mindfulness program, n (%)			
Android	5 (19)	5 (26)	0 (0)
iOS[b]	15 (56)	14 (74)	1 (13)
Nonparticipator	7 (26)	0 (0)	7 (88)

[a] PHQ-8: 8-item Patient Health Questionnaire.

[b] iOS: iPhone OS.

Feasibility Outcomes

Retention

Seventy percent (19/27) of participants completed the baseline and follow-up surveys (Table 1).

Adherence to the Intervention

High rates of engagement were noted among the participants. Of the 27 women who enrolled in the study, 20 (74%) used the Headspace app at least once, with 9 (33%) practicing meditation for at least half of the days during the 6-week study period and 5 (19%) practicing at least 70% of the days. These rates were similar when only women who completed the study were considered. Among the women who completed the study, 100% (19/19) used the Headspace app at least once. Of these 19 participants, 9 (47%) practiced meditation for at least half of the days during the 6-week study period, and 5 (26%) practiced meditation for at least 70% of the days. In addition, 58% (11/19) participants used Headspace at least once during the month after the 6-week intervention period ended.

Acceptability

Of the participants who completed the study, 84% (16/19) participated in a postintervention telephone interview; 11 participants (11/16, 69%) responded that they were either *very* or *extremely* satisfied with the Headspace app, and all of the participants planned to continue their mindfulness practice with Headspace after the study ended. Results from the semistructured interviews supported participant satisfaction, as women reported that they felt it was easy to use, liked the variety of available meditation options, and valued the convenience of an app-based intervention. Whereas some participants liked the male facilitator's voice for meditation, others shared that they would prefer different voice options to guide the meditations, particularly a female voice, an option that Headspace added during the course of the study. Two participants wanted more freedom to explore different meditation sessions rather than to follow the prescribed course required by the study; 1 had meditation experience and would have liked to start at a more advanced level than to follow the basic meditation course; the other was interested in trying shorter (eg, 1 min) meditations because she found it hard to remain uninterrupted with a newborn. Another participant would have liked more phone calls from study staff to assist with initially downloading the app and weekly check-ins to encourage the use of the app.

Participants noted several benefits of using Headspace. The most commonly mentioned benefits were improved stress management, reduced anxiety, improved sleep, and increased physical activity. Many also noted that meditation allowed them to take some time off for themselves. Several participants liked having a structured routine of meditating each day. Others said they would work meditation, particularly the counted breathing technique, into their day while driving, exercising, or in moments of stress (eg, when the baby was crying).

Preliminary Efficacy of Participant-Reported Outcomes

At the 6-week postintervention follow-up assessment, participants experienced significant improvements in depressive symptoms (−3.8, SD 5.0, *P*=.004), perceived stress (−6.0, SD 7.9, *P*=.005), and sleep quality (−2.1, SD 3.4, *P*=.02, indicating less sleep disturbance) compared with baseline (Figure 2). Participants also achieved significantly greater levels of mindfulness in 3 of the 5 mindfulness domains (observing: 2.5, SD 3.9, *P*=.01; describing 2.4, SD 4.2, *P*=.02; nonjudging of inner experience 2.4, SD 4.1, *P*=.02; Figure 3) and significant improvements in overall mindfulness (10.9, SD 16.8, *P*=.01). Although not statistically significant, trends suggest a greater improvement in depression symptoms (−4.6, SD 5.2 vs −3.1, SD 4.9, *P*=.54), stress (−6.7, SD 8.8 vs −5.3, SD 7.3, *P*=.55; Figure 4), and overall mindfulness (13.3, SD 18.4 vs 8.8, SD 15.8, *P*=.57) for women who meditated using the app for at least 50% of the days compared with women who meditated for less than 50% of the days of the 6-week intervention, respectively. No differences were noted in sleep quality between the groups (−1.2, SD 3.9 vs −2.8, SD 2.8, *P*=.32; Figure 4).

Figure 2. Differences in pre- and postintervention patient-reported outcomes in a mobile health mindfulness feasibility study of postpartum women with moderate to moderately severe symptoms of depression in Kaiser Permanente Northern California between 2018 and 2019. **P*<.05 and ***P*<.01.

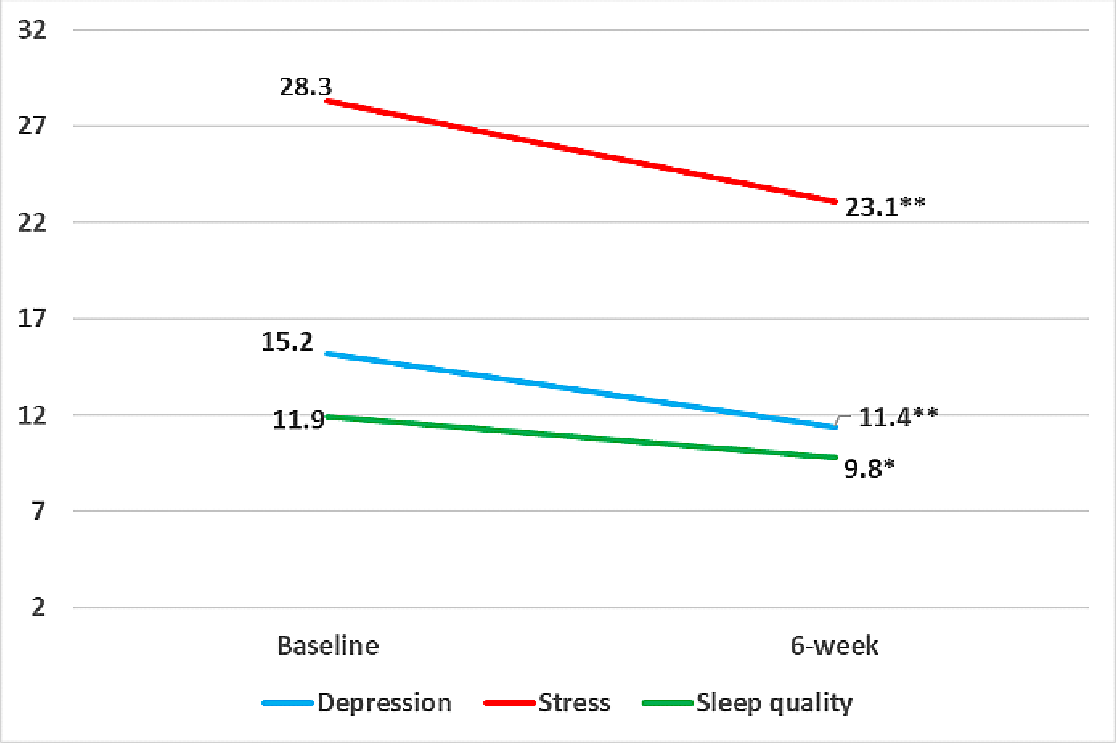


Figure 3. Differences in pre- and postintervention patient-reported mindfulness outcomes in a mobile health mindfulness feasibility study of postpartum women with moderate to moderately severe symptoms of depression in Kaiser Permanente Northern California between 2018 and 2019. **P*<.05 and ***P*<.01.

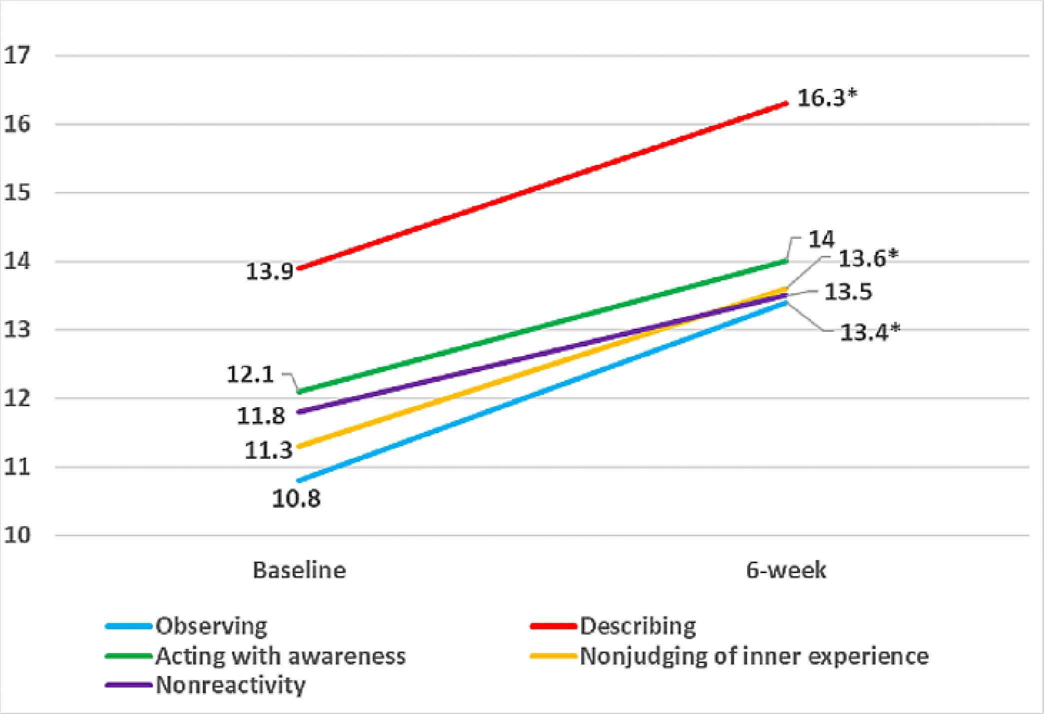
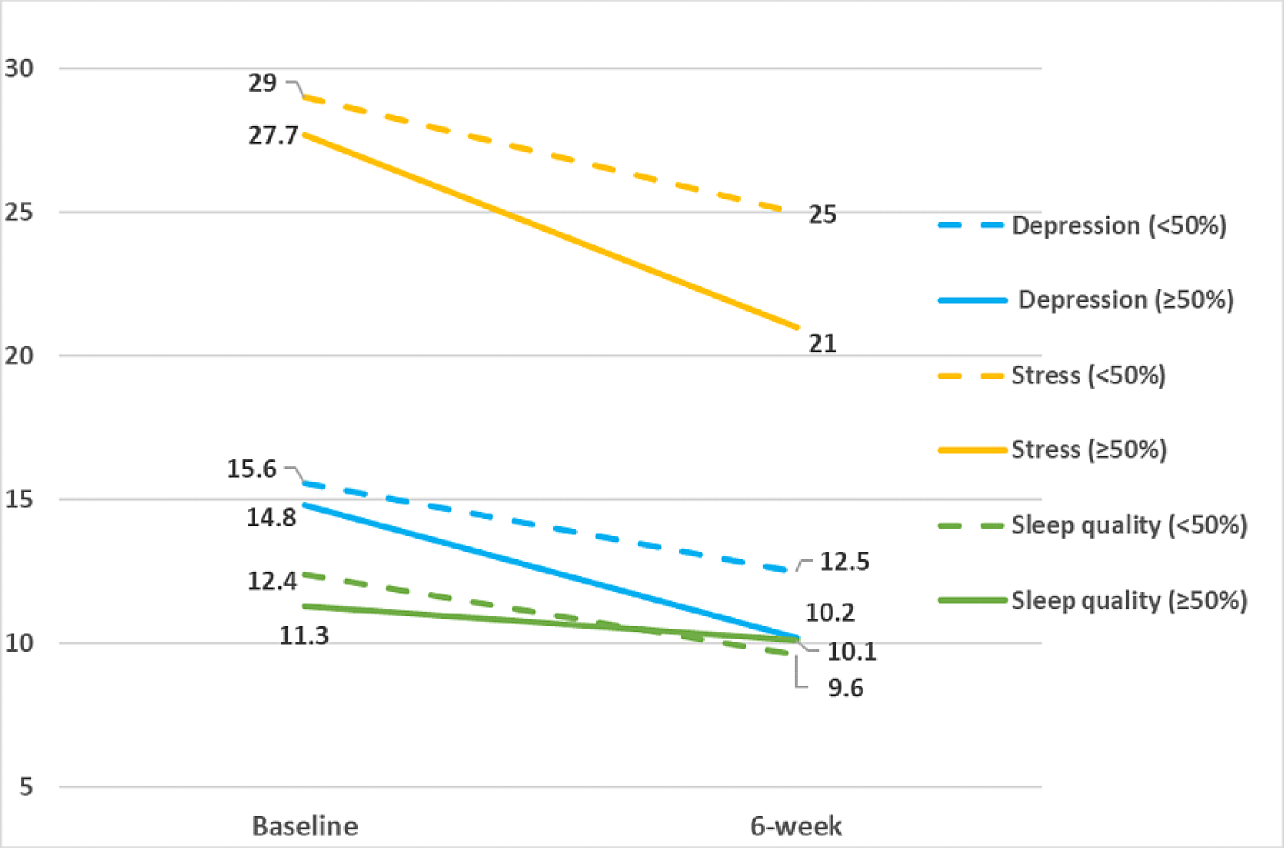


Figure 4. Patient-reported outcomes by percent Headspace intervention adherence in a mobile health mindfulness feasibility study of postpartum women with moderate to moderately severe symptoms of depression in Kaiser Permanente Northern California between 2018 and 2019.



Discussion

Principal Findings

Findings from this study suggest that a self-paced mHealth mindfulness intervention for women with moderate to moderately severe symptoms of PPD is both feasible and acceptable. This study demonstrated our team's ability to identify and recruit postpartum women with significant depression symptoms with good recruitment and retention rates. The participants appreciated the convenience of the intervention: most of the women engaged in the meditation program at least one time and a third of all recruited participants and nearly half of those who completed the study meditated for a majority of the intervention days. The quantitative data and findings from the semistructured interviews suggest preliminary efficacy and improvement in depression symptoms, stress, sleep, and mindfulness in postpartum women with moderate to moderately severe symptoms of PPD, suggesting that a full-powered trial is warranted.

This feasibility study also provided information on study protocols that may be used to improve the efficiency of a larger efficacy trial. Tracking participants' app usage allowed us to reach out when there was an extended period of inactivity (ie, <3 sessions in the previous week). This helped remind some participants to get back on track with the app. Using the reminder features and push notifications built into the app may also be useful as a reminder, given that at least 1 participant shared that they would have liked weekly reminders from the study team to encourage the use of the app. Further, automatic tracking of progress can be used as a source of motivation to continue using the app. Such interactive features of a mobile app can increase its adherence and effectiveness, and ease of use can facilitate widespread, efficient implementation [53]. Additionally, the assessment of patient-reported outcomes provides preliminary effect estimates for informing power in a randomized control trial.

Comparison With Previous Work

The theoretical basis for mindfulness derives from a model based on studies of the influence of mindfulness on brain regions involved in learning and memory processes, emotion regulation, self-referential processing, and perspective taking [54,55]. Recent research suggests that mindfulness programs may serve as a nonpharmacological treatment option for postpartum women with depression. For example, a recent meta-analysis, which included 9 studies of adults with depression (75% women), documented a significantly reduced risk of relapse over a 5-year follow-up period for those who received a mindfulness-based intervention compared with those who did not (hazard ratio, 0.69; 95% CI 0.58-0.82) [56]. Additionally, a meta-analysis of 8 randomized controlled trials of pregnant women concluded that women in the in-person mindfulness arm experienced significant reductions in depression [34].

Despite the positive effects of mindfulness programs on depression symptoms, the requirements of traditional mindfulness programs limit the accessibility and adherence of postpartum women. Traditional mindfulness programs often require more than 30 hours of *in-person* instruction along with 45 min of home practice daily. Postpartum women often face competing priorities such as a child at home or a full-time job, making attendance at in-person sessions challenging. A recent in-person mindfulness pilot study for pregnant women reported that recruitment was challenging for these reasons and that a more practical intervention was needed [32]. These findings may also be generalized to postpartum women. A recent meta-analysis including 65 randomized controlled trials, totaling 5489 participants, demonstrated that brief mindfulness training (ranging from a single-session to multisession interventions lasting up to 2 weeks) was also effective in reducing depression and anxiety [57]. Given the busy lifestyle of new mothers, convenient, frequent, and low dosage mindfulness programs that can be accessed from anywhere are more feasible than the traditional, lengthy, and in-person mindfulness interventions.

In addition to these common barriers, the COVID-19 pandemic has caused a surge in PPD symptoms, increasing the demand for mental health care. A recent study documented a near tripling of perinatal depression, with 15% of women reporting high depression symptoms prepandemic compared with 41% reporting such symptoms since the pandemic started [58]. Our study was conducted before COVID-19, and reasons for women declining participation included already seeking other treatments for depression and being too busy, although being too busy was a barrier this intervention was developed to address. The sheltering-in-place orders have had a dramatic impact on busyness, whereas the COVID-19 pandemic has simultaneously had a significant impact on mental health care resources, decreasing the availability of other options for depression treatment. Further, it is not clear whether the women felt that participating in an intervention study would be time consuming or the actual intervention would be so. Thus, as perinatal depression has increased during this time, the lifestyles of people have changed, and access to in-person delivered health care has greatly decreased [26,27], further highlighting the need for effective, remotely delivered interventions.

Recent studies report that many Americans having depression, stress, or anxiety prefer internet-based mindfulness training over in-person sessions [59,60]. A recent meta-analysis of web-based, webinar-versions of traditional mindfulness-based interventions (eg, Mindfulness-based Stress Reduction) demonstrated a significant beneficial impact on stress, anxiety, depression, and well-being [61]. Although the use of technology for the delivery of mindfulness programs is a major advancement, these studies used web-based versions of the traditional mindfulness interventions in a webinar format with a facilitator, which still required the participants to log on weekly at a specified time, and had the same extensive training duration (eg, 30 or more hours, 2.5 hours per session) and homework requirements (45 min per day) as the in-person versions [62-65]. Although the use of technology can help increase accessibility to the intervention compared with in-person sessions, this type of intervention can still be resource-intensive, and thus not readily scalable, and be still challenging for new mothers who often do not have flexible schedules to attend scheduled sessions. Our study addressed these barriers by offering a web-based, self-paced, and brief mindfulness intervention.

Women in our study valued the variety of options provided by the mindfulness program. They enjoyed both the guided meditations and the breathing techniques and the ability to choose from a male or female voice. Additionally, women also reported the perceived benefits of the program, such as improved stress management, reduced anxiety, improved sleep, and increased physical activity and techniques for managing stress in stressful moments (eg, baby crying). Most of the women contacted for potential recruitment into the study had access to the technology required, and very few participants had trouble downloading or using the app. Although our study was not able to assess an effective daily dose, future studies should consider assessing the effectiveness of various dosages (duration and frequency). However, overall, women enjoyed the intervention.

Our results also suggest the potential for mHealth mindfulness-based interventions to reach women of low socioeconomic status, a population that often does not have access to more traditional treatment options (eg, psychotherapy) for PPD. Nearly a quarter of our sample received Medicaid benefits during the postpartum period. Women of low socioeconomic status are at increased risk of depression, and most women with a depression diagnosis do not receive the treatment they need [66]. Mobile devices are becoming increasingly popular, and approximately 90% of Americans of reproductive age own a smartphone [67]. This type of intervention will also be cost saving for health care systems that have limited resources to offer mental health services, particularly during the COVID-19 pandemic where health care resources are scarce, and distress rates are high.

Limitations

This feasibility study has several limitations. First, we are not able to draw conclusions regarding the efficacy of the mHealth mindfulness intervention on participant-reported outcomes or assess dose-response relationships given the small, single-arm feasibility study design. Without a control group, it is not possible to know whether the observed improvements can be attributed to the intervention. However, we successfully demonstrated the feasibility and acceptability of the intervention, which the study was designed to evaluate. Further, the preliminary efficacy results on patient-reported outcomes assessed immediately after the intervention were promising; however, we acknowledge that these results could be affected by missing data. The positive feasibility, acceptability, and preliminary efficacy findings support the need for a future, larger effectiveness study with longer-term follow-up. Second, the generalizability may be limited because we required the use of technology as part of the intervention. However, as described above, most women who we attempted to recruit owned mobile devices or computers with an internet connection, making the *digital divide* unlikely. The main reasons reported for declining to participate were lack of time or interest and not lack of access to technology. Third, the intervention was only offered in English, and thus non-English speakers were ineligible. Future mHealth mindfulness-based studies should incorporate other languages, given the high burden of depression among minority populations [68]. Of note, Headspace now offers programs in other languages, offering opportunities to conduct more generalizable studies.

Conclusions

Findings from this study demonstrate that postpartum women with moderate to moderately severe depressive symptoms are interested in a mobile-based mindfulness intervention. In addition, we demonstrate that conducting an mHealth mindfulness intervention study in this population is feasible within a large integrated health care system. Larger-scale randomized trials are needed to establish the efficacy and effectiveness of mHealth mindfulness interventions in this population. Highlighting the need for such studies is the recent recommendation by the United States Preventive Services Task Force to refer all women at increased risk of perinatal depression to counseling services. Implementing this recommendation will place a tremendous burden on the health care system, which already has a shortage of mental health care providers [69-74]. Thus, the findings from this study support the need for pragmatic trials, which will provide evidence on the effectiveness of implementing low-cost, technology-based programs for women with moderate to moderately severe PPD symptoms, which are necessary for improving the health of families.

Acknowledgments

This study was funded by Kaiser Permanente Community Benefits and National Institutes of Health (grants number K01MH103444, awarded to LA, and K07 CA166143, awarded to AK).

Authors' Contributions

LA conceptualized and designed the study, obtained funding, and drafted the initial manuscript. AK conceptualized and designed the study and obtained funding. EK, MM, LN prepared study materials, managed study recruitment, collected participant data, and conducted qualitative data analyses. SA extracted the data, created the database for analysis, and analyzed the data. CQ advised on the statistical analyses and interpretation of the results. All authors reviewed and revised the final manuscript.

Conflicts of Interest

None declared.

References

1. Gavin NI, Gaynes BN, Lohr KN, Meltzer-Brody S, Gartlehner G, Swinson T. Perinatal depression: a systematic review of prevalence and incidence. Obstet Gynecol 2005 Nov;106(5 Pt 1):1071-1083. [doi: 10.1097/01.AOG.0000183597.31630.db] [Medline: 16260528]
2. US Preventive Services Task Force, Curry SJ, Krist AH, Owens DK, Barry MJ, Caughey AB, et al. Interventions to prevent perinatal depression: US preventive services task force recommendation statement. J Am Med Assoc 2019 Feb 12;321(6):580-587. [doi: 10.1001/jama.2019.0007] [Medline: 30747971]
3. Gaynes BN, Gavin N, Meltzer-Brody S, Lohr KN, Swinson T, Gartlehner G, et al. Perinatal depression: prevalence, screening accuracy, and screening outcomes. Evid Rep Technol Assess (Summ) 2005 Feb(119):1-8. [doi: 10.1037/e439372005-001] [Medline: 15760246]
4. Committee on Obstetric Practice. The American College of Obstetricians and Gynecologists Committee Opinion no. 630. Screening for perinatal depression. Obstet Gynecol 2015 May;125(5):1268-1271. [doi: 10.1097/01.AOG.0000465192.34779.dc] [Medline: 25932866]
5. Davidson J, Robertson E. A follow-up study of post partum illness, 1946-1978. Acta Psychiatr Scand 1985 May;71(5):451-457. [doi: 10.1111/j.1600-0447.1985.tb05057.x] [Medline: 4013805]
6. Pearlstein T, Howard M, Salisbury A, Zlotnick C. Postpartum depression. Am J Obstet Gynecol 2009 Apr;200(4):357-364 [FREE Full text] [doi: 10.1016/j.ajog.2008.11.033] [Medline: 19318144]
7. Lovejoy JC, Sainsbury A, Stock Conference 2008 Working Group. Sex differences in obesity and the regulation of energy homeostasis. Obes Rev 2009 Mar;10(2):154-167. [doi: 10.1111/j.1467-789X.2008.00529.x] [Medline: 19021872]
8. Beck CT. The effects of postpartum depression on maternal-infant interaction: a meta-analysis. Nurs Res 1995;44(5):298-304. [Medline: 7567486]
9. Wouk K, Stuebe AM, Meltzer-Brody S. Postpartum Mental Health and Breastfeeding Practices: An Analysis Using the 2010-2011 Pregnancy Risk Assessment Monitoring System. Matern Child Health J 2017 Mar;21(3):636-647 [FREE Full text] [doi: 10.1007/s10995-016-2150-6] [Medline: 27449655]
10. Zajicek-Farber ML. Postnatal Depression and Infant Health Practices Among High-Risk Women. J Child Fam Stud 2008 Aug 5;18(2):236-245. [doi: 10.1007/s10826-008-9224-z]
11. Minkovitz CS, Strobino D, Scharfstein D, Hou W, Miller R, Mistry KB, et al. Maternal depressive symptoms and children's receipt of health care in the first 3 years of life. Pediatrics 2005 Feb;115(2):306-314. [doi: 10.1542/peds.2004-0341] [Medline: 15687437]
12. Murray L. The impact of postnatal depression on infant development. J Child Psychol Psychiatry 1992 Mar;33(3):543-561. [doi: 10.1111/j.1469-7610.1992.tb00890.x] [Medline: 1577898]
13. Hay DF, Pawlby S, Sharp D, Asten P, Mills A, Kumar R. Intellectual problems shown by 11-year-old children whose mothers had postnatal depression. J Child Psychol Psychiatry 2001 Oct;42(7):871-889. [doi: 10.1111/1469-7610.00784] [Medline: 11693583]
14. Grace SL, Evindar A, Stewart DE. The effect of postpartum depression on child cognitive development and behavior: a review and critical analysis of the literature. Arch Womens Ment Health 2003 Nov;6(4):263-274. [doi: 10.1007/s00737-003-0024-6] [Medline: 14628179]
15. Ray GT, Croen LA, Habel LA. Mothers of children diagnosed with attention-deficit/hyperactivity disorder: health conditions and medical care utilization in periods before and after birth of the child. Med Care 2009 Jan;47(1):105-114 [FREE Full text] [doi: 10.1097/MLR.0b013e31817e18c0] [Medline: 19106738]
16. Beck CT. The effects of postpartum depression on child development: a meta-analysis. Arch Psychiatr Nurs 1998 Feb;12(1):12-20. [doi: 10.1016/s0883-9417(98)80004-6] [Medline: 9489170]
17. Santos IS, Matijasevich A, Barros AJ, Barros FC. Antenatal and postnatal maternal mood symptoms and psychiatric disorders in pre-school children from the 2004 Pelotas Birth Cohort. J Affect Disord 2014 Aug;164:112-117 [FREE Full text] [doi: 10.1016/j.jad.2014.04.033] [Medline: 24856563]
18. Luca DL, Garlow N, Staatz C, Margiotta C, Zivin K. Societal Costs of Untreated Perinatal Mood and Anxiety Disorders in the United States. Mathematica Policy Research. 2019. URL: https://www.mathematica.org/our-publications-and-findings/publications/societal-costs-of-untreated-perinatal-mood-and-anxiety-disorders-in-the-united-states [accessed 2020-03-15]
19. Siu AL, US Preventive Services Task Force (USPSTF), Bibbins-Domingo K, Grossman DC, Baumann LC, Davidson KW, et al. Screening for Depression in Adults: US Preventive Services Task Force Recommendation Statement. JAMA 2016 Jan 26;315(4):380-387. [doi: 10.1001/jama.2015.18392] [Medline: 26813211]
20. Committee on Obstetric Practice. The American College of Obstetricians and Gynecologists Committee Opinion no. 630. Screening for perinatal depression. Obstet Gynecol 2015 May;125(5):1268-1271. [doi: 10.1097/01.AOG.0000465192.34779.dc] [Medline: 25932866]
21. Earls MF, Committee on Psychosocial Aspects of ChildFamily Health American Academy of Pediatrics. Incorporating recognition and management of perinatal and postpartum depression into pediatric practice. Pediatrics 2010 Nov;126(5):1032-1039. [doi: 10.1542/peds.2010-2348] [Medline: 20974776]
22. Mental Health Care Health Professional Shortage Areas (HPSAs). Kaiser Family Foundation. URL: https://tinyurl.com/y9u2g69b [accessed 2019-09-30]

23. Goodman JH. Women's attitudes, preferences, and perceived barriers to treatment for perinatal depression. Birth 2009 Mar;36(1):60-69. [doi: 10.1111/j.1523-536X.2008.00296.x] [Medline: 19278385]
24. Dimidjian S, Goodman S. Nonpharmacologic intervention and prevention strategies for depression during pregnancy and the postpartum. Clin Obstet Gynecol 2009 Sep;52(3):498-515 [FREE Full text] [doi: 10.1097/GRF.0b013e3181b52da6] [Medline: 19661764]
25. New Poll: Covid-19 Impacting Mental Well-being: Americans Feeling Anxious, Especially for Loved Ones; Older Adults Are Less Anxious2020. American Psychiatric Association. URL: https://www.psychiatry.org/newsroom/news-releases/new-poll-covid-19-impacting-mental-well-being-americans-feeling-anxious-especially-for-loved-ones-older-adults-are-less-anxious [accessed 2020-08-12]
26. The Implications of COVID-19 for Mental Health and Substance Use Issue Brief 2020. KFF - Health Policy Analysis, Polling and Journalism. URL: https://www.kff.org/coronavirus-covid-19/issue-brief/the-implications-of-covid-19-for-mental-health-and-substance-use/ [accessed 2020-08-09]
27. Li L. Challenges and Priorities in Responding to COVID-19 in Inpatient Psychiatry. Psychiatr Serv 2020 Jun 01;71(6):624-626. [doi: 10.1176/appi.ps.202000166] [Medline: 32321388]
28. Kabat-Zinn J. Full Catastrophe Living: Using the Wisdom of Your Body and Mind to Face Stress, Pain, and Illness. New York, USA: Delacorte; 1990.
29. Creswell JD. Mindfulness Interventions. Annu Rev Psychol 2017 Jan 03;68:491-516. [doi: 10.1146/annurev-psych-042716-051139] [Medline: 27687118]
30. Pagnini F, Philips D. Being mindful about mindfulness. Lancet Psychiatry 2015 Apr;2(4):288-289. [doi: 10.1016/S2215-0366(15)00041-3] [Medline: 26360065]
31. Vieten C, Astin J. Effects of a mindfulness-based intervention during pregnancy on prenatal stress and mood: results of a pilot study. Arch Womens Ment Health 2008;11(1):67-74. [doi: 10.1007/s00737-008-0214-3] [Medline: 18317710]
32. Guardino CM, Dunkel Schetter C, Bower JE, Lu MC, Smalley SL. Randomised controlled pilot trial of mindfulness training for stress reduction during pregnancy. Psychol Health 2014;29(3):334-349 [FREE Full text] [doi: 10.1080/08870446.2013.852670] [Medline: 24180264]
33. Dimidjian S, Goodman SH, Felder JN, Gallop R, Brown AP, Beck A. Staying well during pregnancy and the postpartum: A pilot randomized trial of mindfulness-based cognitive therapy for the prevention of depressive relapse/recurrence. J Consult Clin Psychol 2016 Feb;84(2):134-145 [FREE Full text] [doi: 10.1037/ccp0000068] [Medline: 26654212]
34. Lever Taylor B, Cavanagh K, Strauss C. The Effectiveness of Mindfulness-Based Interventions in the Perinatal Period: A Systematic Review and Meta-Analysis. PLoS One 2016;11(5):e0155720 [FREE Full text] [doi: 10.1371/journal.pone.0155720] [Medline: 27182732]
35. Khoury B, Sharma M, Rush SE, Fournier C. Mindfulness-based stress reduction for healthy individuals: A meta-analysis. J Psychosom Res 2015 Jun;78(6):519-528. [doi: 10.1016/j.jpsychores.2015.03.009] [Medline: 25818837]
36. Segal Z, Williams J, Teasdale J. Mindfulness Based Cognitive Therapy for Depression: a New Approach to Preventing Relapse. New York, USA: Guilford Press; 2002.
37. Gordon N. Similarity of the Adult Kaiser Permanente Membership in Northern California to the Insured and General Population in Northern California: Statistics from the 2011 California Health Interview Survey. Division of Research. 2015. URL: https://divisionofresearch.kaiserpermanente.org/projects/memberhealthsurvey/SiteCollectionDocuments/chis_non_kp_2011.pdf [accessed 2020-03-14]
38. Gordon N. A Comparison of Sociodemographic and Health Characteristics of the Kaiser Permanente Northern California Membership Derived from Two Data Sources: The 2008 Member Health Survey and the. Division of Research. 2007. URL: https://divisionofresearch.kaiserpermanente.org/projects/memberhealthsurvey/SiteCollectionDocuments/chis_mhs_comparison_2008.pdf [accessed 2020-03-13]
39. Flanagan T, Avalos LA. Perinatal Obstetric Office Depression Screening and Treatment: Implementation in a Health Care System. Obstet Gynecol 2016 May;127(5):911-915 [FREE Full text] [doi: 10.1097/AOG.0000000000001395] [Medline: 27054937]
40. Avalos LA, Raine-Bennett T, Chen H, Adams AS, Flanagan T. Improved Perinatal Depression Screening, Treatment, and Outcomes With a Universal Obstetric Program. Obstet Gynecol 2016 May;127(5):917-925 [FREE Full text] [doi: 10.1097/AOG.0000000000001403] [Medline: 27054938]
41. Mani M, Kavanagh DJ, Hides L, Stoyanov SR. Review and Evaluation of Mindfulness-Based iPhone Apps. JMIR Mhealth Uhealth 2015 Aug 19;3(3):e82 [FREE Full text] [doi: 10.2196/mhealth.4328] [Medline: 26290327]
42. Economides M, Martman J, Bell MJ, Sanderson B. Improvements in Stress, Affect, and Irritability Following Brief Use of a Mindfulness-based Smartphone App: A Randomized Controlled Trial. Mindfulness (N Y) 2018;9(5):1584-1593 [FREE Full text] [doi: 10.1007/s12671-018-0905-4] [Medline: 30294390]
43. Champion L, Economides M, Chandler C. The efficacy of a brief app-based mindfulness intervention on psychosocial outcomes in healthy adults: A pilot randomised controlled trial. PLoS One 2018;13(12):e0209482 [FREE Full text] [doi: 10.1371/journal.pone.0209482] [Medline: 30596696]
44. Wen L, Sweeney TE, Welton L, Trockel M, Katznelson L. Encouraging Mindfulness in Medical House Staff via Smartphone App: A Pilot Study. Acad Psychiatry 2017 Oct;41(5):646-650. [doi: 10.1007/s40596-017-0768-3] [Medline: 28795335]

45. Howells A, Ivtzan I, Eiroa-Orosa FJ. Putting the ‘app’ in Happiness: A Randomised Controlled Trial of a Smartphone-Based Mindfulness Intervention to Enhance Wellbeing. J Happiness Stud 2014 Oct 29;17(1):163-185. [doi: 10.1007/s10902-014-9589-1]
46. Bennike I, Wieghorst A, Kirk U. Online-based Mindfulness Training Reduces Behavioral Markers of Mind Wandering. J Cogn Enhanc 2017 Apr 25;1(2):172-181. [doi: 10.1007/s41465-017-0020-9]
47. Mistler LA, Ben-Zeev D, Carpenter-Song E, Brunette MF, Friedman MJ. Mobile Mindfulness Intervention on an Acute Psychiatric Unit: Feasibility and Acceptability Study. JMIR Ment Health 2017 Aug 21;4(3):e34 [FREE Full text] [doi: 10.2196/mental.7717] [Medline: 28827214]
48. Forbes G, Newton S, Cantalapiedra Calvete C, Birch J, Dodds J, Steed L, et al. MEMPHIS: a smartphone app using psychological approaches for women with chronic pelvic pain presenting to gynaecology clinics: a randomised feasibility trial. BMJ Open 2020 Mar 12;10(3):e030164. [doi: 10.1136/bmjopen-2019-030164] [Medline: 32165549]
49. Kroenke K, Strine TW, Spitzer RL, Williams JB, Berry JT, Mokdad AH. The PHQ-8 as a measure of current depression in the general population. J Affect Disord 2009 Apr;114(1-3):163-173. [doi: 10.1016/j.jad.2008.06.026] [Medline: 18752852]
50. Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. J Health Soc Behav 1983 Dec;24(4):385-396. [Medline: 6668417]
51. Akman T, Yavuzsen T, Sevgen Z, Ellidokuz H, Yilmaz AU. Evaluation of sleep disorders in cancer patients based on Pittsburgh Sleep Quality Index. Eur J Cancer Care (Engl) 2015 Jul;24(4):553-559. [doi: 10.1111/ecc.12296] [Medline: 25727241]
52. Baer RA, Smith GT, Hopkins J, Krietemeyer J, Toney L. Using self-report assessment methods to explore facets of mindfulness. Assessment 2006 Mar;13(1):27-45. [doi: 10.1177/1073191105283504] [Medline: 16443717]
53. Hebden L, Cook A, van der Ploeg HP, Allman-Farinelli M. Development of smartphone applications for nutrition and physical activity behavior change. JMIR Res Protoc 2012 Aug 22;1(2):e9 [FREE Full text] [doi: 10.2196/resprot.2205] [Medline: 23611892]
54. Hölzel BK, Carmody J, Vangel M, Congleton C, Yerramsetti SM, Gard T, et al. Mindfulness practice leads to increases in regional brain gray matter density. Psychiatry Res 2011 Jan 30;191(1):36-43 [FREE Full text] [doi: 10.1016/j.pscychresns.2010.08.006] [Medline: 21071182]
55. Hölzel BK, Lazar SW, Gard T, Schuman-Olivier Z, Vago DR, Ott U. How Does Mindfulness Meditation Work? Proposing Mechanisms of Action From a Conceptual and Neural Perspective. Perspect Psychol Sci 2011 Nov;6(6):537-559. [doi: 10.1177/1745691611419671] [Medline: 26168376]
56. Kuyken W, Warren FC, Taylor RS, Whalley B, Crane C, Bondolfi G, et al. Efficacy of Mindfulness-Based Cognitive Therapy in Prevention of Depressive Relapse: An Individual Patient Data Meta-analysis From Randomized Trials. JAMA Psychiatry 2016 Jun 01;73(6):565-574 [FREE Full text] [doi: 10.1001/jamapsychiatry.2016.0076] [Medline: 27119968]
57. Schumer MC, Lindsay EK, Creswell JD. Brief mindfulness training for negative affectivity: A systematic review and meta-analysis. J Consult Clin Psychol 2018 Jul;86(7):569-583 [FREE Full text] [doi: 10.1037/ccp0000324] [Medline: 29939051]
58. Davenport MH, Meyer S, Meah VL, Strynadka MC, Khurana R. Moms Are Not OK: COVID-19 and Maternal Mental Health. Front. Glob. Womens Health 2020 Jun 19;1:-. [doi: 10.3389/fgwh.2020.00001]
59. Wahbeh H, Lane JB, Goodrich E, Miller M, Oken BS. One-on-one Mindfulness Meditation Trainings in a Research Setting. Mindfulness (N Y) 2014 Feb 01;5(1):88-99 [FREE Full text] [doi: 10.1007/s12671-012-0155-9] [Medline: 24554970]
60. Wahbeh H, Oken BS. Internet Mindfulness Meditation Intervention for the General Public: Pilot Randomized Controlled Trial. JMIR Ment Health 2016 Aug 08;3(3):e37 [FREE Full text] [doi: 10.2196/mental.5900] [Medline: 27502759]
61. Spijkerman MP, Pots WT, Bohlmeijer ET. Effectiveness of online mindfulness-based interventions in improving mental health: A review and meta-analysis of randomised controlled trials. Clin Psychol Rev 2016 Apr;45:102-114 [FREE Full text] [doi: 10.1016/j.cpr.2016.03.009] [Medline: 27111302]
62. Compen F, Bisseling E, Schellekens M, Donders R, Carlson L, van der Lee M, et al. Face-to-Face and Internet-Based Mindfulness-Based Cognitive Therapy Compared With Treatment as Usual in Reducing Psychological Distress in Patients With Cancer: A Multicenter Randomized Controlled Trial. J Clin Oncol 2018 Aug 10;36(23):2413-2421. [doi: 10.1200/JCO.2017.76.5669] [Medline: 29953304]
63. Lengacher CA, Reich RR, Ramesar S, Alinat CB, Moscoso M, Cousin L, et al. Feasibility of the mobile mindfulness-based stress reduction for breast cancer (mMBSR(BC)) program for symptom improvement among breast cancer survivors. Psychooncology 2018 Feb;27(2):524-531. [doi: 10.1002/pon.4491] [Medline: 28665541]
64. Bruggeman-Everts FZ, Wolvers MD, van de Schoot R, Vollenbroek-Hutten MM, Van der Lee ML. Effectiveness of Two Web-Based Interventions for Chronic Cancer-Related Fatigue Compared to an Active Control Condition: Results of the "Fitter na kanker" Randomized Controlled Trial. J Med Internet Res 2017 Oct 19;19(10):e336 [FREE Full text] [doi: 10.2196/jmir.7180] [Medline: 29051138]
65. Zernicke KA, Campbell TS, Speca M, McCabe-Ruff K, Flowers S, Carlson LE. A randomized wait-list controlled trial of feasibility and efficacy of an online mindfulness-based cancer recovery program: the eTherapy for cancer applying mindfulness trial. Psychosom Med 2014 May;76(4):257-267. [doi: 10.1097/PSY.0000000000000053] [Medline: 24804884]

66. Kozhimannil KB, Trinacty CM, Busch AB, Huskamp HA, Adams AS. Racial and ethnic disparities in postpartum depression care among low-income women. Psychiatr Serv 2011 Jun;62(6):619-625 [FREE Full text] [doi: 10.1176/ps.62.6.pss6206_0619] [Medline: 21632730]
67. Mobile Fact Sheet. Pew Research Center. 2017. URL: http://www.pewinternet.org/fact-sheet/mobile/ [accessed 2020-10-02]
68. Mukherjee S, Trepka MJ, Pierre-Victor D, Bahelah R, Avent T. Racial/Ethnic Disparities in Antenatal Depression in the United States: A Systematic Review. Matern Child Health J 2016 Sep;20(9):1780-1797. [doi: 10.1007/s10995-016-1989-x] [Medline: 27016352]
69. US Preventive Services Task Force, Curry SJ, Krist AH, Owens DK, Barry MJ, Caughey AB, et al. Interventions to Prevent Perinatal Depression: US Preventive Services Task Force Recommendation Statement. JAMA 2019 Feb 12;321(6):580-587. [doi: 10.1001/jama.2019.0007] [Medline: 30747971]
70. O'Connor E, Senger CA, Henninger ML, Coppola E, Gaynes BN. Interventions to Prevent Perinatal Depression: Evidence Report and Systematic Review for the US Preventive Services Task Force. JAMA 2019 Feb 12;321(6):588-601. [doi: 10.1001/jama.2018.20865] [Medline: 30747970]
71. Freeman MP. Perinatal Depression: Recommendations for Prevention and the Challenges of Implementation. JAMA 2019 Feb 12;321(6):550-552. [doi: 10.1001/jama.2018.21247] [Medline: 30747953]
72. Avalos LA, Flanagan T, Li D. Preventing Perinatal Depression to Improve Maternal and Child Health-a Health Care Imperative. JAMA Pediatr 2019 Apr 01;173(4):313-314. [doi: 10.1001/jamapediatrics.2018.5491] [Medline: 30747947]
73. Felder JN. Implementing the USPSTF Recommendations on Prevention of Perinatal Depression-Opportunities and Challenges. JAMA Intern Med 2019 Apr 01;179(4):467-468 [FREE Full text] [doi: 10.1001/jamainternmed.2018.7729] [Medline: 30747945]
74. Wisner KL, Miller ES, Tandon D. Attention to Prevention-Can We Stop Perinatal Depression Before It Starts? JAMA Psychiatry 2019 Apr 01;76(4):355-356 [FREE Full text] [doi: 10.1001/jamapsychiatry.2018.4085] [Medline: 30747943]

Abbreviations

ANOVA: analysis of variance
EHR: electronic health record
KPNC: Kaiser Permanente Northern California
mHealth: mobile health
PHQ-8: 8-item Patient Health Questionnaire
PHQ-9: 9-item Patient Health Questionnaire
PPD: postpartum depression

Edited by G Eysenbach; submitted 31.03.20; peer-reviewed by R Haase, M Schellekens, L Mistler, G Forbes; comments to author 29.06.20; revised version received 21.08.20; accepted 22.09.20; published 12.11.20.

Please cite as:
Avalos LA, Aghaee S, Kurtovich E, Quesenberry Jr C, Nkemere L, McGinnis MK, Kubo A
A Mobile Health Mindfulness Intervention for Women With Moderate to Moderately Severe Postpartum Depressive Symptoms: Feasibility Study
JMIR Ment Health 2020;7(11):e17405
URL: https://mental.jmir.org/2020/11/e17405
doi:10.2196/17405
PMID:33180028

Original Paper

The Reach, Use, and Impact of a Free mHealth Mindfulness App in the General Population: Mobile Data Analysis

Elissa Kozlov[1,2], PhD; Erin Bantum[3], PhD; Ian Pagano[3], PhD; Robyn Walser[4], PhD; Kelly Ramsey[4], BA; Katherine Taylor[4], MPH, PsyD; Beth Jaworski[4], PhD; Jason Owen[4], PhD

[1]Institute for Health, Health Policy and Aging Research, Rutgers University, New Brunswick, NJ, United States
[2]School of Public Health, Rutgers University, New Brunswick, NJ, United States
[3]Cancer Prevention in the Pacific, University of Hawaii Cancer Center, Honolulu, HI, United States
[4]National Center for PTSD Research, VA Palo Alto Healthcare System, Menlo Park, CA, United States

Corresponding Author:
Elissa Kozlov, PhD
Institute for Health, Health Policy and Aging Research
Rutgers University
112 Paterson St
New Brunswick, NJ, 08901
United States
Phone: 1 9147153012
Email: elissa.kozlov@rutgers.edu

Abstract

Background: As smartphones are now used by most Americans, it is increasingly possible for mental health mobile apps to be disseminated to the general public. However, little is known about how mobile mental health apps are used by the general population outside of a controlled research design.

Objective: Our objective is to describe how the general population engages with Mindfulness Coach, an iOS- and Android-based app designed to deliver a mindfulness training course.

Methods: Using anonymous download and analytics data, we characterized the reach, usage, retention, and impact of Mindfulness Coach. We included mobile analytics data from all unique downloads of Mindfulness Coach between August 1, 2018, and April 8, 2019 (N=104,067) as well as starred reviews from all Mindfulness Coach users who provided reviews of the app as of March 1, 2020. Mindfulness characteristics were measured by an in-app assessment using the Five-Facet Mindfulness Questionnaire–Short Form (FFMQ-SF).

Results: Users engaged, on average, in 4.3 visits to the app (SD 8.8; median 2; 90th percentile 8) and associated with an average total of 49.2 interactions with the app (ie, clicks within the app) (SD 113.8; median 19; 90th percentile 105). Users spent an average of 16.2 minutes (SD 63.1) engaged with the app over the full study time period. There were strong linear effects of app engagement on total FFMQ-SF scores. For example, FFMQ-SF scores were associated with more time spent engaged with the app (R^2=.23; P<.001). Mindfulness Coach has been reviewed in the Google Play Store 3415 times, with an average rating of 4.7 out of 5 stars, and over 2000 times in the Apple App Store, with an average rating of 4.8 out of 5 stars.

Conclusions: These findings suggest that Mindfulness Coach has achieved substantial and sustained reach in the general population; however, it was used less frequently by many downloaders than researchers and designers intended. There was a subpopulation of users who engaged in the app regularly over an extended period of time, and there was a clear relationship between app use and improvements in mindfulness. To strengthen Mindfulness Coach's public health impact, more research is needed to understand who is using the app and how, and to design strategies to increase user engagement in order for users to receive a larger dose of mindfulness treatment.

(JMIR Ment Health 2020;7(11):e23377) doi:10.2196/23377

KEYWORDS

mHealth; mindfulness; mHealth psychotherapy; mHealth mindfulness; public health; self-management; mental health

Introduction

A recent National Academy of Medicine report recommended increased emphasis on disseminating and implementing evidence-based psychotherapies [1] in order to have maximal public health impact. Mindfulness therapy (MT) is a promising, nonpharmacological approach to manage various types of psychological distress [2-5]. Using guided meditation, psychoeducation, and targeted exercises, MT teaches people how to pay attention to the present moment without judgment. MT has proven efficacy and effectiveness in reducing anxiety and posttraumatic stress symptoms in diverse populations [2-4,6,7]. While the evidence base for MT is growing, traditional MT (ie, 8 or more sessions of face-to-face treatment with trained providers [2]) is likely not a realistic treatment model for the general population due to the lack of trained personnel, time constraints, reimbursement issues, and patient and provider availability [8,9]. An innovative delivery model is required to overcome these barriers. Mobile health (mHealth) can be useful to deliver behavioral interventions, as it surmounts many obstacles to traditional psychotherapy [10-12].

There is growing evidence that mHealth apps are an effective mechanism to deliver accessible mental health care [11]. Various studies demonstrate that mHealth interventions are both feasible and effective in teaching adults with depression or anxiety skills to manage their symptoms [13-18] and can be helpful to teach skills to manage other conditions, such as chronic pain [19-22]. There is preliminary evidence that apps can be utilized to disseminate aspects of MT so that individuals can learn and practice aspects of mindfulness on their own [23-25]; however, little is known about how mHealth-delivered MT is used by the general population without individualized guidance from practitioners or outside of a controlled research design.

It is essential that we better understand how mHealth strategies to manage psychological distress can be beneficial to the general population. Mobile mental health apps have the potential to reach millions of individuals who are unable to access individualized face-to-face or video therapy services [19,20]. There are many mobile mental health apps currently available, but while research is emerging on how apps work in optimal, highly supported conditions [12,21,22,26,27], less is known about how apps are used by the general population without additional input from providers. For example, what is the most efficacious dosage and duration of use, and how are people inclined to use the app without specific guidance? Without information about general public use patterns, it will be difficult to design apps to reach specific populations and address specific conditions and needs. In order to optimize and tailor user experiences with an app, we must understand how the app is used without instruction or guidance. This information has the potential to inform developers and clinicians about how users naturally engage with apps, offering the opportunity to develop targeted recommendations for enhanced use. Prior research with mental health mobile apps has revealed that many users will download the app and use it only once [28]. Further research on natural use and attrition patterns will allow developers to develop strategies to enhance sustained usage, such as setting in-app reminders to encourage sustained usage.

The purpose of this paper is to describe how the general population currently engages with Mindfulness Coach, an iOS- and Android-based app designed to deliver a mindfulness training course based on the adaptation of several US Department of Veterans Affairs (VA) protocols. Based on anonymous download and analytics data, we aim to characterize the reach, usage, retention, and impact of Mindfulness Coach, a publicly available mobile app.

Methods

Overview

Mindfulness Coach is a native iOS- and Android-based app designed to deliver mindfulness training adapted from several VA protocols. The app is intended to provide a highly engaging introduction to MT and is tailored to users who may be skeptical about meditation practices by providing simple instructions and brief exercises. After downloading the app, users are provided with a brief tutorial that introduces the major features (ie, training plan, practice exercises, learning topics, and tracking). The training plan attempts to provide users with direction on how to use the app by gently introducing the user to each of the components within the app. Users can set push notification reminders in the settings section if they choose. The app delivers 14 sessions (ie, levels), each culminating in a meditative exercise. The app provides a training plan, evidence-based mindfulness audio exercises, assessment using the Five-Facet Mindfulness Questionnaire (FFMQ) [29], and education about MT. The app transmits deidentified usage data to a secure server using methods approved under the VA's Technical Reference Model [30].

Data Sources

Data were derived from 2 sources. First, we included mobile analytics data from all unique downloads of Mindfulness Coach between August 1, 2018, and April 8, 2019 (N=104,067). Second, we included starred reviews from all Mindfulness Coach users who provided reviews of the app on either the Apple App Store or the Google Play Store as of March 1, 2020.

Measures

The Five-Facet Mindfulness Questionnaire–Short Form

The FFMQ is a measure of the 5 facets of the tendency to be mindful in daily life: observing, describing, acting with awareness, nonreactivity to inner experience, and nonjudging of inner experience. As a measure of impact, the app collects data on the FFMQ-Short Form (FFMQ-SF). The app's training plan recommends that users complete the FFMQ-SF at levels 1, 7, and 14, and they are provided prompts to do so. Additionally, participants can take the assessment whenever they want by clicking the *Track my Progress* button. While taking the assessment is recommended, it is not required, as users can close out of the assessment at any time. Mindfulness Coach administers only 4 of the FFMQ-SF's 5 subscales: *being observant*, *acting with awareness*, *nonjudging*, and *nonreactivity*. The *describing* subscale was not used because Mindfulness Coach does not provide tools for improving communication skills. The FFMQ-SF without the *describing* subscale consists of 19 items, each measured on a 5-point Likert scale, ranging from 1 (*Never or very rarely true*) to 5 (*Very often or always true*). Each of the 4 subscales has been shown to have adequate internal consistency (α values from .75 to .83) and comparability to the full 39-item FFMQ scale, with strong divergent and convergent validity and sensitivity to change over time [31,32].

Mobile Analytics

For each platform, iOS or Android, we captured basic user engagement measures (ie, number of downloads, active users, number of events within the app during each visit, visit duration, and number of visits) in addition to 2 primary measures of retention across time. Upon initial use of Mindfulness Coach, the app generates a unique, randomly generated string for that installation that is then associated with user engagement measures. The app did not collect any identifiers, such as IP address, location, device identifiers, or phone numbers, nor any other personal information (eg, battery state, data connections, etc) that could be used to identify an individual user [33].

Return use was calculated as the proportion of users who returned to use the app within 1 week of initial download, 1 month of initial download, 3 months of initial download, 6 months of initial download, and 12 months of initial download. Rolling retention was measured as the number of active days, weeks, and months of use of the app between the time of initial download and final use of the app during the observation period. Finally, sequences of events undertaken within the app were used to capture users' navigation through the content pages of the app across time. Fully nonidentifying, anonymous, and encrypted event sequences were stored using JavaScript Object Notation (JSON) format on a remote Amazon Web Services GovCloud server. The event sequence data contained 9,170,219 records and were parsed using Perl regular expressions in SAS 9.4 (SAS Institute Inc) software. First-time users (ie, those who accepted the end-user license agreement) and returning users were identified. Each session began with the launch of the app and was classified as either a first-time use or a return visit on the basis of whether the end-user license agreement was displayed at launch or not.

For each session type (ie, first-time use vs return visit), specific usage events were tracked, including completing the app orientation, navigation from the home screen to 1 of the 4 primary content areas (ie, training plan, practice now, building expertise, or track progress), and navigation from one content area to another. We capped events at 30 minutes, unless the user had specified a longer value. We defined *visits* as clusters of events less than 30 minutes apart. Each record contained a field for its duration, and we defined *events* as records with nonzero values in the field, indicating that the user spent some amount of time with the app.

Means were determined by first calculating the within-subject means so that each subject's average contributed equally to the grand mean. We considered the alternate method of calculating each visit's sessions, but then users with more visits would contribute more to the grand mean, which could obfuscate potentially relevant data.

Analyses

SAS 9.4 (SAS Institute Inc) software was used to perform all data management and analyses. We calculated descriptive statistics for usage data and the content areas visited. The MIXED procedure ran polynomial (ie, linear, quadratic, and cubic) repeated-measures regression analyses on the FFMQ-SF outcome. The predictor variables were the totals for retention time (ie, time spent in the app), number of visits, number of events, and highest mindfulness level achieved. We ran a separate model for each predictor given a high degree of multicollinearity among the predictor variables.

To further interpret and simplify the data, we created 4 categories of users: those who opened the app only once (*exploratory users*), those who visited the app 2 to 3 times (*limited users*), those who visited the app 4 to 7 times (*moderateusers*), and those who visited the app 8 or more times (*committed users*). Categories were identified based on expected clinical benefit and also provided cut points that provided 4 roughly equal groups in terms of numbers of users in each group.

Results

Reach of Mindfulness Coach

We analyzed 104,067 unique downloads: 62.90% (65,458/104,067) on Android devices and 37.10% (38,609/104,067) on iOS devices. The app has been downloaded by an average of 6720 users per month since its release to the public on January 2019, with the number of users increasing steadily over time (ie, average of 9737 each month since December 2019). The total number of downloads was 278,606 since the app's release on iOS in January 2014 and 147,535 since its release on Android in February 2018. Data available from the Apple App Store for iOS devices suggest that 94% of users accessed the app from a phone and 6% accessed the app from a tablet device (eg, an iPad).

Satisfaction

Mindfulness Coach has been reviewed in the Google Play Store 3415 times, with an average rating of 4.7 out of 5 stars, and over 2000 times in the Apple App Store, with an average rating of 4.8 out of 5 stars.

Use of the App

Elapsed time between first and final uses of the Mindfulness Coach app averaged 4.1 weeks (SD 6.7; median 0; 90th percentile 15). Users engaged, on average, in 4.3 visits to the app (SD 8.8; median 2; 90th percentile 8) and associated with an average total of 49.2 interactions with the app (ie, clicks within the app) (SD 113.8; median 19; 90th percentile 105). Users spent an average of 16.2 minutes (SD 63.1) engaged with the app over the full study time period. See Table 1 for the breakdown of use by user category and Table 2 for distribution of use by user category.

Table 1. Descriptive statistics of use variables by user visits to the app.

Visits, n	Downloads, n	License accepted, n (%)	Level achieved, mean (SD), range	Retention weeks, mean (SD), range	Active minutes, mean (SD), range	Visits, mean (SD), range	Average visit, mean (SD), range	Events, mean (SD), range
All	104,067	Total: 91,371 (87.80) Android: 66,499 (63.90) iOS: 38,609 (37.10)	1.4 (1.4), 1-14	4.1 (6.7), 0-34	16.2 (63.1), 9-5006	4.3 (9.5), 1-506	3.2 (4.4), 0-155	49.2 (113.8), 1-7597
1	40,544	Total: 34,908 (86.10) Android: 25,543 (63.00) iOS: 15,001 (37.00)	1.1 (0.3), 1-10	0 (0), N/A[a]	3.2 (5.3), 0-155	1 (0), N/A	3.2 (5.3), 0-155	15.4 (18.1), 1-275
2-3	32,962	Total: 29,402 (89.20) Android: 20,832 (63.20) iOS: 12,130 (36.80)	1.1 (0.5), 1-14	3.8 (5.6), 0-31	7.3 (9.4), 0-203	2.4 (0.5), 2-3	3.1 (3.9), 0-101	29.8 (28.9), 1-617
4-7	18,528	Total: 16,620 (89.70) Android: 11,543 (62.30) iOS: 6985 (37.70)	1.4 (0.9), 1-14	7.8 (7.3), 0-32	16.3 (17.7), 0-340	5.1 (1.1), 4-7	3.2 (3.4), 0-82	57.1 (48.8), 1-2003
≥8	12,033	Total: 10,421 (86.60) Android: 7557 (62.80) iOS: 4476 (37.20)	2.9 (3.3), 1-14	12.8 (8.4), 0-34	84.8 (167.6), 0-5006	19.3 (22.3), 8-506	4 (3.6), 0-46	204.1 (275.7), 1-7597

[a] N/A: not applicable; all values of the data set are the same, hence, there is no range.

Table 2. Distribution of use variables by user groups.

Use variables	Downloads, n	25th percentile	50th percentile	75th percentile	90th percentile
Level achieved					
All	104,067	1	1	1	2
1 visit	40,544	1	1	1	1
2-3 visits	32,962	1	1	1	2
4-7 visits	18,528	1	1	1	2
≥8 visits	12,033	1	1	3	7
Retention weeks					
All	104,067	0	0	5	15
1 visit	40,544	0	0	0	0
2-3 visits	32,962	0	1	5	12
4-7 visits	18,528	2	5	12	19
≥8 visits	12,033	5	12	20	25
Active minutes					
All	104,067	1	4	13	32
1 visit	40,544	0	1	4	9
2-3 visits	32,962	1	4	10	18
4-7 visits	18,528	4	11	22	37
≥8 visits	12,033	18	40	89	184
Visits					
All	104,067	1	2	4	8
1 visit	40,544	1	1	1	1
2-3 visits	32,962	2	2	3	3
4-7 visits	18,528	4	5	6	7
≥8 visits	12,033	9	12	20	36
Events					
All	104,067	7	19	53	105
1 visit	40,544	4	8	18	42
2-3 visits	32,962	9	19	45	66
4-7 visits	18,528	22	48	79	115
≥8 visits	12,033	69	130	240	440

Retention of Users

We used mobile analytics to characterize average use patterns as well as patterns of use over time. First, we investigated the proportion of users who actively used the app over time. Among individuals who downloaded and opened Mindfulness Coach between August 1, 2018, and April 8, 2019, 54.20% (56,404/104,067) used the app at least once after the first day it was installed, 43.20% (44,957/104,067) used the app at least once beyond the first week when it was downloaded, 30.40% (31,636/104,067) used the app after 1 month from the date it was installed, 17.40% (10,108/104,067) used the app after 3 months from the date it was installed, and 5.60% (5828/104,067) used the app after 6 months from the date it was installed.

Click stream data were analyzed to better understand how all users engaged with the app and to evaluate whether usage patterns differed between first-time and returning users (see Table 3). Among those using the app for the first time and who visited a content area (N=73,119), the first content area visited was *mindfulness training* at 63.23% (46,236/73,119), *practice now* at 23.03% (16,837/73,119), *build expertise* at 4.31% (3152/73,119), *track progress* at 2.35% (1719/73,119), or *other* at 6.28% (4594/73,119). Upon initially visiting the app, 29.74% (30,948/104,067) did not visit any of the key content areas of the app, 47.19% (49,110/104,067) visited only a single content area, and 23.07% (24,009/104,067) visited 2 or more content areas. Across all returning visits to the app (N=210,177), the first content areas visited were *mindfulness training* (89,433/210,177, 42.55%), *practice now* (96,026/210,177, 45.69%), *track progress* (10,728/210,177, 5.10%), and *build expertise* (9591/210,177, 4.56%) (see Table 3).

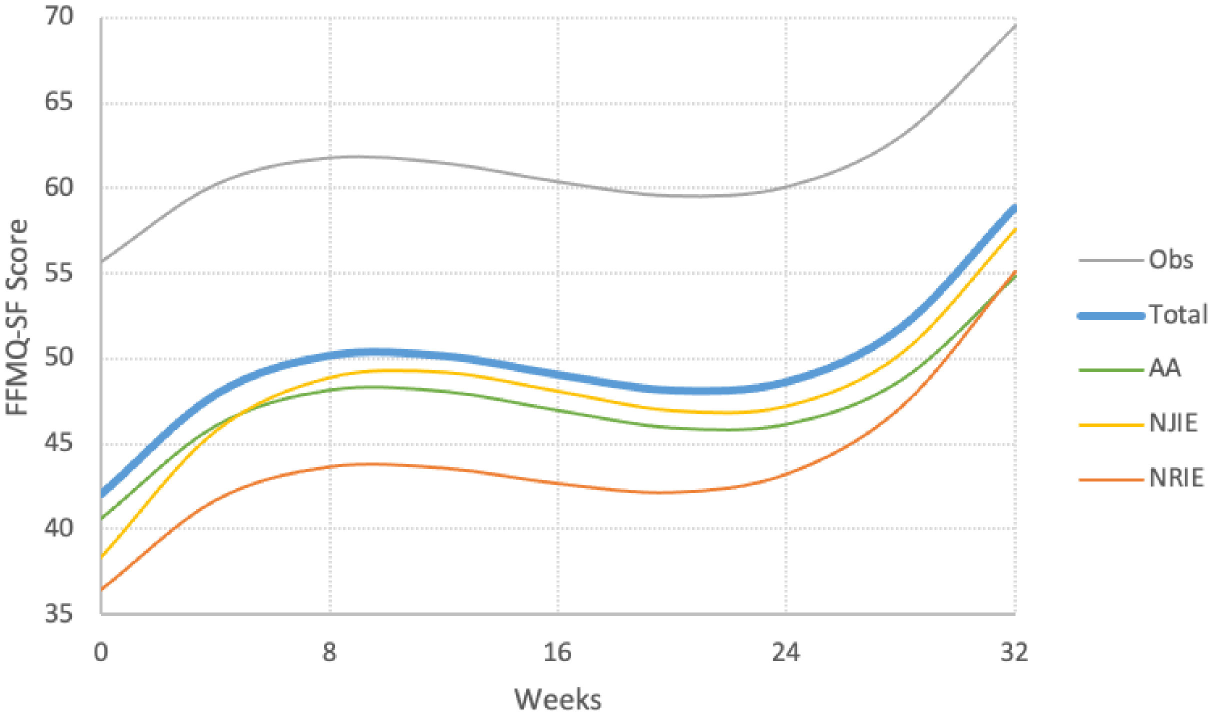
Table 3. Detailed session analysis for Mindfulness Coach by first-time users and returning users, from August 2018 to April 2019.

Session analysis	First visit users, n (%)	Returning visit users, n (%)	Between-group differences, *P* value
First content area visited (first visit users: N=73,119; returning visit users: N=210,177)			**<.001**
Badges	581 (0.79)	1430 (0.68)	
Build expertise	3152 (4.31)	9591 (4.56)	
Mindfulness training	46,236 (63.23)	89,433 (42.55)	
Other	4594 (6.28)	2969 (1.41)	
Practice now	16,837 (23.03)	96,026 (45.69)	
Track progress	1719 (2.35)	10,728 (5.10)	
Any content area visited (first visit users: N=107,430; returning visit users: N=272,682)			**<.001**
Badges	1481 (1.38)	3399 (1.25)	
Build expertise	9266 (8.63)	21,709 (7.96)	
Mindfulness training	52,975 (49.31)	101,636 (37.27)	
Other	6888 (6.41)	6303 (2.31)	
Practice now	29,131 (27.12)	115,693 (42.43)	
Track progress	7689 (7.16)	23,942 (8.78)	
Number of content areas visited (first visit users: N=104,067; returning visit users: N=340,955)			**<.001**
0	30,948 (29.74)	130,779 (38.36)	
1	49,110 (47.19)	164,507 (48.25)	
2	16,542 (15.90)	33,084 (9.70)	
3	5065 (4.87)	8857 (2.60)	
4	1969 (1.89)	3210 (0.94)	
5	433 (0.42)	518 (0.15)	

Change in Mindfulness Mastery Over Time

There was a significant cubic effect of time on FFMQ-SF (R^2=.16, P<.001) and each of the 4 subscales: *being observant* (R^2=.06, P<.001), *acting with awareness* (R^2=.06, P<.001), *nonjudging* (R^2=.08, P<.001), and *nonreactivity* (R^2=.07, P<.001). Slopes for FFMQ-SF over time rose rapidly between 0 and 8 weeks of app use, leveled off from 8 to 24 weeks, and began to rise again between 25 and 32 weeks after initial use of the app (see Figure 1).

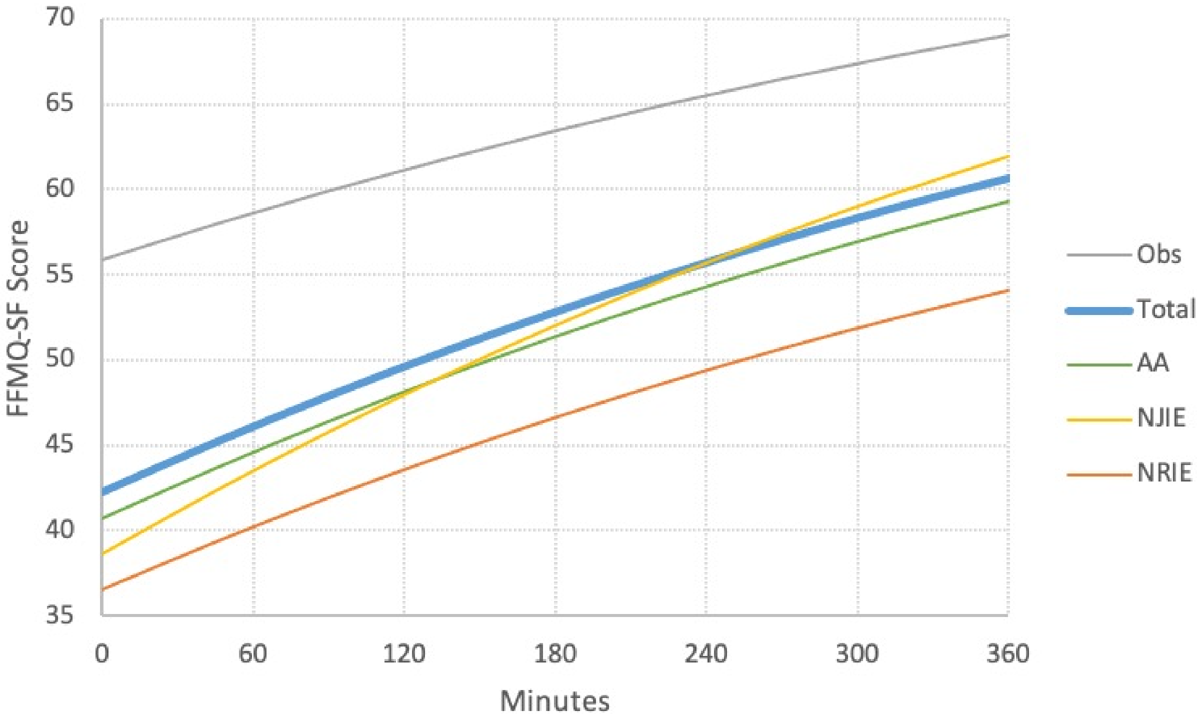
Figure 1. Association between time since installation and being observant (Obs), acting with awareness (AA), nonjudging of inner experience (NJIE), and nonreactivity to inner experience (NRIE). FFMQ-SF: Five-Facet Mindfulness Questionnaire-Short Form.



Mindfulness Mastery as a Function of Engagement With the App

There were also strong linear effects of app engagement on total FFMQ-SF scores. FFMQ-SF scores were associated with more hours spent engaged with the app (R^2=.23, P<.001) (see Figure 2), total number of visits to the app (R^2=.25, P<.001), and number of interactions (ie, events) with the app (R^2=.28, P<.001).

Figure 2. Association between duration of app use and being observant (Obs), acting with awareness (AA), nonjudging of inner experience (NJIE), and nonreactivity to inner experience (NRIE). FFMQ-SF: Five-Facet Mindfulness Questionnaire-Short Form.



Discussion

Principal Findings

This natural-use investigation of an mHealth MT app revealed that users vary tremendously in how they use the app in a natural setting. The app was reviewed favorably by users who chose to leave a review, and increased engagement with the app was associated with improved scores on a measure of mindfulness mastery. Because Mindfulness Coach and other similar mHealth mental health apps will be used by many more people over time, this study provides useful information about how the app is used “in the wild.”

Our results revealed the typical use case for mobile app users, which needs to be considered when planning and implementing app-related interventions. Understanding typical use allows mobile interventionists to consider strategies to enhance reach and adherence when creating interventions in this format. While we do not yet know the optimal frequency and duration of use of Mindfulness Coach, this paper helps us understand that many users’ natural inclination is to use the app infrequently and for a short period of time. Mindfulness Coach is currently being examined in open clinical trials, so more information to help understand optimal dose and duration will be forthcoming.

For the highest-engagement group, committed users, the app reached 12,033 users over the span of 8 months. High-engagement users averaged 84 minutes of in-app mindfulness-based training and practice. This suggests that for at least some individuals, the app is highly engaging. Importantly, this study was not able to measure if users were practicing mindfulness and meditating outside of the app. It is possible that mindfulness skills are being transferred and practiced without the app, so our metrics of use are potentially an underestimation of amount of time per week users spend engaged in some type of mindfulness practice (eg, breathing, meditation, additional readings, etc).

Similarly, we found a significant positive relationship between app use and FFMQ-SF scores, suggesting that dedicated users experience improvements in mindfulness characteristics, which may in turn convey improved mental health [2,4]. Other areas of improvement that regular use of the app may be related to needs further research. For example, face-to-face MT has demonstrated benefits for improving pain, depression, anxiety, and quality of life [2]. Does Mindfulness Coach, when used regularly, confer similar benefits? Future research is needed to elucidate the dose response of Mindfulness Coach and the associated benefits of using the mobile app.

Our study indicated that Mindfulness Coach is not being used frequently enough or for a long enough duration by many users. Given that increased Mindfulness Coach use is associated with improved FFMQ-SF scores, it is likely that limited Mindfulness Coach use is not maximally impactful. There was an extreme positive skew in engagement, and nearly 30% of first-time users and 38% of returning users access *only* the home screen, suggesting that users are opening but not using the app. Given that the app is currently used in very low doses by many users, it may be advantageous to couple Mindfulness Coach with face-to-face care, which allows therapy to extend beyond the traditional session. The app, with guidance from a therapist, could be used as a tool to practice and learn mindfulness skills in between face-to-face sessions or, alternatively, as the primary intervention with a therapist checking in less frequently. More research is needed to explore app engagement and corresponding effects as adjuncts to face-to-face therapy. Another use of the app in an intervention could be for someone to check in and prompt people to use the app. Ways to provide structure and to tailor the intervention could help participants receive the full benefits of the app.

Lastly, our study revealed that Mindfulness Coach received excellent ratings in the Google Play Store (Android) and Apple App Store (iOS), signaling that people are very satisfied with the app. Though this is a limited subsample of the user population, it indicates that the intervention is well received by at least some proportion of users. Because an intervention would not be useful or engaging if people were not satisfied with the app, high satisfaction scores are likely necessary but not sufficient in determining whether the app is a beneficial intervention.

Limitations

While this study provides information on the natural, unguided, and untailored use of the app, we have no information regarding demographics of users. It would be helpful to know who is using the app and if there are populations for whom the app could be more beneficial. Furthermore, we provided a cursory look at the relationship between app use and mindfulness scores as measured by the FFMQ-SF; however, we were unable to explore how app use impacts the well-being, both mental and physical, of the users. Future studies are needed to further explore downstream effects of using Mindfulness Coach, such as quality of life, depression, anxiety, pain, sleep, etc. Additionally, because completing the FFMQ-SF is optional, it is possible that there is bias in the responders who chose to complete the FFMQ-SF. The app does prompt all users to complete the FFMQ-SF assessment, but users are able to close out of the assessment if they choose. Further research will need to validate and confirm this finding in a less potentially biased sample. Another limitation is that assessments were not routinely administered based on time passed, but rather on level achieved or when a user chose to take an assessment. This lack of uniformity in when and if assessments occurred introduces potential for bias. Future studies should look at more uniform assessments at regular time intervals as well as the longitudinal effects of using Mindfulness Coach.

Conclusions

MT is associated with a broad range of improvements in quality of life and well-being within a range of populations. Mobile apps are accessed regularly by the general population and, thus, represent a potentially ideal way to expand the reach of mental health interventions. The National Center for Posttraumatic Stress Disorder researchers created Mindfulness Coach as a vehicle to deliver mindfulness training, and this study was a first step toward understanding how the app is used, without in-person guidance, by the general population. We looked at data from over 100,000 people and found that the app was used less frequently than the developers and researchers intended by the majority of downloaders, though there was a subpopulation of users (5.60%) who engaged in the app regularly over an extended period of time. We also found a clear relationship between use and improvements in mindfulness as measured by the FFMQ-SF. Future research is needed to understand more specifically who is using the app and how, ways in which we can improve the use of the app, and how to design the app in a way where more users can receive a larger dose of mindfulness treatment.

Conflicts of Interest

None declared.

References

1. Committee on Developing Evidence-Based Standards for Psychosocial Interventions for Mental Disorders, Board on Health Sciences Policy, Institute of Medicine. In: England MJ, Butler AS, Gonzalez ML, editors. Psychosocial Interventions for Mental and Substance Use Disorders: A Framework for Establishing Evidence-Based Standards. Washington, DC: The National Academies Press; 2015.
2. Hofmann SG, Sawyer AT, Witt AA, Oh D. The effect of mindfulness-based therapy on anxiety and depression: A meta-analytic review. J Consult Clin Psychol 2010 Apr;78(2):169-183 [FREE Full text] [doi: 10.1037/a0018555] [Medline: 20350028]
3. Bajaj JS, Ellwood M, Ainger T, Burroughs T, Fagan A, Gavis EA, et al. Mindfulness-based stress reduction therapy improves patient and caregiver-reported outcomes in cirrhosis. Clin Transl Gastroenterol 2017 Jul 27;8(7):e108 [FREE Full text] [doi: 10.1038/ctg.2017.38] [Medline: 28749453]
4. Grossman P, Niemann L, Schmidt S, Walach H. Mindfulness-based stress reduction and health benefits. J Psychosom Res 2004 Jul;57(1):35-43. [doi: 10.1016/s0022-3999(03)00573-7] [Medline: 15256293]
5. Kim YW, Lee S, Choi TK, Suh SY, Kim B, Kim CM, et al. Effectiveness of mindfulness-based cognitive therapy as an adjuvant to pharmacotherapy in patients with panic disorder or generalized anxiety disorder. Depress Anxiety 2009;26(7):601-606. [doi: 10.1002/da.20552] [Medline: 19242985]
6. Possemato K, Kuhn E, Johnson E, Hoffman JE, Owen JE, Kanuri N, et al. Using PTSD Coach in primary care with and without clinician support: A pilot randomized controlled trial. Gen Hosp Psychiatry 2016;38:94-98. [doi: 10.1016/j.genhosppsych.2015.09.005] [Medline: 26589765]
7. Possemato K, Bergen-Cico D, Treatman S, Allen C, Wade M, Pigeon W. A randomized clinical trial of primary care brief mindfulness training for veterans with PTSD. J Clin Psychol 2016 Mar;72(3):179-193. [doi: 10.1002/jclp.22241] [Medline: 26613203]
8. Mohr DC, Hart SL, Howard I, Julian L, Vella L, Catledge C, et al. Barriers to psychotherapy among depressed and nondepressed primary care patients. Ann Behav Med 2006 Dec;32(3):254-258. [doi: 10.1207/s15324796abm3203_12] [Medline: 17107299]
9. Pepin R, Segal DL, Coolidge FL. Intrinsic and extrinsic barriers to mental health care among community-dwelling younger and older adults. Aging Ment Health 2009 Sep;13(5):769-777. [doi: 10.1080/13607860902918231] [Medline: 19882416]
10. 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 05;16(6):e146 [FREE Full text] [doi: 10.2196/jmir.3077] [Medline: 24905070]
11. Price M, Yuen EK, Goetter EM, Herbert JD, Forman EM, Acierno R, et al. mHealth: A mechanism to deliver more accessible, more effective mental health care. Clin Psychol Psychother 2014;21(5):427-436 [FREE Full text] [doi: 10.1002/cpp.1855] [Medline: 23918764]
12. Ben-Zeev D, Schueller SM, Begale M, Duffecy J, Kane JM, Mohr DC. Strategies for mHealth research: Lessons from 3 mobile intervention studies. Adm Policy Ment Health 2015 Mar;42(2):157-167 [FREE Full text] [doi: 10.1007/s10488-014-0556-2] [Medline: 24824311]
13. Dear BF, Zou J, Titov N, Lorian C, Johnston L, Spence J, et al. Internet-delivered cognitive behavioural therapy for depression: A feasibility open trial for older adults. Aust N Z J Psychiatry 2013 Feb;47(2):169-176. [doi: 10.1177/0004867412466154] [Medline: 23152358]
14. Crabb R, Cavanagh K, Proudfoot J, Learmonth D, Rafie S, Weingardt K. Is computerized cognitive-behavioural therapy a treatment option for depression in late-life? A systematic review. Br J Clin Psychol 2012 Nov;51(4):459-464. [doi: 10.1111/j.2044-8260.2012.02038.x] [Medline: 23078214]
15. Jones SL, Hadjistavropoulos HD, Soucy JN. A randomized controlled trial of guided internet-delivered cognitive behaviour therapy for older adults with generalized anxiety. J Anxiety Disord 2016 Jan;37:1-9. [doi: 10.1016/j.janxdis.2015.10.006] [Medline: 26561733]
16. Olthuis JV, Watt MC, Bailey K, Hayden JA, Stewart SH. Therapist-supported internet cognitive behavioural therapy for anxiety disorders in adults. Cochrane Database Syst Rev 2016 Mar 12;3:CD011565 [FREE Full text] [doi: 10.1002/14651858.CD011565.pub2] [Medline: 26968204]

17. Firth J, Torous J, Nicholas J, Carney R, Rosenbaum S, Sarris J. Can smartphone mental health interventions reduce symptoms of anxiety? A meta-analysis of randomized controlled trials. J Affect Disord 2017 Aug 15;218:15-22 [FREE Full text] [doi: 10.1016/j.jad.2017.04.046] [Medline: 28456072]
18. Donker T, Petrie K, Proudfoot J, Clarke J, Birch M, Christensen H. Smartphones for smarter delivery of mental health programs: A systematic review. J Med Internet Res 2013 Nov 15;15(11):e247 [FREE Full text] [doi: 10.2196/jmir.2791] [Medline: 24240579]
19. Clough BA, Casey LM. The smart therapist: A look to the future of smartphones and mHealth technologies in psychotherapy. Prof Psychol Res Pr 2015 Jun;46(3):147-153. [doi: 10.1037/pro0000011]
20. Patrick K, Griswold WG, Raab F, Intille SS. Health and the mobile phone. Am J Prev Med 2008 Aug;35(2):177-181 [FREE Full text] [doi: 10.1016/j.amepre.2008.05.001] [Medline: 18550322]
21. Mikolasek M, Witt CM, Barth J. Adherence to a mindfulness and relaxation self-care app for cancer patients: Mixed-methods feasibility study. JMIR Mhealth Uhealth 2018 Dec 06;6(12):e11271 [FREE Full text] [doi: 10.2196/11271] [Medline: 30522990]
22. Mattila E, Lappalainen R, Välkkynen 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] [Medline: 27468653]
23. Mani M, Kavanagh DJ, Hides L, Stoyanov SR. Review and evaluation of mindfulness-based iPhone apps. JMIR Mhealth Uhealth 2015 Aug 19;3(3):e82 [FREE Full text] [doi: 10.2196/mhealth.4328] [Medline: 26290327]
24. Chittaro L, Vianello A. Computer-supported mindfulness: Evaluation of a mobile thought distancing application on naive meditators. Int J Hum Comput Stud 2014 Mar;72(3):337-348. [doi: 10.1016/j.ijhcs.2013.11.001]
25. Howells A, Ivtzan I, Eiroa-Orosa FJ. Putting the ‘app’ in happiness: A randomised controlled trial of a smartphone-based mindfulness intervention to enhance well-being. J Happiness Stud 2014 Oct 29;17(1):163-185. [doi: 10.1007/s10902-014-9589-1]
26. Crandall A, Cheung A, Young A, Hooper AP. Theory-based predictors of mindfulness meditation mobile app usage: A survey and cohort study. JMIR Mhealth Uhealth 2019 Mar 22;7(3):e10794 [FREE Full text] [doi: 10.2196/10794] [Medline: 30900992]
27. Proudfoot J, Clarke J, Birch M, Whitton AE, Parker G, Manicavasagar V, et al. Impact of a mobile phone and web program on symptom and functional outcomes for people with mild-to-moderate depression, anxiety and stress: A randomised controlled trial. BMC Psychiatry 2013 Nov 18;13:312 [FREE Full text] [doi: 10.1186/1471-244X-13-312] [Medline: 24237617]
28. Owen JE, Jaworski BK, Kuhn E, Makin-Byrd KN, Ramsey KM, Hoffman JE. mHealth in the wild: Using novel data to examine the reach, use, and impact of PTSD Coach. JMIR Ment Health 2015;2(1):e7 [FREE Full text] [doi: 10.2196/mental.3935] [Medline: 26543913]
29. Baer RA, Smith GT, Hopkins J, Krietemeyer J, Toney L. Using self-report assessment methods to explore facets of mindfulness. Assessment 2006 Mar;13(1):27-45. [doi: 10.1177/1073191105283504] [Medline: 16443717]
30. VA Technical Reference Model v 20.10. Office of Information and Technology (OIT). Washington, DC: US Department of Veterans Affairs URL: https://www.oit.va.gov/Services/TRM/TRMHomePage.aspx [accessed 2020-11-02]
31. Christopher MS, Neuser NJ, Michael PG, Baitmangalkar A. Exploring the psychometric properties of the Five-Facet Mindfulness Questionnaire. Mindfulness 2012 Jan 6;3(2):124-131. [doi: 10.1007/s12671-011-0086-x]
32. Bohlmeijer E, ten Klooster PM, Fledderus M, Veehof M, Baer R. Psychometric properties of the five facet mindfulness questionnaire in depressed adults and development of a short form. Assessment 2011 Sep;18(3):308-320. [doi: 10.1177/1073191111408231] [Medline: 21586480]
33. de Montjoye Y, Gambs S, Blondel V, Canright G, de Cordes N, Deletaille S, et al. On the privacy-conscientious use of mobile phone data. Sci Data 2018 Dec 11;5:180286 [FREE Full text] [doi: 10.1038/sdata.2018.286] [Medline: 30532052]

Abbreviations

FFMQ: Five-Facet Mindfulness Questionnaire
FFMQ-SF: Five-Facet Mindfulness Questionnaire–Short Form
JSON: JavaScript Object Notation
mHealth: mobile health
MT: mindfulness therapy
VA: Veterans Affairs

Edited by J Torous; submitted 10.08.20; peer-reviewed by J Duffecy, J Ruzek; comments to author 07.10.20; revised version received 16.10.20; accepted 27.10.20; published 27.11.20.

Please cite as:
Kozlov E, Bantum E, Pagano I, Walser R, Ramsey K, Taylor K, Jaworski B, Owen J
The Reach, Use, and Impact of a Free mHealth Mindfulness App in the General Population: Mobile Data Analysis
JMIR Ment Health 2020;7(11):e23377
URL: http://mental.jmir.org/2020/11/e23377/
doi:10.2196/23377
PMID:33245289

Original Paper

Testing the Efficacy of a Multicomponent, Self-Guided, Smartphone-Based Meditation App: Three-Armed Randomized Controlled Trial

Simon B Goldberg[1,2], PhD; Theodore Imhoff-Smith[1], BSc; Daniel M Bolt[3], PhD; Christine D Wilson-Mendenhall[1], PhD; Cortland J Dahl[1,4], PhD; Richard J Davidson[1,5,6], PhD; Melissa A Rosenkranz[1,6], PhD

[1]Center for Healthy Minds, University of Wisconsin, Madison, WI, United States
[2]Department of Counseling Psychology, University of Wisconsin-Madison, Madison, WI, United States
[3]Department of Educational Psychology, University of Wisconsin-Madison, Madison, WI, United States
[4]Healthy Minds Innovations Inc, Madison, WI, United States
[5]Department of Psychology, University of Wisconsin-Madison, Madison, WI, United States
[6]Department of Psychiatry, University of Wisconsin-Madison, Madison, WI, United States

Corresponding Author:
Simon B Goldberg, PhD
Center for Healthy Minds
University of Wisconsin
625 W Washington Ave
University of Wisconsin-Madison
Madison, WI, 53703
United States
Phone: 1 6082658986
Fax: 1 6082654174
Email: sbgoldberg@wisc.edu

Abstract

Background: A growing number of randomized controlled trials (RCTs) suggest psychological benefits associated with meditation training delivered via mobile health. However, research in this area has primarily focused on mindfulness, only one of many meditative techniques.

Objective: This study aims to evaluate the efficacy of 2 versions of a self-guided, smartphone-based meditation app—the Healthy Minds Program (HMP)—which includes training in mindfulness (Awareness), along with practices designed to cultivate positive relationships (Connection) or insight into the nature of the self (Insight).

Methods: A three-arm, fully remote RCT compared 8 weeks of one of 2 HMP conditions (Awareness+Connection and Awareness+Insight) with a waitlist control. Adults (≥18 years) without extensive previous meditation experience were eligible. The primary outcome was psychological distress (depression, anxiety, and stress). Secondary outcomes were social connection, empathy, compassion, self-reflection, insight, rumination, defusion, and mindfulness. Measures were completed at pretest, midtreatment, and posttest between October 2019 and April 2020. Longitudinal data were analyzed using intention-to-treat principles with maximum likelihood.

Results: A total of 343 participants were randomized and 186 (54.2%) completed at least one posttest assessment. The majority (166/228, 72.8%) of those assigned to HMP conditions downloaded the app. The 2 HMP conditions did not differ from one another in terms of changes in any outcome. Relative to the waitlist control, the HMP conditions showed larger improvements in distress, social connectedness, mindfulness, and measures theoretically linked to insight training (d=–0.28 to 0.41; Ps≤.02), despite modest exposure to connection- and insight-related practice. The results were robust to some assumptions about nonrandom patterns of missing data. Improvements in distress were associated with days of use. Candidate mediators (social connection, insight, rumination, defusion, and mindfulness) and moderators (baseline rumination, defusion, and empathy) of changes in distress were identified.

Conclusions: This study provides initial evidence of efficacy for the HMP app in reducing distress and improving outcomes related to well-being, including social connectedness. Future studies should attempt to increase study retention and user engagement.

Trial Registration: ClinicalTrials.gov NCT04139005; https://clinicaltrials.gov/ct2/show/NCT04139005

(JMIR Ment Health 2020;7(11):e23825) doi:10.2196/23825

KEYWORDS

meditation; mindfulness; compassion; mobile health; loneliness; randomized controlled trial; mobile phone

Introduction

Background

Mindfulness and meditation have become household words for many people in the United States and across the globe in the past 20 years. Derived from Buddhist and Hindu contemplative traditions [1], secularized meditative practices are being taught in schools, recommended by health care providers, and employed by businesses [2-4]. The use of meditation tripled in the United States between 2012 and 2017 (from 4.1% to 14.2%) [5]. Meta-analyses involving hundreds of randomized controlled trials (RCTs) suggest that meditation training can decrease psychological symptoms (eg, depression, anxiety, stress) and increase aspects of well-being and positive functioning (eg, meaning in life, compassion, prosocial behavior) [6-15].

To date, the vast majority of research on meditation has focused on interventions delivered in person. Standardized mindfulness-based interventions such as mindfulness-based stress reduction (MBSR) [16] and mindfulness-based cognitive therapy (MBCT) [17] were explicitly designed as group-based interventions delivered by trained instructors, and these interpersonal elements are viewed as central ingredients (eg, group format) [18]. Despite some meditation-based interventions (MBIs) being recommended as first-line treatments (eg, MBCT for depression relapse prevention) [19-21], their availability remains limited [22]. Barriers for the dissemination of MBIs are similar to those facing other evidence-based psychotherapies (eg, lack of available providers, cost, logistical challenges) [23-25].

Delivering interventions through mobile technology has been proposed as a solution for increasing access to psychological interventions, including MBIs [26,27]. Web- and smartphone-based interventions have obvious advantages over traditional in-person delivery in terms of cost and scalability. Furthermore, mobile health (mHealth) interventions can, in theory, do things that in-person interventions typically never do, such as providing access 24 hours a day or customizing content based on passively sensed data (eg, location) [28]. Among mHealth delivery platforms, smartphone-based interventions may be particularly promising, with these devices often kept within arm's reach, charged, turned on, and being owned by the vast majority of the population [29].

There has been a dramatic increase in the past five years in RCTs testing smartphone-based interventions that include training in meditation [30]. These studies have begun to examine efficacy in various clinical and nonclinical populations [31-38]. Although preliminary, available evidence suggests that smartphone-based interventions that include training in meditation and mindfulness may provide psychological benefits that are similar to in-person MBIs (eg, decreased psychological symptoms, increased positive functioning), albeit smaller in magnitude [30,39-42].

Similar to the in-person MBI literature, RCTs testing the mobile delivery of MBIs have focused almost entirely on mindfulness. The term *mindfulness* is derived from the Pali word *sati*, which in Buddhism refers to the cultivation of receptive, present-moment awareness [43]. In the scientific literature, mindfulness can refer to a mental state, trait, or faculty amenable to training [44-47]. Mindfulness-based interventions commonly adopt the definition by Kabat-Zinn [48]: "paying attention in a particular way: on purpose, in the present moment, and nonjudgmentally." Meta-analyses suggest that dispositional mindfulness along with both short-term (eg, mindfulness inductions) and long-term (eg, mindfulness-based interventions such as MBSR) training are associated with decreased psychiatric symptoms, negative affect, substance use, and neuroticism [9,49-53].

Importantly, mindfulness training represents the implementation of primarily one meditative approach drawn from rich contemplative traditions [54]. Although largely untested, it is possible that a variety of meditative techniques may serve as valuable complements or alternatives to mindfulness. Different meditation practices have shown distinct neural signatures [55,56] and can produce different psychological effects [57]. Dahl et al [54] provide a useful typology for situating mindfulness training within the broader contemplative practice landscape. Using a family resemblance approach, they describe *attentional*, *constructive*, and *deconstructive* families. Mindfulness meditation, as implemented in MBSR, falls primarily within the *attentional family*, with training focused on regulating attention. The *constructive family* includes practices designed to strengthen psychological habits conducive to psychosocial health. This includes connection-based practices that involve cultivating feelings of warmth and friendliness toward oneself and others (eg, gratitude, loving kindness, and compassion practices) [58,59]. Experimental evidence suggests that connection practices increase well-being and decrease psychological symptoms [8,60]. The *deconstructive family* includes practices designed to modify unhelpful cognitive patterns, particularly regarding one's view of self and others. Practices in this family involve intentional self-inquiry into the dynamics of conscious experience and the nature of the self with the goal of generating an understanding of cognitive patterns (ie, insight). Deconstructive elements are present in MBCT and cognitive therapy more generally (eg, seeing thoughts as thoughts) [17,61]. However, research on deconstructive meditative practices has been limited.

Smartphone-based meditation interventions have almost exclusively focused on mindfulness training [30,62], although several studies have investigated internet-based interventions that include connection-related practices [63-65]. Although some interventions include connection-based practice as one of several guided practices within a general mindfulness framework [32], RCTs primarily examining constructive or deconstructive practices are rare (with some promising exceptions) [66,67].

This Study

This study sought to investigate the effects of a self-guided, smartphone-based meditation intervention that included explicit training in constructive and deconstructive families of practices [54]. In a three-arm RCT, we compared training in mindful awareness, paired with Connection or Insight practices, with a waitlist control. As both arms included the Awareness module first, we refer to them by their unique module (ie, Connection or Insight, rather than Awareness+Connection and Awareness+Insight). We included outcome measures designed to detect global effects (psychological distress) and practice-specific effects (eg, social connection, shift in relationship to one's thoughts). Our primary hypothesis was that participants in both active conditions would show reduced psychological distress relative to the waitlist control. In addition, we expected those randomized to connection practices to show larger improvements in connection-related measures and those randomized to insight practice to show larger improvements in insight-related measures. We had several exploratory secondary hypotheses. We hypothesized that app usage would be positively associated with reduced distress. We hypothesized that improvements in connection- and insight-related measures would mediate effects on distress for those in the Connection and Insight arms, respectively. We hypothesized that those lower in mindfulness at baseline would show larger improvements in the active conditions and that those lower in connection- and insight-related measures would show larger improvements in the Connection and Insight arms, respectively. These hypotheses were preregistered at the Open Science Framework [68].

Methods

Procedure

We conducted an 8-week, fully remote RCT comparing 2 active smartphone-based meditation interventions with a waitlist control. Participants were recruited through emails sent to faculty, staff, and students at the University of Wisconsin-Madison and through a database of individuals who had previously expressed interest in research at the Center for Healthy Minds. All screening procedures and data collection were web-based and carried out using REDCap [69]. Participants completed a screening protocol to determine eligibility and received their group assignment via an automated email following the completion of baseline questionnaires. Randomization was achieved by automatically allocating participants to groups based on sequentially assigned participant identification numbers (ie, 1:1:1 randomization ratio). Participants were contacted by email to complete questionnaires 4 and 8 weeks postbaseline.

Progress through the material in the Healthy Minds Program (HMP) app was self-guided. There was minimal contact with the study staff. Participants were provided with a study email address to contact for technical support or study-related questions. All procedures were approved by the institutional review board. The study was registered at ClinicalTrials.gov (NCT04139005).

Participants

Eligible participants were aged ≥18 years, had access to a smartphone or other device capable of running the intervention app (Android or iOS), and did not have extensive previous meditation experience defined as meditation retreat experience, meditation practice weekly for >1 year or daily practice within the previous 6 months; or previous training under the instruction of a meditation teacher, other than an introductory course. Participants received US $25 for completing the assessments.

Intervention

Participants assigned to one of the 2 active intervention arms were instructed to download the HMP app through the Google Play or Apple App Store. The full HMP app includes 4 modules with practices designed to cultivate categories of mental and emotional skills linked to both hedonic and eudaimonic well-being [70,71]. These include the cultivation of mindful attention (Awareness), positive relationships with self and others (Connection), insight into the nature of self and internal experience (Insight), and purpose, values, and meaning in life (Purpose). In this study, the 2 active interventions included 4 weeks of Awareness training, followed by 4 weeks of either Connection or Insight training. This design was predicated on the view that training in the stabilization of attention is foundational to skills trained by Connection and Insight [72]. Each module included brief, podcast-style didactic material along with guided meditation practices. Didactic content included discussion of the scientific bases of the practices. Participants were encouraged to follow a prespecified sequence while going through the material. Participants could select the length of the guided practices (5-30 min) and a variety of practices were available in each module. For example, the Awareness module included practices focused on awareness of breathing and mindfulness of sound. The Connection module included gratitude and kindness practices. The Insight module included practices involving noticing the changing nature of the phenomenon (ie, impermanence) and examining how thoughts and emotions influence perception. Participants in the waitlist condition received access to the full HMP app (ie, all 4 modules) at the conclusion of the study.

Measures

A demographic questionnaire was completed at baseline. App usage was measured objectively using the HMP app. Additional information about the psychometric properties and theoretical relevance of the included measures is provided in Multimedia Appendix 1 [73-101].

Psychological Distress

A psychological distress composite score was created from measures of depression, anxiety, and stress. We computed the mean across scaled (z-transformed) scores for each measure. The 8-item Patient-Reported Outcome Measures Information System Depression and Anxiety Scales [73] assessed depression and anxiety. Items are rated on a 5-point scale (1=never; 5=always), with higher scores indicating greater severity in the past 7 days. T scores ≥55 suggest mild or greater severity [102,103]. Internal consistency was high (=.93 to .94).

The 14-item Perceived Stress Scale [74] assessed psychological stress. Items are rated on a 5-point scale (0=never; 4=very often), with higher scores indicating greater stress in the past month. Internal consistency was high (=.89).

Measures Related to the Connection Module

The 20-item Social Connectedness Scale-Revised [75] assessed interpersonal connections. Items are rated on a 6-point scale (1=strongly disagree; 6=strongly agree), with higher scores indicating higher social connectedness. Internal consistency was high (=.95).

The 28-item Interpersonal Reactivity Index (IRI) [76] assessed empathy. Items are rated based on how well they describe the respondent on a 5-point scale (0=not well; 4=very well), with higher scores indicating greater empathy. Internal consistency was high for the total score (=.83).

The 21-item Compassionate Love Scale [77] assesses feelings of compassion. Items are rated on a 7-point scale (1=not at all true of me; 7=very true of me). Higher scores indicate greater feelings of compassion. Internal consistency was high (=.95).

Measures Related to the Insight Module

The 20-item Self-Reflection and Insight Scale (SRIS) [78] assessed participants' tendency toward self-reflection (eg, "I frequently examine my feelings") and self-understanding or insight (eg, "I usually know why I feel the way I do"). Items are rated on a 6-point scale (1=strongly disagree; 6=strongly agree) and yield subscales for self-reflection and insight, with higher scores indicating greater self-reflection or insight. Internal consistency was high (=.88 to .92).

The 15-item Perseverative Thinking Questionnaire (PTQ) [79] assessed rumination. For simplicity, we used the term "rumination" to refer to repetitive negative thinking as captured by the PTQ, although it captures both rumination and worry. Items are rated on a 5-point scale (0=never; 4=almost always), with higher scores indicating greater rumination. Internal consistency was high (=.96).

The 10-item Drexel Defusion Scale [80] assessed the ability to achieve psychological distance from internal experiences (ie, defusion). Items are rated on a 6-point scale (0=not at all; 5=very much), with higher scores indicating greater defusion. Internal consistency was high (=.89).

Mindfulness

The 39-item Five Facet Mindfulness Questionnaire [81] assessed mindfulness. Items are rated on a 5-point scale (1=never or very rarely true; 5=very often or always true), with higher scores indicating greater mindfulness. Internal consistency was high for the total score (=.94).

Data Analysis

Results from all preregistered primary and secondary measures are reported. For deviations made from the preregistered data analytic plan, see Multimedia Appendix 1.

Data were analyzed using intention-to-treat principles (ie, participants were not excluded based on engagement) [104]. Primary analyses used multilevel models (MLMs [82] with restricted information maximum likelihood estimation in the *lme4* package [105] in R [106]. MLMs with a maximum likelihood estimator are generally robust to data that are missing at random (MAR) [83]. For each outcome, an MLM was specified in which a linear change (coded as 0, 1, 2, for pre-, mid-, and posttest, respectively) in outcome was assumed over time, with participant-level random intercepts. Intervention effects were evaluated by the interaction between linear growth and group status, with contrasts comparing the 2 active conditions (ie, Connection, Insight) as well as the combined active conditions relative to waitlist control (see Multimedia Appendix 1 for the model). A subsequent sensitivity analysis restricted the sample to participants above the clinical cut-off for depression or anxiety at baseline (T≥55) [103]. Sensitivity analyses were also conducted with outliers (ie, 3 SD from the mean) and each participant sequentially removed [84].

Additional analyses assessed the potential impact of attrition, which is common in fully remote RCTs [107]. In this study, it is plausible that missingness was related to the unobserved value itself (ie, missing not at random [MNAR]). For example, individuals who failed to benefit from the HMP app may have been less likely to complete the study and would have shown worse outcomes had they been observed. Therefore, we relaxed our MAR assumptions to evaluate the degree to which intervention effects would be maintained under MNAR assumptions. We examined intervention effects in the presence of different assumed outcomes for dropout-missing observations, focusing on residualized change scores (from baseline to posttest) to simplify the study of missingness implications. We coded outcomes for dropout missingness at different levels, ranging from no difference in outcomes (relative to those that remained in the study) to all dropout-missing values being the worst possible outcome of those in the study. As, operationally, it becomes easier to study this range of conditions using outcome ranks as opposed to retaining the metrics of the studied measures, we applied a nonparametric Wilcoxon rank sum test to compare the active conditions against the waitlist control under different missingness assumptions.

To test our exploratory mediation hypotheses, we used the *mediation* package in R [108]. In these models, active group status (Connection or Insight) served as the independent variable; pre-post changes in mindfulness or connection- or insight-related measures served as the mediators; and posttest distress (controlling for pretest) served as the dependent variable. Pre-post changes were examined as mediators as unique Connection and Insight content was provided after the midtreatment assessments. We used MLMs to examine the effect of app usage, testing the time×usage interaction with usage operationalized as the median split of days of use. As noted in Multimedia Appendix 1, a median split was used because of deviations from normality in usage metrics. To assess baseline characteristics as moderators of change in distress, we tested 3-way interactions between time, group, and baseline characteristics within MLMs. False discovery rate (FDR) adjustment [85] was applied to all analyses to control for inflation of a type I error.

Sample Size and Power

We planned to recruit 300 participants (100 per group), which would allow the detection of small-to-moderate differences between any 2 groups (d=0.40) and between the active and waitlist control conditions (d=0.34) at 80% power and P=.05. Power was estimated using the *pwr.t.test* and *pwr.t2n.test* functions in the *pwr* package in R [109].

Results

Recruitment and Participant Characteristics

A total of 954 potential participants were assessed for eligibility, of which 343 met the inclusion criteria and were randomized to Connection (n=121), Insight (n=107), or waitlist (n=115; Figure 1). Demographics are reported in Table 1. The sample was predominantly White (280/343, 81.6%), female (290/343, 84.5%), and with graduate-level education (190/343, 55.4%). Income was more variable (89/343, 25.9% earned US $50,000 or less). The mean age was 41.74 (SD 12.52) years.

Figure 1. CONSORT (Consolidated Standards of Reporting Trials) diagram. Mid=week 4 assessment; Post=week 8 assessment.

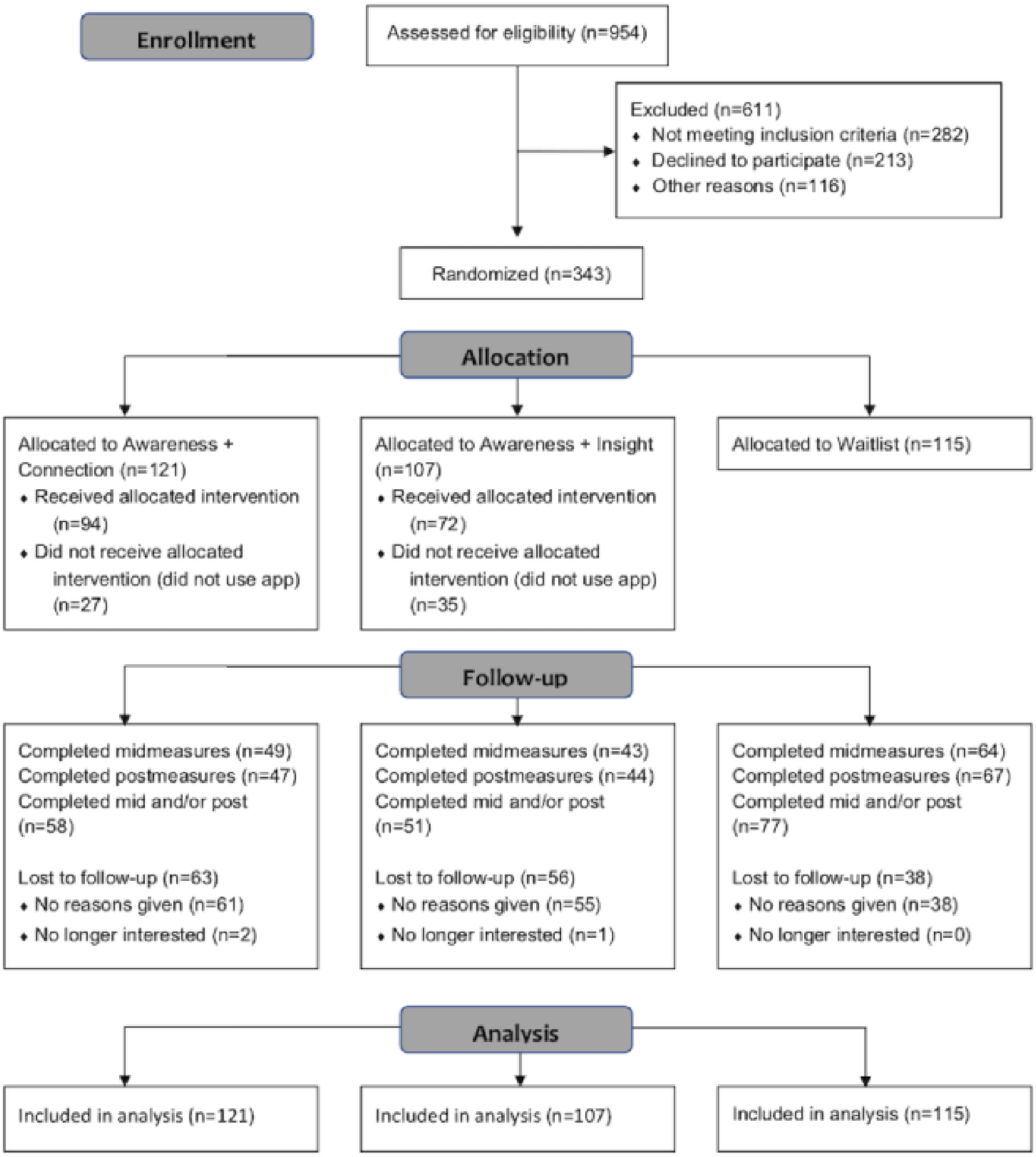


Table 1. Sample demographics.

Variable	Overall (n=343)	Connection (n=121)	Insight (n=107)	Waitlist (n=115)	*P* value[a]
Race and ethnicity, n (%)					.91
White	280 (81.6)	99 (81.8)	86 (80.4)	95 (82.6)	
Black	6 (1.7)	3 (2.5)	1 (0.9)	2 (1.7)	
Latinx	4 (1.2)	0 (0.0)	4 (3.7)	0 (0.0)	
Asian	18 (5.2)	7 (5.8)	6 (5.6)	5 (4.3)	
Multiracial	33 (9.6)	12 (9.9)	9 (8.4)	12 (10.4)	
Not reported	2 (0.6)	0 (0.0)	1 (0.9)	1 (0.9)	
Gender, n (%)					.93
Female	290 (84.5)	101 (83.5)	89 (83.2)	98 (85.2)	
Male	51 (14.9)	20 (16.5)	15 (14.0)	16 (13.9)	
Nonbinary	2 (0.6)	0 (0.0)	1 (0.9)	1 (0.9)	
Not reported	2 (0.6)	0 (0.0)	2 (1.9)	0 (0.0)	
Income (US $), n (%)					.18
≤50,000	89 (25.9)	33 (27.3)	26 (24.3)	30 (26.1)	
50,000-100,000	120 (35.0)	35 (28.9)	37 (34.6)	48 (41.7)	
100,000-150,000	76 (22.2)	32 (26.4)	22 (20.6)	22 (19.1)	
>150,000	57 (16.6)	21 (17.4)	21 (19.6)	15 (13.0)	
Not reported	1 (0.3)	0 (0.0)	1 (0.9)	0 (0.0)	
Education, n (%)					.45
Some high school	1 (0.3)	0 (0.0)	1 (0.9)	0 (0.0)	
High school graduate	6 (1.7)	3 (2.5)	0 (0.0)	3 (2.6)	
Some college	32 (9.3)	9 (7.4)	11 (10.3)	12 (10.4)	
College graduate	114 (33.2)	37 (30.6)	36 (33.6)	41 (35.7)	
Graduate school	190 (55.4)	72 (59.5)	59 (55.1)	59 (51.3)	
Age (years), mean (SD)	41.74 (12.52)	42.31 (12.8)	43.21 (12.39)	39.78 (12.2)	.10
Elevated symptoms[b], n (%)	252 (73.5)	86 (71.1)	81 (75.7)	85 (73.9)	.73

[a] *P* values based on a one-way analysis of variance with group (Connection, Insight, or waitlist) predicting demographics (White, female, high income [≥US $100,000], and graduate school).

[b] Elevated symptoms: Patient-Reported Outcomes Measurement Information System (PROMIS) Depression or PROMIS Anxiety in the mild or higher range (T≥55).

Utilization

Of those randomized to one of the 2 active conditions, 77.7% (94/121) of Connection participants and 67.3% (72/107) of Insight participants downloaded and used the HMP app at least once. By assigning values of zero to those who did not use the app, we found that average utilization was 10.52 days (SD 13.31; median 4), with 18.09 activities within the app (SD 23.30; median 7), 9.45 meditation practices (SD 13.34; median 3), and 102.16 total min of meditation practice (SD 187.74; median 26). All usage metrics were highly zero inflated (Multimedia Appendix 1). Days of use had the lowest skewness (1.34) and kurtosis (0.91), so a median split of days of use was used in the analyses. The median survival time (ie, time before last use) was 12 days. Group status (Connection vs Insight) was not associated with usage (*P*=.15), and survival time did not differ between groups (Multimedia Appendix 1; *P*=.24). As Connection or Insight content was provided at week 5 of the program, 32.2% (39/121) of Connection and 23.4% (25/107) of Insight participants engaged with the unique content. This proportion did not differ between groups (OR 0.64, 95% CI 0.35-1.15; *P*=.14).

Attrition Analysis

We examined baseline demographic and outcome variables as predictors of attrition. We constructed logistic regression models predicting the presence of any follow-up data (ie, midtreatment or posttest). Participants were invited to complete the posttest measures even if they had not completed the midtreatment measures. The average completion of at least one follow-up assessment (mid- or posttreatment) was 54.2% (186/343). Waitlist participants were more likely to complete follow-up assessments (77/115, 67.0% vs 109/228, 47.8%; waitlist n=77; Connection and Insight combined n=109; OR 2.21, 95% CI 1.39-3.56; *P*<.001). Completion of follow-up assessments did not differ between the Connection and Insight groups (OR 0.99, 95% CI 0.59-1.67; *P*=.97). However, participants who used the app at least once were more likely to complete the follow-up assessments (OR 3.66, 95% CI 1.95-7.16]; *P*<.001). Completion of follow-up was not associated with demographics (White, female, high income [≥US $100,000], and graduate education) or outcome measures at baseline (*P*s≥.15), with one exception. Participants with higher empathy scores (IRI) at baseline were more likely to complete follow-up assessments (OR 1.02, 95% CI 1.00-1.04; *P*=.04).

Primary Analyses

Correlations between outcomes are reported in Multimedia Appendix 1. The 3 groups did not differ in any demographic or outcome measures at baseline (*P*s≥.10; Tables 1 and 2). Within- and between-group effect sizes (Cohen *d*) and *P* values from MLMs are reported in Table 3. The 2 active conditions did not differ from one another in terms of change over time for distress or any secondary outcomes (time×group, *P*s≥.29). Therefore, all subsequent analyses combined the 2 active groups. When compared with the waitlist control, the active conditions showed greater decreases in distress (*d*=0.28) and rumination (*d*=0.18) and greater increases in social connectedness, self-reflection, insight, defusion, and mindfulness (*d*=0.13 to 0.41; FDR-adjusted *P*s≤.02; Figure 2). The active conditions did not differ from the waitlist on changes in empathy (*d*=0.02) or compassion (*d*=0.12). Significance tests for time×group interactions did not change when restricting to those with elevated symptoms at baseline (Table 2), when excluding outliers (with the exception of Self-Reflection, *P*=.05; Multimedia Appendix 1), nor when each case was excluded sequentially.

A larger proportion of participants in the active conditions showed a minimally important decrease in distress (*d*≤–0.30) [86] relative to the waitlist condition (70% vs 49%; Connection and Insight combined n=64/91; waitlist n=33/67; OR 2.44 [95% CI 1.27-4.75]; *P*=.008). A smaller proportion in the active condition showed a minimally important increase in distress (ie, deterioration, *d*≥0.30) relative to the waitlist condition (3% vs 16%; Connection and Insight combined n=3/91; waitlist n=11/67; OR 0.17 [0.04, 0.58]; *P*=.009).

Table 2. Descriptive statistics for repeated measures by group and timepoint.

Group and outcome	Pretest		Midtreatment		Posttest		*P* value[a]
	n	Mean (SD)	n	Mean (SD)	n	Mean (SD)	
CO[b]							
Psychological distress[c]	121	0.08 (0.90)	49	0.74 (0.80)	47	0.78 (0.78)	.49
Social connection[d]	121	83.41 (21.43)	48	90.62 (17.29)	46	92.32 (18.30)	.73
Empathy[e]	121	67.67 (12.33)	48	67.42 (12.55)	47	66.01 (12.90)	.41
Compassion[f]	121	99.18 (22.19)	48	102.52 (21.71)	46	101.68 (21.64)	.57
Self-reflection subscale[g]	121	56.31 (11.13)	48	58.45 (10.29)	46	57.49 (11.10)	.97
Insight subscale[g]	121	34.12 (7.46)	48	37.06 (6.17)	46	37.54 (6.20)	.46
Rumination[h]	121	31.12 (12.65)	46	25.31 (9.63)	46	25.36 (10.71)	.90
Defusion[i]	121	22.95 (9.37)	48	28.54 (8.37)	47	30.75 (7.81)	.86
Mindfulness[j]	121	122.16 (20.41)	49	136.28 (16.37)	47	139.63 (19.33)	.77
IN[k]							
Psychological distress	107	0.04 (0.94)	43	0.52 (0.88)	44	0.62 (0.88)	N/A[l]
Social connection	107	83.28 (19.04)	41	87.57 (19.28)	44	90.20 (20.64)	N/A
Empathy	107	69.77 (12.45)	41	70.24 (11.76)	44	69.49 (11.05)	N/A
Compassion	107	101.57 (21.44)	41	101.07 (24.52)	44	108.88 (23.29)	N/A
Self-reflection subscale	107	55.99 (10.40)	41	55.90 (9.00)	44	57.92 (9.45)	N/A
Insight subscale	107	32.84 (8.42)	41	33.67 (8.28)	44	36.14 (8.44)	N/A
Rumination	107	30.49 (11.71)	41	27.21 (10.79)	44	23.97 (10.68)	N/A
Defusion	107	23.62 (10.36)	41	27.36 (10.01)	44	30.49 (10.23)	N/A
Mindfulness	107	121.55 (24.96)	43	128.56 (22.96)	44	139.11 (19.75)	N/A
WL[m]							
Psychological distress	115	0.05 (0.88)	64	0.23 (1.01)	67	0.36 (0.91)	N/A
Social connection	115	81.53 (19.68)	60	82.78 (21.10)	63	84.63 (20.24)	N/A
Empathy	115	69.31 (12.85)	60	71.09 (13.49)	63	67.96 (13.26)	N/A
Compassion	115	98.62 (22.39)	58	99.53 (22.64)	63	100.85 (22.42)	N/A
Self-reflection subscale	115	56.34 (10.78)	57	56.87 (11.14)	63	56.23 (11.26)	N/A
Insight subscale	115	33.22 (8.02)	57	35.40 (8.29)	63	35.55 (7.83)	N/A
Rumination	115	31.14 (11.56)	56	29.51 (13.04)	62	27.41 (11.92)	N/A
Defusion	115	23.49 (9.58)	60	24.88 (9.95)	64	26.71 (9.86)	N/A
Mindfulness	115	120.16 (18.93)	62	125.15 (20.21)	65	128.88 (20.23)	N/A

[a] *P* value from a one-way analysis of variance predicting baseline values for outcome measures by group status.

[b] CO: Awareness+Connection.

[c] Composite of Patient-Reported Outcomes Measurement Information System (PROMIS) Depression, PROMIS Anxiety, and Perceived Stress Scale.

[d] Social Connectedness Scale.

[e] Interpersonal Reactivity Index.

[f] Compassionate Love Scale.

[g] Subscales of the Self-Reflection and Insight Scale.

[h] Perseverative Thinking Questionnaire.

[i] Drexel Defusion Scale.

[j] Total score of Five Facet Mindfulness Questionnaire.

[k]IN: Awareness+Insight.

[l]N/A: not applicable.

[m]WL: waitlist.

Table 3. Results of multilevel models assessing differential change over time.

Outcome	CO[a] versus IN[b]					Active versus WL[c]						
	d_{CO}[d]	d_{IN}	d_{diff}	*P* value[e]	P_{FDR}[f]	d_{Active}[g]	d_{WL}	d_{diff}	*P* value	P_{FDR}	Elev *P*[h]	Elev P_{FDR}
Psychological distress[i]	–0.77	–0.70	–0.07	.86	.97	–0.74	–0.46	–0.28	<.001	<.001	<.001	<.001
Social connection[j]	0.42	0.36	0.06	.54	.82	0.39	0.16	0.23	.003	.007	.01	.02
Empathy[k]	–0.14	–0.02	–0.12	.37	.82	–0.08	–0.10	0.02	.63	.63	.48	.48
Compassion[l]	0.11	0.34	–0.23	.29	.82	0.22	0.10	0.12	.14	.16	.18	.20
Self-reflection subscale[m]	0.11	0.18	–0.07	.51	.82	0.14	–0.01	0.15	.007	.01	.02	.03
Insight subscale[m]	0.46	0.39	0.07	.98	.98	0.42	0.29	0.13	.02	.02	.001	.002
Rumination[n]	–0.45	–0.56	0.11	.32	.82	–0.5	–0.32	–0.18	.01	.02	.007	.01
Defusion[o]	0.83	0.66	0.17	.78	.97	0.75	0.34	0.41	<.001	<.001	<.001	<.001
Mindfulness[p]	0.86	0.70	0.16	.55	.82	0.77	0.46	0.31	<.001	<.001	<.001	<.001

[a]CO: Awareness+Connection.

[b]IN: Awareness+Insight.

[c]WL: waitlist.

[d]Cohen *d* calculated as pre-post for within-group effects and the difference between within-group effects (Connection−Insight, active−waitlist) for d_{diff}. For within-group, subscripted CO (ie, d_{CO}), IN, Active, and WL refer to subgroups noted.

[e]*P* value from time×group interaction from multilevel models.

[f]FDR: false discovery rate adjusted *P* values.

[g]Combined Awareness+Connection and Awareness+Insight.

[h]Active versus waitlist time×group interaction restricted to sample with elevated depression or anxiety at baseline (T≥55).

[i]Composite of Patient-Reported Outcomes Measurement Information System (PROMIS) Depression, PROMIS Anxiety, and Perceived Stress Scale.

[j]Social connection: Social Connectedness Scale.

[k]Interpersonal Reactivity Index.

[l]Compassionate Love Scale.

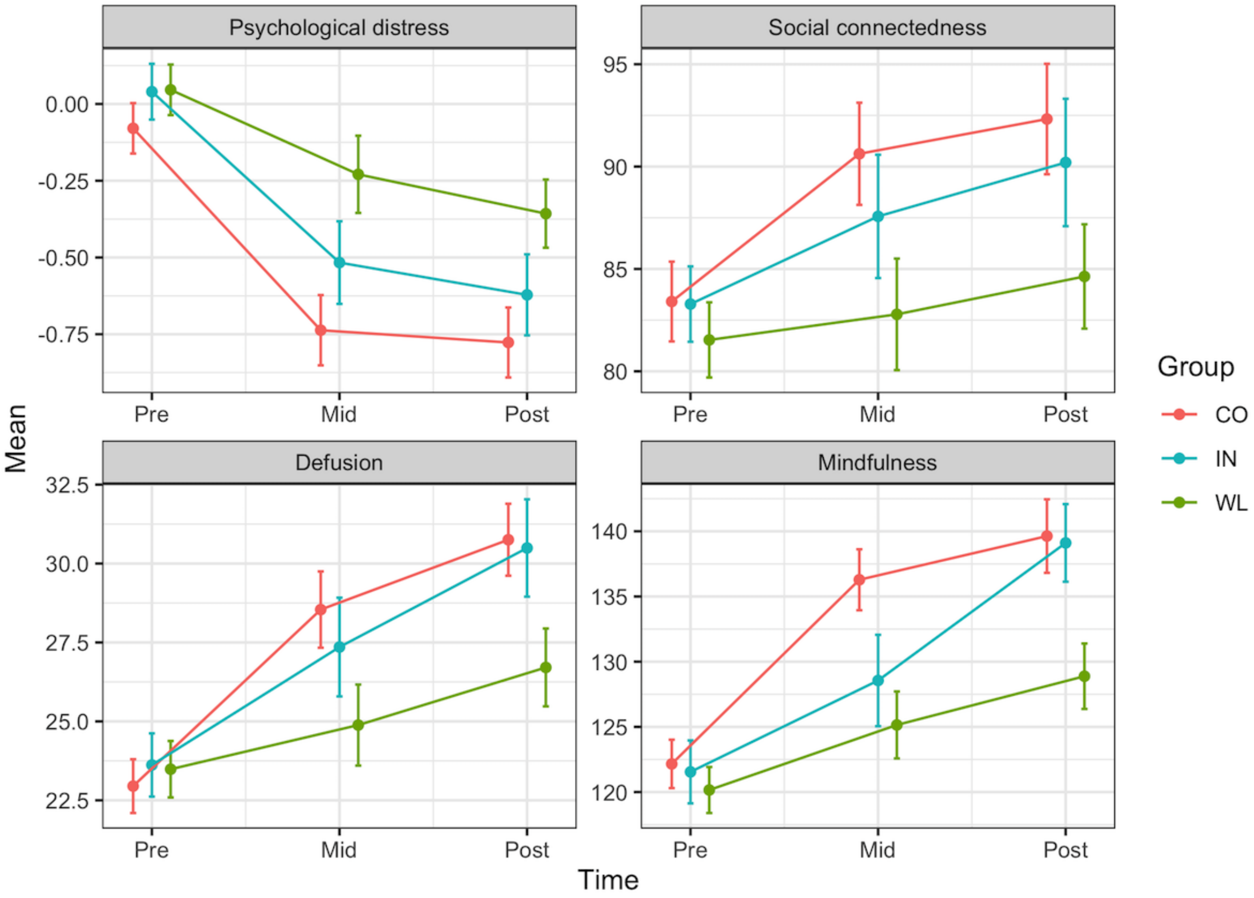
[m]Subscales of the Self-Reflection and Insight Scale.

[n]Perseverative Thinking Questionnaire.

[o]Drexel Defusion Scale.

[p]Total score of Five Facet Mindfulness Questionnaire.

Figure 2. Longitudinal changes in psychological distress, social connectedness, defusion, and mindfulness by group. The figure displays observed means and SEs (error bars=1 SE) based on all available data (n=343). CO: Awareness+Connection; IN: Awareness+Insight; WL: waitlist.



Robustness Check: Sensitivity Analyses

Although maximum likelihood is robust to data MAR [83], subsequent analyses evaluated treatment effects based on varying assumptions under MNAR conditions. Using the completer sample, a Wilcoxon rank sum test on the residualized gain score mirrored the MLM results, with larger improvements in the active conditions relative to the waitlist on several outcomes (FDR-adjusted *P*s≤.047; Multimedia Appendix 1). In the worst-case scenario model in which missing reflects the worst possible outcome across both active and waitlist groups, the groups did not differ, although the direction of the mean rank favored the waitlist group for all outcomes. Thus, we examined the results in between these extreme conditions to understand where significance goes away and where the direction of intervention effect reverses. When we assumed that missing values are on average 0.25 SD above the mean (implying worse than average outcomes for the missing observations), the results continued to favor the active conditions for changes in distress, social connectedness, defusion, and mindfulness (FDR-adjusted *P*s≤.04; Multimedia Appendix 1). When we assumed that missing values were on average 0.50 SD above the mean, the differences between groups were not statistically significant for any outcome. The difference remained nonsignificant when we assumed that missing values are on average 0.75 SD above the mean. Thus, it appears that our results are robust to MNAR up to a point, specifically that missing outcomes are no more than 0.25 SD above the mean on average, under the assumption that missingness implies comparable outcomes for both the active and waitlist groups.

Secondary Analyses

The results of the usage analyses are reported in Table 4. HMP use above the median number of days was associated with larger improvements in distress, insight, defusion, and mindfulness (FDR-adjusted *P*s≤.03; Figure 3).

The results of the mediation analyses are reported in Table 5. Changes in 5 candidate mechanisms showed a significant average causal mediation effect (FDR-adjusted *P*s≤.04) in the expected direction (ie, improvements in social connection, insight, rumination, defusion, and mindfulness mediated improvements in distress). Changes in mindfulness were associated with the largest proportion mediated (0.45).

Table 4. Results of multilevel models predicting changes in outcomes from Healthy Minds Program app usage (n=228).

Outcome	Time×usage *B*[a]	*t* test (*df*)[b]	*P* value[c]	P_{FDR}[d]
Psychological distress[e]	–0.17	–2.46 (207)	.02	.03
Social connection[f]	2.42	1.61 (202)	.11	.17
Empathy[g]	–0.47	–0.58 (200)	.56	.56
Compassion[h]	–1.16	–0.76 (201)	.45	.50
Self-reflection subscale[i]	0.80	0.87 (218)	.38	.49
Insight subscale[i]	1.55	2.46 (211)	.02	.03
Rumination[j]	–1.85	–2.13 (197)	.03	.06
Defusion[k]	2.40	2.64 (229)	.009	.03
Mindfulness[l]	5.17	2.93 (212)	.004	.03

[a] Multilevel model regression coefficient. Usage: days of use split into high (median or above) and low (below median) groups.

[b] *t* statistic for time×usage interaction with associated degrees of freedom (*df*).

[c] *P* value for time×usage interaction.

[d] FDR: false discovery rate adjusted *P* values.

[e] Composite of Patient-Reported Outcomes Measurement Information System (PROMIS) Depression, PROMIS Anxiety, and Perceived Stress Scale.

[f] Social Connectedness Scale.

[g] Interpersonal Reactivity Index.

[h] Compassionate Love Scale.

[i] Subscales of the Self-Reflection and Insight Scale.

[j] Perseverative Thinking Questionnaire.

[k] Drexel Defusion Scale.

[l] Total score of Five Facet Mindfulness Questionnaire.

Figure 3. Healthy Minds Program app usage predicting longitudinal changes in psychological distress, defusion, insight, and mindfulness in active conditions (FDR-corrected Ps≤.04). Usage=median split of days of use (n=228). HMP: Healthy Minds Program; WL: waitlist.

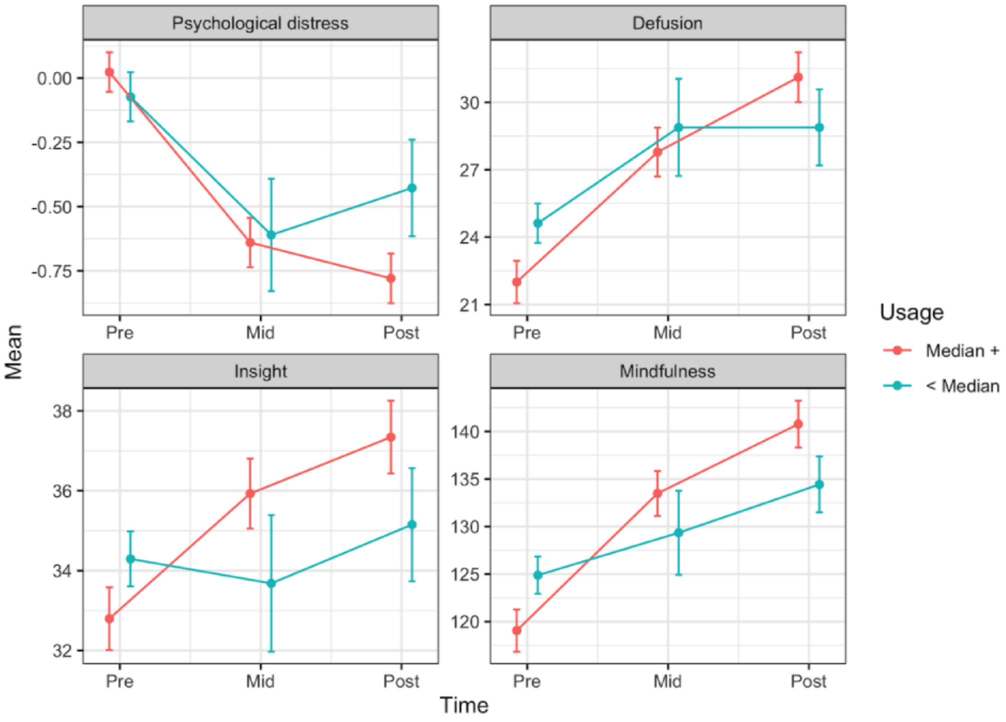


Table 5. Results of mediation analyses predicting changes in psychological distress.

Outcome	ACME[a]	ADE[b]	Prop mediated[c]	*P* value[d]	*P* FDR [e]
Social connection[f]	–0.10	–0.32	0.24	.002	.005
Empathy[g]	0.02	–0.44	–0.03	.40	.47
Compassion[h]	0.01	–0.43	–0.01	.59	.59
Self-reflection subscale[i]	0.02	–0.44	–0.04	.41	.47
Insight subscale[i]	–0.06	–0.37	0.13	.03	.04
Rumination[j]	–0.10	–0.32	0.23	.02	.04
Defusion[k]	–0.10	–0.32	0.23	.002	.005
Mindfulness[l]	–0.18	–0.22	0.45	<.001	<.001

[a] ACME: average causal mediation effect (ie, indirect effect).

[b] ADE: average direct effect (ie, from active to posttreatment distress controlling for pretreatment distress, when active=1 and waitlist=0).

[c] Proportion mediated computed as indirect effect (ie, ACME) divided by total effect [108].

[d] *P* value based on quasi-Bayesian CIs.

[e] FDR: false discovery rate adjusted *P* values. Models examining pre-post change in constructs related to Awareness, Connection, and Insight modules as mediators of pre-post change in (composite of PROMIS Depression, PROMIS Anxiety, and Perceived Stress Scale). Proportion mediated can be negative in instances where direct effect and indirect effect have opposite signs.

[f] Social connection: Social Connectedness Scale.

[g] Interpersonal Reactivity Index.

[h] Compassionate Love Scale.

[i] Subscales of the Self-Reflection and Insight Scale.

[j] Perseverative Thinking Questionnaire.

[k] Drexel Defusion Scale.

[l] Total score of Five Facet Mindfulness Questionnaire.

The results of the baseline moderation analyses are reported in Table 6. A total of 3 baseline variables showed significant time×group×baseline interactions after FDR adjustment. Psychological vulnerability, as indicated by 2 outcomes (higher rumination and lower defusion) at baseline, was associated with significant improvements in distress in the HMP conditions relative to the waitlist condition. Baseline empathy showed the opposite pattern, with those higher at baseline showing significant improvements in distress in HMP relative to the waitlist condition (Multimedia Appendix 1).

Table 6. Baseline outcomes as moderators of longitudinal changes in psychological distress.

Outcome	Time×group B[a]	Time×group×baseline B	t test (df)[b]	P value[c]	P_{FDR}[d]
Social connection[e]	–0.64	0.01	2.16 (363)	.03	.06
Empathy[f]	0.67	–0.01	–3.23 (343)	.001	.008
Compassion[g]	0.07	0.00	–1.20 (343)	.23	.29
Self-reflection subscale[h]	–0.59	0.01	1.55 (339)	.12	.19
Insight subscale[h]	–0.44	0.01	1.14 (359)	.25	.29
Rumination[i]	0.21	–0.01	–3.15 (390)	.002	.008
Defusion[j]	–0.54	0.01	2.86 (361)	.004	.01
Mindfulness[k]	–0.44	0.00	0.82 (372)	.41	.41

[a] Multilevel model regression coefficient.

[b] *t* test: *t* statistic for time×group×baseline (with group coded as active=1, waitlist=0) with associated degrees of freedom (df).

[c] *P* value for time×group×baseline.

[d] FDR: false discovery rate adjusted *P* values.

[e] Social Connectedness Scale.

[f] Interpersonal Reactivity Index.

[g] Compassionate Love Scale.

[h] Subscales of the Self-Reflection and Insight Scale.

[i] Perseverative Thinking Questionnaire.

[j] Drexel Defusion Scale.

[k] Total score of Five Facet Mindfulness Questionnaire.

Discussion

Principal Findings

This study sought to expand the scientific understanding of the impact of smartphone-delivered meditation training beyond mindfulness. To do so, we evaluated the effects of mindfulness training (Awareness) paired with practices designed to cultivate kindness toward oneself and others (Connection) or insight into the nature of self and internal experience (Insight). We assessed the effects on psychological distress and constructs theoretically linked to connection- and insight-based training [54].

Contrary to our expectations, there was no indication that training in connection produced differential effects relative to insight-related practices. There are several potential reasons for this. One likely explanation is that the actual content completed by each group was largely overlapping. Both groups began with foundational mindfulness training. Given the modest engagement (a perennial concern in mHealth interventions) [110,111], most participants did not engage with the unique Connection or Insight modules. It is also possible that meditation training produces similar effects for novices, regardless of the specific type of training. Novices may spend much of their initial meditation practice simply regathering a wandering attention, regardless of the actual practice instructions. Indeed, studies showing distinct neural signatures associated with various forms of meditation practice have primarily been conducted with long-term practitioners with thousands of hours of experience [55]. A third possibility is that various forms of meditation training contain common ingredients (eg, acceptance, curiosity) that may, especially early in training, be more potent than style-specific ingredients.

Despite the absence of differential effects, the results suggest that meditation delivered via smartphones produced small reductions in psychological distress (d=–0.28) and improvements in several candidate mechanisms relative to a waitlist control (d=–0.18 to 0.41). These results are generally consistent with meta-analyses of the broader mHealth and mHealth MBI literature, which has shown small benefits of self-guided smartphone apps on depression and anxiety symptoms (g=0.21 to 0.23) and measures of mindfulness and acceptance (g=0.27) [30]. These effects are considerably smaller than those produced by in-person MBIs (eg, d=0.55 vs waitlist) [9]. It is likely that mHealth MBIs may be less potent than in-person interventions, indicating trade-offs between scalability, cost, and potency. On the basis of those completing posttreatment measures, HMP appears safe in that the rates of clinically significant increases in distress were rare (3%) and were less common than the rates in the control condition (16%). This finding is consistent with a recent large-scale evaluation of the deterioration in MBSR [87].

One important caveat for interpreting our findings is high attrition, particularly within the active conditions. Both high attrition and differential attrition are common in mHealth research [107,112]. Our overall attrition rate was almost identical to that typically found in RCTs testing smartphone interventions without telephone or in-person enrollment (45.8% in this study and 43.4% in the meta-analysis) [107]. In addition to employing maximum likelihood estimation in all MLMs (which is robust to MAR) [83], we conducted a series of sensitivity analyses to assess the effects of various MNAR assumptions. Most effects were robust to noncompleters having outcomes slightly worse than completers (SD 0.25). However, the effects did not persist with larger deviations (SD ≥0.50). It is impossible to directly test which of these scenarios is most likely (as is the case for other MNAR approaches) [113]. Future studies should include items specifically to predict missingness (eg, "how likely are you to drop out of this study") [114]. Responses can then be included as auxiliary variables to improve the performance of MAR methods (effectively converting MNAR to MAR) [114].

In light of the degree of attrition, secondary analyses should be interpreted as exploratory. However, these models provide tentative possibilities to be examined further. We found evidence that higher usage (median or above days of use) was associated with larger improvements in distress and several other outcomes. This mirrors the dosage-outcome associations seen in the in-person MBI literature [115]. Mediation analyses suggest candidate mechanisms theoretically linked to each HMP module that may indirectly contribute to decreased distress (ie, mindfulness, social connectedness, defusion, rumination). This also mirrors reviews of the in-person MBI literature that have found changes in mindfulness and rumination mediate effects [116]. The possibility that social connectedness also plays a role should be explored further, particularly as it has been associated with numerous psychological and physical health outcomes [117,118]. Moderation analyses indicated larger improvements among those showing higher rumination or empathy and lower defusion at baseline. These are somewhat conflicting findings, with the rumination and defusion associations suggesting that HMP may be most effective for those with deficits at baseline, whereas the association with empathy suggesting a higher baseline level may be necessary to benefit most. Given that mHealth interventions could, in theory, be easily adapted to participant characteristics (eg, participants routed to receive a particular version based on baseline questionnaires), future experimental work can specifically examine who is likely to benefit from which kind of training (eg, randomizing to adapted vs nonadapted versions). The scalability of mHealth RCTs may allow recruitment of the sample sizes necessary for adequately powered tests of moderation [119].

Limitations and Future Directions

As noted, high and differential attrition are limitations of this study. Although attrition is commonplace in mHealth research [110], future studies should include study design features that decrease attrition (eg, telephone enrollment) [107]. Relatedly, engagement with the HMP app was relatively modest. Similar to attrition, low engagement is a rule rather than an exception for mHealth [120]. Presumably low engagement diminished any potential effects of the HMP app, making estimates of efficacy more ambiguous. Future studies could evaluate engagement strategies by randomizing participants to receive approaches found to improve adherence to medical regimens (eg, modifying dosage recommendations, providing reminders) [121]. Microrandomized trials could assess the impact of a variety of small manipulations on short-term engagement outcomes [122].

The lack of a follow-up assessment is another limitation, making it unclear whether any benefits were sustained. Furthermore, the lack of active control conditions makes it impossible to rule out effects due to a digital placebo [123]. Similarly, the included self-report measures are vulnerable to social desirability bias, although this may be less of an issue within a fully remote RCT [124]. Sampling procedures and sample demographics raise questions regarding generalizability, especially to racial/ethnic minority populations and those with lower levels of education. Participants in the Center for Healthy Minds database may have been particularly amenable to the HMP app (although those with prior meditation experience would have been excluded).

An obvious future study would assign participants to receive only Connection or Insight module content. This could clarify the unique effects of these types of practices. As we observed effects on distress that persisted when restricted to those with elevated symptoms at baseline, it would be worthwhile replicating this study within a clinical sample. For this, it could be important or even necessary, for safety reasons, to include some amount of professional guidance [41], perhaps telephone or text-based support [125]. Other extensions of this study could include the use of non–self-report measures, both to rule out social desirability as well as to clarify underlying mechanistic processes. Numerous biological and behavioral mechanisms have been assessed for in-person MBIs (eg, changes in attention, body awareness, stress physiology) [126-128] and may be operant within mHealth MBIs. A wide variety of extensions could also be made to the HMP app itself. For example, intervention components could be provided in response to passively assessed mood states (eg, using data streams routinely gathered through phone sensors). This would require not only the validation of passive measures [129,130] but also studies that clarify the optimal pairing of intervention components to mood. Microrandomized trials may be an ideal design for this purpose. The amount of engagement necessary for clinical benefits (ie, recommended dosage) should be clarified in future studies (eg, by random assignment to dosage conditions). RCTs using active control conditions can help clarify the role that nonspecific factors may play in mHealth MBIs. On the basis of the robust association between working alliances and outcomes within in-person interventions [131], future studies should assess its digital corollary [132]. Finally, a critical future direction is investigating the efficacy of mHealth MBIs specifically among (and ideally tailored for) [133,134] racial/ethnic minorities. Racial/ethnic minority populations are at increased risk for racism-related negative psychological and physical health consequences [135] and have been historically underrepresented in research on mindfulness [136,137].

Conclusions

Research into the mobile delivery of meditation training has expanded rapidly in the past five years. However, the vast majority of this work has focused exclusively on training mindfulness. We found preliminary evidence that 2 versions of a novel smartphone app that includes training in mindfulness, in addition to skills specifically targeted to improve social connection and intrapersonal insight, are associated with reduced psychological distress, increased social connectedness, and improvements in candidate mechanisms believed to underlie the beneficial effects of MBIs. Future studies, particularly those focusing on decreasing study dropout and increasing intervention engagement, are warranted.

Conflicts of Interest

RD is the founder, president, and serves on the board of directors for the nonprofit organization, Healthy Minds Innovations, Inc. CD is Chief Contemplative Officer for Healthy Minds Innovations, Inc. CW has served as a paid consultant and content contributor for Healthy Minds Innovations, Inc.

Multimedia Appendix 1
Healthy Minds Program randomized controlled trial supplemental materials.
[DOCX File , 113 KB - mental_v7i11e23825_app1.docx]

Multimedia Appendix 2
CONSORT-eHealth (V 1.6.1).
[PDF File (Adobe PDF File), 3037 KB - mental_v7i11e23825_app2.pdf]

References

1. Harrington A, Dunne J. When mindfulness is therapy: ethical qualms, historical perspectives. Am Psychol 2015 Oct;70(7):621-631. [doi: 10.1037/a0039460] [Medline: 26436312]
2. Dunning DL, Griffiths K, Kuyken W, Crane C, Foulkes L, Parker J, et al. Research review: the effects of mindfulness-based interventions on cognition and mental health in children and adolescents - a meta-analysis of randomized controlled trials. J Child Psychol Psychiatry 2019 Mar;60(3):244-258 [FREE Full text] [doi: 10.1111/jcpp.12980] [Medline: 30345511]
3. Jamieson SD, Tuckey MR. Mindfulness interventions in the workplace: a critique of the current state of the literature. J Occup Health Psychol 2017 Apr;22(2):180-193. [doi: 10.1037/ocp0000048] [Medline: 27643606]
4. Meditation: A Simple, Fast Way to Reduce Stress. Mayo Clinic Staff. URL: https://www.mayoclinic.org/tests-procedures/meditation/in-depth/meditation/art-20045858 [accessed 2020-03-07]
5. Clarke TC, Barnes PM, Black LI, Stussman BJ, Nahin RL. Use of yoga, meditation, and chiropractors among US adults aged 18 and over. NCHS Data Brief 2018 Nov(325):1-8 [FREE Full text] [Medline: 30475686]
6. Chu ST, Mak WW. How mindfulness enhances meaning in life: a meta-analysis of correlational studies and randomized controlled trials. Mindfulness 2019 Nov 28;11(1):177-193. [doi: 10.1007/s12671-019-01258-9]
7. Donald JN, Sahdra BK, Van Zanden B, Duineveld JJ, Atkins PW, Marshall SL, et al. Does your mindfulness benefit others? A systematic review and meta-analysis of the link between mindfulness and prosocial behaviour. Br J Psychol 2019 Feb;110(1):101-125. [doi: 10.1111/bjop.12338] [Medline: 30094812]
8. Galante J, Galante I, Bekkers M, Gallacher J. Effect of kindness-based meditation on health and well-being: a systematic review and meta-analysis. J Consult Clin Psychol 2014 Dec;82(6):1101-1114. [doi: 10.1037/a0037249] [Medline: 24979314]
9. Goldberg SB, Tucker RP, Greene PA, Davidson RJ, Wampold BE, Kearney DJ, et al. Mindfulness-based interventions for psychiatric disorders: a systematic review and meta-analysis. Clin Psychol Rev 2018 Feb;59:52-60 [FREE Full text] [doi: 10.1016/j.cpr.2017.10.011] [Medline: 29126747]
10. Goyal M, Singh S, Sibinga EM, Gould NF, Rowland-Seymour A, Sharma R, et al. Meditation programs for psychological stress and well-being: a systematic review and meta-analysis. JAMA Intern Med 2014 Mar;174(3):357-368 [FREE Full text] [doi: 10.1001/jamainternmed.2013.13018] [Medline: 24395196]
11. Kuyken W, Warren FC, Taylor RS, Whalley B, Crane C, Bondolfi G, et al. Efficacy of mindfulness-based cognitive therapy in prevention of depressive relapse: an individual patient data meta-analysis from randomized trials. JAMA Psychiatry 2016 Jun 1;73(6):565-574 [FREE Full text] [doi: 10.1001/jamapsychiatry.2016.0076] [Medline: 27119968]
12. Luberto CM, Shinday N, Song R, Philpotts LL, Park ER, Fricchione GL, et al. A systematic review and meta-analysis of the effects of meditation on empathy, compassion, and prosocial behaviors. Mindfulness (N Y) 2018 Jun;9(3):708-724 [FREE Full text] [doi: 10.1007/s12671-017-0841-8] [Medline: 30100929]
13. Sedlmeier P, Eberth J, Schwarz M, Zimmermann D, Haarig F, Jaeger S, et al. The psychological effects of meditation: a meta-analysis. Psychol Bull 2012 Nov;138(6):1139-1171. [doi: 10.1037/a0028168] [Medline: 22582738]
14. Strohmaier S. The relationship between doses of mindfulness-based programs and depression, anxiety, stress, and mindfulness: a dose-response meta-regression of randomized controlled trials. Mindfulness 2020 Mar 2;11(6):1315-1335. [doi: 10.1007/s12671-020-01319-4]
15. Goldberg S, Riordan K, Sun S, Davidson R. The empirical status of mindfulness-based interventions: a systematic review of 44 meta-analyses of randomized controlled trials. Perspect Psychol Sci 2021 Jan 1:- (forthcoming).
16. Kabat-Zinn J. Full Catastrophe Living: How to Cope With Stress, Pain and Illness Using Mindfulness Meditation. Revised Edition. London, UK: Piatkus; 2013.

17. Segal Z, Williams J, Teasdale J. Mindfulness-Based Cognitive Therapy for Depression. Second Edition. New York, USA: Guilford Press; 2013.
18. Kabat-Zinn J. An outpatient program in behavioral medicine for chronic pain patients based on the practice of mindfulness meditation: theoretical considerations and preliminary results. Gen Hosp Psychiatry 1982 Apr;4(1):33-47. [doi: 10.1016/0163-8343(82)90026-3] [Medline: 7042457]
19. Goldberg S, Segal Z. Mindfulness-based Cognitive Therapy. American Psychological Association. URL: https://www.div12.org/treatment/mindfulness-based-cognitive-therapy/ [accessed 2020-03-07]
20. Goldberg SB, Tucker RP, Greene PA, Davidson RJ, Kearney DJ, Simpson TL. Mindfulness-based cognitive therapy for the treatment of current depressive symptoms: a meta-analysis. Cogn Behav Ther 2019 Nov;48(6):445-462. [doi: 10.1080/16506073.2018.1556330] [Medline: 30732534]
21. Depression in Adults: Recognition and Management. National Institute for Health and Care Excellence (NICE). 2009. URL: https://www.nice.org.uk/guidance/cg90 [accessed 2020-03-07]
22. Rycroft-Malone J, Gradinger F, Griffiths H, Crane R, Gibson A, Mercer S, et al. Accessibility and implementation in the UK NHS services of an effective depression relapse prevention programme: Learning from mindfulness-based cognitive therapy through a mixed-methods study. Health Serv Deliv Res 2017;5(14):-. [doi: 10.3310/hsdr05140] [Medline: 28368561]
23. Andrade LH, Alonso J, Mneimneh Z, Wells JE, Al-Hamzawi A, Borges G, et al. Barriers to mental health treatment: results from the WHO World Mental Health surveys. Psychol Med 2014 Apr;44(6):1303-1317 [FREE Full text] [doi: 10.1017/S0033291713001943] [Medline: 23931656]
24. Beidas RS, Kendall PC. Training therapists in evidence-based practice: a critical review of studies from a systems-contextual perspective. Clin Psychol (New York) 2010 Mar;17(1):1-30 [FREE Full text] [doi: 10.1111/j.1468-2850.2009.01187.x] [Medline: 20877441]
25. Kazdin AE. Annual research review: expanding mental health services through novel models of intervention delivery. J Child Psychol Psychiatry 2019 Apr;60(4):455-472. [doi: 10.1111/jcpp.12937] [Medline: 29900543]
26. Fairburn CG, Patel V. The impact of digital technology on psychological treatments and their dissemination. Behav Res Ther 2017 Jan;88:19-25 [FREE Full text] [doi: 10.1016/j.brat.2016.08.012] [Medline: 28110672]
27. Segal Z, Dimidjian S, Vanderkruik R, Levy J. A maturing mindfulness-based cognitive therapy reflects on two critical issues. Curr Opin Psychol 2019 Aug;28:218-222 [FREE Full text] [doi: 10.1016/j.copsyc.2019.01.015] [Medline: 30798103]
28. Gustafson DH, McTavish FM, Chih M, 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] [Medline: 24671165]
29. 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 2020-11-19]
30. Linardon J. Can acceptance, mindfulness, and self-compassion be learned by smartphone apps? A systematic and meta-analytic review of randomized controlled trials. Behav Ther 2020 Jul;51(4):646-658. [doi: 10.1016/j.beth.2019.10.002] [Medline: 32586436]
31. Cox CE, Hough CL, Jones DM, Ungar A, Reagan W, Key MD, et al. Effects of mindfulness training programmes delivered by a self-directed mobile app and by telephone compared with an education programme for survivors of critical illness: a pilot randomised clinical trial. Thorax 2019 Jan;74(1):33-42 [FREE Full text] [doi: 10.1136/thoraxjnl-2017-211264] [Medline: 29793970]
32. Garrison KA, Pal P, O'Malley SS, Pittman BP, Gueorguieva R, Rojiani R, et al. Craving to quit: a randomized controlled trial of smartphone app-based mindfulness training for smoking cessation. Nicotine Tob Res 2020 Mar 16;22(3):324-331 [FREE Full text] [doi: 10.1093/ntr/nty126] [Medline: 29917096]
33. Noone C, Hogan MJ. A randomised active-controlled trial to examine the effects of an online mindfulness intervention on executive control, critical thinking and key thinking dispositions in a university student sample. BMC Psychol 2018 Apr 5;6(1):13 [FREE Full text] [doi: 10.1186/s40359-018-0226-3] [Medline: 29622047]
34. Rosen KD, Paniagua SM, Kazanis W, Jones S, Potter JS. Quality of life among women diagnosed with breast cancer: a randomized waitlist controlled trial of commercially available mobile app-delivered mindfulness training. Psychooncology 2018 Aug;27(8):2023-2030. [doi: 10.1002/pon.4764] [Medline: 29766596]
35. Segal ZV, Dimidjian S, Beck A, Boggs JM, Vanderkruik R, Metcalf CA, et al. Outcomes of online mindfulness-based cognitive therapy for patients with residual depressive symptoms: a randomized clinical trial. JAMA Psychiatry 2020 Jun 1;77(6):563-573. [doi: 10.1001/jamapsychiatry.2019.4693] [Medline: 31995132]
36. Yang E, Schamber E, Meyer RM, Gold JI. Happier healers: randomized controlled trial of mobile mindfulness for stress management. J Altern Complement Med 2018 May;24(5):505-513. [doi: 10.1089/acm.2015.0301] [Medline: 29420050]
37. Spears CA, Abroms LC, Glass CR, Hedeker D, Eriksen MP, Cottrell-Daniels C, et al. Mindfulness-based smoking cessation enhanced with mobile technology (iQuit mindfully): pilot randomized controlled trial. JMIR Mhealth Uhealth 2019 Jun 24;7(6):e13059 [FREE Full text] [doi: 10.2196/13059] [Medline: 31237242]

38. Huberty J, Green J, Glissmann C, Larkey L, Puzia M, Lee C. Efficacy of the mindfulness meditation mobile app 'calm' to reduce stress among college students: randomized controlled trial. JMIR Mhealth Uhealth 2019 Jun 25;7(6):e14273 [FREE Full text] [doi: 10.2196/14273] [Medline: 31237569]
39. Firth J, Torous J, Nicholas J, Carney R, Rosenbaum S, Sarris J. Can smartphone mental health interventions reduce symptoms of anxiety? A meta-analysis of randomized controlled trials. J Affect Disord 2017 Aug 15;218:15-22 [FREE Full text] [doi: 10.1016/j.jad.2017.04.046] [Medline: 28456072]
40. 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]
41. Linardon J, Cuijpers P, Carlbring P, Messer M, Fuller-Tyszkiewicz M. The efficacy of app-supported smartphone interventions for mental health problems: a meta-analysis of randomized controlled trials. World Psychiatry 2019 Oct;18(3):325-336 [FREE Full text] [doi: 10.1002/wps.20673] [Medline: 31496095]
42. Spijkerman MP, Pots WT, Bohlmeijer ET. Effectiveness of online mindfulness-based interventions in improving mental health: a review and meta-analysis of randomised controlled trials. Clin Psychol Rev 2016 Apr;45:102-114 [FREE Full text] [doi: 10.1016/j.cpr.2016.03.009] [Medline: 27111302]
43. Goldstein J. Mindfulness: a Practical Guide to Awakening. Boulder, CO: Sounds True; 2013.
44. Davidson RJ, Kaszniak AW. Conceptual and methodological issues in research on mindfulness and meditation. Am Psychol 2015 Oct;70(7):581-592 [FREE Full text] [doi: 10.1037/a0039512] [Medline: 26436310]
45. Grossman P, Van Dam NT. Mindfulness, by any other name…: trials and tribulations of in western psychology and science. Contemporary Buddhism 2011 Jun 14;12(1):219-239. [doi: 10.1080/14639947.2011.564841]
46. Lutz A, Jha AP, Dunne JD, Saron CD. Investigating the phenomenological matrix of mindfulness-related practices from a neurocognitive perspective. Am Psychol 2015 Oct;70(7):632-658 [FREE Full text] [doi: 10.1037/a0039585] [Medline: 26436313]
47. van Dam NT, van Vugt MK, Vago DR, Schmalzl L, Saron CD, Olendzki A, et al. Mind the hype: a critical evaluation and prescriptive agenda for research on mindfulness and meditation. Perspect Psychol Sci 2018 Jan;13(1):36-61 [FREE Full text] [doi: 10.1177/1745691617709589] [Medline: 29016274]
48. Kabat-Zinn J. Wherever You Go, There You Are. New York, USA: Hyperion; 1994.
49. Carpenter JK, Conroy K, Gomez AF, Curren LC, Hofmann SG. The relationship between trait mindfulness and affective symptoms: a meta-analysis of the five facet mindfulness questionnaire (FFMQ). Clin Psychol Rev 2019 Dec;74:101785. [doi: 10.1016/j.cpr.2019.101785] [Medline: 31751877]
50. Giluk TL. Mindfulness, big five personality, and affect: a meta-analysis. Personal Individ Differ 2009 Dec;47(8):805-811. [doi: 10.1016/j.paid.2009.06.026]
51. Karyadi KA, VanderVeen JD, Cyders MA. A meta-analysis of the relationship between trait mindfulness and substance use behaviors. Drug Alcohol Depend 2014 Oct 1;143:1-10 [FREE Full text] [doi: 10.1016/j.drugalcdep.2014.07.014] [Medline: 25113434]
52. Mattes J. Systematic review and meta-analysis of correlates of FFMQ mindfulness facets. Front Psychol 2019;10:2684 [FREE Full text] [doi: 10.3389/fpsyg.2019.02684] [Medline: 31866899]
53. Schumer MC, Lindsay EK, Creswell JD. Brief mindfulness training for negative affectivity: a systematic review and meta-analysis. J Consult Clin Psychol 2018 Jul;86(7):569-583 [FREE Full text] [doi: 10.1037/ccp0000324] [Medline: 29939051]
54. Dahl CJ, Lutz A, Davidson RJ. Reconstructing and deconstructing the self: cognitive mechanisms in meditation practice. Trends Cogn Sci 2015 Sep;19(9):515-523 [FREE Full text] [doi: 10.1016/j.tics.2015.07.001] [Medline: 26231761]
55. Fox KC, Dixon ML, Nijeboer S, Girn M, Floman JL, Lifshitz M, et al. Functional neuroanatomy of meditation: a review and meta-analysis of 78 functional neuroimaging investigations. Neurosci Biobehav Rev 2016 Jun;65:208-228. [doi: 10.1016/j.neubiorev.2016.03.021] [Medline: 27032724]
56. Singer T, Engert V. It matters what you practice: differential training effects on subjective experience, behavior, brain and body in the ReSource Project. Curr Opin Psychol 2019 Aug;28:151-158. [doi: 10.1016/j.copsyc.2018.12.005] [Medline: 30684917]
57. Kok BE, Singer T. Phenomenological fingerprints of four meditations: differential state changes in affect, mind-wandering, meta-cognition, and interoception before and after daily practice across 9 months of training. Mindfulness (N Y) 2017;8(1):218-231 [FREE Full text] [doi: 10.1007/s12671-016-0594-9] [Medline: 28163798]
58. Hutcherson CA, Seppala EM, Gross JJ. Loving-kindness meditation increases social connectedness. Emotion 2008 Oct;8(5):720-724. [doi: 10.1037/a0013237] [Medline: 18837623]
59. Salzberg S. Lovingkindness: the Revolutionary Art of Happiness. Boston, UK: Shambhala; 2004.
60. Kirby JN, Tellegen CL, Steindl SR. A meta-analysis of compassion-based interventions: current state of knowledge and future directions. Behav Ther 2017 Nov;48(6):778-792. [doi: 10.1016/j.beth.2017.06.003] [Medline: 29029675]
61. Beck A, Rush A, Shaw B, Emery G. Cognitive Therapy of Depression. New York, USA: Guilford Press; 1979.
62. Mani M, Kavanagh DJ, Hides L, Stoyanov SR. Review and evaluation of mindfulness-based iPhone apps. JMIR Mhealth Uhealth 2015 Aug 19;3(3):e82 [FREE Full text] [doi: 10.2196/mhealth.4328] [Medline: 26290327]

63. Galante J, Bekkers M, Mitchell C, Gallacher J. Loving-kindness meditation effects on well-being and altruism: a mixed-methods online RCT. Appl Psychol Health Well Being 2016 Nov;8(3):322-350. [doi: 10.1111/aphw.12074] [Medline: 27333950]
64. Krieger T, Reber F, von Glutz B, Urech A, Moser CT, Schulz A, et al. An internet-based compassion-focused intervention for increased self-criticism: a randomized controlled trial. Behav Ther 2019 Mar;50(2):430-445. [doi: 10.1016/j.beth.2018.08.003] [Medline: 30824257]
65. Sommers-Spijkerman MP, Trompetter HR, Schreurs KM, Bohlmeijer ET. Compassion-focused therapy as guided self-help for enhancing public mental health: a randomized controlled trial. J Consult Clin Psychol 2018 Feb;86(2):101-115. [doi: 10.1037/ccp0000268] [Medline: 29265836]
66. Mak WW, Tong AC, Yip SY, Lui WW, Chio FH, Chan AT, et al. Efficacy and moderation of mobile app-based programs for mindfulness-based training, self-compassion training, and cognitive behavioral psychoeducation on mental health: randomized controlled noninferiority trial. JMIR Ment Health 2018 Oct 11;5(4):e60 [FREE Full text] [doi: 10.2196/mental.8597] [Medline: 30309837]
67. Rodgers RF, Donovan E, Cousineau T, Yates K, McGowan K, Cook E, et al. BodiMojo: efficacy of a mobile-based intervention in improving body image and self-compassion among adolescents. J Youth Adolesc 2018 Jul;47(7):1363-1372. [doi: 10.1007/s10964-017-0804-3] [Medline: 29349593]
68. Imhoff-Smith T, Rosenkranz M, Goldberg S. Impact of Targeted Mental Exercises on Awareness, Connection, Insight: a Pilot RCT. Open Science Framework. 2019. URL: https://osf.io/rdey9 [accessed 2020-03-07]
69. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, REDCap Consortium. The REDCap consortium: building an international community of software platform partners. J Biomed Inform 2019 Jul;95:103208 [FREE Full text] [doi: 10.1016/j.jbi.2019.103208] [Medline: 31078660]
70. Diener E, Suh EM, Lucas RE, Smith HL. Subjective well-being: three decades of progress. Psychol Bull 1999;125(2):276-302. [doi: 10.1037/0033-2909.125.2.276]
71. Ryff CD. Psychological well-being revisited: advances in the science and practice of eudaimonia. Psychother Psychosom 2014;83(1):10-28 [FREE Full text] [doi: 10.1159/000353263] [Medline: 24281296]
72. Lutz A, Slagter HA, Dunne JD, Davidson RJ. Attention regulation and monitoring in meditation. Trends Cogn Sci 2008 Apr;12(4):163-169 [FREE Full text] [doi: 10.1016/j.tics.2008.01.005] [Medline: 18329323]
73. Pilkonis PA, Choi SW, Reise SP, Stover AM, Riley WT, Cella D, PROMIS Cooperative Group. Item banks for measuring emotional distress from the patient-reported outcomes measurement information system (PROMIS): depression, anxiety, and anger. Assessment 2011 Sep;18(3):263-283 [FREE Full text] [doi: 10.1177/1073191111411667] [Medline: 21697139]
74. Cohen S, Williamson GM. Perceived stress in a probability sample of the United States. In: The Social Psychology of Health: Claremont Symposium on Applied Social Psychology. Newbury Park, CA: Sage; 1988.
75. Lee RM, Draper M, Lee S. Social connectedness, dysfunctional interpersonal behaviors, and psychological distress: testing a mediator model. J Couns Psychol 2001;48(3):310-318. [doi: 10.1037/0022-0167.48.3.310]
76. Davis MH. Measuring individual differences in empathy: evidence for a multidimensional approach. J Pers Soc Psychol 1983 Jan;44(1):113-126. [doi: 10.1037/0022-3514.44.1.113]
77. Sprecher S, Fehr B. Compassionate love for close others and humanity. J Soc Pers Relatsh 2016 Jun 30;22(5):629-651. [doi: 10.1177/0265407505056439]
78. Grant AM, Franklin J, Langford P. The self-reflection and insight scale: a new measure of private self-consciousness. Soc Behav Pers 2002 Jan 1;30(8):821-835. [doi: 10.2224/sbp.2002.30.8.821]
79. Ehring T, Zetsche U, Weidacker K, Wahl K, Schönfeld S, Ehlers A. The perseverative thinking questionnaire (PTQ): validation of a content-independent measure of repetitive negative thinking. J Behav Ther Exp Psychiatry 2011 Jun;42(2):225-232 [FREE Full text] [doi: 10.1016/j.jbtep.2010.12.003] [Medline: 21315886]
80. Forman EM, Herbert JD, Juarascio AS, Yeomans PD, Zebell JA, Goetter EM, et al. The Drexel defusion scale: a new measure of experiential distancing. J Context Behav Sci 2012 Dec;1(1-2):55-65. [doi: 10.1016/j.jcbs.2012.09.001]
81. Baer RA, Smith GT, Hopkins J, Krietemeyer J, Toney L. Using self-report assessment methods to explore facets of mindfulness. Assessment 2006 Mar;13(1):27-45. [doi: 10.1177/1073191105283504] [Medline: 16443717]
82. Snijders T, Bosker R. Multilevel Analysis: an Introduction to Basic and Advance Multilevel Modeling. Second Edition. London, UK: Sage; 2012.
83. Graham JW. Missing data analysis: making it work in the real world. Annu Rev Psychol 2009;60:549-576. [doi: 10.1146/annurev.psych.58.110405.085530] [Medline: 18652544]
84. Nieuwenhuis R, Grotenhuis M, Pelzer B. Influence.me: tools for detecting influential data in mixed effects models. R J 2012;4(2):38. [doi: 10.32614/rj-2012-011]
85. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Ser B Methodol 2018 Dec 5;57(1):289-300. [doi: 10.1111/j.2517-6161.1995.tb02031.x]
86. Schalet BD, Pilkonis PA, Yu L, Dodds N, Johnston KL, Yount S, et al. Clinical validity of PROMIS depression, anxiety, and anger across diverse clinical samples. J Clin Epidemiol 2016 May;73:119-127 [FREE Full text] [doi: 10.1016/j.jclinepi.2015.08.036] [Medline: 26931289]

87. Hirshberg MJ, Goldberg SB, Rosenkranz M, Davidson RJ. Prevalence of harm in mindfulness-based stress reduction. Psychol Med 2020 Aug 18:1-9. [doi: 10.1017/S0033291720002834] [Medline: 32807249]
88. Choi SW, Schalet B, Cook KF, Cella D. Establishing a common metric for depressive symptoms: linking the BDI-II, CES-D, and PHQ-9 to PROMIS depression. Psychol Assess 2014 Jun;26(2):513-527 [FREE Full text] [doi: 10.1037/a0035768] [Medline: 24548149]
89. Schalet BD, Cook KF, Choi SW, Cella D. Establishing a common metric for self-reported anxiety: linking the MASQ, PANAS, and GAD-7 to PROMIS Anxiety. J Anxiety Disord 2014 Jan;28(1):88-96 [FREE Full text] [doi: 10.1016/j.janxdis.2013.11.006] [Medline: 24508596]
90. Roberti JW, Harrington LN, Storch EA. Further psychometric support for the 10 - item version of the Perceived Stress Scale. J Coll Couns 2006;9(2):135-147. [doi: 10.1002/j.2161-1882.2006.tb00100.x]
91. Lee RM, Dean BL, Jung K. Social connectedness, extraversion, and subjective well-being: Testing a mediation model. Personal Individ Differ 2008 Oct;45(5):414-419. [doi: 10.1016/j.paid.2008.05.017]
92. Cordier R, Milbourn B, Martin R, Buchanan A, Chung D, Speyer R. A systematic review evaluating the psychometric properties of measures of social inclusion. PLoS One 2017;12(6):e0179109 [FREE Full text] [doi: 10.1371/journal.pone.0179109] [Medline: 28598984]
93. Lang AJ, Malaktaris AL, Casmar P, Baca SA, Golshan S, Harrison T, et al. Compassion meditation for posttraumatic stress disorder in veterans: a randomized proof of concept study. J Trauma Stress 2019 Apr;32(2):299-309. [doi: 10.1002/jts.22397] [Medline: 30929283]
94. Cliffordson C. The hierarchical structure of empathy: dimensional organization and relations to social functioning. Scand J Psychol 2002 Feb;43(1):49-59. [doi: 10.1111/1467-9450.00268] [Medline: 11885760]
95. Hojat M, Mangione S, Kane GC, Gonnella JS. Relationships between scores of the Jefferson Scale of Physician Empathy (JSPE) and the Interpersonal Reactivity Index (IRI). Med Teach 2005 Nov;27(7):625-628. [doi: 10.1080/01421590500069744] [Medline: 16332555]
96. Harrington R, Loffredo DA, Perz CA. Dispositional mindfulness as a positive predictor of psychological well-being and the role of the private self-consciousness insight factor. Personal Individ Differ 2014 Dec;71:15-18. [doi: 10.1016/j.paid.2014.06.050]
97. Van Dam NT, Earleywine M, Danoff-Burg S. Differential item function across meditators and non-meditators on the Five Facet Mindfulness Questionnaire. Personal Individ Differ 2009 Oct;47(5):516-521. [doi: 10.1016/j.paid.2009.05.005]
98. Goldberg SB, Del Re AC, Hoyt WT, Davis JM. The secret ingredient in mindfulness interventions? A case for practice quality over quantity. J Couns Psychol 2014 Jul;61(3):491-497 [FREE Full text] [doi: 10.1037/cou0000032] [Medline: 25019551]
99. Lykins EL, Baer RA. Psychological functioning in a sample of long-term practitioners of mindfulness meditation. J Cogn Psychother 2009 Aug 01;23(3):226-241. [doi: 10.1891/0889-8391.23.3.226]
100. Baron RM, Kenny DA. The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. J Pers Soc Psychol 1986 Dec;51(6):1173-1182. [doi: 10.1037//0022-3514.51.6.1173] [Medline: 3806354]
101. Imai K, Keele L, Tingley D. A general approach to causal mediation analysis. Psychol Methods 2010 Dec;15(4):309-334. [doi: 10.1037/a0020761] [Medline: 20954780]
102. Cella D, Riley W, Stone A, Rothrock N, Reeve B, Yount S, PROMIS Cooperative Group. The patient-reported outcomes measurement information system (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005-2008. J Clin Epidemiol 2010 Nov;63(11):1179-1194. [doi: 10.1016/j.jclinepi.2010.04.011] [Medline: 20685078]
103. PROMIS Score Cut Points. Health Measures. URL: https://www.healthmeasures.net/score-and-interpret/interpret-scores/promis/promis-score-cut-points [accessed 2020-03-07]
104. Polit DF, Gillespie BM. Intention-to-treat in randomized controlled trials: recommendations for a total trial strategy. Res Nurs Health 2010 Aug;33(4):355-368. [doi: 10.1002/nur.20386] [Medline: 20645423]
105. Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. J Stat Soft 2015;67(1):-. [doi: 10.18637/jss.v067.i01]
106. R: a Language and Environment for Statistical Computing. R Foundation for Statistical Computing. 2018. URL: http://www.R-project.org/ [accessed 2020-03-07]
107. Linardon J, Fuller-Tyszkiewicz M. Attrition and adherence in smartphone-delivered interventions for mental health problems: a systematic and meta-analytic review. J Consult Clin Psychol 2020 Jan;88(1):1-13. [doi: 10.1037/ccp0000459] [Medline: 31697093]
108. Tingley D, Yamamoto T, Hirose K, Keele L, Imai K. Package for causal mediation analysis. J Stat Soft 2014;59(5):-. [doi: 10.18637/jss.v059.i05]
109. Champley S. PWR: Basic Functions for Power Analysis. The Comprehensive R Archive Network. URL: https://CRAN.R-project.org/package=pwr [accessed 2020-03-07]
110. Eysenbach G. The law of attrition. J Med Internet Res 2005 Mar 31;7(1):e11 [FREE Full text] [doi: 10.2196/jmir.7.1.e11] [Medline: 15829473]

111. Blankers M, Koeter MW, Schippers GM. Missing data approaches in eHealth research: simulation study and a tutorial for nonmathematically inclined researchers. J Med Internet Res 2010 Dec 19;12(5):e54. [doi: 10.2196/jmir.1448] [Medline: 21169167]
112. 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] [Medline: 19403466]
113. Enders CK. Applied Missing Data Analysis. New York: Guilford Press; 2010.
114. Schafer JL, Graham JW. Missing data: our view of the state of the art. Psychol Methods 2002 Jun;7(2):147-177. [Medline: 12090408]
115. Parsons CE, Crane C, Parsons LJ, Fjorback LO, Kuyken W. Home practice in mindfulness-based cognitive therapy and mindfulness-based stress reduction: a systematic review and meta-analysis of participants' mindfulness practice and its association with outcomes. Behav Res Ther 2017 Aug;95:29-41 [FREE Full text] [doi: 10.1016/j.brat.2017.05.004] [Medline: 28527330]
116. Gu J, Strauss C, Bond R, Cavanagh K. How do mindfulness-based cognitive therapy and mindfulness-based stress reduction improve mental health and wellbeing? A systematic review and meta-analysis of mediation studies. Clin Psychol Rev 2015 Apr;37:1-12. [doi: 10.1016/j.cpr.2015.01.006] [Medline: 25689576]
117. Holt-Lunstad J, Smith TB, Baker M, Harris T, Stephenson D. Loneliness and social isolation as risk factors for mortality: a meta-analytic review. Perspect Psychol Sci 2015 Mar;10(2):227-237. [doi: 10.1177/1745691614568352] [Medline: 25910392]
118. Leigh-Hunt N, Bagguley D, Bash K, Turner V, Turnbull S, Valtorta N, et al. An overview of systematic reviews on the public health consequences of social isolation and loneliness. Public Health 2017 Nov;152:157-171. [doi: 10.1016/j.puhe.2017.07.035] [Medline: 28915435]
119. Cohen J, Cohen P, West S, Aiken L. Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences. Third Edition. Mahwah, New Jersey: Lawrence Erlbaum Associates Publishers; 2003.
120. Michie S, Yardley L, West R, Patrick K, Greaves F. Developing and evaluating digital interventions to promote behavior change in health and health care: recommendations resulting from an international workshop. J Med Internet Res 2017 Jun 29;19(6):e232 [FREE Full text] [doi: 10.2196/jmir.7126] [Medline: 28663162]
121. van Dulmen S, Sluijs E, van Dijk L, de Ridder D, Heerdink R, Bensing J. Patient adherence to medical treatment: a review of reviews. BMC Health Serv Res 2007 Apr 17;7:55 [FREE Full text] [doi: 10.1186/1472-6963-7-55] [Medline: 17439645]
122. Klasnja P, Hekler EB, Shiffman S, Boruvka A, Almirall D, Tewari A, et al. Microrandomized trials: an experimental design for developing just-in-time adaptive interventions. Health Psychol 2015 Dec;34S:1220-1228 [FREE Full text] [doi: 10.1037/hea0000305] [Medline: 26651463]
123. Torous J, Firth J. The digital placebo effect: mobile mental health meets clinical psychiatry. Lancet Psychiatry 2016 Feb;3(2):100-102. [doi: 10.1016/S2215-0366(15)00565-9] [Medline: 26851322]
124. Tracey TJ. A note on socially desirable responding. J Couns Psychol 2016 Mar;63(2):224-232. [doi: 10.1037/cou0000135] [Medline: 26689626]
125. Ly KH, Trüschel A, Jarl L, Magnusson S, Windahl T, Johansson R, et al. Behavioural activation versus mindfulness-based guided self-help treatment administered through a smartphone application: a randomised controlled trial. BMJ Open 2014 Jan 9;4(1):e003440 [FREE Full text] [doi: 10.1136/bmjopen-2013-003440] [Medline: 24413342]
126. Chiesa A, Calati R, Serretti A. Does mindfulness training improve cognitive abilities? A systematic review of neuropsychological findings. Clin Psychol Rev 2011 Apr;31(3):449-464. [doi: 10.1016/j.cpr.2010.11.003] [Medline: 21183265]
127. Pascoe MC, Thompson DR, Jenkins ZM, Ski CF. Mindfulness mediates the physiological markers of stress: systematic review and meta-analysis. J Psychiatr Res 2017 Dec;95:156-178. [doi: 10.1016/j.jpsychires.2017.08.004] [Medline: 28863392]
128. Treves IN, Tello LY, Davidson RJ, Goldberg SB. The relationship between mindfulness and objective measures of body awareness: a meta-analysis. Sci Rep 2019 Nov 22;9(1):17386 [FREE Full text] [doi: 10.1038/s41598-019-53978-6] [Medline: 31758073]
129. Stachl C, Au Q, Schoedel R, Gosling SD, Harari GM, Buschek D, et al. Predicting personality from patterns of behavior collected with smartphones. Proc Natl Acad Sci U S A 2020 Jul 28;117(30):17680-17687 [FREE Full text] [doi: 10.1073/pnas.1920484117] [Medline: 32665436]
130. Pratap A, Atkins DC, Renn BN, Tanana MJ, Mooney SD, Anguera JA, et al. The accuracy of passive phone sensors in predicting daily mood. Depress Anxiety 2019 Jan;36(1):72-81. [doi: 10.1002/da.22822] [Medline: 30129691]
131. Flückiger C, Del Re AC, Wampold BE, Horvath AO. The alliance in adult psychotherapy: a meta-analytic synthesis. Psychotherapy (Chic) 2018 Dec;55(4):316-340. [doi: 10.1037/pst0000172] [Medline: 29792475]
132. Henson P, Wisniewski H, Hollis C, Keshavan M, Torous J. Digital mental health apps and the therapeutic alliance: initial review. BJPsych Open 2019 Jan;5(1):e15 [FREE Full text] [doi: 10.1192/bjo.2018.86] [Medline: 30762511]
133. Benish SG, Quintana S, Wampold BE. Culturally adapted psychotherapy and the legitimacy of myth: a direct-comparison meta-analysis. J Couns Psychol 2011 Jul;58(3):279-289. [doi: 10.1037/a0023626] [Medline: 21604860]

134. Griner D, Smith TB. Culturally adapted mental health intervention: a meta-analytic review. Psychotherapy (Chic) 2006;43(4):531-548. [doi: 10.1037/0033-3204.43.4.531] [Medline: 22122142]
135. Paradies Y, Ben J, Denson N, Elias A, Priest N, Pieterse A, et al. Racism as a determinant of health: a systematic review and meta-analysis. PLoS One 2015;10(9):e0138511 [FREE Full text] [doi: 10.1371/journal.pone.0138511] [Medline: 26398658]
136. DeLuca SM, Kelman AR, Waelde LC. A systematic review of ethnoracial representation and cultural adaptation of mindfulness- and meditation-based interventions. Psychol Stud 2018 May 22;63(2):117-129. [doi: 10.1007/s12646-018-0452-z]
137. Waldron EM, Hong S, Moskowitz JT, Burnett-Zeigler I. A systematic review of the demographic characteristics of participants in US-based randomized controlled trials of mindfulness-based interventions. Mindfulness 2018 Mar 12;9(6):1671-1692. [doi: 10.1007/s12671-018-0920-5]

Abbreviations

FDR: false discovery rate
HMP: Healthy Minds Program
IRI: Interpersonal Reactivity Index
MAR: missing at random
MBCT: mindfulness-based cognitive therapy
MBI: meditation-based intervention
MBSR: mindfulness-based stress reduction
mHealth: mobile health
MLM: multilevel model
MNAR: missing not at random
PTQ: Perseverative Thinking Questionnaire
RCT: randomized controlled trial

Edited by J Torous; submitted 28.08.20; peer-reviewed by A Ferguson, T Krieger, O Bur; comments to author 19.10.20; revised version received 23.10.20; accepted 26.10.20; published 27.11.20.

Please cite as:
Goldberg SB, Imhoff-Smith T, Bolt DM, Wilson-Mendenhall CD, Dahl CJ, Davidson RJ, Rosenkranz MA
Testing the Efficacy of a Multicomponent, Self-Guided, Smartphone-Based Meditation App: Three-Armed Randomized Controlled Trial
JMIR Ment Health 2020;7(11):e23825
URL: http://mental.jmir.org/2020/11/e23825/
doi:10.2196/23825
PMID:33245288

Original Paper

Behavior of Callers to a Crisis Helpline Before and During the COVID-19 Pandemic: Quantitative Data Analysis

Robin Turkington[1], BSc, MSc; Maurice Mulvenna[1], BSc, MPhil, PhD; Raymond Bond[1], BCompSc, PhD; Edel Ennis[2], PhD; Courtney Potts[1], BSc, MSc; Ciaran Moore[3], BCom; Louise Hamra[3], BA; Jacqui Morrissey[4], BA, MSc; Mette Isaksen[4], BA, MA; Elizabeth Scowcroft[4], BSc, MSc, PhD; Siobhan O'Neill[2], BA, MPsychSc, PhD

[1]School of Computing, Ulster University, Newtownabbey, United Kingdom
[2]School of Psychology, Ulster University, Coleraine, United Kingdom
[3]Samaritans Ireland, Dublin, Ireland
[4]Samaritans UK, Ewell, United Kingdom

Corresponding Author:
Robin Turkington, BSc, MSc
School of Computing
Ulster University
Shore Road
Newtownabbey
United Kingdom
Phone: 44 28 9036 6129
Email: Turkington-R@ulster.ac.uk

Abstract

Background: The World Health Organization declared the outbreak of COVID-19 to be an international pandemic in March 2020. While numbers of new confirmed cases of the disease and death tolls are rising at an alarming rate on a daily basis, there is concern that the pandemic and the measures taken to counteract it could cause an increase in distress among the public. Hence, there could be an increase in need for emotional support within the population, which is complicated further by the reduction of existing face-to-face mental health services as a result of measures taken to limit the spread of the virus.

Objective: The objective of this study was to determine whether the COVID-19 pandemic has had any influence on the calls made to Samaritans Ireland, a national crisis helpline within the Republic of Ireland.

Methods: This study presents an analysis of calls made to Samaritans Ireland in a four-week period before the first confirmed case of COVID-19 (calls=41,648, callers=3752) and calls made to the service within a four-week period after a restrictive lockdown was imposed by the government of the Republic of Ireland (calls=46,043, callers=3147). Statistical analysis was conducted to explore any differences between the duration of calls in the two periods at a global level and at an hourly level. We performed k-means clustering to determine the types of callers who used the helpline based on their helpline call usage behavior and to assess the impact of the pandemic on the caller type usage patterns.

Results: The analysis revealed that calls were of a longer duration in the postlockdown period in comparison with the pre–COVID-19 period. There were changes in the behavior of individuals in the cluster types defined by caller behavior, where some caller types tended to make longer calls to the service in the postlockdown period. There were also changes in caller behavior patterns with regard to the time of day of the call; variations were observed in the duration of calls at particular times of day, where average call durations increased in the early hours of the morning.

Conclusions: The results of this study highlight the impact of COVID-19 on a national crisis helpline service. Statistical differences were observed in caller behavior between the prelockdown and active lockdown periods. The findings suggest that service users relied on crisis helpline services more during the lockdown period due to an increased sense of isolation, worsening of underlying mental illness due to the pandemic, and reduction or overall removal of access to other support resources. Practical implications and limitations are discussed.

(JMIR Ment Health 2020;7(11):e22984) doi:10.2196/22984

KEYWORDS

COVID-19; coronavirus; pandemic; mental health; crisis helplines; machine learning; clustering; caller behavior

Introduction

COVID-19 has spread globally; by May 1, 2020, the disease had reached over 215 countries and territories worldwide, with over 3.1 million confirmed cases and 224,172 confirmed deaths [1]. On January 30, 2020, the World Health Organization (WHO) declared the COVID-19 outbreak to be a Public Health Emergency of International Concern (PHEIC); this was only the sixth time the WHO had declared a PHEIC since 2005. Since this declaration, many governments have imposed lockdown measures preventing people from mixing and attending work or school and have instructed the general public to adhere to “social distancing” or self-isolate to slow the spread of the disease. There is concern that COVID-19 has had a negative impact on the mental well-being of individuals, particularly with the implementation of strict lockdown measures.

Data indicate that symptoms of anxiety and depression increased as a result of the pandemic and peaked at the time of government announcements regarding restrictions to curb the spread of the virus [2,3]. Furthermore, evidence indicates that the people at highest risk of having symptoms of mental illness are those in lower income households, who are required to self-isolate because of their risk of having an adverse outcome if they contract the virus, and who have existing mental health problems [2]. Although the public relies on news and media reports to assess the evolving state of the crisis, repeated exposure to negative news coverage of the pandemic can enhance psychological distress, leading to the development and worsening of mental illness symptoms [4,5]. The need to self-isolate may cause people to feel isolated from friends and family, and the impact is greater for people with underlying psychological vulnerability or mental illness who rely on social and support networks to stay well. Moreover, many mental health services and the availability of face-to-face support were disrupted as a result of the restrictions, leaving people who depend on such services without the support groups and resources that benefit them [6]. Some individuals may be reluctant to seek help and support from face-to-face mental health services due to concern that such services are being overwhelmed or out of fear of contracting COVID-19 in a face-to-face appointment setting [5]. These individuals may therefore rely more on help from additional remote services, such as suicide prevention and crisis helplines [6,7].

Suicide prevention and crisis helplines provide support to people who are experiencing a crisis, which is defined as a state of psychological disequilibrium where the individual’s coping mechanisms are no longer effective [8,9]. Samaritans Ireland is a crisis support and suicide prevention helpline that provides free confidential support to people, many of whom are highly distressed, suicidal, and may have underlying mental illnesses [10]. Callers to Samaritans speak to trained volunteers who provide respectful and nonjudgmental active listening. Callers may contact the service on a single occasion or repeatedly; for a systematic review on repeat callers, see Middleton et al [11]. Samaritans is one of the oldest helplines in operation within the United Kingdom and Ireland, and it provides free confidential support 24 hours per day every day of the year [10].

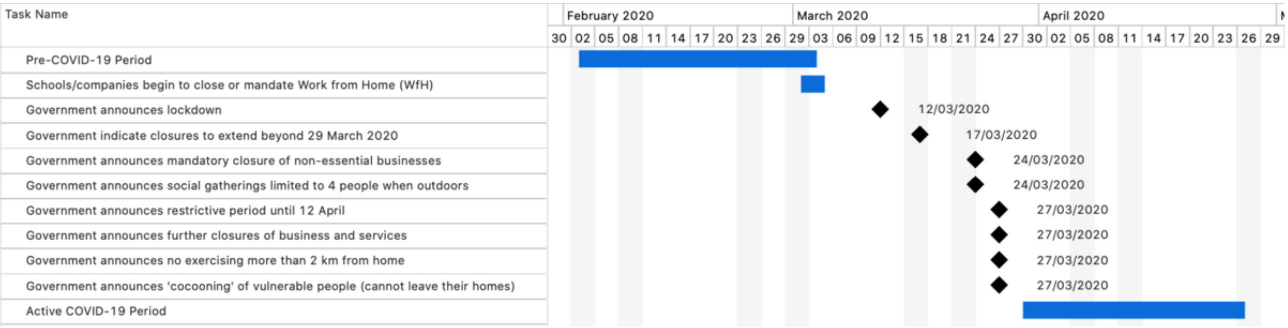
Understanding a caller’s needs based on contact patterns and how they interact with the service can be valuable for operational purposes and for understanding how suicide prevention and crisis helplines can be used in population-level well-being and suicide prevention programs [12-14]. Patterns of calls to crisis or suicide prevention helplines may also reflect the impact of COVID-19 on suicidal distress, mental health, and well-being. An understanding of caller patterns will also help inform population-level support planning and the guidance provided regarding accessing emotional support. Crisis lines have been identified as an important means of supporting people who are at risk of mental illness in a time where face-to-face contact must be avoided [15]. It is therefore important to examine patterns of help seeking and use of crisis helplines so that services can respond accordingly.

The objective of the current study is to analyze the potential impact of the COVID-19 pandemic on the behavior of callers to Samaritans Ireland. The study analyzed caller behavior from a four-week period prior and up to the first confirmed case of COVID-19 in the Republic of Ireland and compared it to caller behavior from a similar four-week period after the introduction of the lockdown restrictions by the Irish government (Figure 1).

This research addresses the following research questions:

- Which aspects of caller behavior have changed as a result of the COVID-19 pandemic?
- Which cohorts of callers demonstrate changes in behavior in response to the restrictions imposed to address the pandemic?

Figure 1. Timelines of the Pre–COVID-19 Period, dates of key government messages, and the Active COVID-19 Period.



Methods

Data Background

Calls made to Samaritans Ireland are made from, and answered, within the Republic of Ireland. Each call that is made to the service is represented electronically. The fields that were used for the analysis were the caller identifier (which was modified to anonymize the caller); the date and time stamp of the call; and the duration of the call (in seconds).

Call detail records were retrieved from a dedicated application programming interface (API) created by Samaritans. This enabled the creation of a real-time anonymized data stream for analysis. There are no identifiable aspects within the call data, nor are there any data fields that contain complementary information about the caller's condition or any indications of whether the caller is living with any physical or mental illness, if they are already a service user, the level of distress the caller is in, or whether the caller is experiencing a crisis at the time of the call.

Analysis of Key Dates in the Timeline Before and After the COVID-19 Lockdown

Using the dedicated API, call data from January 1, 2019, to May 11, 2020, were retrieved, which equated to 1,054,089 calls by 30,659 callers. After the data were subjected to data cleansing (ie, normalization of time and date stamps, inspection and removal of anomalous data entries), two new call data frame subsets were derived from the original data set. One of these new data frames consisted of all calls that were made to the service within a four-week period prior to the first confirmed case of COVID-19 in the Republic of Ireland (Week 6 to Week 9 of 2020; calls=41,648, callers=3752); this is referred to as the Pre–COVID-19 period. The other new data frame consisted of all calls that were made to the service within a four-week period after the commencement of the lockdown (Week 14 to Week 17 of 2020; (calls=46,043, callers=3147); this is referred to as the Active COVID-19 period.

Unsupervised Machine Learning Using k-Means Clustering

The call data were subjected to k-means clustering to discover the types of callers that used the service. In k-means clustering, data points are grouped together based on their closeness by Euclidean distance. In other words, the aim is to find *k* groups in *n* objects based on the similarity of their characteristics, where the characteristics in one group show high similarity with each other but low similarity with other groups [16,17]. To determine the types of callers that use the service, three attributes of caller behavior were selected for clustering: the number of calls made by each caller, the mean duration of the calls by each caller, and the standard deviation of the duration of the calls made by each caller.

These attributes of caller behavior were chosen due to their explanatory power; the number of calls that the caller makes to the service indicates the frequency of help-seeking, the average duration of the calls indicates how complex the calls may be, and the standard deviation of the call duration indicates the consistency (or inconsistency) of the call durations. New data sets that contained numerical summaries of these attributes for callers in each period were created. Each attribute was then scaled for standardization, which is an appropriate prerequisite for k-means clustering. The next stage was to specify the value of *k*, which specifies the number of groups into which the data are to be clustered. Based on previous research that used k-means clustering to identify caller types [17-19], *k* was set to 5, meaning that 5 caller types were discovered as a result of clustering.

The 5 caller types can be described as follows:

1. Typical callers: These callers make approximately 5 calls on average to the helpline. Calls last approximately 5 minutes on average and are consistent in duration; this group is the largest in size.
2. High Frequency callers: These callers make the most calls on average to the helpline, averaging hundreds of calls. Calls are very short in duration but can be highly variable; this group is the smallest in size.
3. Regular callers: These callers make the second highest average number of calls to the service. They can make upwards of a hundred calls to the service on average; however, this number can be greater or smaller depending on the period of the data set being analyzed. Calls can last approximately 10 minutes on average, although the duration of the calls may be much longer or shorter.
4. Unpredictable callers: These callers make approximately 8-12 calls on average. Calls can be upwards of 25 minutes long; however, the call duration is the most variable of all the cluster types.
5. Single Lengthy callers: These callers make 1 to 2 calls on average. The call duration is the longest and most consistent of all the caller types.

Clustering was conducted on both the Pre–COVID-19 and Active COVID-19 periods, and any changes in caller *archetypes* (cluster types defined by caller behavior) will be discussed.

Data Analytics Materials

R 3.5.1 (the R Project) was used in all aspects of analysis. The ggplot2 package [20] was used to create data visualizations, while base R functions were used to conduct k-means clustering analysis on the call data and other statistical analyses. The unpaired Wilcoxon rank sum test was conducted to compare differences in each hourly mean duration between the Pre–COVID-19 and Active COVID-19 periods.

Results

Differences in Call Duration Between the Pre–COVID-19 and Active COVID-19 Periods

There was an increase in the mean and median duration of calls by hour of day across all but one hour from the Pre–COVID-19 period (Figure 2A and Figure 2B; mean=620 seconds/10.33 minutes; median=250 seconds/4.17 minutes) to the Active COVID-19 period (mean=709 seconds/11.82 minutes; median=388 seconds/6.47 minutes); a *t* test found a significant difference (t_{51434}=11.94, *P*<.001) in call duration between the

Pre–COVID-19 period and the Active COVID-19 period. An unpaired Wilcoxon rank sum test was conducted to compare differences between each hourly mean duration between the Pre–COVID-19 and Active COVID-19 periods; 22 of the 24 hours yielded a statistically significant difference between the average durations in both periods. There was a lower density (the term *density* refers to the distribution of calls over a continuous interval; in other words, the distribution of calls based on their duration) of calls with a shorter duration from the Pre–COVID-19 period to the Active COVID-19 period (Figure 2C and Figure 2D). In contrast, there was an increase in the density of calls of a longer duration from the Pre–COVID-19 period to the Active COVID-19 period.

Figure 2. Differences in (A) call duration for answered calls between the Pre–COVID-19 and Active COVID-19 periods; (B) mean duration of calls by hour of day; and (C,D) density of call duration.

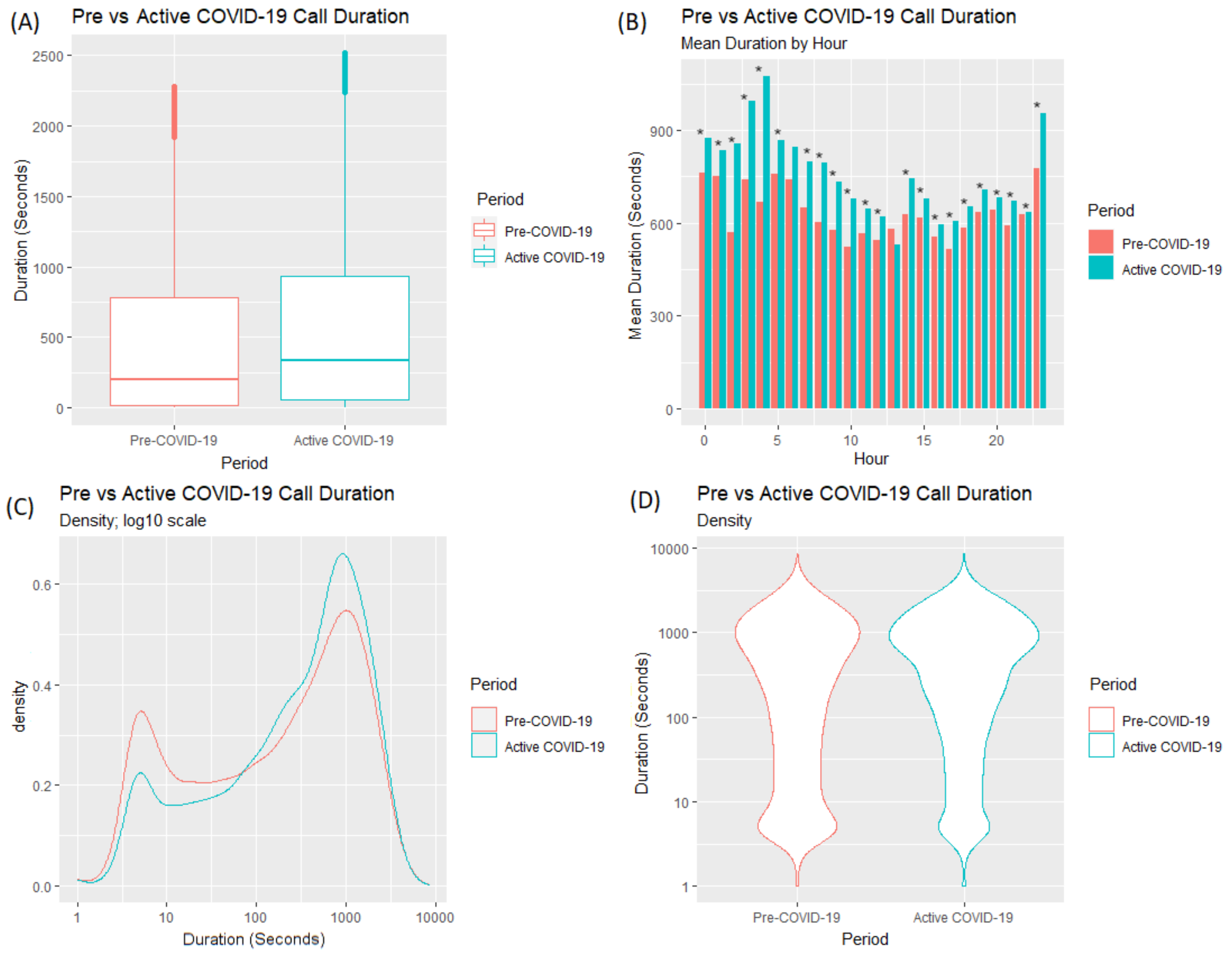
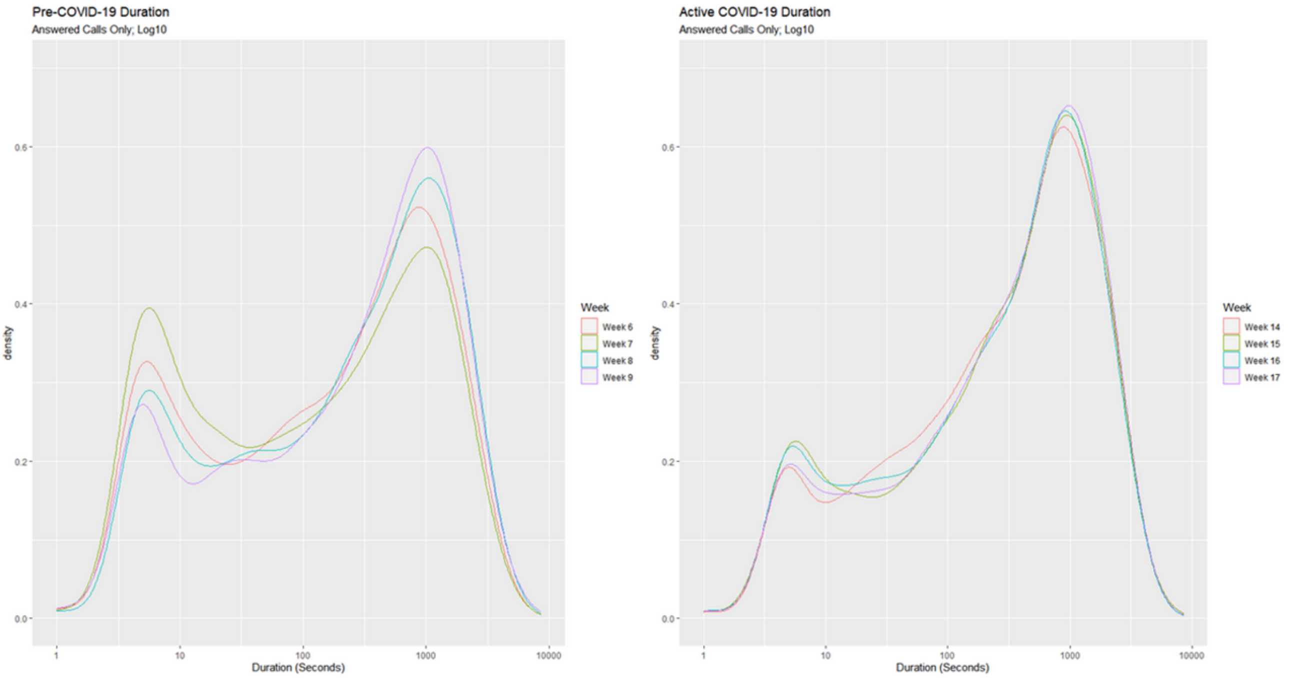


Figure 3 displays the changes in the density of call durations for answered calls across each of the consecutive weeks in the Pre–COVID-19 and Active COVID-19 periods. Within the Pre–COVID-19 period, there was a considerable amount of variation in call duration density across the weeks. In the Active COVID-19 period, the distribution of the data appeared to be stable across Week 14, Week 15, Week 16, and Week 17.

Figure 3. Densities of call durations of answered calls by week within the Pre–COVID-19 period (left) and Active COVID-19 period (right).



Compared to the Pre–COVID-19 period, all four Active COVID-19 weeks display a lower density of calls with short durations, with Week 14 displaying the lowest density of calls with short durations and Week 15 displaying the highest. There was little variation in the density of calls with long durations between Week 14, Week 15, and Week 16; all four Active COVID-19 weeks show a higher density of calls with longer durations than all the Pre–COVID-19 weeks. The same analysis was conducted on calls made in comparative weeks in 2019. A similar bimodal profile was found; however, the trends in Weeks 14-17 in 2019 were not as significant as those observed in the Active COVID-19 weeks in 2020. In 2019, the percentage increase in long calls (ie, calls over 600 seconds/10 minutes in duration) from the Pre–COVID-19 period to the Active COVID-19 period was 2.1; in 2020, the percentage increase in long calls from the Pre–COVID-19 period to the Active COVID-19 period was 6.1.

A one-way test was conducted to determine if there were any significant differences in the duration of calls between the weeks within both periods. In the Pre–COVID-19 period, there were significant differences across the four weeks ($F_{3,15528}$=50.19, *P*<.001; Table 1). In the Active COVID-19 period, there were no significant differences across the four weeks ($F_{3,13544}$=2.1026, *P*=.09; Table 2).

Table 1. Statistical comparisons between weeks in the Pre–COVID-19 period.

Period	Difference	Confidence interval	*t* value	Degrees of freedom	*P* value	Adjusted *P* value
Week 7-Week 6	–51	–85to –17	3.88	15,081	<.001	<.001
Week 8-Week 6	84	48 to 121	5.93	14,165	<.001	<.001
Week 9-Week 6	97	59 to 135	6.6	12,968	<.001	<.001
Week 8-Week 7	136	100 to 171	9.75	14,183	<.001	<.001
Week 9-Week 7	148	111 to 186	10.3	12,808	<.001	<.001
Week 9-Week 8	13	–27 to 52	0.83	13,099	.84	.84

Table 2. Statistical comparisons between weeks in the Active COVID-19 period.

Period	Difference	Confidence interval	*t* value	Degrees of freedom	*P* value	Adjusted *P* value
Week 15-Week 14	18.7	–21.7 to 59	1.19	12,159	.63	.95
Week 16-Week 14	–3.3	–42.7 to 36	0.21	12,247	>.99	>.99
Week 17-Week 14	30.5	–10 to 71	1.93	11,951	.21	.64
Week 16-Week 15	–22	–61.6 to 18	1.43	12,405	.48	.95
Week 17-Week 15	11.7	–28.9 to 52	0.74	12,101	.88	>.99
Week 17-Week 16	33.8	–5.8 to 73	2.19	12,166	.13	.64

Clustering Analysis: Differences in Caller Characteristics Between the Pre–COVID-19 and Active COVID-19 Period

Table 3 and Table 4 show the cluster characteristics of the callers who contacted the service within the Pre–COVID-19 period and Active COVID-19 period, respectively. Some notable fluctuations in cluster means are noticeable in the High Frequency caller clusters in relation to the number of calls made to the service. Otherwise, the cluster centroids remain stable between the Pre–COVID-19 and Active COVID-19 periods.

Table 3. Cluster centroids of callers within the Pre–COVID-19 period.

Caller type	Average number of calls	Mean duration	Standard deviation of duration	Cluster size	Within sum of squares
Typical	3.685639	213.3336	48.00642	2284	501.3016
High Frequency	424.8	294.3486	419.1129	35	868.5136
Regular	22.039326	675.4679	725.19471	712	658.9827
Unpredictable	9.686636	1586.4138	1626.68238	217	430.4528
Single Lengthy	1.126984	1977.2117	33.53513	504	600.5583

Table 4. Cluster centroids of callers within the Active COVID-19 period.

Caller type	Average number of calls	Mean duration	Standard deviation of duration	Cluster size	Within sum of squares
Typical	3.953930	183.4235	48.05530	1845	379.3097
High Frequency	396.113636	314.5596	456.93319	44	694.0469
Regular	26.471942	655.0175	658.17619	695	576.7514
Unpredictable	12.257282	1336.8626	1493.99101	206	328.7420
Single Lengthy	1.109244	1950.0461	25.32212	357	568.8687

Figure 4 displays the distributions of the call duration data for the five caller archetypes. The five caller types show similar distributions from the Pre–COVID-19 period to the Active COVID-19 period, with some changes for Typical callers, High Frequency callers, and Single Lengthy callers. Typical callers begin to trend toward longer calls while showing a reduced density of calls with shorter durations; a similar trend is also noticeable with High Frequency Callers.

Single Lengthy callers exhibited a higher density of call frequency, with longer call durations in the Active COVID-19 period compared to the Pre–COVID-19 period. The Wilcox ranked sum test was conducted to determine whether the caller types differed in call duration between periods. Statistically significant differences were observed within the Typical caller type (W=709173, *P*<.001), High Frequency caller type (W=42237518, *P*<.001), and Regular caller type (W=49030484, *P*<.001).

Figure 5 displays the differences in the mean durations at each hour of the day at the cluster level. The Wilcox ranked sum test was conducted to compare the differences in mean duration (in seconds) at each hourly interval from the Pre–COVID-19 period to the Active COVID-19 period for each cluster.

There were visible differences in mean call durations by hour between the Pre–COVID-19 and Active COVID-19 periods for High Frequency callers. In contrast, Single Lengthy and Unpredictable callers showed similar patterns during both periods.

Figure 4. Densities of call durations for the five caller types from the Pre–COVID-19 period to the Active COVID-19 period.

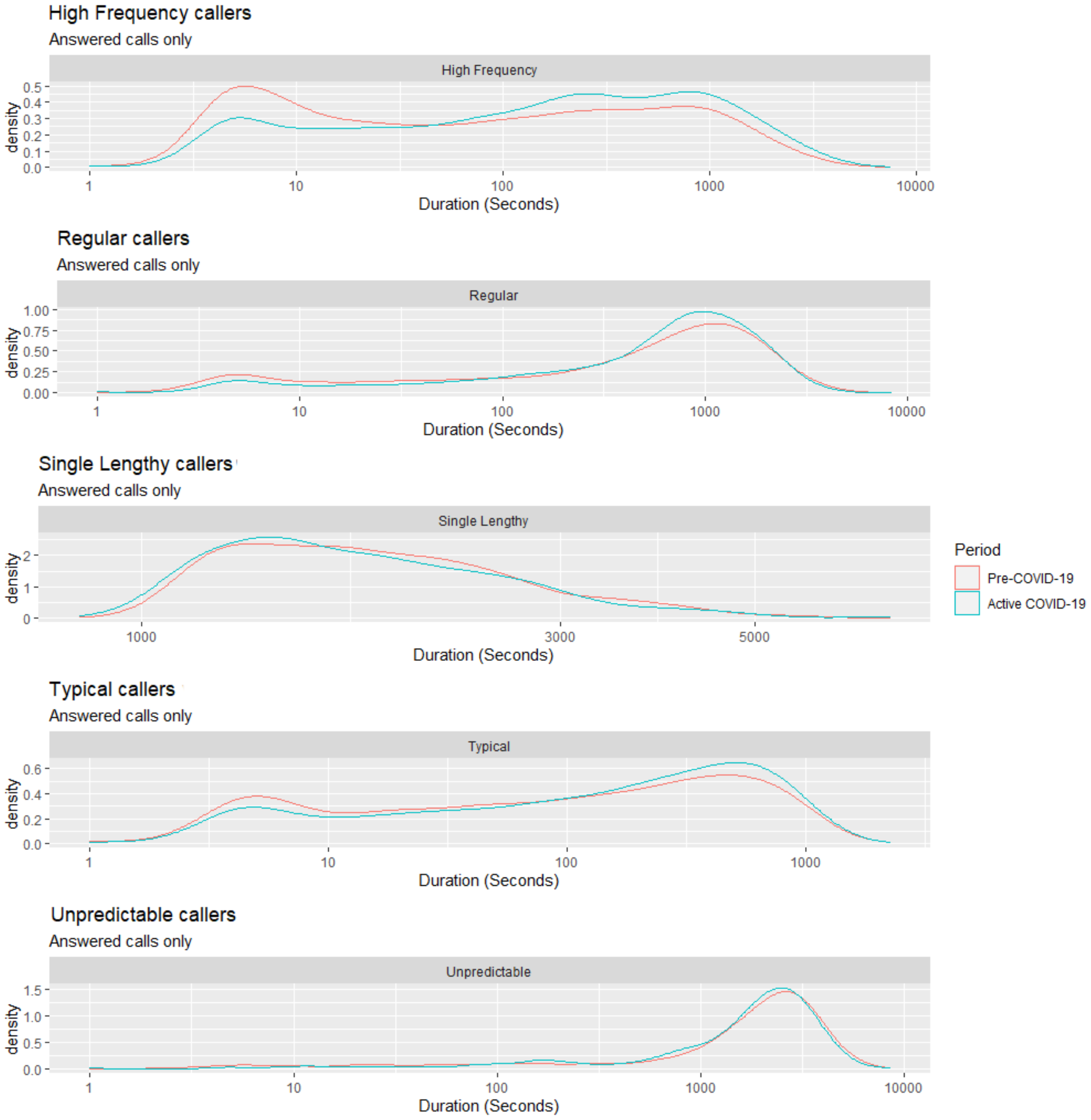
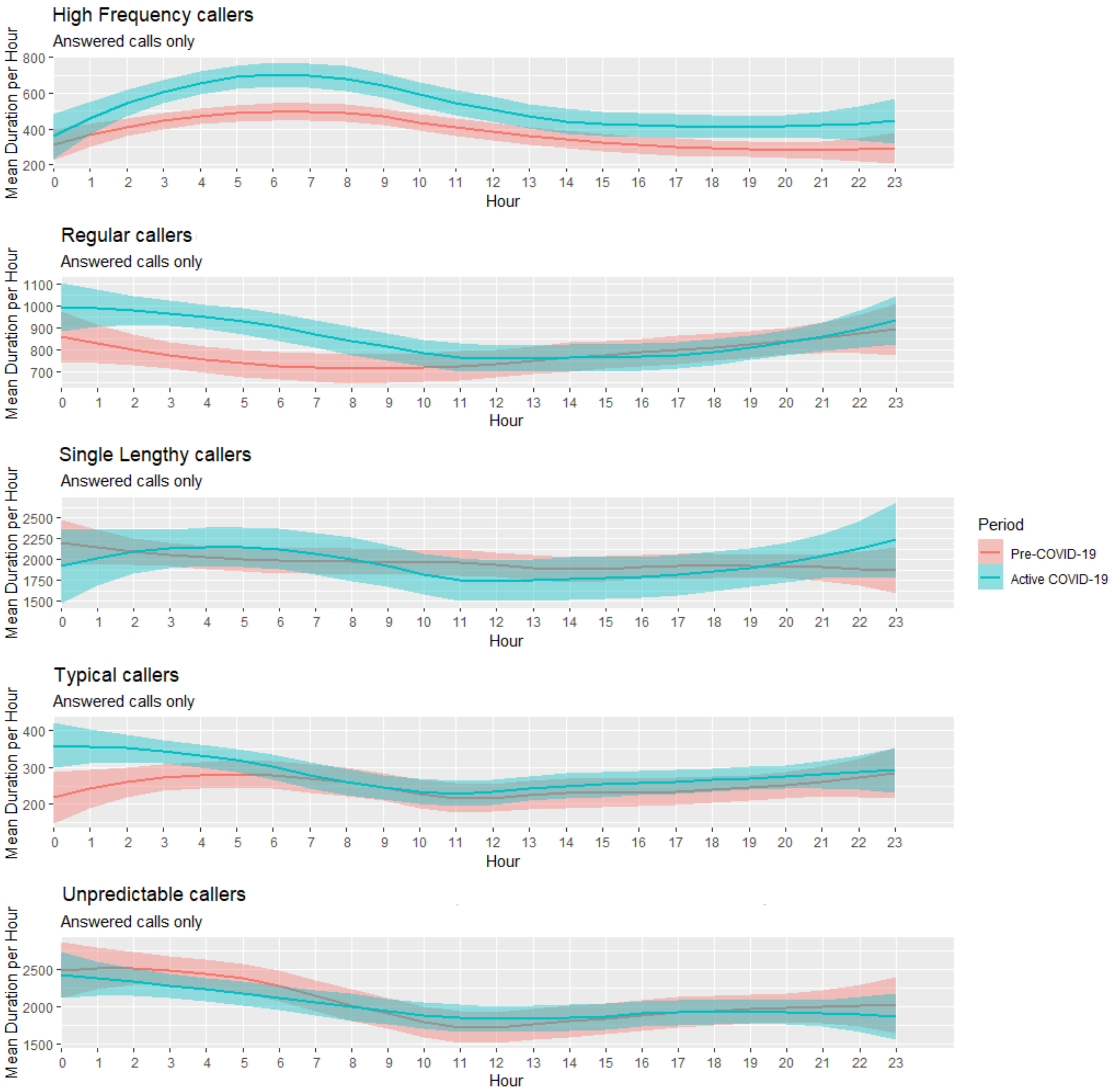


Figure 5. Smoothed conditional means plots displaying the mean call durations in seconds for the five caller types across a 24-hour period between the Pre-COVID-19 and Active COVID-19 periods.



Common Callers

This phase of analysis focuses on the callers who contacted the service in both the Pre–COVID-19 period and the Active COVID-19 period, termed *common callers*.

Figure 6 displays the distributions of the call duration of answered calls for callers who contacted the service in both the Pre–COVID-19 and Active COVID-19 periods. The mean and median call duration (Figure 6A) increased from the Pre–COVID-19 period (mean=628 seconds/10.47 minutes; median=279 seconds/4.65 minutes) to the Active COVID-19 period (mean=689 seconds/11.48 minutes; median=360 seconds/6 minutes); a *t* test found a significant difference between periods with regards to call duration (t_{40657}=–7.2291, *P*<.001). Similar to the findings regarding call duration from all callers (Figure 2C and 2D), the density of call durations for shorter calls decreased from the Pre–COVID-19 period to the Active COVID-19 period, while the density of call durations for longer calls increased from the Pre–COVID-19 period to the Active COVID-19 period (Figure 6C and Figure 6D). The same analysis was conducted on common callers within the comparative weeks in 2019. In 2019, the percentage increase in long calls (ie, calls over 600 seconds/10 minutes) from the Pre–COVID-19 period to the Active COVID-19 period was 2.1%; in 2020, the percentage increase in long calls from the Pre–COVID-19 period to Active COVID-19 period was 4%.

Figure 7 displays the changes in the density of call durations for answered calls across each of the consecutive weeks for the callers who contacted the service in both the Pre–COVID-19 and Active COVID-19 periods. There was a similar trend in the density of call durations for all callers (Figure 3). Within the Pre–COVID-19 period, we observed a variation in call duration density across the weeks.

Figure 6. Differences in (A) call duration for answered calls by common callers in the Pre–COVID-19 and Active COVID-19 periods; (B) mean duration of calls by hour of day; and (C,D) density of call duration.

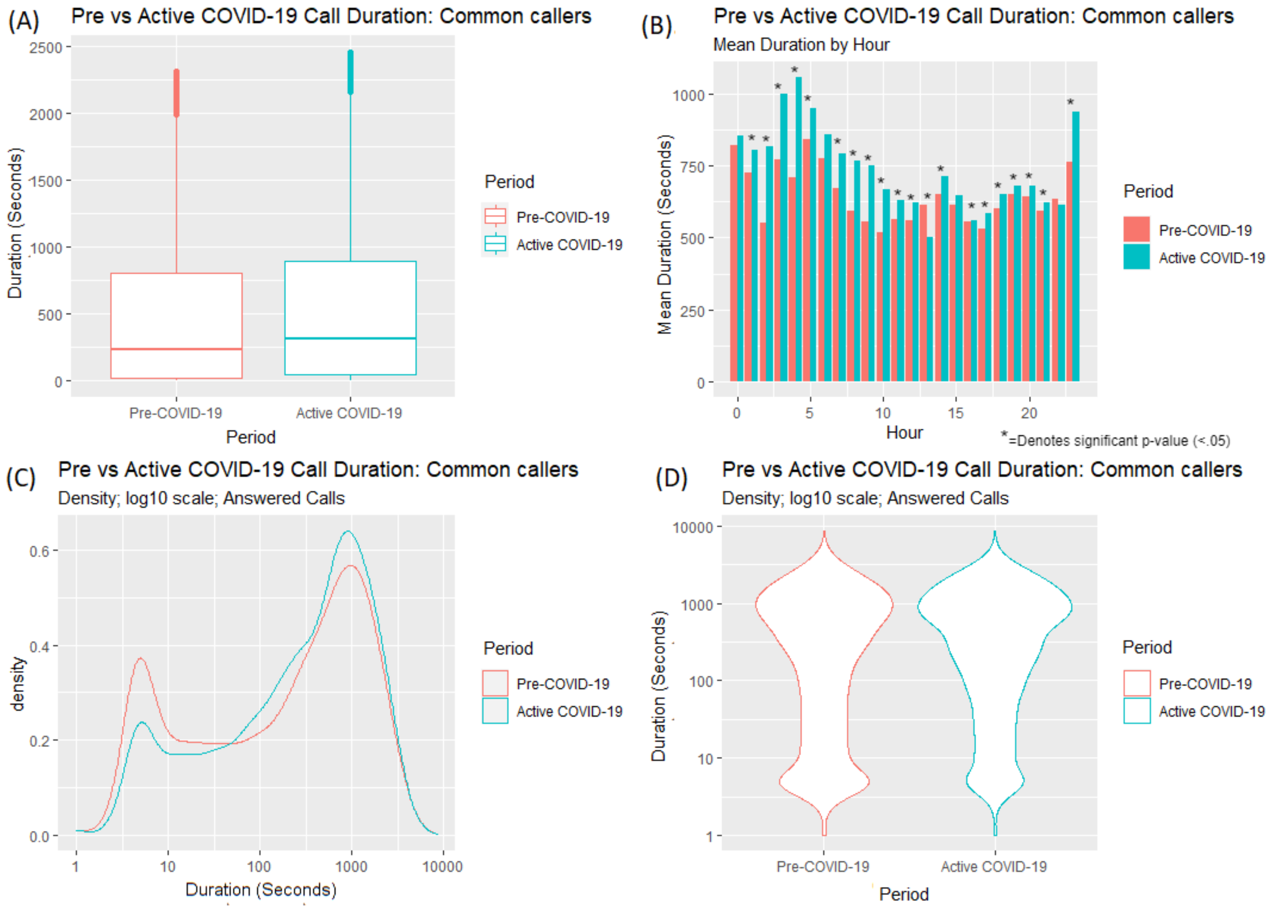
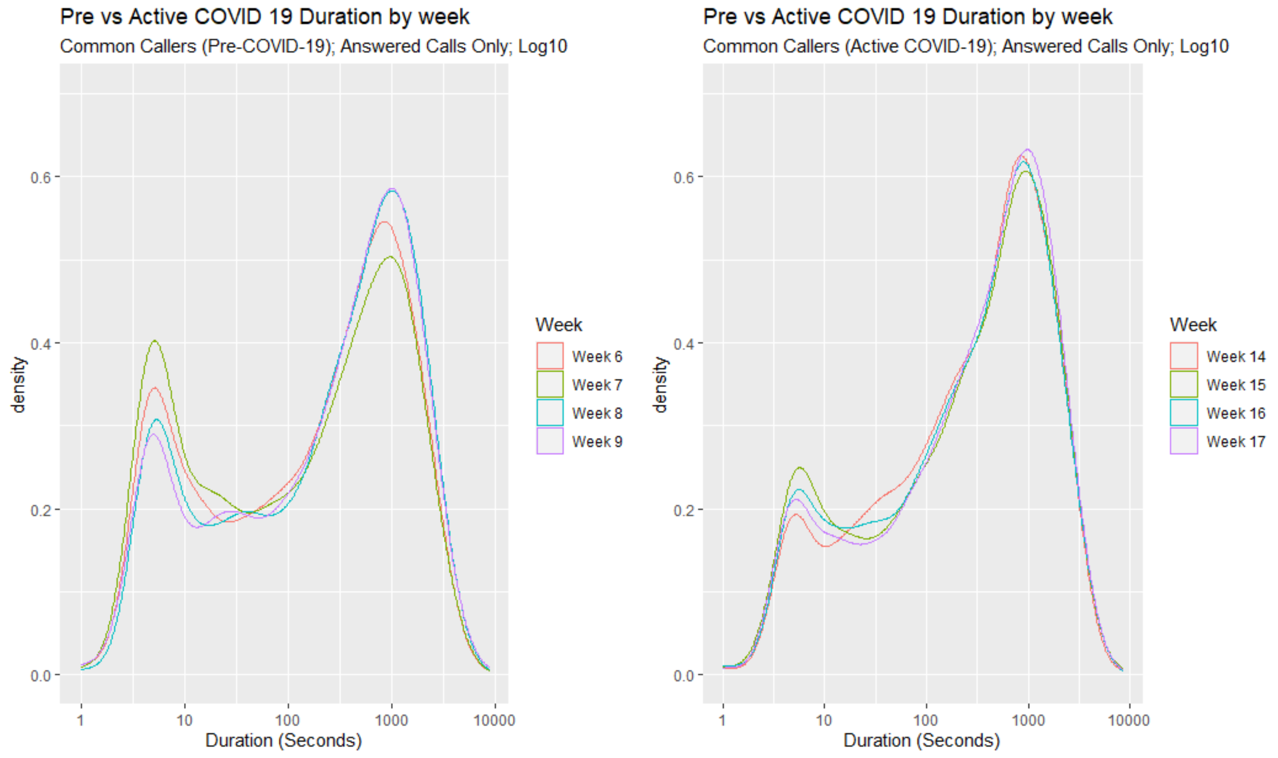


Figure 7. Density of call durations for answered calls by each week by common callers in the Pre–COVID-19 period (left) and Active COVID-19 period (right).



In the Active COVID-19 period, there was little variation in call duration density across the weeks. Compared to the Pre–COVID-19 period, all four Active COVID-19 weeks displayed a lower density of calls with shorter call durations in total. The density of call durations was also analyzed for the same weeks in 2019, and while a similar trend in the density of call durations was observed, the trend was not as significant as those observed in the 2020 weeks.

A one-way test showed that in the Pre–COVID-19 period, there were significant differences across the four weeks ($F_{3,12285}$=28.8, *P*<.001; Table 5). In the Active COVID-19 period, there were no significant differences across the four weeks ($F_{3,10726}$=1.8009, *P*=.14; Table 6).

Table 5. Comparisons between weeks in the Pre–COVID-19 period (2020) for common callers.

Period	Difference	Confidence interval	*t* value	Degrees of freedom	*P* value	Adjusted *P* value
Week 7-Week 6	–25.2	–64 to 13	1.68	11,466	.34	.4
Week 8-Week 6	91.6	51 to 133	5.75	10,923	<.001	<.001
Week 9-Week 6	88.9	47 to 131	5.47	10,635	<.001	<.001
Week 8-Week 7	116.8	76 to 157	7.43	11,121	<.001	<.001
Week 9-Week 7	114.1	73 to 155	7.11	10,766	<.001	<.001
Week 9-Week 8	–2.7	–46 to 41	0.16	10,719	>.99	>.99

Table 6. Comparisons between weeks in the Active COVID-19 period (2020) for common callers.

Period	Difference	Confidence interval	*t* value	Degrees of freedom	*P* value	Adjusted *P* value
Week 15-Week 14	13.5	–31.1 to 58	0.78	9607	.86	>.99
Week 16-Week 14	2.5	–41 to 46	0.15	9815	>.99	>.99
Week 17-Week 14	36.4	–8.4 to 81	2.09	9407	.16	.61
Week 16-Week 15	–11	–55.4 to 33	0.64	9854	.92	>.99
Week 17-Week 15	–22.9	–22.7 to 69	1.29	9504	.57	>.99
Week 17-Week 16	34	–10.6 to 78	1.96	9634	.20	.61

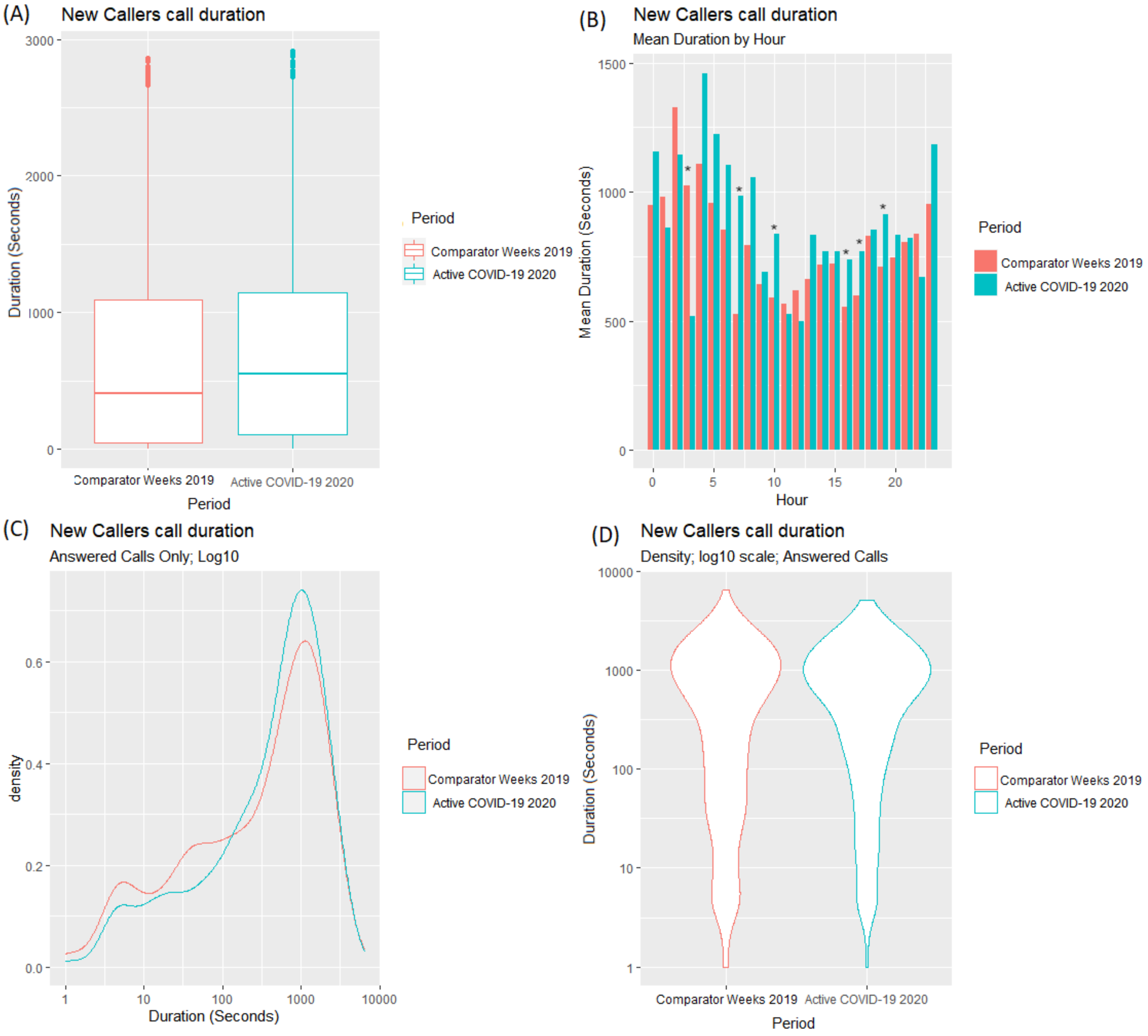
New Callers

This section examines the behavior of callers who contacted the service for the first time in the Active COVID-19 period and do not appear in any other record within the data (as far back as January 1, 2019). For comparison, callers who contacted the service for the first time (again, since January 1, 2019) in weeks 14-17 in 2019 (the same time period in 2019 as the Active COVID-19 period in 2020) were also analyzed (see Figure 8). Both these cohorts are termed *new callers.*

Figure 8 displays the distributions of the call duration for answered calls in the comparative Weeks 14-17 in 2019 and the same weeks in 2020 (the Active COVID-19 period).

The mean and median call durations (Figure 8A) increased from the comparative 2019 period (mean=766 seconds/12.77 minutes; median= 450 seconds/7.5 minutes) to the Active COVID-19 period in 2020 (mean=831 seconds/13.85 minutes; median=582 seconds/9.7 minutes). A *t* test found a significant difference between periods with regards to call duration (t_{3212}=–2.1943, *P*=.03).

Figure 8. Differences in (A) call duration for answered calls by new callers in the Pre–COVID-19 and Active COVID-19 periods; (B) mean duration of calls by hour of day; and (C,D) density of call duration.



Discussion

Principal Findings

The aims of this study were to examine which aspects of caller behavior to a national crisis helpline were impacted as a result of the COVID-19 pandemic and to determine which cohorts of callers were impacted the most. Call detail records spanning two periods of four weeks were compared; one period spanned the four weeks prior to the first confirmed case of COVID-19 in Ireland, and the other spanned the four weeks after the implementation of the restrictive lockdown by the Irish government. We also used k-means clustering to identify the types of callers who contacted the service based on their usage patterns. These data may help us understand which cohorts of callers required more support as a result of lockdown restrictions applied to manage the spread of the virus. Differences in call durations were analyzed for all callers, at a cluster level, and at a level that included “common callers” who had contacted the service in both time periods.

The findings suggest that caller behavior changed as a result of the COVID-19 pandemic and that due to the pandemic, callers made more calls of a longer duration and made fewer calls of a shorter duration. There were fewer calls around 5 minutes in duration and more calls lasting 30 minutes and longer. There were also differences in average duration at an hourly level between both periods, with a statistically significant difference in mean call duration in 22 of 24 hours. In the Active COVID-19 time period, average call durations peaked at around 3 AM and 4 AM (Figure 2). However, in the Active COVID-19 time period, while call durations across the day were greater for each hour, this increase was relatively small (approximately 3-5 minutes). Therefore, consideration must be given as to whether these differences at an hourly level are actually practically significant.

Differences in call duration were then analyzed on a week-by-week basis to determine if any progressive changes were evident. There was a variation in call duration density across the weeks in the Pre–COVID-19 period, meaning that the call duration varied from week to week. In contrast, within the Active COVID-19 weeks, very little variation in call duration density was observed, with no significant differences observed across these weeks. There was a trend toward more calls with longer durations and fewer calls with shorter durations across the Active COVID-19 weeks, with a sustained high density of longer duration calls across the four weeks at this time.

The analysis demonstrated significant differences in the behaviors of some of the caller types that were identified through clustering. There were changes within three particular cohorts of callers. High Frequency callers, Single Lengthy callers, and Typical callers, who previously all exhibited a mix of 5-minute and 30-minute calls, now tended to make longer calls. These changes appear to be more profound between the hours of 1 AM and 6 AM, where these callers spent much longer on the telephone on average.

Due to the lockdown restrictions, many of the existing mental health support groups have either had their services limited or have been removed altogether, particularly if these services relied on face-to-face appointments [6]. For instance, Samaritans Ireland halted all its face-to-face services from the commencement of the lockdown in Ireland but continued to offer telephone and email service [10], and the Health Service Executive of Ireland provided contact references for web-based and telephone support in the absence of face-to-face services [21]. These web-based and telephone services are solutions to the removal of face-to-face services and are considered to be effective in reducing levels of anxiety and depression [6,22]. These resources are vital at this time given that people are more likely to suffer from symptoms of mental illness, particularly anxiety and depression [2,3]. Callers who used the service because they already felt isolated may have an increased sense of isolation due to the restrictions imposed as a result of the pandemic. These findings attest to the impact of the pandemic on mental health and the need for additional support. They may well also reflect increased levels of distress amongst some high risk callers compared with the other caller types, such as the High Frequency callers, Single Lengthy callers, and Typical callers.

Suicide rates have been known to increase as a result of historical pandemics, such as the influenza epidemic in the United States between 1918 and 1919 [23] and the severe acute respiratory syndrome (SARS) epidemic [5,24]. Individuals with pre-existing mental illnesses will likely see their symptoms become exacerbated due to the pandemic. Furthermore, individuals with no pre-existing disorders may begin to develop a disorder, such as depression, anxiety, or posttraumatic stress disorder; these disorders may be more prevalent in essential and frontline health care workers, particularly medical personnel, due to the physical and emotional stress caused by extreme workloads and by experiencing traumatic events in the workplace [5].

There is agreement within the literature that many aspects of an individual's daily routine may have changed as a result of the pandemic [5]. Individuals are staying at home and working from home or may have lost their jobs. People are experiencing reduced social interactions with others during this lockdown and are creating new routines to facilitate family needs. Home confinement has been stressful for many and has disturbed daily routines. Altena et al [25] summarized how confinement imposed by the COVID-19 pandemic can disrupt sleeping patterns by highlighting the factors that impact the stress-sleep link. Individuals who are more susceptible to stress-related sleep disruption are more likely to experience impacts to their circadian rhythms and develop insomnia as a result [25,26]. If the pandemic is causing people to feel more stressed and disrupts sleep and circadian rhythms as a result, this may explain why there has been a shift to a higher frequency of early morning calls, with the longer call durations reflecting daily routines that have been altered as a result of the pandemic.

Future Research

One suggestion for future research is to analyze an additional period of data beyond the Active COVID-19 period to determine whether the aspects of caller behavior analyzed within this study returned to a Pre–COVID-19 period norm as a result of the easing of the lockdown restrictions. As new positive cases decline as a result of social distancing and lockdown measures, this may parallel a decrease in distress within the population, resulting in fewer emotional support calls being made to the service. If a decrease in the number of emotional support calls parallels the easing of lockdown restrictions, this may also represent habituation to a new norm caused by the pandemic or indicate that individuals have developed appropriate coping strategies to alleviate pandemic-related distress [27].

Implications for Policy and Practice

These findings provide an indication of the impact of the pandemic on the behavior of people who use crisis line services. They may also point to the need for high risk individuals to have increased support to mitigate the impact of the virus and measures taken to minimize spread on their well-being and mental health. The changes in the times that people used the service can inform service planning and volunteer scheduling to ensure that more calls could be answered at the new peak times. Although this was not directly assessed within the current study, the increase in the trend toward longer calls being made to the service may have taken up more service capacity. If this is the case, it may be necessary to alter volunteer scheduling to meet this change in demand. Volunteers may have to undergo retraining to prepare for new presenting reasons associated with the pandemic, such as bereavement due to COVID-19, becoming unemployed or furloughed, and increased isolation. Lastly, as call durations have increased over the Active COVID-19 period, callers may require further training in preparation for longer conversations.

This work highlights the need for mental health and well-being and suicide prevention support services, such as crisis helplines, to be provided with the appropriate support and funding to mitigate the impact of the pandemic on the mental well-being of the population. If services such as crisis helplines are funded appropriately, this may also help prevent further worsening of mental health within the population and thus relieve strain on national health services [28,29]. It is important to recognize that while the reproductive rate of COVID-19 is gradually decreasing over time, there is a possibility that the virus can return as a second wave or local outbreak. The findings in this study may be used to inform similar services of how the behavior of their clients may change as a result of the pandemic, which affords services the opportunity to change aspects of their service to mitigate the impact of the pandemic.

Limitations

There are some inherent limitations to consider when interpreting this type of call data. Each caller who contacts the service is represented in the call data by an anonymized identifier based on the telephone number used. It is not possible to know whether the same telephone has been used by multiple individuals to contact the service. This may be the case in a small minority of cases and also in residential settings, and it would lead to misclassification of that caller. In addition, callers who had insufficient contact with the service to accurately classify them as high-frequency users may have been misclassified by the clustering algorithm [17].

One limitation of the current study is that no demographic information about the callers was available for analysis. The Samaritans Ireland service is entirely confidential. Volunteers may ask for the name of the caller, as this is a natural element of conversation; however, callers can remain anonymous if they wish. Moreover, caller demographic information was not available for analysis. However, if caller demographic information were available for analysis, it may be of interest to learn which cohorts of the population were most likely to contact the service during both periods within the study and determine which demographic was most impacted.

Conclusions

This study investigated the impact of the COVID-19 pandemic on aspects of the behavior of callers to Samaritans Ireland, a national crisis helpline. Aspects of behavior from callers who contacted the service within two comparable time periods, a Pre–COVID-19 period and an Active COVID-19 period, were analyzed.

Visible differences were observed in caller behavior from the Pre–COVID-19 to Active COVID-19 periods. Callers made fewer calls of a short duration and trended toward making more calls of a longer duration. Callers also appeared to make longer calls across all but one hour of the day. At a weekly level, the density of call durations was highly variable across the four individual weeks within the Pre–COVID-19 period. In contrast, the density of call durations was highly stable with very few differences between each individual week within the Active COVID-19 period. Moreover, callers trended toward making more and longer calls to the service across the four individual weeks in the Active COVID-19 period. At a cluster level, there were statistical differences between three of five caller types in relation to call duration density; these callers trended toward fewer shorter calls and more longer calls. Changes in the mean duration were observed at hourly intervals, with the most pronounced changes between the hours of 1 AM and 6 AM.

This work provides evidence of the impact of the COVID-19 pandemic on mental well-being within the population and its impact on support-seeking and help-seeking behavior. The patterns identified in this research suggest that callers have additional mental health and suicide prevention support needs as a result of the effects of the pandemic and that helplines can play a vital role in helping to meet these needs.

Acknowledgments

Financial support for this study was provided in part by Samaritans Ireland together with Ireland’s National Office for Suicide Prevention. The funding agreement ensured the authors’ independence in designing the study, interpreting the data, and writing and publishing the manuscript. The authors would like to thank Samaritans Ireland for the provision of the data analyzed in this study.

Conflicts of Interest

None declared.

References

1. Coronavirus disease (COVID-19) pandemic. World Health Organization. URL: https://www.who.int/emergencies/diseases/novel-coronavirus-2019 [accessed 2020-05-01]
2. Stress related to catching COVID-19 is highest among 30-59 year-olds. University College London. 2020 Apr 15. URL: https://www.ucl.ac.uk/news/2020/apr/stress-related-catching-covid-19-highest-among-30-59-year-olds [accessed 2020-06-04]
3. COVID-19 Psychological Research Consortium Study. Ulster University. URL: https://www.ulster.ac.uk/coronavirus/research/impact/psychology-study [accessed 2020-06-04]
4. Garfin DR, Silver RC, Holman EA. The novel coronavirus (COVID-2019) outbreak: Amplification of public health consequences by media exposure. Health Psychol 2020 May;39(5):355-357. [doi: 10.1037/hea0000875] [Medline: 32202824]
5. Gunnell D, Appleby L, Arensman E, Hawton K, John A, Kapur N, COVID-19 Suicide Prevention Research Collaboration. Suicide risk and prevention during the COVID-19 pandemic. Lancet Psychiatry 2020 Jun;7(6):468-471 [FREE Full text] [doi: 10.1016/S2215-0366(20)30171-1] [Medline: 32330430]

6. Wind TR, Rijkeboer M, Andersson G, Riper H. The COVID-19 pandemic: The 'black swan' for mental health care and a turning point for e-health. Internet Interv 2020 Apr;20:100317 [FREE Full text] [doi: 10.1016/j.invent.2020.100317] [Medline: 32289019]
7. Pereira-Sanchez V, Adiukwu F, El Hayek S, Bytyçi DG, Gonzalez-Diaz JM, Kundadak GK, et al. COVID-19 effect on mental health: patients and workforce. Lancet Psychiatry 2020 Jun;7(6):e29-e30. [doi: 10.1016/s2215-0366(20)30153-x]
8. Kalafat J, Gould MS, Munfakh JLH, Kleinman M. An evaluation of crisis hotline outcomes. Part 1: Nonsuicidal crisis callers. Suicide Life Threat Behav 2007 Jun;37(3):322-337. [doi: 10.1521/suli.2007.37.3.322] [Medline: 17579544]
9. Spittal MJ, Fedyszyn I, Middleton A, Bassilios B, Gunn J, Woodward A, et al. Frequent callers to crisis helplines: who are they and why do they call? Aust N Z J Psychiatry 2015 Jan 27;49(1):54-64. [doi: 10.1177/0004867414541154] [Medline: 24972602]
10. Samaritans. URL: https://www.samaritans.org/ [accessed 2020-06-03]
11. Middleton A, Gunn J, Bassilios B, Pirkis J. Systematic review of research into frequent callers to crisis helplines. J Telemed Telecare 2014 Feb 11;20(2):89-98. [doi: 10.1177/1357633x14524156]
12. Mulvenna M, Bond R, Grigorash A, O'Neill S, Ryan A. HILDA - A Health Interaction Log Data Analysis Workflow to Aid Understanding of Usage Patterns and Behaviours. 2018 Feb Presented at: 2nd Symposium on Social Interactions in Complex Intelligent Systems (SICIS) at Artificial Intelligence and Simulation of Behaviour Convention; April 4-6, 2018; Liverpool, UK URL: https://pure.ulster.ac.uk/ws/portalfiles/portal/11666886/MULVENNA-HILDA-SICIS-20180305.pdf
13. Turkington R, Mulvenna M, Bond R, O'Neill S, Armour C. The Application of User Event Log Data for Mental Health and Wellbeing Analysis. In: Proceedings of the 32nd International BCS Human Computer Interaction Conference (HCI). 2018 Jul Presented at: 32nd International BCS Human Computer Interaction Conference (HCI); July 4-6, 2018; Belfast, UK. [doi: 10.14236/ewic/hci2018.4]
14. Turkington R, Mulvenna MD, Bond RR, O'Neill S, Potts C, Armour C, et al. Why do people call crisis helplines? Identifying taxonomies of presenting reasons and discovering associations between these reasons. Health Informatics J 2020 Apr 18:1460458220913429. [doi: 10.1177/1460458220913429] [Medline: 32306837]
15. Holmes EA, O'Connor RC, Perry VH, Tracey I, Wessely S, Arseneault L, et al. Multidisciplinary research priorities for the COVID-19 pandemic: a call for action for mental health science. Lancet Psychiatry 2020 Jun;7(6):547-560. [doi: 10.1016/s2215-0366(20)30168-1]
16. Jain AK. Data clustering: 50 years beyond K-means. Pattern Recognit Lett 2010 Jun;31(8):651-666. [doi: 10.1016/j.patrec.2009.09.011]
17. Turkington R, Bond R, Mulvenna MD, Potts C, O'Neill S, Armour C. User Archetype Discovery By Cluster Analysis of Caller Log Data: Tenure Evolution is Stable as Time Period Reduces. In: ECCE 2019: Proceedings of the 31st European Conference on Cognitive Ergonomics. 2019 Presented at: 31st European Conference on Cognitive Ergonomics; September 11-13, 2019; Belfast, UK p. 196-202. [doi: 10.1145/3335082.3335090]
18. O'Neill S, Bond RR, Grigorash A, Ramsey C, Armour C, Mulvenna MD. Data analytics of call log data to identify caller behaviour patterns from a mental health and well-being helpline. Health Informatics J 2019 Dec 17;25(4):1722-1738. [doi: 10.1177/1460458218792668] [Medline: 30222034]
19. Grigorash A, Bond RR, Mulvenna MD, O’Neill S, Armour C, Ramsey C. Frequency domain analysis of telephone helpline call data. Data Science and Knowledge Engineering for Sensing Decision Support 2018:1267-1272. [doi: 10.1142/9789813273238_0158]
20. Wickham H. Data Analysis. In: ggplot2: Elegant Graphics for Data Analysis. New York, NY: Springer; 2016.
21. Mental health supports and services during COVID-19. Health Service Executive. URL: https://www2.hse.ie/services/mental-health-supports-and-services-during-coronavirus/ [accessed 2020-06-04]
22. Berryhill MB, Culmer N, Williams N, Halli-Tierney A, Betancourt A, Roberts H, et al. Videoconferencing Psychotherapy and Depression: A Systematic Review. Telemed J E Health 2019 Jun;25(6):435-446. [doi: 10.1089/tmj.2018.0058] [Medline: 30048211]
23. Wasserman I. The impact of epidemic, war, prohibition and media on suicide: United States, 1910-1920. Suicide Life Threat Behav 1992;22(2):240-254. [Medline: 1626335]
24. Cheung YT, Chau PH, Yip PSF. A revisit on older adults suicides and Severe Acute Respiratory Syndrome (SARS) epidemic in Hong Kong. Int J Geriatr Psychiatry 2008 Dec;23(12):1231-1238. [doi: 10.1002/gps.2056] [Medline: 18500689]
25. Altena E, Baglioni C, Espie CA, Ellis J, Gavriloff D, Holzinger B, et al. Dealing with sleep problems during home confinement due to the COVID-19 outbreak: Practical recommendations from a task force of the European CBT-I Academy. J Sleep Res 2020 Apr 04;29(4):e13052. [doi: 10.1111/jsr.13052] [Medline: 32246787]
26. Akerstedt T, Kecklund G, Axelsson J. Impaired sleep after bedtime stress and worries. Biol Psychol 2007 Oct;76(3):170-173. [doi: 10.1016/j.biopsycho.2007.07.010] [Medline: 17884278]
27. Fullana MA, Hidalgo-Mazzei D, Vieta E, Radua J. Coping behaviors associated with decreased anxiety and depressive symptoms during the COVID-19 pandemic and lockdown. J Affect Disord 2020 Oct 01;275:80-81 [FREE Full text] [doi: 10.1016/j.jad.2020.06.027] [Medline: 32658829]

28. Torous J, Jän Myrick K, Rauseo-Ricupero N, Firth J. Digital Mental Health and COVID-19: Using Technology Today to Accelerate the Curve on Access and Quality Tomorrow. JMIR Ment Health 2020 Mar 26;7(3):e18848 [FREE Full text] [doi: 10.2196/18848] [Medline: 32213476]
29. Inkster B, O'Brien R, Selby E, Joshi S, Subramanian V, Kadaba M, et al. Digital Health Management During and Beyond the COVID-19 Pandemic: Opportunities, Barriers, and Recommendations. JMIR Ment Health 2020 Jul 6;7(7):e19246. [doi: 10.2196/19246]

Abbreviations

API: application programming interface
PHEIC: Public Health Emergency of International Concern
SARS: severe acute respiratory syndrome
WHO: World Health Organization

Edited by J Torous; submitted 29.07.20; peer-reviewed by S Stjernswärd, G Dosovitsky; comments to author 28.08.20; revised version received 16.09.20; accepted 29.09.20; published 06.11.20.

Please cite as:
Turkington R, Mulvenna M, Bond R, Ennis E, Potts C, Moore C, Hamra L, Morrissey J, Isaksen M, Scowcroft E, O'Neill S
Behavior of Callers to a Crisis Helpline Before and During the COVID-19 Pandemic: Quantitative Data Analysis
JMIR Ment Health 2020;7(11):e22984
URL: http://mental.jmir.org/2020/11/e22984/
doi:10.2196/22984
PMID:33112759

Original Paper

The Relationships of Deteriorating Depression and Anxiety With Longitudinal Behavioral Changes in Google and YouTube Use During COVID-19: Observational Study

Boyu Zhang[1*], BSc; Anis Zaman[1*], PhD; Vincent Silenzio[2], MPH, MD; Henry Kautz[1], PhD; Ehsan Hoque[1], PhD

[1]Department of Computer Science, University of Rochester, Rochester, NY, United States

[2]Department of Urban-Global Public Health, Rutgers University, Piscataway and Newark, NJ, United States

[*]these authors contributed equally

Corresponding Author:
Anis Zaman, PhD
Department of Computer Science
University of Rochester
2513 Wegmans Hall
Rochester, NY, 14627
United States
Phone: 1 6262981861
Email: azaman2@cs.rochester.edu

Abstract

Background: Depression and anxiety disorders among the global population have worsened during the COVID-19 pandemic. Yet, current methods for screening these two issues rely on in-person interviews, which can be expensive, time-consuming, and blocked by social stigma and quarantines. Meanwhile, how individuals engage with online platforms such as Google Search and YouTube has undergone drastic shifts due to COVID-19 and subsequent lockdowns. Such ubiquitous daily behaviors on online platforms have the potential to capture and correlate with clinically alarming deteriorations in depression and anxiety profiles of users in a noninvasive manner.

Objective: The goal of this study is to examine, among college students in the United States, the relationships of deteriorating depression and anxiety conditions with the changes in user behaviors when engaging with Google Search and YouTube during COVID-19.

Methods: This study recruited a cohort of undergraduate students (N=49) from a US college campus during January 2020 (prior to the pandemic) and measured the anxiety and depression levels of each participant. The anxiety level was assessed via the General Anxiety Disorder-7 (GAD-7). The depression level was assessed via the Patient Health Questionnaire-9 (PHQ-9). This study followed up with the same cohort during May 2020 (during the pandemic), and the anxiety and depression levels were assessed again. The longitudinal Google Search and YouTube history data of all participants were anonymized and collected. From individual-level Google Search and YouTube histories, we developed 5 features that can quantify shifts in online behaviors during the pandemic. We then assessed the correlations of deteriorating depression and anxiety profiles with each of these features. We finally demonstrated the feasibility of using the proposed features to build predictive machine learning models.

Results: Of the 49 participants, 49% (n=24) of them reported an increase in the PHQ-9 depression scores; 53% (n=26) of them reported an increase in the GAD-7 anxiety scores. The results showed that a number of online behavior features were significantly correlated with deteriorations in the PHQ-9 scores (*r* ranging between –0.37 and 0.75, all *P* values less than or equal to .03) and the GAD-7 scores (*r* ranging between –0.47 and 0.74, all *P* values less than or equal to .03). Simple machine learning models were shown to be useful in predicting the change in anxiety and depression scores (mean squared error ranging between 2.37 and 4.22, R^2 ranging between 0.68 and 0.84) with the proposed features.

Conclusions: The results suggested that deteriorating depression and anxiety conditions have strong correlations with behavioral changes in Google Search and YouTube use during the COVID-19 pandemic. Though further studies are required, our results demonstrate the feasibility of using pervasive online data to establish noninvasive surveillance systems for mental health conditions that bypasses many disadvantages of existing screening methods.

(JMIR Ment Health 2020;7(11):e24012) doi:10.2196/24012

KEYWORDS

mental health; anxiety; depression; Google Search; YouTube; pandemic; COVID-19

Introduction

Background

Worldwide mental health problems such as depression, anxiety, and suicidal ideation have severely worsened during the COVID-19 pandemic [1-3], specifically for college students [4-7]. Yet, current methods for screening mental health issues and identifying vulnerable individuals rely on in-person interviews. Such assessments can be expensive, time-consuming, and blocked by social stigma, not to mention the reluctancy induced by travel restrictions and exposure risks. It has been reported that few patients in need were correctly identified and received proper mental health treatments on time under the current health care system [8,9]. Even with emerging telehealth technologies and online surveys, the screening requires patients to actively reach out to care providers.

At the same time, because of the lockdown caused by the global pandemic outbreak, people's engagements with online platforms underwent notable changes, particularly in search engine trends [10-12], exposures to media reports [13,14], and through quotidian smartphone use for COVID-19 information [5]. Reliance on the internet has significantly increased due to the overnight change in lifestyles, for example, remote working and learning, imposed by the pandemic on society. The sort of content consumed, the time and duration spent online, and the purpose of online engagements may be influenced by COVID-19. Furthermore, the digital footprints left by online interactions may reveal information about these changes in user behaviors.

Most importantly, such ubiquitous online footprints may provide useful signals of deteriorating mental health profiles (eg, depression and anxiety) of users during COVID-19. They may capture insights into what was going on in the mind of the user through a noninvasive manner, especially since Google and YouTube searches are short and succinct, and can be quite rich in providing the real-time cognitive state of a person. On one hand, online engagements can cause fluctuations in mental health. On the other hand, having certain mental health conditions can cause certain types of online behaviors. This opens up possibilities for potential health care frameworks that leverage pervasive computing approaches to monitor mental health conditions and deliver interventions on time. However, the findings of this study do not imply any causal relationship between specific types of online activities and one's level of anxiety or depression at a given point in time.

Prior Work

Extensive research has been conducted on a population level, correlating mental health problems with user behaviors on social platforms [15,16], especially among young adolescents. Researchers monitored Twitter to understand mental health profiles of the general population, such as suicidal ideation [17] and depression [18]. Similar research has been done with Reddit, where anxiety [19], suicidal ideation [17], and other general disorders were studied [20,21]. Another popular public platform is Facebook, and experiments have been done studying anxiety, depression, body-shaming, and stress online [22,23]. In addition, it has been shown that college student communities rely heavily on YouTube for both academic and entertainment purposes [24,25]. Yet, abundant use may lead to compulsive YouTube engagements [26], and researchers have found that social anxiety is associated with YouTube consumption in a complex way [27].

During COVID-19, multiple studies have reported deteriorating mental health conditions in various communities [1-3,28] such as nationwide [29,30], across the health care industry [31,32], and among existing mental health patients [33]. Recently, it has been shown that greater use of social media during COVID-19 may induce increasing levels of anxiety and depression at both population and individual levels [14,34]. In addition, online behaviors during COVID-19 have been explored, especially for web searches related to the pandemic [10-12] and abnormal TV consumption during the lockdown [13]. Many of the behavioral studies also discussed the effects of online interactions on the spread, misinformation, knowledge, and protective measures of COVID-19, including the roles of YouTube [35-37] and other platforms [38]. Lyu et al [39] investigated hate speech targeting the Chinese and Asian communities on Twitter during COVID-19. A study in 2009 showed the opposite effect in mental health risk factors: a communitywide crisis may reduce self-harm ideation behaviors [40].

Ubiquitous data has been proved to be useful in detecting mental health conditions. Mobile sensor data such as GPS logs [41,42]; electrodermal activity; and sleep behavior, motion, and phone use patterns [43,44] have been applied in investigating depressive symptoms. Zaman et al [45] found that individual private Google Search histories can be used to detect low self-esteem conditions among college students. Huckins et al [5] examined the longitudinal changes in mental health and smartphone use through ecological momentary assessments during COVID-19 among college populations. Although studies exploring anxiety and depression have been conducted in the past, none of them have leveraged individual-level Google Search and YouTube activity logs to examine the effect of COVID-19 on college students.

Goal of This Study

It has been shown that online platforms preserve useful information about the mental health conditions of users, and COVID-19 is jeopardizing the mental well-being of the global community. Thus, we demonstrate the richness of online engagement logs and how they can be leveraged to uncover alarming mental health conditions during COVID-19. In this study, we aim to examine whether the changes in user behaviors during COVID-19 have a relationship with deteriorating depression and anxiety profiles. We focused on Google Search and YouTube use, and we investigated if the behavior shifts when engaging with these two platforms signify worsened mental health conditions.

The scope of the study covers undergraduate students in the United States. We envision this project as a pilot study; it may lay a foundation for mental health surveillance and help delivery frameworks based on pervasive computing and ubiquitous online data. Compared to traditional interviews and surveys, such a noninvasive system may be cheaper and efficient, and avoid being blocked by social stigma while notifying caregivers on time about individuals at risk.

Methods

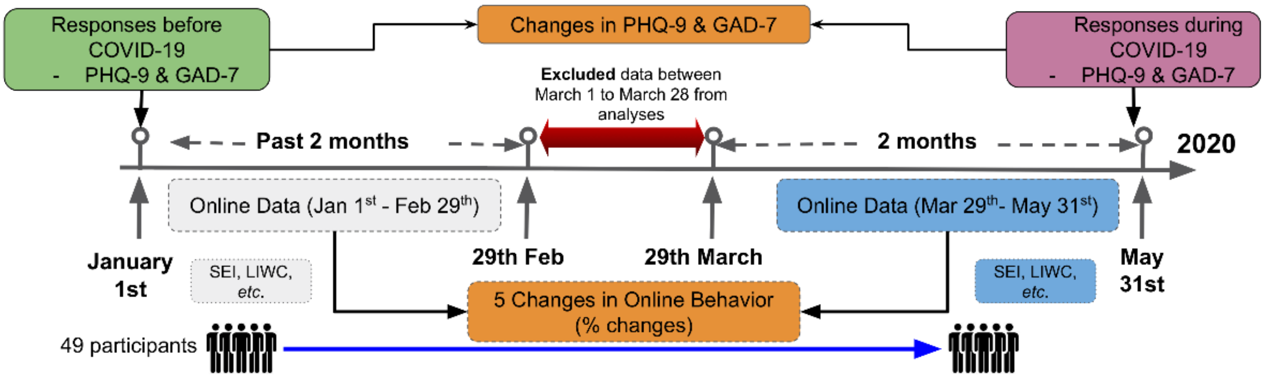
Recruitment and Study Design

We recruited a cohort of undergraduate students, all of whom were at least 18 years of age and have an active Google account for at least 2 years, from the University of Rochester River Campus, Rochester, NY. Participation was voluntary, and individuals had the option to opt out of the study at any time, although we did not encounter any such cases. We collected individual-level longitudinal online data (Google Search and YouTube) in the form of private history logs from the participants. For every participant, we measured the depression and anxiety levels via the clinically validated Patient Health Questionnaire-9 (PHQ-9) and General Anxiety Disorder-7 (GAD-7), respectively. Basic demographic information was also recorded. There were in total two rounds of data collection: the first round during January 2020 (prior to the pandemic) and the second round during May 2020 (during the pandemic). During each round and for each participant, the anxiety and depression scores were assessed, and the change in mental health conditions was calculated in the end. The entire individual online history data up until the date of participation was also collected in both rounds from the participants. Figure 1 gives an illustration of the recruitment timeline and two rounds of data collections. All individuals participated in both rounds and were compensated with US $10 Amazon gift cards during each round of participation.

Given the sensitivity and proprietary nature of private Google Search and YouTube histories, we leveraged the Google Takeout web interface [46] to share the data with the research team. Prior to any data cleaning and analysis, all sensitive information such as the name, email, phone number, social security number, and credit card information was automatically removed via the Data Loss Prevention application programming interface (API) [47] of Google Cloud. For online data and survey response storage, we used a Health Insurance Portability and Accountability Act–compliant cloud-based secure storing pipeline. The whole study design, pipelines, and survey measurements involved were similar to our previous setup [45] and have been approved by the Institutional Review Board of the University of Rochester.

To address participation bias, the study was advertised among the college population via campus wide digital announcements. The text in the study advertisements and consent materials were generic with text such as “help uncover mental health understanding via your online activities.” There was no explicit mention of anxiety or depression in the advertisement. Participation was voluntary with the option to opt out of the study anytime, and their data would not be part of the research study. The intent of the study was clearly explained at the beginning of the recruitment process via one-on-one interviews with the recruiter. We did not have anyone declining to participate or withdrawing in the middle of the study.

Figure 1. The study recruitment procedure and feature development process. All of the participants moved to remote learning on March 7, 2020, the same day a state of emergency was declared in New York State. To avoid any acute behavior during the transition to remote learning, we excluded the data from March 1 to 28. GAD-7: General Anxiety Disorder-7; LWIC: Linguistic Inquiry and Word Count; PHQ-9: Patient Health Questionnaire-9; SEI: short event interval.



Online Data Processing and Feature Extractions

The Google Takeout platform enables users to share the entire private history logs associated with their Google accounts, and as long as the account of the user was logged-in, all histories would be recorded regardless of which device the individual was using. Each activity in Google Search and YouTube engagement logs were time stamped, signifying when the activity happened to the precision of seconds. Furthermore, for each Google Search, the history log contained the query text input by the user. It also recorded the URL if the user directly inputted a website address to the search engine. For each YouTube video watched by the user, the history log contained the URL to the video. If the individual directly searched with one or more keywords on the YouTube platform, the history log also recorded the URL to the search results.

To capture the change in online behaviors for the participants, we first introduced a set of features that quantified certain aspects of how individuals interact with Google Search and YouTube. The set of features was calculated for each participant separately. Individual-level behavior changes were then obtained by examining the variations of the feature between January to March 1, 2020 (a week before the state of emergency in New York State), and March 28 to May 2020 (after the outbreak, following the lockdown and mandated social distancing).

We excluded the online data generated between March 1, 2020, and March 28, 2020, to account for any acute or temporal behavior changes concentrated around the initial lockdown or due to adapting to remote work. We focused on the persistent and stabilized online behaviors throughout the time after the lockdown. Furthermore, the spring break at our institution started on March 7, and the state of emergency in New York State was issued on the same day. All students were asked to leave campus at the start of spring break and complete the rest of the semester remotely.

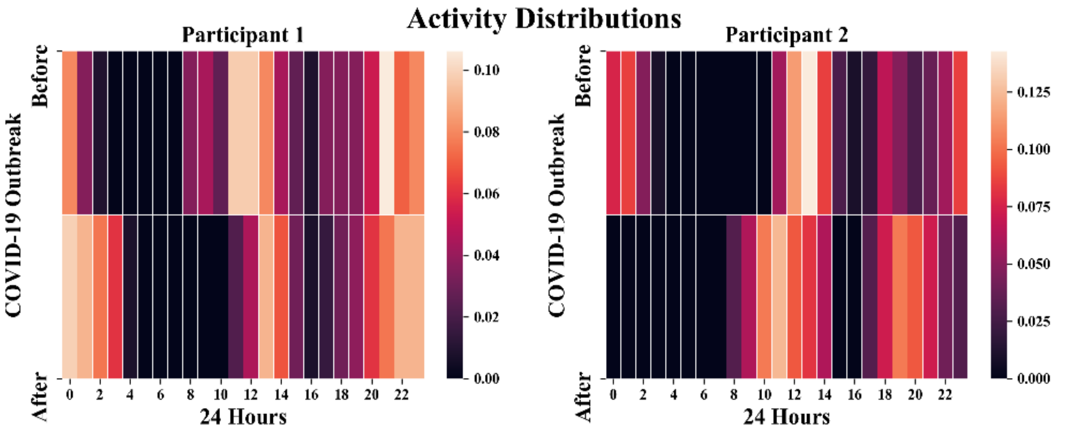
Concretely, we defined 5 features and cut the longitudinal data of each participant into two segments: (1) from January 1 to February 29, 2020, and (2) from March 29 to May 31. Each segment spanned 2 months. We excluded online data from March 1 to March 28 to account for the fact that not all individuals transitioned into work from home on a specific date or practiced a strict social distancing lifestyle, although all of our participants are residents of New York State. The same feature was extracted from both segments of data, and the change was calculated. Such change was referred to as the behavior shifts during the pandemic and lockdown. Figure 1 gives an illustration of data segmentations and feature development pipelines.

Online Activity Distributions

We considered, for each participant, how the Google Search and YouTube activities were distributed across the 24 hours of a day before and after the lockdown, given the previously defined dates. For each trimmed data segment, we cumulated the total number of activities, regardless of Google Search or YouTube, that happened in each of the 24 hours. Thus, we obtained two 24-bin histograms, representing the activity distributions before (D_{before}) and after (D_{after}) the lockdown.

Figure 2 showcases the normalized distributions before and after the outbreak for two participants, each cumulates 2 months of data. For participant one (PHQ-9 increased by 8 and GAD-7 increased by 3) before the outbreak, a few activities started to appear at 8 AM. After the outbreak, these early morning activities disappeared. In addition, a considerable amount of online activities appeared during late-night hours. These patterns most likely indicated a delay in bedtime. For participant two (PHQ-9 decreased by 2 and GAD-7 decreased by 6), there were several activities during late night hours before the lockdown. Followed by a long absence from Google Search and YouTube, the next event usually appeared around noon. After the lockdown, the first activity of the day started to appear in the early morning, and those late-night activities disappeared. Similarly, participant two may also have had afternoon classes at around 3 PM-4 PM. Notice that these two random cases were chosen simply to represent the fact that study participants reacted nonuniformly to the lockdown.

Figure 2. The normalized activity distributions over 24 hours before and after the outbreak of COVID-19 for two example participants.



After that, for each of the 24 hours (h) of a user, we calculated the percentage (ie, relative) change of online activities before and after the lockdown:

For the rest of the study, any mentioned percentage or relative changes of features were calculated in this way.

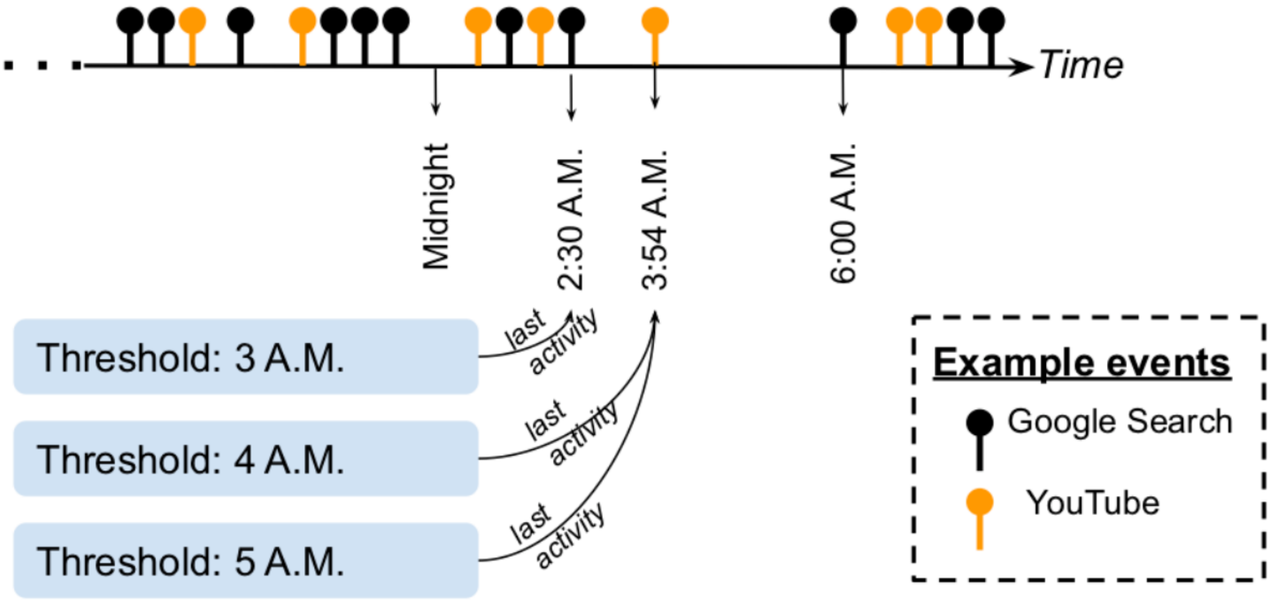
Last Seen Activities

We further considered the last seen activity, regardless of Google Search or YouTube, of each user in a day. It is reasonable to assume that, given the nature of our college student population, the last event before they go to bed does not necessarily happen before midnight. Strictly speaking, our goal was to capture the last event before they went to bed. Therefore, we set a threshold at late night or early morning and considered the last online activity before it. Since a discrete threshold was used, we tried several cutoff hours to perform sensitivity analyses. For our study population, we observed that the hourly volume of Google Search and YouTube activities started to decrease after midnight, and it reached the minimum at 5 AM. This pattern was periodic and persistent across our longitudinal data. Motivated by this observation, we tried a cutoff hour of midnight, 1 AM, 2 AM, 3 AM, 4 AM, and 5 AM, and counted the last events before these thresholds for each participant. Different from the aforementioned *online activity distributions*, which measures the volume of activities on Google Search and

YouTube hourly, the last seen events focused solely on participants staying up late. An example illustration of the *last seen activities* is provided in Figure 3.

With each threshold, we obtained two distributions of the last seen event time stamps before and after the lockdown from each participant. On a continuous scale, we then picked the two medians of the last seen event time stamps before and after the lockdown and took the difference. For example, a difference of 1.5 hours means that the median time of last seen events shifted 1.5 hours later after the lockdown. A difference of –0.3 hours means the median time of last seen events shifted 0.3 hours earlier after the lockdown. All the time differences are in the unit of hours. There is no need to distinguish between Google Search or YouTube for this feature as we are merely looking for the last event, which could be either.

Figure 3. An example to demonstrate how the last seen activities are selected for different threshold hours.



Short Event Intervals

We defined a short event interval (SEI) as the period of time that is less than a certain threshold (eg, 5 minutes) between two adjacent events. It usually occurs when one is consuming several related YouTube videos or is searching for similar content. Taking into consideration that YouTube and Google Search may have different thresholds to define a user session, we adapted the method in Halfaker et al's [48] study to identify proper thresholds for consecutive activities on each of the platforms. After obtaining the session thresholds through mixture models, we counted the total numbers of such SEIs for each participant before (SEI_{before}) and after (SEI_{after}) the outbreak. We calculated the relative change of SEI the same way as in Equation 1 and used it as a behavioral feature.

Linguistic Inquiry and Word Count Attributes

The Linguistic Inquiry and Word Count (LIWC) is a toolkit used to analyze various emotions, cognitive processes, social concerns, and psychological dimensions in a given text by counting the numbers of specific words [49]. It has been widely applied in research involving social media and mental health. For the complete list of linguistic and psychological dimensions LIWC measures, see [49]. We segmented the data log for each participant by the previously mentioned dates as two blobs of texts and analyzed the words using LIWC.

Since the contexts and linguistic properties of Google Search and YouTube may be distinct, we extracted the LIWC features from them separately. For Google Search, we inputted the raw query text; for YouTube, we inputted the video title and the YouTube query text, if any. There were in total 51 different LIWC attributes. LIWC outputted the count of words falling in each dimension among the whole text. We quantified the shift in behavior by calculating the percentage change of words in each dimension after the outbreak.

Google Search and YouTube Categories

We labeled each Google Search query with a category using the Google natural language processing (NLP) API [50]. We used the official YouTube API to retrieve the information of videos watched by the participants, including the title, duration, number of likes and dislikes, and default YouTube category tags. For a comprehensive list of Google NLP category labels and default YouTube category tags, please refer to [51,52]. There were several categories overlapping with the LIWC dimensions, such as *health* and *finance*, and we regarded the LIWC dimensions as a more well-studied standard. Instead, we focused on the number of activities belonging to the *adult* (specifically originating from Google Search logs) and *news* categories, which were not presented in the LIWC.

Concretely, activities such as visiting a porn site (identified via the URL) and searching explicitly for information related to porn and mature content were labelled as *adult*. There was no other ambiguous nonpornographic material being categorized as *adult*. We used Google Cloud Content Classification API for labeling the search queries and used the Webshrinker [53] API to categorize the domain of every URL an individual visited. We calculated the relative changes of activities in these two categories as the behavior shifts for each participant (the same as Equation 1).

We now present a qualitative example of behavior changes in Google Search and YouTube categories. Textbox 1 showcases, for a single example participant, the top five Google Search and YouTube categories before and after the lockdown, defined by the percentages out of the total activity volume. For Google Search, we observed the disappearances of *food and drinks* (including searching for restaurants) and *shopping* from the list. In contrast, the numbers of searches related to the *beauty and fitness*, *home* (including kitchen and cooking subcategories), and *health* topics increased during the quarantine. The *reference* category was largely composed of academic content such as dictionaries, humanity and history references, and scientific proceedings. For YouTube, videos belonging to the *education* category boosted during the remote learning period after the lockdown, as did *film and animation*. The *travel and events*, and *sports* topics vanished from the list. Note that this was merely a single example, and the traits reflected here may be personal, uncorrelated to anxiety or depression, or prevalent among everyone.

Textbox 1. The top five Google Search and YouTube categories for an example participant before and after the lockdown.

Top five Google Search categories

- Before lockdown
  - Art and entertainment
  - Reference
  - Food and drinks
  - Shopping
  - Beauty and fitness
  - Finance
- After lockdown
  - Art and entertainment
  - Reference
  - Beauty and fitness
  - Home
  - Health
  - Food and drinks

Top five YouTube video categories

- Before lockdown
  - Music
  - Travel and events
  - Sports
  - News and politics
  - Education
  - Film and animation
- After lockdown
  - Music
  - Education
  - Film and animation
  - News and politics
  - Pets and animals
  - Comedy

Measurement Outcomes

Measurements for Changes in Online Behaviors

There were in total 5 scalar continuous dependent variables measuring various aspects of the changes in online behavior for each participant, as previously defined. These variables were extracted from two segments of the online data logs, namely, the data before and after the pandemic outbreak. All of the measurements were in percentage changes. For the *online activity distributions*, there were 24 measurements for each hour of a day. For the *last seen events*, there were several thresholds for sensitivity analyses. For the *SEIs*, Google Search and YouTube activities were considered separately with their own fitted session intervals.

Measurements for Mental Health Conditions

For both rounds of the data collection, anxiety levels were assessed using the GAD-7 survey, and depression levels were assessed using the PHQ-9 survey. With two rounds of surveys reported before and after the outbreak, the change in mental health conditions of each participant was obtained. According to Spitzer et al [54] and Rutter and Brown [55], an increase greater than or equal to 5 in the GAD-7 score may be clinically alarming. Similarly, as stated by Kroenke [56], an increase greater than or equal to 5 in the PHQ-9 score may indicate the need for medical interventions.

Demographics

In addition to the online data and mental health surveys, we also collected basic demographic information such as school year, gender, and nationality.

Statistical Analysis

Before any analysis of mental health conditions, to eliminate the possibility of annual confounding factors interfering with the shifts in online behaviors, two-tailed paired independent *t* tests were performed. We inspected, in terms of the five quantitative features, whether the online behavior changes happened every year, such as due to seasonal factors, or only during COVID-19 for the whole study population. As previously mentioned, we collected the entire Google history log back to the registration date of the Google accounts of all participants.

We now use the example of *SEIs* to illustrate the idea. For each participant, we obtained 4 *SEIs* counts from 4 periods of time: January 1, 2019, to February 28, 2019; March 29, 2019, to May 31, 2019; January 1, 2020, to February 29, 2020; and March 29, 2020, to May 31, 2020. These counts are represented as 4 points on a Cartesian coordinate plane, where the y-axis represents the counts and the x-axis represents the time. We then calculated the slope (S) of the line connecting the two points from the same year. With the aforementioned process, we achieved two measurements for each participant, namely, S_{2019} and S_{2020}. Viewing all the participants as a cohort, we computed the S_{2019} and S_{2020} for all features and performed multiple paired *t* tests. This enabled us to estimate the seasonal confounding factors. We could not perform the paired *t* tests on the changes of behavioral features directly because it may only validate a change in the intercept (baseline) amount of activity while ignoring the slope for each feature.

In the main experiments, with each of the aforementioned 5 features and various thresholds, we investigated the correlation of online behavioral changes with deteriorations in the GAD-7 and PHQ-9 scores, which did not require arbitrary discretization decisions. The dependent variables were the 5 behavior changes extracted from the longitudinal individual online data. Experiments were carried out in a one-on-one fashion: anxiety or depression condition was the single independent variable, and one of the five online behavior changes was the single dependent variable each time. Both of them were continuous variables.

Results

Study Population Statistics

We recruited 49 participants in total, and all of them participated in both rounds of the study (100% response rate). On average, each participant made 2357 (95% CI 2106.28 to 2433.45) Google Searches and 2901 (95% CI 2556.92 to 3248.67) YouTube interactions from January to February 29, 2020, and 2497 (95% CI 2069.45 to 2901.34) Google Searches and 3105 (95% CI 2702.48 to 3487.56) YouTube interactions from March 29 to the end of May. Of the 49 participants, 49% (n=24) of them reported an increase in the PHQ-9 score, and 53% (n=26) of them reported an increase in the GAD-7 score. An increase in the PHQ-9 score≥5 was reported by 41% (n=20) of participants, and 45% (n=22) of them reported an increase in the GAD-7 score≥5.

Figure 4 shows the baseline (collected on January 1, 2020) and follow-up postlockdown (collected on May 31, 2020) distributions of depression and anxiety scores in our sample student population. The PHQ-9 scores are shown on the left, ranging from 0 to 27. The GAD-7 scores are shown on the right, ranging from 0 to 21. Each dot represents a participant. The x-axis represents the baseline score in January, and the y-axis represents the follow-up score during the lockdown in May. Figure 5 shows the distributions of the change in PHQ-9 depression and GAD-7 anxiety scores before and after the lockdown. Again, the PHQ-9 scores are shown on the left, and the GAD-7 scores are shown on the right. The changes were calculated as the follow-up scores in May subtracted by the baseline scores in January. Putting the pandemic into context, the deterioration in anxiety or depression levels may have been triggered by the fear of getting infected, loss of jobs, the death of family members or friends, and many other negative impacts from COVID-19. Particularly for college students, other major reasons may be the pressure of online learning, loss of financial aids, and living alone. In contrast, students that underwent quarantines with their families safely may not have shown signals of deteriorating anxiety or depression, compared to the high stress levels during normal school days.

Figure 4. The distributions of PHQ-9 depression and GAD-7 anxiety scores before and after the lockdown. The PHQ-9 scores are shown on the left, and the GAD-7 scores are shown on the right. GAD-7: General Anxiety Disorder-7; PHQ-9: Patient Health Questionnaire-9.

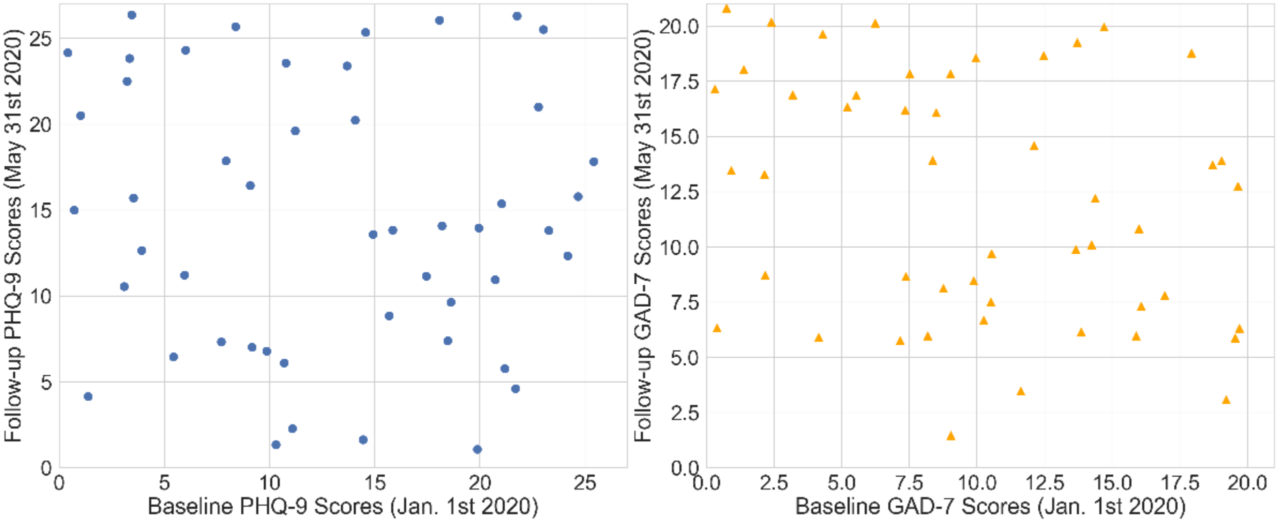
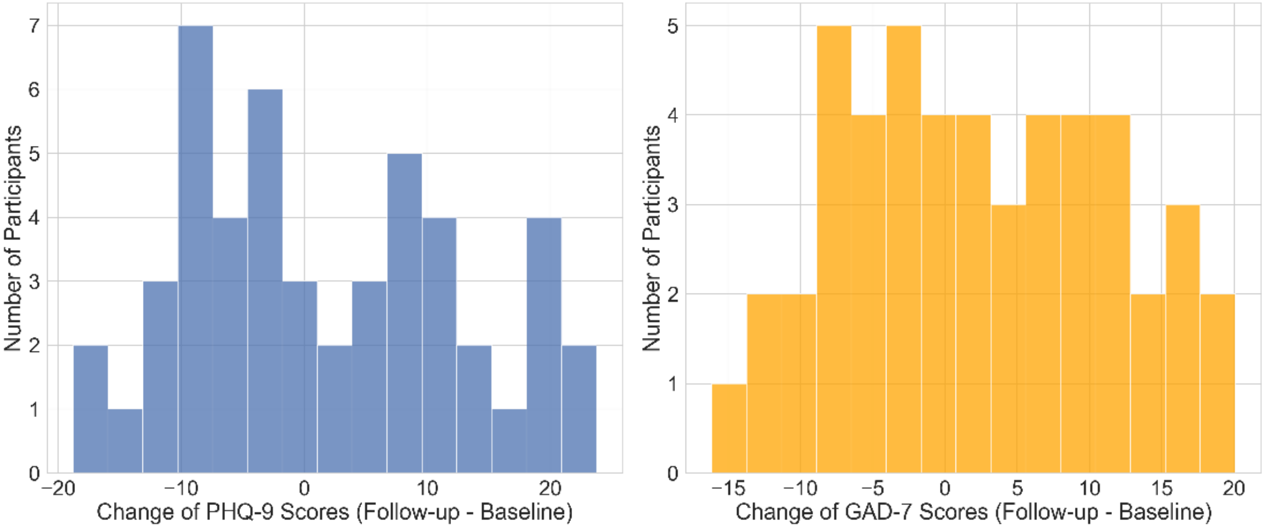


Figure 5. The distributions of the changes in PHQ-9 depression and GAD-7 anxiety scores before and after the lockdown. The PHQ-9 scores are shown on the left, and the GAD-7 scores are shown on the right. GAD-7: General Anxiety Disorder-7; PHQ-9: Patient Health Questionnaire-9.



Of the 49 participants, 61% (n=30) of the them were female, 35% (n=17) were male, and 4% (n=2) reported nonbinary genders. First-, second-, third-, and forth-year students occupied 22% (n=11), 41% (n=20), 31% (n=15), and 6% (n=3) of the whole cohort, respectively. A total of 80% (n=39) of the participants were US citizens, and the rest (n=10) were international students. A complete breakdown of demographics with respect to the deteriorating anxiety and depressive disorders are given in Table 1.

Table 1. Demographics of the study population.

Demographic	Increased PHQ-9[a] (n=24), n (%)	Increased GAD-7[b] (n=26), n (%)
Female	18 (75)	22 (85)
US citizen	18 (75)	20 (77)
First-year students	5 (21)	4 (15)
Second-year students	11 (46)	12 (46)
Third-year students	7 (29)	8 (31)
Fourth-year students	1 (4)	2 (8)

[a] PHQ-9: Patient Health Questionnaire-9.

[b] GAD-7: General Anxiety Disorder-7.

Evaluation Outcomes

The two-tailed paired independent *t* tests mentioned at the beginning of the Statistical Analysis section were designed to rule out seasonal factors in online behavior changes and focus on COVID-19 before any of the main experiments. All features had *P* values less than .003. Hence, the presence of annual or seasonal factors accountable for online behavior changes was neglectable, and it was safe to carry out the following main experiment. This is consistent with one of the main conclusions of Huckins et al [5] that, when comparing the longitudinal data between different years, behaviors during COVID-19 shifted drastically.

We calculated the Pearson product-moment correlations, *r*, of online behavior shifts with deteriorations in anxiety and depression levels. We reported the correlation coefficients with *P* values and 95% CIs obtained for each of the aforementioned features.

Online Activity Distributions

For *online activity distributions*, we calculated the percentage changes in the volume of activities for all 24-hour bins separately. We obtained 24 measurements for each participant. We then evaluated for each hour of a day the correlation between the relative hourly activity change and the change in the PHQ-9 depression scores. Significant correlations were found for all hours between 10 PM and 6 AM of the next day, inclusively. All of them were positive correlations (*r* ranging between 0.32 and 0.75, all *P* values were less than or equal to .04). The correlations started increasing after 10 PM, reached the maximum at 3 AM, and started to decrease afterwards. This suggests that greater late-night online activity volumes after the lockdown may be a signal of deteriorating (increasing) depressive levels.

Next, we observed similar results from the analysis between the hourly activity change and the change in the GAD-7 anxiety scores. Significant correlations were found for all hours between 11 PM and 6 AM of the next day, inclusively. All of them were positive correlations (*r* ranging between 0.39 and 0.74, all *P* values were less than or equal to .006). The correlations started increasing after 11 PM, reached the maximum at 3 AM, and started to decrease afterwards. This implies that greater late-night online activity volumes after the lockdown may represent deteriorating (increasing) anxiety levels. For the detailed hourly correlations, 95% CIs, and comparisons between the deteriorating PHQ-9 and GAD-7 groups, see Table 2.

Table 2. The correlation coefficients between the change in the volume of online activities of each hour in a day and deteriorating PHQ-9 and GAD-7 scores.

Hour	Deteriorating PHQ-9[a]		Deteriorating GAD-7[b]	
	Correlation coefficient, *r* (95% CI)	*P* value	Correlation coefficient, *r* (95% CI)	*P* value
12 AM	*0.54*[c] (0.38 to 0.72)	*<.001*	*0.45* (0.20 to 0.65)	*.001*
1 AM	*0.64* (0.45 to 0.81)	*<.001*	*0.52* (0.28 to 0.70)	*<.001*
2 AM	*0.66* (0.47 to 0.85)	*<.001*	*0.60* (0.39 to 0.75)	*<.001*
3 AM	*0.75* (0.59 to 0.94)	*<.001*	*0.74* (0.58 to 0.92)	*<.001*
4 AM	*0.63* (0.43 to 0.80)	*<.001*	*0.63* (0.43 to 0.77)	*<.001*
5 AM	*0.45* (0.29 to 0.65)	*.001*	*0.58* (0.35 to 0.74)	*<.001*
6 AM	*0.34* (0.21 to 0.54)	*.02*	*0.39* (0.12 to 0.60)	*.006*
7 AM	0.27 (0.01 to 0.54)	.06	0.24 (–0.05 to 0.48)	.10
8 AM	–0.26 (–0.50 to 0.02)	.07	–0.28 (–0.52 to 0.00)	.05
9 AM	–0.28 (–0.52 to 0.00)	.05	–0.26 (–0.50 to 0.02)	.07
10 AM	–0.27 (–0.51 to 0.01)	.06	–0.21 (–0.46 to 0.07)	.05
11 AM	–0.22 (–0.47 to 0.07)	.13	–0.25 (–0.49 to 0.03)	.08
12 PM	–0.19 (–0.44 to 0.09)	.19	–0.20 (–0.45 to 0.08)	.16
1 PM	0.23 (–0.05 to 0.48)	.10	–0.10 (–0.36 to 0.48)	.51
2 PM	0.17 (–0.12 to 0.43)	.24	0.17 (–0.12 to 0.19)	.24
3 PM	0.24 (–0.04 to 0.49)	.09	–0.22 (–0.47 to 0.06)	.12
4 PM	0.18 (–0.11 to 0.43)	.23	0.16 (–0.12 to 0.42)	.26
5 PM	0.15 (–0.14 to 0.41)	.31	0.15 (–0.14 to 0.41)	.31
6 PM	–0.16 (–0.42 to 0.13)	.27	–0.15 (–0.41 to 0.13)	.30
7 PM	–0.12 (–0.39 to 0.16)	.39	–0.24 (–0.49 to 0.04)	.10
8 PM	–0.23 (–0.47 to 0.06)	.12	–0.23 (–0.47 to 0.06)	.11
9 PM	0.25 (–0.03 to 0.50)	.08	0.24 (–0.04 to 0.49)	.09
10 PM	*0.32* (0.02 to 0.55)	*.04*	0.27 (–0.01 to 0.51)	.06
11 PM	*0.41* (0.23 to 0.62)	*.003*	*0.39* (0.13 to 0.61)	*.005*

[a] PHQ-9: Patient Health Questionnaire-9.

[b] GAD-7: General Anxiety Disorder-7.

[c] Italics indicate significant results.

Last Seen Activities

For last seen activities, we measured the shift in the median time of last seen activities after the lockdown. Last seen events were determined by the last activity performed before the threshold hour. First, we calculated correlation coefficients between the time shifts and changes in the PHQ-9 depression scores with different cutoff threshold hours. There was a positive correlation overall between the shift of the median time and deteriorating PHQ-9 scores. Thus, the more positive the shift (ie, the median time of the last events moved to later hours), the greater the deterioration. Specifically, for cutoff hours at 2 AM, 3 AM, 4 AM, and 5 AM, the correlations were significant (*r* ranging between 0.35 and 0.59, all *P* values were less than or equal to .01). The correlation was strongest for last seen events before 5 AM. These cutoff values were motivated by the periodic hourly online activity volumes we observed from the study population, as described in the Last Seen Activities section.

Next, we observed similar results when exploring the deterioration in GAD-7 anxiety scores. The positive correlation shows that staying up late tends to signal deteriorations in anxiety levels. For cutoff hours at 2 AM, 3 AM, 4 AM, and 5 AM, the correlations were significant (*r* ranging between 0.30 and 0.57, all *P* values were less than or equal to .03). The last seen events before 4 AM showed the most significant correlation. For the detailed hourly correlations, 95% CIs, and comparisons between the deteriorating PHQ-9 and GAD-7 groups, see Table 3.

Table 3. The correlation coefficients between the change in the median time of last seen events and deteriorating PHQ-9 and GAD-7 scores at different threshold hours.

Threshold	Deteriorating PHQ-9[a]		Deteriorating GAD-7[b]	
	Correlation coefficient, *r* (95% CI)	*P* value	Correlation coefficient, *r* (95% CI)	*P* value
12 AM	0.21 (–0.07 to 0.46)	.15	0.25 (–0.03 to 0.50)	.08
1 AM	0.26 (–0.02 to 0.50)	.07	0.25 (–0.03 to 0.49)	.09
2 AM	*0.35*[c] (0.08 to 0.58)	*.01*	*0.30* (0.03 to 0.54)	*.03*
3 AM	*0.42* (0.15 to 0.65)	*.002*	*0.45* (0.18 to 0.67)	*.001*
4 AM	*0.58* (0.31 to 0.79)	*<.001*	*0.57* (0.30 to 0.78)	*<.001*
5 AM	*0.59* (0.33 to 0.80)	*<.001*	*0.56* (0.29 to 0.78)	*<.001*

[a] PHQ-9: Patient Health Questionnaire-9.

[b] GAD-7: General Anxiety Disorder-7.

[c] Italics indicate significant results.

SEIs

By considering Google Search and YouTube activities separately, we found different short interevent session thresholds using the method established by Halfaker et al [48]. For Google Search, the boundary between the in-session and between-session mixtures was 1 hour, which is consistent with the finding in the original paper. Thus, we set the threshold of 1 hour and considered all consecutive searches on Google within 1 hour as in the same session. By counting the numbers of adjacent searches within 1 hour before and after the lockdown, and calculating the percentage change, we did not find significant correlations between the number of short Google Search intervals and deteriorating PHQ-9 depression scores (*r*=–0.22, 95% CI –0.43 to 0.01, *P*=.06). Nor did we find significant correlations between the number of short Google Search intervals and deteriorating GAD-7 anxiety scores (*r*=–0.21, 95% CI –0.40 to 0.02, *P*=.08).

In contrast, we found that the threshold for the in-session and between-session mixtures for YouTube watching histories was 3.2 minutes, much shorter than that of Google Search. The average YouTube video interval time was 21 minutes. This threshold indicated that two adjacent videos consumed within 3.2 minutes of idle time should be considered as in the same session (ie, consecutive consumption). We then calculated the relative change of the number of such short YouTube intervals after the lockdown. We found a significant positive correlation between the increase of YouTube short intervals and deteriorating PHQ-9 scores (*r*=0.57, 95% CI 0.38 to 0.76, *P*<.001). Similarly, a significant positive correlation was found between the increase of YouTube short intervals and deteriorating GAD-7 scores (*r*=0.41, 95% CI 0.20 to 0.62, *P*=.001).

In addition, we include visualizations (Figures 6 and 7) of the behavioral measurements for *SEIs* over time. As mentioned in the Measurements for Mental Health Conditions section, an increase ≥5 in PHQ-9 or GAD-7 scores is clinically alarming. Thus, we first separated the samples by groups with and without an increase ≥5 in the PHQ-9 depression scores. We then plotted the 7-day moving average total number of short YouTube event intervals (ie, consecutive video consumption) of the two groups as a function of dates. We overlayed the series with the activity data from the same time period (January 1 to May 31) in 2019 in dashed lines for contrast. As shown in Figure 6, after the lockdown in mid-March, both groups were having increasing amounts of short YouTube intervals. The participants with severe deteriorating depression outran others significantly. A similar trend was found when we separated the groups by an increase ≥5 in the GAD-7 anxiety scores, shown in Figure 7. The group with significantly deteriorated anxiety disorders tended to have higher numbers of consecutive YouTube consumption. The shaded area represents 1 SD. Most importantly, these patterns were stabilized throughout the time after the lockdown, reflecting a meaningful behavioral shift instead of mere acute or temporal observations. No such phenomenon was observed in 2019 for any group, and we further argue the fact that seasonal factors are accountable for the behavioral difference.

Figure 6. The 7-day moving average time series of the total amount of short YouTube activity intervals between groups with and without significant increases in the PHQ-9 depression scores. PHQ-9: Patient Health Questionnaire-9.

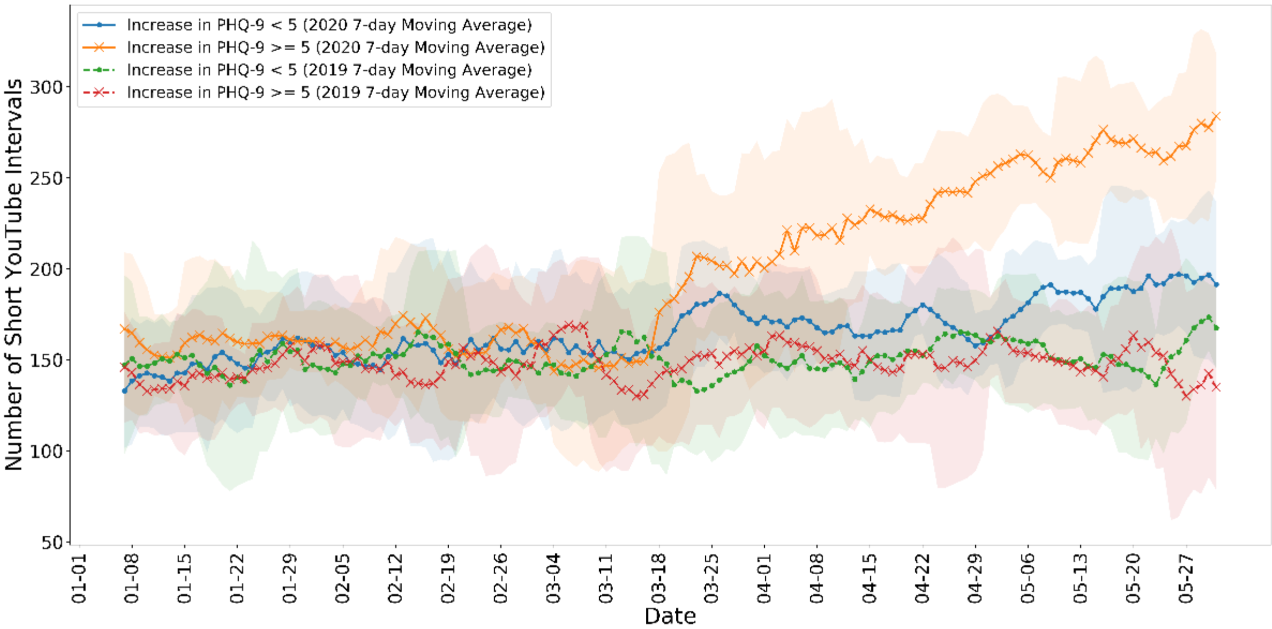
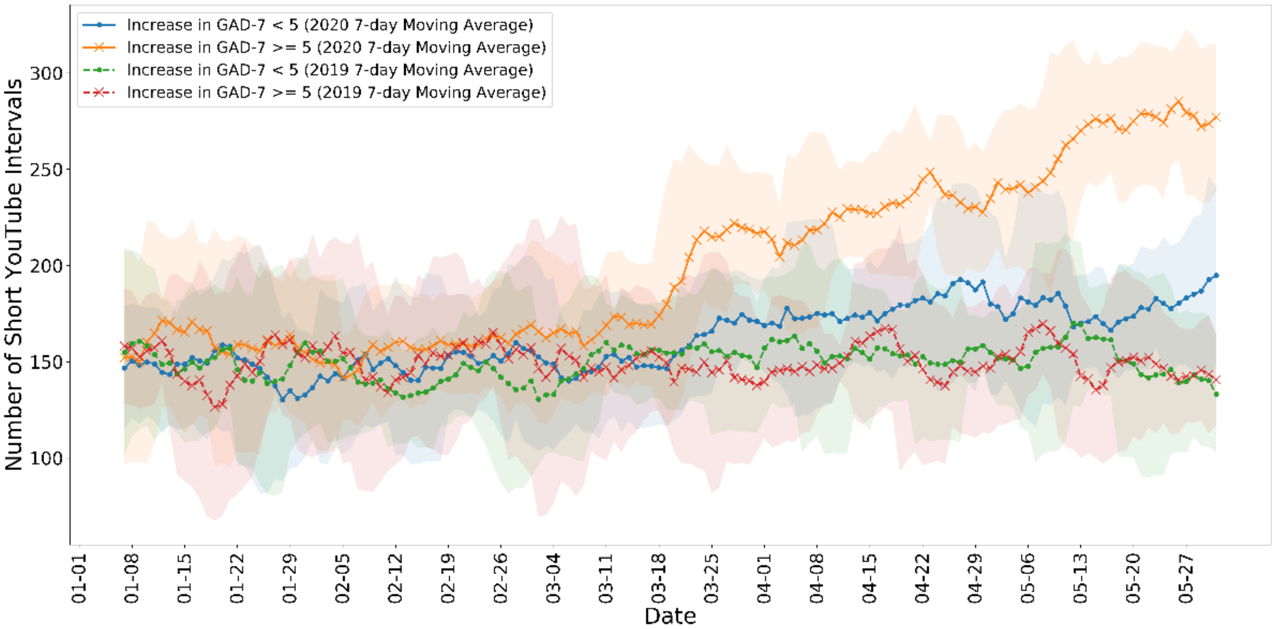


Figure 7. The 7-day moving average time series of the total amount of short YouTube activity intervals between groups with and without significant increases in the GAD-7 anxiety scores. GAD-7: General Anxiety Disorder-7.



LIWC Attributes

As the contexts and linguistic properties of Google Search and YouTube may be different, we extracted the LIWC features from them separately. For Google Search, we found that search queries in the *work*, *money*, and *death* categories under the *personal concerns* dimension showed significant positive correlation with deteriorating PHQ-9 depression scores (*r* ranging between 0.33 and 0.59, all *P* values less than or equal to .02). Similar results were found for deteriorating GAD-7 anxiety scores (*r* ranging between 0.41 and 0.51, all *P* values were less than or equal to .003).

Moreover, for both depression (*r*=0.31, 95% CI 0.08 to 0.49, *P*=.03) and anxiety (*r*=0.34, 95% CI 0.10 to 0.52, *P*=.02) deteriorations, significant correlations were found in the *authentic* scores of the search queries under the *summary variable*. The *authentic* scores were developed by Newman et al [57] to provide a continuous scale for measuring how honest and genuine a given piece of text is. The higher the score, the more authentic the language. It was shown that when individuals address themselves in an authentic manner, they tend to be personal and vulnerable. For instance, a randomly chosen participant searched: “What not to do during at-home exams under proctor surveillance?” For the detailed correlations coefficients and 95% CIs, see Table 4.

Table 4. The significant LIWC features from Google Search histories and their correlation coefficients with deteriorating PHQ-9 and GAD-7 scores.

LIWC[a] attributes	Deteriorating PHQ-9[b]		Deteriorating GAD-7[c]	
	r (95% CI)	*P* value	*r* (95% CI)	*P* value
Personal concerns				
Work	*0.33*[d] (0.09-0.53)	*.02*	*0.43* (0.22-0.60)	*.002*
Money	*0.47* (0.28-0.60)	*.001*	*0.51* (0.33-0.63)	*<.001*
Death	*0.59* (0.42-0.71)	*<.001*	*0.41* (0.21-0.56)	*.003*
Summary variable				
Authentic	*0.31* (0.08-0.49)	*.03*	*0.34* (0.10-0.52)	*.02*

[a] LIWC: Linguistic Inquiry and Word Count.

[b] PHQ-9: Patient Health Questionnaire-9.

[c] GAD-7: General Anxiety Disorder-7.

[d] Italics indicate significant results.

For YouTube histories, we found that videos containing *anxiety* and *sadness* keywords under the *negative emotion* dimension showed significant positive correlation with deteriorating PHQ-9 depression scores (*r* ranging between 0.50 and 0.52, all *P* values less than or equal to .001). Similar results were found for deteriorating GAD-7 anxiety scores (*r* ranging between 0.55 and 0.57, all *P* values less than or equal to .001). In addition, the *friends* keywords from the *social words* dimension showed a significant positive correlation with worsened GAD-7 (*r*=0.41, 95% CI 0.21 to 0.56, *P*=.003) but not PHQ-9 (*r*=0.27, 95% CI 0.02 to 0.47, *P*=.06).

Moreover, for both depression (*r*=–0.37, 95% CI –0.16 to –0.52, *P*=.009) and anxiety (*r*=–0.47, 95% CI –0.28 to –0.60, *P*=.001) deteriorations, significant negative correlations were found in the *emotional tone* scores of the videos, under the *summary variable*. The *emotional tone* scores were developed by Cohn et al [58] to provide a continuous scale for measuring the positivity of a given piece of text and vice versa. The higher the score, the more positive the text. For the detailed correlation coefficients and 95% CIs, see Table 5.

Table 5. The significant LIWC features from YouTube videos and their correlation coefficients with deteriorating PHQ-9 and GAD-7 scores.

LIWC[a] attributes	Deteriorating PHQ-9[b]		Deteriorating GAD-7[c]	
	r (95% CI)	*P* value	*r* (95 CI%)	*P* value
Negative emotion				
Anxiety	*0.52*[d] (0.35 to 0.64)	*<.001*	*0.57* (0.42 to 0.66)	*<.001*
Sadness	*0.50* (0.31 to 0.63)	*<.001*	*0.55* (0.38 to 0.67)	*<.001*
Social words				
Friends	*0.42* (0.22 to 0.57)	*.002*	0.27 (0.02 to 0.47)	.06
Summary variable				
Emotional tone	*–0.37* (–0.16 to –0.52)	*.009*	*–0.47* (–0.28 to –0.60)	*.001*

[a] LIWC: Linguistic Inquiry and Word Count.

[b] PHQ-9: Patient Health Questionnaire-9.

[c] GAD-7: General Anxiety Disorder-7.

[d] Italics indicate significant results.

Google Search and YouTube Categories

The *adult* category consists of explicit browser histories of porn-related content and visiting porn sites (identified via the URL). On one hand, the percentage change of the *adult* content showed a significant positive correlation with deteriorating depression levels (*r*=0.56, 95% CI 0.35 to 0.77, *P*<.001). On the other hand, the percentage change of the *adult* content did not show a significant correlation with deteriorating anxiety levels (*r*=0.29, 95% CI 0.02 to 0.56, *P*=.08). The *news* content did not show any significant correlation with deteriorating depression (*r*=0.25, 95% CI 0.04 to 0.46, *P*=.13) nor anxiety (*r*=0.14, 95% CI –0.01 to 0.29, *P*=.21).

Predictive Modeling

In this section, we probe the feasibility of using common machine learning models to predict the change (eg, deterioration) in PHQ-9 and GAD-7 scores. We framed the task as a supervised regression problem and treated the aforementioned behavioral changes as feature inputs to the model. The goal was to predict the change in the PHQ-9 or GAD-7 scores. Given our small sample size, we evaluated the model performance by multiple (N=49) leave-one-out train (n=48) and test (n=1) splits. For each of the splits, we tuned model hyperparameters with another complete leave-one-out cross-validation on the 48 training samples.

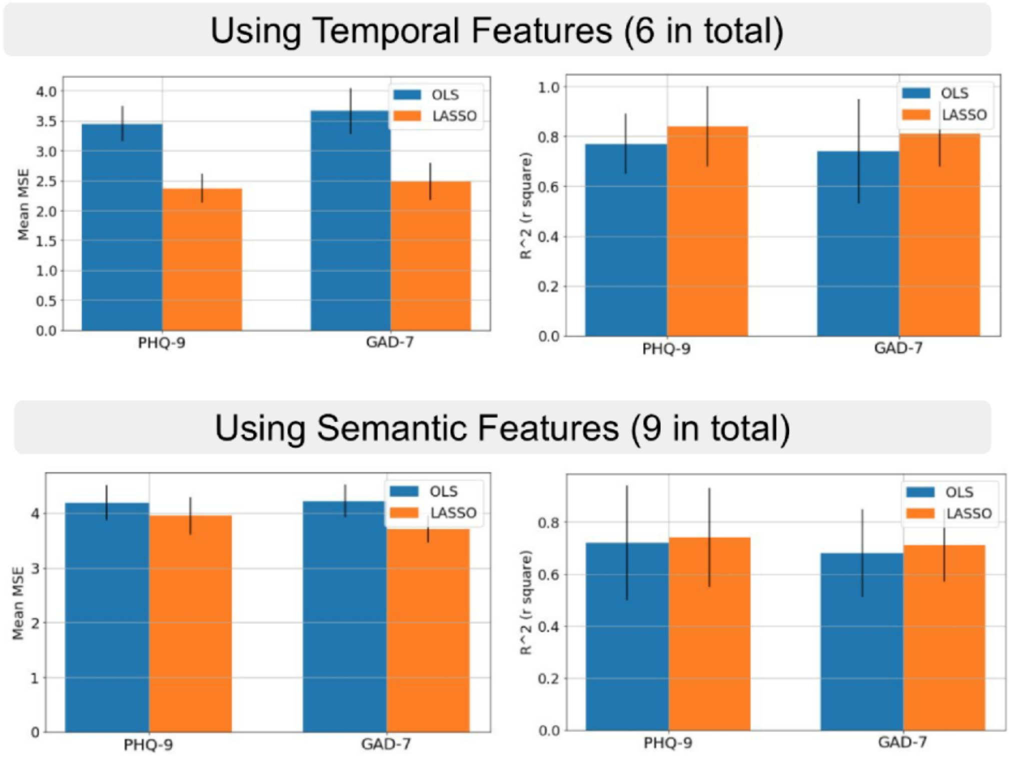
Feature Vectors

We developed the feature vector by concatenating some of the significant behavioral shifts (scalars) previously mentioned. Nonetheless, given the small sample size, it was reasonable to avoid high-dimensional feature vectors. Thus, we first separated the aforementioned 5 behavioral changes into 2 groups, namely, the *temporal* and *semantic* features. Specifically, we picked the *online activity distributions* at 2 AM, 3 AM, and 4 AM; the *last seen activities* before 4 AM and 5 AM; and the *short YouTube intervals* for the *temporal* feature vector, and hence, it had a dimensionality of 6. We then concatenated the 4 significant LIWC categories for Google Search, the 4 significant LIWC categories for YouTube, and the *adult* Google Search content as the *semantic* feature vector. The dimensionality was 9.

Regression Models

We experimented with two of the most common linear models: ordinary least square regression and lasso regression. In Figure 8, we present the performances of the regression models. We reported the mean squared error (MSE) and the average coefficient of determination (R^2). The mean and SDs were calculated from the 49 leave-one-out splits. Overall, *temporal* features performed better than *semantic* ones, regardless of predicting the change in PHQ-9 or GAD-7. The best average performance in predicting the change of PHQ-9 was achieved by the *temporal* features (MSE=2.37, R^2=0.84). The best average performance in predicting the change of GAD-7 was also achieved by the *temporal* features (MSE=2.48, R^2=0.81).

Figure 8. The performances of the two regression models averaged across 49 leave-one-out splits. GAD-7: General Anxiety Disorder-7; LASSO: lasso regression; MSE: mean squared error; OLS: ordinary least square regression; PHQ-9: Patient Health Questionnaire-9.



Model Examination

We examined the coefficient weights from the linear models for PHQ-9 and GAD-7 predictions. Since lasso regression performed the best in all cases, we looked at the weights from the fitted lasso models. For the 6 temporal features, as shown in Table 6, *online activity distributions* at 4 AM had the most significant importance in both prediction tasks. For the 9 semantic features, as shown in Table 7, keywords related to the *anxiety* dimension in the LIWC from YouTube histories had the most significant importance. Moreover, as the aforementioned statistical tests have shown that neither the *friends* dimension in LIWC nor the *adult* Google Search category were significant for deteriorating GAD-7, the lasso regression indeed assigned zero weights to these two features. We used an alpha value of .0001 for the regularization term in the lasso model.

Table 6. The coefficients of 6 temporal features in lasso regression.

Temporal features	Coefficient in PHQ-9[a] prediction	Coefficient in GAD-7[b] prediction
Online activity distributions		
2 AM	0.08	0.11
3 AM	0.34	0.28
4 AM	0.68	0.71
Last seen activities		
4 AM	0.21	0.03
5 AM	0.64	0.54
Short YouTube intervals	0.00	0.02

[a] PHQ-9: Patient Health Questionnaire-9.

[b] GAD-7: General Anxiety Disorder-7.

Table 7. The coefficients of 9 semantic features in lasso regression.

Semantic features	Coefficient in PHQ-9[a] prediction	Coefficient in GAD-7[b] prediction
LIWC[c]		
Anxiety	0.63	0.52
Sadness	0.48	0.47
Friends	0.02	0.00
Emotional tone	–0.23	–0.19
Work	0.04	0.13
Money	0.18	0.26
Death	0.01	0.07
Authentic	0.02	0.15
Google Search categories		
Adult	0.05	0.00

[a] PHQ-9: Patient Health Questionnaire-9.

[b] GAD-7: General Anxiety Disorder-7.

[c] LIWC: Linguistic Inquiry and Word Count.

Discussion

Principal Results

In this study, we collected longitudinal individual-level Google Search and YouTube data from college students, and we measured their anxiety (GAD-7) and depression (PHQ-9) levels before and after the outbreak of COVID-19. We then developed explainable features from Google Search and YouTube logs and quantified various online behavior shifts of the participants during the pandemic. We also calculated the change in mental health conditions for all participants. Our experiment examined the correlations of online behavior features with deteriorating anxiety and depression levels. We finally demonstrated the feasibility of building simple predictive machine learning models with the proposed behavioral signals. To the best of our knowledge, we are the first to conduct observational studies on how anxiety and depression problems and Google Search and YouTube use of college students are related during COVID-19.

Our results showed that online behavior changes have significant correlations with worsened depression and anxiety profiles during the pandemic. The features we developed based on online activities were all explainable and preserved certain levels of interpretability. For example, the *SEIs* and *online activity distributions* measured the consecutive use and hourly activity volumes of Google Search and YouTube, which were inspired by previous studies on excessive YouTube use [26], internet addictions [59], and positive associations with social anxiety among college students [27]. Our results indicated that individuals with increasing anxiety or depressive disorders during the pandemic tended to have long use sessions (multiple consecutive activities with short time intervals) when engaging with Google Search and YouTube.

Moreover, the increasing activities during late night hours in *online activity distributions* and the positive shifts of medians of *last seen events* corresponded with previous studies in sleep deprivation and subsequent positive correlations with mental health deteriorations [60,61]. Our results demonstrated that individuals with worsened anxiety or depressive symptoms during the pandemic were indeed likely to stay up late and engage more online. The aforementioned three features captured the temporal aspects of user online behaviors, and they generated the best performance in the regression tasks.

Additionally, our analysis found that the amount of porn consumption had significant correlations with deteriorating depression, which adheres to previous findings that people with depression and loneliness are likely to consume excessive pornography [62,63]. For the LIWC features, participants with significant increases in anxiety watched more videos with *anxiety*, *sadness*, and negative tone, and previous research has shown that negative YouTube videos tended to receive more attention from vulnerable individuals [64]. They also consumed more videos related to social activities and *friends* keywords. The *friends* keywords did not show any significance in the depression analysis. This is consistent with studies on patterns of social withdrawal and depression [42,65,66], and social interactions and isolations have been recognized by Leigh-Hunt et al [67] as one of the priorities in mental illness prevention, especially during COVID-19 [30]. For Google Search, both the participants with significant increases in anxiety and depression searched more content related to *work*, *money*, and *death*, focusing on real-life practices. None of the emotional dimensions were significant in Google Search logs. Instead, LIWC considered the search queries from depressed and anxious individuals more honest and vulnerable (eg, asking for help) after the lockdown, given the *authentic* score. Although prior research has shown that individuals living with depression tend to use more first-person languages [68], we did not observe any similar pattern. This is probably due to the fact that search queries are more succinct, imperative, and functional, which leaves less necessity for personal references. These attributes captured the semantic aspect of user online behaviors. The prevalence of personal affair, social activity, and negative keywords as well as porn consumption has shown statistically significant correlations.

Many researchers have reported that there has been a significant boost in health- and news-related topics at the population level in various online platforms during COVID-19. This is partly due to additional measures taken by individuals, various stakeholders, and agencies with regard to preventive measures [11,36,37], daily statistics [10,12,13], and health care information and misinformation [35,37,38]. However, unlike many, our investigation was carried out considering individual-level Google Search and YouTube engagement logs, and our analysis did not reveal any significant spikes in the *news* and *health and illness* categories among individuals with deteriorating anxiety and depression during the pandemic. One possible explanation for such observation could be due to the target population (college students) of our study, who may prefer to follow news from other popular platforms such as social media.

Finally, COVID-19 has forced us to alter daily lifestyles. The world was not ready for such a viral outbreak. Since there is no cure for COVID-19, it, or an even more deadly viral disease, may resurface at different capacities in the near future. Society may be forced to rely on technologies even more and employ remote learning, working, and socializing for a longer period of time. It is important that we learn from our experience of living through the initial COVID-19 outbreak and take necessary measures to uncover the changes in online behaviors, investigating how that can be leveraged to understand and monitor various mental health conditions of individuals in the least invasive manner. Furthermore, we hope our work paves the path for technology stakeholders to consider incorporating various mental health assessment monitoring systems using user engagements, following users' consent in a privacy-preserving manner. They can periodically share the mental health monitoring assessment report with users based on their online activities and education, and inform users about their current mental health. This can eventually encourage individuals to acknowledge the importance of mental health and take better care of themselves.

Limitations

First, although most of the online behavioral features we developed showed significant correlations, our study cohort only represented a small portion of the whole population with mental health difficulties. Therefore, further studies are required to investigate if the significant behavioral changes still hold among more general communities not limited to college students. It is possible that the relationship of worsened anxiety and depression with online activities on Google Search and YouTube of our college population is different from that of other populations whose education and social backgrounds may vary. There may also be differences in mental health for substudent populations, such as those living in harmful environments and those depending on financial aid, all of whom may experience more physical and economic crises during COVID-19. Nonetheless, we argue that the explainable features we constructed, such as late-night activities, continuous use, inactivity, pornography, and certain keywords, can remain behaviorally representative and be applied universally across experiments exploring the relationship of anxiety and depression with online activities during the pandemic.

Second, in this study, we explored the relationship between user online behaviors and the fluctuations in anxiety and depression conditions during COVID-19. Any causal relationship between online behavior and mental disorders is beyond the scope of this study. As one can readily imagine, online behavioral changes could both contribute to or be caused by deteriorating anxiety or depressive disorders.

Third, we acknowledge that it may be impossible to obtain data without noise because one may seldom, or even never, search on Google or watch YouTube videos. Such concealed information makes it impossible for the proposed model to flag alarming symptoms that are reflected in the PHQ-9 and GAD-7 questionnaires.

Fourth, though we included preliminary demographic information as covariates, there remains the possibility of other confounding factors. In fact, both the shifts in online behaviors and deteriorating mental health profiles may be due to common factors such as living conditions, financial difficulties, and other health problems during the pandemic. There was also no causal direction implied between COVID-19 and online behavior changes, which was introduced in the first paragraph of the Statistical Analysis section as a precaution before the main experiments.

Ethical and Privacy Concerns

Despite being a pilot study, our results indicate that it is possible to build an anxiety and depression surveillance system based on passively collected private Google data histories during COVID-19. Such noninvasive systems shall be subject to rigorous data security and anonymity checks. Necessary measures need to be in place to ensure personal safety and privacy concerns when collecting sensitive and proprietary data such as Google Search logs and YouTube histories. Even in pilot studies, participants should have full rights over their data; they may choose to opt out of the study at any stage and remove any data shared in the system.

Moreover, anonymity and systematic bias elimination should be enforced. As an automatic medical screening system based on pervasive data, it has been shown that such frameworks are prone to implicit machine learning bias during data collection or training phases [69-71]. Black-box methods should be avoided, as they are known to be vulnerable to adversarial attacks and produce unexplainable distributional representations [72,73]. Anonymizing data and obscuring identity information should be the first step in data debiasing.

In the end, to what extent should caregivers trust a clinical decision made by machines remains an open question. We believe that possible pervasive computing frameworks shall play the role of a smart assistant, at most, to the care providers. Any final intervention or help delivery decision should be made by health care professionals who understand both the mental health problems and the limitations of automatic detection systems in clinical settings.

Acknowledgments

This research was supported in part by grant W911NF-15-1-0542 and W911NF-19-1-0029 with the US Defense Advanced Research Projects Agency and the Army Research Office. We acknowledge the contributions by Michael Giardino, Adira Blumenthal, and Ariel Hirschhorn at the beginning phase of the project.

Conflicts of Interest

None declared.

References

1. 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;66(4):317-320. [doi: 10.1177/0020764020915212] [Medline: 32233719]
2. Pfefferbaum B, North CS. Mental Health and the Covid-19 Pandemic. N Engl J Med 2020 Aug 06;383(6):510-512. [doi: 10.1056/NEJMp2008017] [Medline: 32283003]
3. Rajkumar RP. COVID-19 and mental health: a review of the existing literature. Asian J Psychiatr 2020 Aug;52:102066 [FREE Full text] [doi: 10.1016/j.ajp.2020.102066] [Medline: 32302935]
4. Elmer T, Mepham K, Stadtfeld C. Students under lockdown: comparisons of students' social networks and mental health before and during the COVID-19 crisis in Switzerland. PLoS One 2020;15(7):e0236337 [FREE Full text] [doi: 10.1371/journal.pone.0236337] [Medline: 32702065]
5. Huckins J, daSilva AW, Wang W, Hedlund E, Rogers C, Nepal SK, et al. Mental health and behavior of college students during the early phases of the COVID-19 pandemic: longitudinal smartphone and Ecological Momentary Assessment study. J Med Internet Res 2020 Jun 17;22(6):e20185 [FREE Full text] [doi: 10.2196/20185] [Medline: 32519963]
6. Liu X, Liu J, Zhong X. Psychological state of college students during COVID-19 epidemic. SSRN J 2020. [doi: 10.2139/ssrn.3552814]
7. Cao W, Fang Z, Hou G, Han M, Xu X, Dong J, et al. The psychological impact of the COVID-19 epidemic on college students in China. Psychiatry Res 2020 May;287:112934 [FREE Full text] [doi: 10.1016/j.psychres.2020.112934] [Medline: 32229390]
8. Kessler RC, Chiu WT, Demler O, Merikangas KR, Walters EE. Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. Arch Gen Psychiatry 2005 Jun;62(6):617-627 [FREE Full text] [doi: 10.1001/archpsyc.62.6.617] [Medline: 15939839]
9. Wang PS, Berglund P, Olfson M, Pincus HA, Wells KB, Kessler RC. Failure and delay in initial treatment contact after first onset of mental disorders in the National Comorbidity Survey Replication. Arch Gen Psychiatry 2005 Jun;62(6):603-613. [doi: 10.1001/archpsyc.62.6.603] [Medline: 15939838]
10. Mavragani A. Tracking COVID-19 in Europe: infodemiology approach. JMIR Public Health Surveill 2020 Apr 20;6(2):e18941 [FREE Full text] [doi: 10.2196/18941] [Medline: 32250957]
11. Walker A, Hopkins C, Surda P. Use of Google Trends to investigate loss-of-smell-related searches during the COVID-19 outbreak. Int Forum Allergy Rhinol 2020 Jul;10(7):839-847 [FREE Full text] [doi: 10.1002/alr.22580] [Medline: 32279437]
12. Rovetta A, Bhagavathula AS. COVID-19-related web search behaviors and infodemic attitudes in Italy: infodemiological study. JMIR Public Health Surveill 2020 May 05;6(2):e19374 [FREE Full text] [doi: 10.2196/19374] [Medline: 32338613]

13. Dixit A, Marthoenis M, Arafat SY, Sharma P, Kar SK. Binge watching behavior during COVID 19 pandemic: a cross-sectional, cross-national online survey. Psychiatry Res 2020 Jul;289:113089 [FREE Full text] [doi: 10.1016/j.psychres.2020.113089] [Medline: 32434095]
14. Gao J, Zheng P, Jia Y, Chen H, Mao Y, Chen S, et al. Mental health problems and social media exposure during COVID-19 outbreak. PLoS One 2020;15(4):e0231924 [FREE Full text] [doi: 10.1371/journal.pone.0231924] [Medline: 32298385]
15. Seabrook EM, Kern ML, Rickard NS. Social networking sites, depression, and anxiety: a systematic review. JMIR Ment Health 2016 Nov 23;3(4):e50 [FREE Full text] [doi: 10.2196/mental.5842] [Medline: 27881357]
16. Lin LY, Sidani JE, Shensa A, Radovic A, Miller E, Colditz JB, et al. Association between social media use and depression among US young adults. Depress Anxiety 2016 Apr;33(4):323-331 [FREE Full text] [doi: 10.1002/da.22466] [Medline: 26783723]
17. De Choudhury M, Kiciman E, Dredze M, Coppersmith G, Kumar M. Discovering shifts to suicidal ideation from mental health content in social media. Proc SIGCHI Conf Hum Factor Comput Syst 2016 May;2016:2098-2110 [FREE Full text] [doi: 10.1145/2858036.2858207] [Medline: 29082385]
18. De Choudhury M, Counts S, Horvitz E. Social media as a measurement tool of depression in populations. In: Proceedings of the 5th Annual ACM Web Science Conference. 2013 Presented at: WebSci '13; May 2013; Paris, France p. 47. [doi: 10.1145/2464464.2464480]
19. Shen JH, Rudzicz F. Detecting anxiety through Reddit. In: Proceedings of the Fourth Workshop on Computational Linguistics and Clinical Psychology — From Linguistic Signal to Clinical Reality. 2017 Presented at: Fourth Workshop on Computational Linguistics and Clinical Psychology — From Linguistic Signal to Clinical Reality; August 2017; Vancouver, BC p. 58. [doi: 10.18653/v1/w17-3107]
20. Silveira Fraga B, Couto da Silva AP, Murai F. Online social networks in health care: a study of mental disorders on Reddit. 2018 Presented at: 2018 IEEE/WIC/ACM International Conference on Web Intelligence (WI); December 2018; Santiago, Chile p. 568-573. [doi: 10.1109/wi.2018.00-36]
21. Gaur M, Kursuncu U, Alambo A, Sheth A, Daniulaityte R, Thirunarayan K, et al. "Let Me Tell You About Your Mental Health!": contextualized classification of Reddit posts to DSM-5 for web-based intervention. In: Proceedings of the 27th ACM International Conference on Information and Knowledge Management. 2018 Presented at: CIKM '18; October 2018; Torino, Italy p. 753-762. [doi: 10.1145/3269206.3271732]
22. Frost RL, Rickwood DJ. A systematic review of the mental health outcomes associated with Facebook use. Comput Hum Behav 2017 Nov;76:576-600. [doi: 10.1016/j.chb.2017.08.001]
23. Zhang R. The stress-buffering effect of self-disclosure on Facebook: an examination of stressful life events, social support, and mental health among college students. Comput Hum Behav 2017 Oct;75:527-537. [doi: 10.1016/j.chb.2017.05.043]
24. Fleck BKB, Beckman LM, Sterns JL, Hussey HD. YouTube in the classroom: helpful tips and student perceptions. J Effective Teaching 2014;14(3):21-37.
25. Moghavvemi S, Sulaiman A, Jaafar NI, Kasem N. Social media as a complementary learning tool for teaching and learning: the case of YouTube. Int J Manage Education 2018 Mar;16(1):37-42. [doi: 10.1016/j.ijme.2017.12.001]
26. Klobas JE, McGill TJ, Moghavvemi S, Paramanathan T. Compulsive YouTube usage: a comparison of use motivation and personality effects. Comput Hum Behav 2018 Oct;87:129-139. [doi: 10.1016/j.chb.2018.05.038]
27. de Bérail P, Guillon M, Bungener C. The relations between YouTube addiction, social anxiety and parasocial relationships with YouTubers: a moderated-mediation model based on a cognitive-behavioral framework. Comput Hum Behav 2019 Oct;99:190-204. [doi: 10.1016/j.chb.2019.05.007]
28. Gunnell D, Appleby L, Arensman E, Hawton K, John A, Kapur N, COVID-19 Suicide Prevention Research Collaboration. Suicide risk and prevention during the COVID-19 pandemic. Lancet Psychiatry 2020 Jun;7(6):468-471 [FREE Full text] [doi: 10.1016/S2215-0366(20)30171-1] [Medline: 32330430]
29. van Agteren J, Bartholomaeus J, Fassnacht D, Iasiello M, Ali K, Lo L, et al. Using internet-based psychological measurement to capture the deteriorating community mental health profile during COVID-19: observational study. JMIR Ment Health 2020 Jun 11;7(6):e20696 [FREE Full text] [doi: 10.2196/20696] [Medline: 32490845]
30. Holmes EA, O'Connor RC, Perry VH, Tracey I, Wessely S, Arseneault L, et al. Multidisciplinary research priorities for the COVID-19 pandemic: a call for action for mental health science. Lancet Psychiatry 2020 Jun;7(6):547-560 [FREE Full text] [doi: 10.1016/S2215-0366(20)30168-1] [Medline: 32304649]
31. Greenberg N, Docherty M, Gnanapragasam S, Wessely S. Managing mental health challenges faced by healthcare workers during covid-19 pandemic. BMJ 2020 Mar 26;368:m1211. [doi: 10.1136/bmj.m1211] [Medline: 32217624]
32. Chen Q, Liang M, Li Y, Guo J, Fei D, Wang L, et al. Mental health care for medical staff in China during the COVID-19 outbreak. Lancet Psychiatry 2020 Apr;7(4):e15-e16 [FREE Full text] [doi: 10.1016/S2215-0366(20)30078-X] [Medline: 32085839]
33. Yao H, Chen J, Xu Y. Patients with mental health disorders in the COVID-19 epidemic. Lancet Psychiatry 2020 Apr;7(4):e21 [FREE Full text] [doi: 10.1016/S2215-0366(20)30090-0] [Medline: 32199510]
34. Ni MY, Yang L, Leung CMC, Li N, Yao XI, Wang Y, et al. Mental health, risk factors, and social media use during the COVID-19 epidemic and cordon sanitaire among the community and health professionals in Wuhan, China: cross-sectional survey. JMIR Ment Health 2020 May 12;7(5):e19009 [FREE Full text] [doi: 10.2196/19009] [Medline: 32365044]

35. Li HO, Bailey A, Huynh D, Chan J. YouTube as a source of information on COVID-19: a pandemic of misinformation? BMJ Glob Health 2020 May;5(5) [FREE Full text] [doi: 10.1136/bmjgh-2020-002604] [Medline: 32409327]
36. Basch CE, Basch CH, Hillyer GC, Jaime C. The role of YouTube and the entertainment industry in saving lives by educating and mobilizing the public to adopt behaviors for community mitigation of COVID-19: successive sampling design study. JMIR Public Health Surveill 2020 Apr 21;6(2):e19145 [FREE Full text] [doi: 10.2196/19145] [Medline: 32297593]
37. Basch C, Hillyer G, Meleo-Erwin Z, Jaime C, Mohlman J, Basch C. Preventive behaviors conveyed on YouTube to mitigate transmission of COVID-19: cross-sectional study. JMIR Public Health Surveill 2020 Apr 02;6(2):e18807 [FREE Full text] [doi: 10.2196/18807] [Medline: 32240096]
38. Pennycook G, McPhetres J, Zhang Y, Lu J, Rand D. Fighting COVID-19 misinformation on social media: experimental evidence for a scalable accuracy-nudge intervention. Psychol Sci 2020 Jul;31(7):770-780 [FREE Full text] [doi: 10.1177/0956797620939054] [Medline: 32603243]
39. Lyu H, Chen L, Wang Y, Luo J. Sense and sensibility: characterizing social media users regarding the use of controversial terms for COVID-19. IEEE Trans Big Data 2020:1. [doi: 10.1109/tbdata.2020.2996401]
40. Gordon K, Bresin K, Dombeck J, Routledge C, Wonderlich J. The impact of the 2009 Red River Flood on interpersonal risk factors for suicide. Crisis 2011;32(1):52-55. [doi: 10.1027/0227-5910/a000051] [Medline: 21371971]
41. Saeb S, Lattie EG, Schueller SM, Kording KP, Mohr DC. The relationship between mobile phone location sensor data and depressive symptom severity. PeerJ 2016;4:e2537. [doi: 10.7717/peerj.2537] [Medline: 28344895]
42. Saeb S, Zhang M, Kwasny MM, Karr CJ, Kording K, Mohr DC. The relationship between clinical, momentary, and sensor-based assessment of depression. Int Conf Pervasive Comput Technol Healthc 2015 Aug;2015 [FREE Full text] [doi: 10.4108/icst.pervasivehealth.2015.259034] [Medline: 26640739]
43. Ghandeharioun A, Fedor S, Sangermano L, Ionescu D, Alpert J, Dale C, et al. Objective assessment of depressive symptoms with machine learning and wearable sensors data. 2017 Presented at: 2017 Seventh International Conference on Affective Computing and Intelligent Interaction (ACII); October 2017; San Antonio, TX. [doi: 10.1109/acii.2017.8273620]
44. Wang R, Wang W, daSilva A, Huckins JF, Kelley WM, Heatherton TF, et al. Tracking depression dynamics in college students using mobile phone and wearable sensing. Proc ACM Interact Mob Wearable Ubiquitous Technol 2018 Mar 26;2(1):1-26. [doi: 10.1145/3191775]
45. Zaman A, Acharyya R, Kautz H, Silenzio V. Detecting low self-esteem in youths from web search data. 2019 Presented at: WWW '19: The World Wide Web Conference; May 2019; San Francisco, CA p. 2270-2280. [doi: 10.1145/3308558.3313557]
46. Google Takeout. URL: https://takeout.google.com/ [accessed 2020-08-07]
47. Cloud data loss prevention. Google Cloud. URL: https://cloud.google.com/dlp [accessed 2020-08-07]
48. Halfaker A, Keyes O, Kluver D, Thebault-Spieker J, Nguyen T, Shores K, et al. User session identification based on strong regularities in inter-activity time. In: Proceedings of the 24th International Conference on World Wide Web. 2015 Presented at: WWW '15; May 2015; Florence, Italy p. 410-418. [doi: 10.1145/2736277.2741117]
49. 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/handle/2152/31333 [accessed 2020-08-02]
50. Classifying content. Google Cloud. URL: https://cloud.google.com/natural-language/docs/classifying-text [accessed 2020-08-07]
51. Content categories. Google Cloud. URL: https://cloud.google.com/natural-language/docs/categories [accessed 2020-08-07]
52. VideoCategories. Google Developers. URL: https://developers.google.com/youtube/v3/docs/videoCategories [accessed 2020-08-07]
53. Domain APIs. Webshrinker. URL: https://www.webshrinker.com/apis/ [accessed 2020-10-16]
54. 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]
55. Rutter LA, Brown TA. Psychometric properties of the Generalized Anxiety Disorder Scale-7 (GAD-7) in outpatients with anxiety and mood disorders. J Psychopathol Behav Assess 2017 Mar;39(1):140-146 [FREE Full text] [doi: 10.1007/s10862-016-9571-9] [Medline: 28260835]
56. Kroenke K. Enhancing the clinical utility of depression screening. CMAJ 2012 Feb 21;184(3):281-282 [FREE Full text] [doi: 10.1503/cmaj.112004] [Medline: 22231681]
57. Newman ML, Pennebaker JW, Berry DS, Richards JM. Lying words: predicting deception from linguistic styles. Pers Soc Psychol Bull 2003 May;29(5):665-675. [doi: 10.1177/0146167203029005010] [Medline: 15272998]
58. Cohn M, Mehl M, Pennebaker J. Linguistic markers of psychological change surrounding September 11, 2001. Psychol Sci 2004 Oct;15(10):687-693. [doi: 10.1111/j.0956-7976.2004.00741.x] [Medline: 15447640]
59. Kuss DJ, Griffiths MD, Binder JF. Internet addiction in students: prevalence and risk factors. Comput Hum Behav 2013 May;29(3):959-966. [doi: 10.1016/j.chb.2012.12.024]
60. Tarokh L, Saletin JM, Carskadon MA. Sleep in adolescence: physiology, cognition and mental health. Neurosci Biobehav Rev 2016 Nov;70:182-188 [FREE Full text] [doi: 10.1016/j.neubiorev.2016.08.008] [Medline: 27531236]

61. Konjarski M, Murray G, Lee VV, Jackson ML. Reciprocal relationships between daily sleep and mood: a systematic review of naturalistic prospective studies. Sleep Med Rev 2018 Dec;42:47-58. [doi: 10.1016/j.smrv.2018.05.005] [Medline: 30404728]
62. Willoughby BJ, Busby DM, Young-Petersen B. Understanding associations between personal definitions of pornography, using pornography, and depression. Sex Res Soc Policy 2018 Jun 30;16(3):342-356. [doi: 10.1007/s13178-018-0345-x]
63. Yoder VC, Virden TB, Amin K. Internet pornography and loneliness: an association? Sex Addict Compulsivity 2005 Jan;12(1):19-44. [doi: 10.1080/10720160590933653]
64. Keelan J, Pavri-Garcia V, Tomlinson G, Wilson K. YouTube as a source of information on immunization: a content analysis. JAMA 2007 Dec 05;298(21):2482-2484. [doi: 10.1001/jama.298.21.2482] [Medline: 18056901]
65. Girard JM, Cohn JF, Mahoor MH, Mavadati SM, Hammal Z, Rosenwald DP. Nonverbal social withdrawal in depression: evidence from manual and automatic analysis. Image Vis Comput 2014 Oct;32(10):641-647 [FREE Full text] [doi: 10.1016/j.imavis.2013.12.007] [Medline: 25378765]
66. Katz SJ, Conway CC, Hammen CL, Brennan PA, Najman JM. Childhood social withdrawal, interpersonal impairment, and young adult depression: a mediational model. J Abnorm Child Psychol 2011 Nov;39(8):1227-1238 [FREE Full text] [doi: 10.1007/s10802-011-9537-z] [Medline: 21744059]
67. Leigh-Hunt N, Bagguley D, Bash K, Turner V, Turnbull S, Valtorta N, et al. An overview of systematic reviews on the public health consequences of social isolation and loneliness. Public Health 2017 Nov;152:157-171. [doi: 10.1016/j.puhe.2017.07.035] [Medline: 28915435]
68. Huang X, Zhang L, Chiu D, Liu T, Li X, Zhu T. Detecting suicidal ideation in Chinese microblogs with psychological lexicons. 2014 Presented at: 2014 IEEE 11th Intl Conf on Ubiquitous Intelligence and Computing and 2014 IEEE 11th Intl Conf on Autonomic and Trusted Computing and 2014 IEEE 14th Intl Conf on Scalable Computing and Communications and Its Associated Workshops; December 2014; Bali, Indonesia p. 844. [doi: 10.1109/uic-atc-scalcom.2014.48]
69. Challen R, Denny J, Pitt M, Gompels L, Edwards T, Tsaneva-Atanasova K. Artificial intelligence, bias and clinical safety. BMJ Qual Saf 2019 Mar;28(3):231-237 [FREE Full text] [doi: 10.1136/bmjqs-2018-008370] [Medline: 30636200]
70. Gianfrancesco MA, Tamang S, Yazdany J, Schmajuk G. Potential biases in machine learning algorithms using electronic health record data. JAMA Intern Med 2018 Nov 01;178(11):1544-1547 [FREE Full text] [doi: 10.1001/jamainternmed.2018.3763] [Medline: 30128552]
71. Wang W, Kiik M, Peek N, Curcin V, Marshall IJ, Rudd AG, et al. A systematic review of machine learning models for predicting outcomes of stroke with structured data. PLoS One 2020;15(6):e0234722 [FREE Full text] [doi: 10.1371/journal.pone.0234722] [Medline: 32530947]
72. Kim H, Jung DC, Choi BW. Exploiting the vulnerability of deep learning-based artificial intelligence models in medical imaging: adversarial attacks. J Korean Soc Radiol 2019;80(2):259. [doi: 10.3348/jksr.2019.80.2.259]
73. Ma X, Niu Y, Gu L, Wang Y, Zhao Y, Bailey J, et al. Understanding adversarial attacks on deep learning based medical image analysis systems. Pattern Recognition 2020 Feb;110:107332. [doi: 10.1016/j.patcog.2020.107332]

Abbreviations

API: application programming interface
GAD-7: General Anxiety Disorder-7
LIWC: Linguistic Inquiry and Word Count
MSE: mean squared error
NLP: natural language processing
PHQ-9: Patient Health Questionnaire-9
S: slope
SEI: short event interval

Edited by J Torous, G Eysenbach; submitted 31.08.20; peer-reviewed by S Abdullah, K Harrigian; comments to author 28.09.20; revised version received 19.10.20; accepted 09.11.20; published 23.11.20.

Please cite as:
Zhang B, Zaman A, Silenzio V, Kautz H, Hoque E
The Relationships of Deteriorating Depression and Anxiety With Longitudinal Behavioral Changes in Google and YouTube Use During COVID-19: Observational Study
JMIR Ment Health 2020;7(11):e24012
URL: http://mental.jmir.org/2020/11/e24012/
doi:10.2196/24012
PMID:33180743

Industry Perspective

Leveraging the Power of Nondisruptive Technologies to Optimize Mental Health Treatment: Case Study

Shiri Sadeh-Sharvit[1,2], PhD; Steven D Hollon[3], PhD

[1]Eleos Health, Cambridge, MD, United States
[2]Center for m2Health, Palo Alto University, Palo Alto, CA, United States
[3]Vanderbilt University, Nashville, TN, United States

Corresponding Author:
Shiri Sadeh-Sharvit, PhD
Eleos Health
1 Broadway
Cambridge, MD, 02141
United States
Phone: 1 5109848132
Email: shiri@eleos.health

Abstract

Regular assessment of the effectiveness of behavioral interventions is a potent tool for improving their relevance to patients. However, poor provider and patient adherence characterize most measurement-based care tools. Therefore, a new approach for measuring intervention effects and communicating them to providers in a seamless manner is warranted. This paper provides a brief overview of the available research evidence on novel ways to measure the effects of behavioral treatments, integrating both objective and subjective data. We highlight the importance of analyzing therapeutic conversations through natural language processing. We then suggest a conceptual framework for capitalizing on data captured through directly collected and nondisruptive methodologies to describe the client's characteristics and needs and inform clinical decision-making. We then apply this context in exploring a new tool to integrate the content of therapeutic conversations and patients' self-reports. We present a case study of how both subjective and objective measures of treatment effects were implemented in cognitive-behavioral treatment for depression and anxiety and then utilized in treatment planning, delivery, and termination. In this tool, called Eleos, the patient completes standardized measures of depression and anxiety. The content of the treatment sessions was evaluated using nondisruptive, independent measures of conversation content, fidelity to the treatment model, and the back-and-forth of client-therapist dialogue. Innovative applications of advances in digital health are needed to disseminate empirically supported interventions and measure them in a noncumbersome way. Eleos appears to be a feasible, sustainable, and effective way to assess behavioral health care.

(JMIR Ment Health 2020;7(11):e20646) doi:10.2196/20646

KEYWORDS

anxiety; behavioral health; depression; digital health; Eleos Health; mental health; natural language processing

Introduction

Background

Psychotherapy is on the edge of a transformation driven by technology. More than 50% of adults in high- and middle-income countries will experience a mental disorder in their lifetimes [1]. However, while psychological treatments for most mental disorders have demonstrated efficacy, the quality and effectiveness of mental health care delivery remain inadequate due to multiple reasons, including access, cost, paucity of clinicians trained in empirically-supported models, and the absence of objective and systematic methods for assessing treatments during their delivery [2,3]. Further, health insurance models often limit the number of sessions patients can receive, thereby requiring sustainable outcomes within a relatively short period of time. Treatment delivery models need to be improved to bridge this gap.

One avenue to improving the efficacy and cost-effectiveness of mental health care and increasing patient satisfaction and engagement is integrating technology into clinical practice through patient evaluations and monitoring of the treatment process [4]. Recent digital developments can provide clinicians with nuanced, real-time information to assist their decision-making capacity. Cutting edge technologies can be incorporated into clinical practice for this purpose. Data can be presented to the provider prior to each session to inform them of the patient's progress, summarize key points from the last session, and prepare them for the upcoming meeting [5]. These data can also inform supervisors, program directors, and other stakeholders to guide clinical decisions, resource allocation, and training.

The goal of this paper is to examine and provide an update on state-of-the-art techniques and methodologies to assess behavioral interventions. We also provide a conceptual framework for collecting and integrating client and treatment data, summarizing, analyzing, and visualizing information to accurately capture the client's progress and needs. We then demonstrate the utility of a platform incorporating such instruments in the treatment for anxiety and depression.

Standardized Assessments of Treatment Effects

To formulate a case and assess the impacts of a prescribed intervention, one must rely on data [6]. Technology can be successfully leveraged to provide measurement-based care (MBC), defined as the practice of grounding clinical care in patient data collected throughout treatment [7]. MBC is superior to usual care because it offers several benefits that optimize treatments, such as providing insight into treatment progress, early detection of symptom relapse, and improving outcomes [8]. Evidence indicates that patients whose providers use MBC achieve greater and faster treatment response and symptomatic remission [9]. For instance, 2 validated symptom assessment measures, The Patient Health Questionnaire-9 (PHQ-9 [10]) and the Generalized Anxiety Disorder-7 (GAD-7 [11]), are widely used in practice and have been utilized in thousands of studies and clinical settings globally. As their names suggest, they are very brief, totaling 16 items. They provide an estimation of the patient's depression and anxiety levels and whether these are within the clinically severe, moderate, or normal ranges. Such self-report tools can help clarify the patient's experience and progress, and provide a real-time signal as to whether the treatment helps them get where they want to be [12].

Despite being the simplest, most cost-effective way of assessing patients' clinical status, self-report measures in clinical practice have several caveats. Patient compliance with self-assessment is limited, with frequent complaints that questionnaires are cumbersome to complete and repetitive [13,14]. Patients may exaggerate or minimize reports to continue receiving services or avoiding potential consequences of their condition. Therefore, a combination of self-report instruments, clinician ratings, and data from additional resources could likely depict the client's progress more accurately [15]. In order to make treatments more effective, precise, and relevant to the individuals seeking them, additional measures need to complement surveys and questionnaires. Mental health care must develop innovative technologies that could transform behavioral treatment without disrupting their delivery (eg, consuming time or interrupting the conversation flow) [16,17].

Machine Learning and Human-Centered Designs in Behavioral Medicine

Timely data are crucial for planning treatment and assessing its effects. Is the client occasionally preoccupied with troubling ideas or does she meet criteria for an obsessive-compulsive disorder? Has the veteran in treatment for insomnia been able to sleep better following the intervention? Does the teen in adolescent-focused therapy feel connected and valued enough in treatment to share their suicidal ideation with their therapist? Clinicians use various methods to collect these data and gauge their predictive value. However, new models should integrate numerous data sources to provide health care that is person-centered, efficient, and targeted to meet patients' unique needs [18]. Technological developments can ensure information collecting is contextualized, optimized, and translated into clinical insights and actionable decisions by providers and stakeholders [19].

Machine learning (ML) provides unparalleled precision and accuracy in predicting treatment outcomes based on data collected early in treatment, and determining the most successful targets for interventions [20]. ML algorithms can integrate many sources of information to predict the client's functioning in treatment, such as text used in the therapeutic conversation, the proportion of time each participant talked, conversation turn-taking, and the client's self-report measures over time [21]. ML can also be used to assess treatment fidelity and the therapeutic relationship, which until recently relied on exhaustive manual work, mostly coded in research trials but not applied to community settings [22]. These capabilities enable timelier identification of trends in patients' symptoms and the issues troubling them, as well as changes in the therapeutic alliance and the therapeutic relationship [23]. Greater awareness of nuanced changes over the course of treatment, both in the content expressed during the sessions and changes in symptoms throughout the treatment period, has the potential to better inform and better prepare the therapist to provide effective interventions, seek out consultation and support, or complement the current intervention with additional treatment modules as needed.

Increased adherence to evidence-based care predicts improved treatment outcomes [24]. However, ongoing evaluation of treatment progress remains a challenge [25]. Nondisruptive measures were introduced in behavioral medicine over half a century ago. There is documentation as early as the 1970s of audio recordings or videotapes of treatment sessions being routinely used in training, supervision, practice, and consultation [26]. Recordings are widely used in behavioral medicine [27] as video or audio recordings have been mandated in many clinical training programs and routine care [28,29]. Most patients report positive attitudes toward the use of recordings in their treatment [30]. For example, 71% of patients in a recent study were open to considering audio or video recordings of their treatment sessions. The patients' comfort with recording was not associated with treatment refusal, duration, or outcomes [31].

Sophisticated algorithms for voice analysis through natural language processing (NLP) allow for the detection of trends in the conversation's sentiment, content, and synchrony of participants. Insights from session recordings can inspire behavior change, informing the clinician of metrics relevant for both the process and the content of the intervention. Of note, data collected passively during the regular course of treatment in a nondisruptive manner can inform with respect to therapist variables as well. These data provide a broader, more nuanced consideration of how patient, treatment, and clinician variables interact to achieve treatment effects [32]. As such, methods that passively assess and integrate multiple variables at the same time can significantly affect treatment engagement, adherence, and outcomes. When ML models integrate several data resources and are scientifically robust, they can be implemented in a platform offering artificial intelligence (AI) information and prediction [33].

A plethora of empirical evidence suggests that within-session process variables predict the patient-therapist connection, the patient’s mood and anxiety, and the content of the patient’s interests, strengths, concerns, and dilemmas [34,35]. As such, data derived from session recordings can serve as objective markers of the treatment process and inform the clinician where to head next [36]. The patient/therapist listening ratio, number of cross-talks, and silences all reflect the nature of the clinical relationship, therapeutic alliance, and the patient’s engagement with the treatment. Evidence shows that treatments in which patients speak about themselves and are engaged in the conversation are particularly likely to maintain momentum [35]. Patient involvement in the session and therapist active listening can be observed by conversational interaction when the patient and the therapist take turns speaking, when there are few extended pauses, and when neither party overrides the other. Silences are effective when they are used in later sessions for brief periods. Long silences by the patient reflect a lesser sense of connection, affecting attrition, adherence, and outcomes [34].

Within-session content variables, such as data on most commonly used themes and affective tone, can inform the therapist, stakeholders, and policymakers in addressing underlying perpetuating factors. Integrating sentiment and the themes discussed can further inform stakeholders regarding treatment progress [37]. Patients that express less emotional content tend to rate the therapy as less helpful and their connection with their therapist as weaker [35]. Further, greater therapist insight into the interests, concerns, and experiences of the patient predicts whether they will reach the outcomes desired by each party [38,39].

Current Implementation of Digital Tools to Augment Behavioral Medicine Outcomes

Although their potential role in optimizing treatment delivery has been proposed, in-session and between-session data are not collected regularly, nor have they been integrated into mental health care services as of yet. Barriers often cited are limited clinician time to administer, collect, and analyze data, and concerns that the administration of measures would interfere with rapport and the therapeutic alliance [40]. For a digital tool to be maximally effective, it needs to collect information passively, without increasing therapist burden or reducing face-to-face communication [41]. Additionally, these data should be provided to therapists via a platform that is straightforward and easy to use, with clear visualization and comparison to earlier sessions. Such information presented in a timely fashion can inform decisions regarding resource allocation, such as increasing treatment dose, revisiting the level of care, and team consultation [7]. Further, aggregated data on treatment progress and outcomes can be used by providers and clinic directors for quality assessment.

The Eleos Health Platform

Description

Eleos Health is a therapy intelligence engine designed to provide intervention insights and inform clinical decision-making. The platform collects key metrics from treatment sessions and integrates them with standardized assessment scales, leveraging insights developed through ML and NLP analysis of large treatment datasets. The Eleos platform integrates subjective and objective measures of the treatment process, the patient-therapist communication, and outcomes into AI software. In the following case illustration, we demonstrate how objective measures of treatment process and content derived from session recordings can be integrated with patient self-assessments in real time to shape clinical insights and decisions in a positive direction. Eleos Health is a digital platform designed to integrate multiple patient data points to present providers with a comprehensive picture of the client’s progress in treatment. The platform is used in an app that can be used on a mobile device or desktop computer for in-person meetings or embedded within teleconferencing programs. The platform applies voice analysis to describe and summarize events throughout the treatment meeting, including the language used by the therapist and the patient. These data are complemented with weekly outcome monitoring through self-report assessments.

Case Example

The following case example illustrates the pilot use of an AI platform hosted by Eleos Health to collect and analyze the content of treatment sessions. The patient described signed consent for using de-identified treatment data prior to beginning the treatment and gave permission for the following text. All identifying information has been changed.

Kyle was a 24-year-old Latino man. He was born and raised in a small suburb next to a big metropolis. The youngest of 4 children, he described a very warm and strongly connected family growing up. He had graduated from a liberal arts college the previous year and found a job in a small startup company. He moved to a new city, where he lived with 3 roommates. Kyle described that in the past 18 months, since his junior year in college, he felt concerned about his future and unsure about which career path to choose. His friends and family recommended that he seek counseling, but after calling several therapists who did not have a slot, he did not begin treatment. Instead, he focused on his eating and physical activity, and thought that his new routine supported his transition into his new job. Kyle recently decided to seek treatment again after receiving a promotion at work. His boss moved out of town and the CEO of the company offered Kyle her position. Kyle accepted the promotion but became very anxious and had a hard time concentrating at work. He also reported sleep problems and that the healthy lifestyle he had worked hard to develop had been derailed.

When Kyle reached out to Dr. Davis, the therapist suggested incorporating a platform in his treatment that could record and analyze their treatment sessions. The therapist also suggested that the system send Kyle weekly assessments to enhance the therapist's understanding of how best to help him. Kyle was skeptical that a digital platform could inform his therapist beyond the treatment session per se, but decided he had little to lose by trying it out. He signed a consent form that the therapist had sent him, which included an authorization to use a HIPAA-compliant platform named Eleos. Treatment was conducted in a blended fashion, integrating in-person and remotely delivered meetings. The Eleos platform provides voice analysis of the sessions, regardless of their delivery method.

Kyle received a text message from the system before the first session, requesting that he complete the PHQ-9 and the GAD-7. He completed the 16 items on his phone within a few minutes, which served as his baseline scores. When he sat in the therapist's office in the first session, she pointed at a mobile device that she would use to record and analyze their meeting. Kyle's therapist was very welcoming and helped him solidify his goals, which were to feel less anxious at work and happier after work hours. The therapist mentioned that Kyle's PHQ-9 score was 18 and that his GAD-7 score was 15, indicating that his depression and anxiety were both in the severe range [10,11]. Since Kyle's insurance covered only 12 sessions, the therapist explained that she would provide cognitive behavioral therapy, focusing on his interpretations of ongoing events and how these beliefs affected his emotions and behaviors. Given the severity of his symptoms, she also referred him for medication evaluation.

After their second session, Dr. Davis reviewed the reports she received from Eleos, the therapeutic intelligence platform she had been using. She was surprised to see that in their last session, she spoke in only 20% of the session and that she was speaking much slower than Kyle (Figure 1). The speech rate difference indicated to Dr. Davis that Kyle was speaking hurriedly and reflected her attempt to "slow him down" to reflect on his maladaptive assumptions and ingrained interactional patterns. However, she was wondering whether she could share these observations with Kyle more explicitly. These data helped the therapist understand that Kyle was experiencing significant mood, anxiety, and stress symptoms, but that in order to help Kyle achieve his goals, their synchrony during the session would have to be substantially improved [35]. A review of the process metrics provided insight into the issues troubling Kyle and the inaccurate beliefs and maladaptive information processing strategies that precipitated and maintained his cognitions [42]. A review of his most frequently used words indicated that he tended to use verbs, adjectives, and expressions associated with negative self-esteem, such as "failed," "not as good as," "disappointed," and "messed it up," when he talked about his work. However, when he talked about his relationships, he tended to use anxiety words, such as "stressed," "pressured," "overwhelmed," and "toxic." The system flagged these words and phrases, as they approximate depressive and anxiety symptoms and may reflect Kyle's subjective experience, thereby enriching the self-report data collected between sessions. In addition, 3 of Kyle's most used phrases were "shoulds" (eg, "have to," "I must," "should have known better"), which the system automatically analyzed and flagged. The therapist realized that Kyle was experiencing 2 distinct phenomena: At work, he felt like an imposter and worried about his functioning, whereas outside of work, he was distressed by blurred boundaries in his interpersonal relationships. Dr. Davis also realized that she was using little Socratic questioning, which may have reduced Kyle's ability to re-examine his assumptions and information-processing skills [43]. However, she did observe her use of reflective listening methods often, mirroring and reframing what Kyle had said, which she was content with [44,45]. Dr. Davis brought the case to her weekly group consultation meeting and received feedback and advice from her peers that she intended to implement in the next few sessions.

Figure 1. Screenshots illustrating some of the Eleos Health platform process features. CBT: cognitive behavioral therapy; Speaking Ratio: proportion (%) of time spent speaking during the session; Speech Rate: speed of words per minute; Techniques Used: intervention strategies employed by the therapist during the session and automatically identified by the platform. Techniques used 3 times or more are indicated with a black checkmark, techniques used once or twice are denoted with a grey checkmark, and interventions not employed in the session do not have a checkmark.

Session 2

SPEAKING RATIO (TALK TIME)

WHO	RATIO (%)
Patient	80
Therapist	20

SPEECH RATE

WHO	WORDS/MIN
Patient	107.7
Therapist	129.5

TECHNIQUES USED
- ☑ Mood Check
- Bridge
- ✓ Agenda setting
- Homework review
- ☑ CBT skills
- ☑ Action plan
- Session summary
- ☑ Socratic questioning
- ✓ Reflective listening
- ✓ Empathic Communication

Session 5

SPEAKING RATIO (TALK TIME)

WHO	RATIO (%)
Patient	64
Therapist	36

SPEECH RATE

WHO	WORDS/MIN
Patient	110.1
Therapist	118.7

TECHNIQUES USED
- ☑ Mood Check
- ✓ Bridge
- ✓ Agenda setting
- ☑ Homework review
- ☑ CBT skills
- ✓ Action plan
- ☑ Session summary
- ✓ Socratic questioning
- ☑ Reflective listening
- ✓ Empathic Communication

Session 9

SPEAKING RATIO (TALK TIME)

WHO	RATIO (%)
Patient	61
Therapist	39

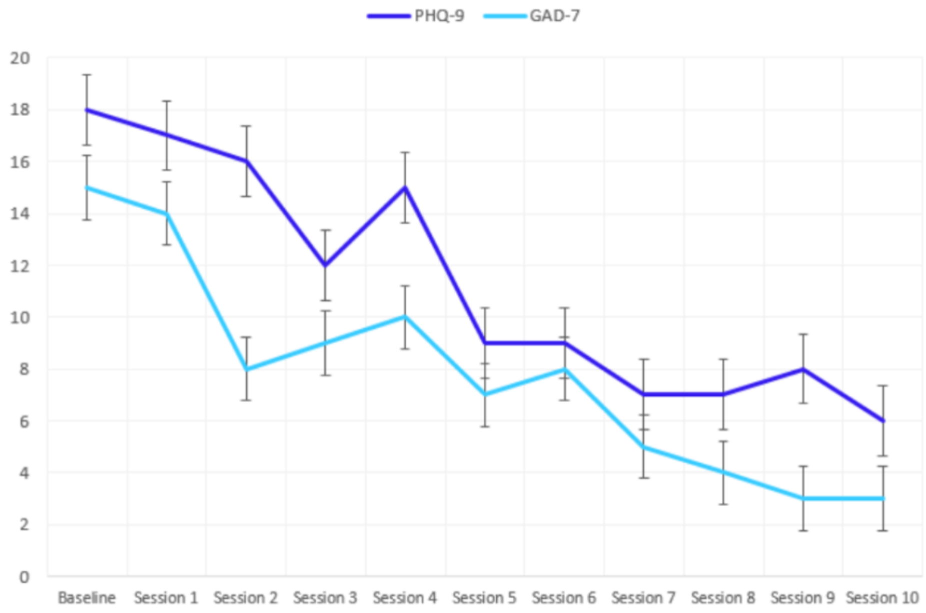
SPEECH RATE

WHO	WORDS/MIN
Patient	112.1
Therapist	119.3

TECHNIQUES USED
- Mood Check
- Bridge
- ✓ Agenda setting
- ☑ Homework review
- ✓ CBT skills
- ✓ Action plan
- ☑ Session summary
- ☑ Socratic questioning
- ☑ Reflective listening
- ☑ Empathic Communication

In the following meeting, Dr. Davis shared with Kyle the insights she had gleaned from the platform. She described the 2 clinical issues—his depression regarding work and his anxiety regarding relationships—and asked him to prioritize his goals. Dr. Davis's presentation of his challenges helped Kyle reflect on them, and he decided to first focus on his self-esteem at work. He learned how to observe, identify, and challenge negative perceptions of himself, and reported very quick improvements in his mood. Next, he was able to undertake the same process regarding several relationships that he felt had not been reciprocal and gratifying. He was pleased with his progress in treatment and felt happier and more relaxed. Kyle also liked that he could complete short assessments on his phone and got into the habit of doing this on his commute home from work when he had a few minutes to spare. Similarly, Dr. Davis appreciated the symptom-tracking feature, which let both her and her patient easily see how he was doing symptom-wise (Figure 2). In the session analytics reports, Dr. Davis also observed an improvement in the therapeutic alliance: She was able to incorporate more open-ended questions, Kyle was more receptive of her questions and comments, and their speech rate was more in sync with one another (Figure 1).

Figure 2. Patient self-monitoring data graphed on the Eleos platform. GAD-7: Generalized Anxiety Disorder-7; PHQ-9: Patient Health Questionnaire-9.



After 10 sessions, Kyle was not only feeling less depressed and anxious but was also receiving praise from his coworkers and his boss. He was able to strengthen his friendships and practice new skills for boundary-setting and interpersonal effectiveness. As treatment termination neared, the last sessions focused on brainstorming strategies for maintaining his progress, particularly since the workload at his company continued to increase.

Conclusions

Psychological interventions for mental disorders were found effective in numerous research trials. However, gaps in training, availability, access, dissemination, and cost impede the successful delivery of these interventions [46]. These are enhanced by privacy and data storage regulations, which are key for maintaining clients' rights and trust [47]. The ever-growing demand for mental health care requires optimizing clinician decision-making using data collected passively [48]. Our goal was to highlight key issues for optimizing available mental health services and to demonstrate how a therapeutic intelligence platform can support this process.

This paper presents a novel approach to collecting comprehensive data on treatment progress. The implementation of ML models and AI in behavioral health care is a rapidly moving and innovative field, with the potential to significantly improve screening and clinical outcomes [49,50]. Accurate data can provide more information about the patient and can be translated into clinical decisions faster. Experts strive to base their decisions on data. Primary physicians, heart surgeons, physiotherapists, and other health professionals all function better when they review the patient's most recent tests and laboratory results prior to the appointment. Therefore, using real-time data is equally important in enhancing the work of mental health professionals [51]. Technologies that are scalable, cost-effective, and that enhance quality without burden can help therapists harness their efforts into providing more effective interventions with increased fidelity to data-informed treatments. Setting out clear performance targets in the training, provision, and implementation of evidence-enhanced treatments will enable health care services to continually improve. Nondisruptive measures are poised to ensure nearly effortless data collection, and innovative methods that inform clinicians and other stakeholders of the patient's progress will likely make treatments more relevant and engaging [52].

Some caveats to the model described here should be mentioned. The Eleos platform was illustrated through its use in an outpatient setting, with a client who was relatively high functioning and a therapist that was tech-savvy and interested in using novel digital programs. Usability testing, user reviews, and long-term engagement with any product are key to realizing the practicality and helpfulness of new tools over time [53]. Further, evidence from more case reports, randomized controlled trials, and meta-analyses is needed to render these technologies pertinent, empirically supported, and easily applied in clinical settings.

The effective implementation of mental health care requires new approaches for developing, implementing, and evaluating interventions. A person-centered approach that capitalizes on greater data insights will certainly enhance the therapeutic process. Technology can help make efforts in this direction scalable and more efficient, thereby increasing the effects of behavioral interventions and reducing the burden of mental health problems worldwide.

Acknowledgments

The authors would like to thank the clinicians and patients who have been piloting the Eleos platform. We also thank patient “Kyle,” who read and approved this case illustration.

Conflicts of Interest

SSS is the Chief Clinical Officer of the commercial entity Eleos Health Inc that created the platform that is the subject of this case report. SDH is an unpaid advisor to Eleos Health Inc.

References

1. Trautmann S, Rehm J, Wittchen H. The economic costs of mental disorders: Do our societies react appropriately to the burden of mental disorders? EMBO Rep 2016 Sep 04;17(9):1245-1249 [FREE Full text] [doi: 10.15252/embr.201642951] [Medline: 27491723]
2. Ewbank MP, Cummins R, Tablan V, Bateup S, Catarino A, Martin AJ, et al. Quantifying the Association Between Psychotherapy Content and Clinical Outcomes Using Deep Learning. JAMA Psychiatry 2020 Jan 01;77(1):35-43 [FREE Full text] [doi: 10.1001/jamapsychiatry.2019.2664] [Medline: 31436785]
3. Kazdin A. Addressing the treatment gap: A key challenge for extending evidence-based psychosocial interventions. Behav Res Ther 2017 Jan;88:7-18 [FREE Full text] [doi: 10.1016/j.brat.2016.06.004] [Medline: 28110678]
4. Sadeh-Sharvit S. Use of Technology in the Assessment and Treatment of Eating Disorders in Youth. Child Adolesc Psychiatr Clin N Am 2019 Oct;28(4):653-661. [doi: 10.1016/j.chc.2019.05.011] [Medline: 31443882]
5. Topol E. Deep medicine: how artificial intelligence can make healthcare human again. New York, NY: Basic Books; 2019.
6. Scott K, Lewis CC. Using Measurement-Based Care to Enhance Any Treatment. Cogn Behav Pract 2015 Feb;22(1):49-59 [FREE Full text] [doi: 10.1016/j.cbpra.2014.01.010] [Medline: 27330267]
7. Fortney JC, Unützer J, Wrenn G, Pyne JM, Smith GR, Schoenbaum M, et al. A Tipping Point for Measurement-Based Care. PS 2017 Feb;68(2):179-188. [doi: 10.1176/appi.ps.201500439]
8. Nasrallah H, Aboraya A. It's time to implement measurement-based care in psychiatric practice. Current Psychiatry 2019;18(6):6-8 [FREE Full text]
9. Guo T, Xiang Y, Xiao L, Hu C, Chiu HFK, Ungvari GS, et al. Measurement-Based Care Versus Standard Care for Major Depression: A Randomized Controlled Trial With Blind Raters. Am J Psychiatry 2015 Oct;172(10):1004-1013. [doi: 10.1176/appi.ajp.2015.14050652] [Medline: 26315978]
10. Kroenke K, Spitzer RL, Williams JBW. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-613 [FREE Full text] [doi: 10.1046/j.1525-1497.2001.016009606.x] [Medline: 11556941]
11. 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]
12. Darcy A, Sadeh-Sharvit S. Mobile Device Applications for the AssessmentTreatment of Eating Disorders. In: Agras SW, Robinson A, editors. The Oxford Handbook of Eating Disorders, Second Edition. Oxford, UK: Oxford University Press; 2017:492-504.
13. Anatchkova M, Donelson SM, Skalicky AM, McHorney CA, Jagun D, Whiteley J. Exploring the implementation of patient-reported outcome measures in cancer care: need for more real-world evidence results in the peer reviewed literature. J Patient Rep Outcomes 2018 Dec 27;2(1):64 [FREE Full text] [doi: 10.1186/s41687-018-0091-0] [Medline: 30588562]
14. Haun JN, Alman AC, Melillo C, Standifer M, McMahon-Grenz J, Shin M, et al. Using Electronic Data Collection Platforms to Assess Complementary and Integrative Health Patient-Reported Outcomes: Feasibility Project. JMIR Med Inform 2020 Jun 26;8(6):e15609. [doi: 10.2196/15609]
15. Cuijpers P, Li J, Hofmann SG, Andersson G. Self-reported versus clinician-rated symptoms of depression as outcome measures in psychotherapy research on depression: a meta-analysis. Clin Psychol Rev 2010 Aug;30(6):768-778. [doi: 10.1016/j.cpr.2010.06.001] [Medline: 20619943]
16. Briffault X, Morgiève M, Courtet P. From e-Health to i-Health: Prospective Reflexions on the Use of Intelligent Systems in Mental Health Care. Brain Sci 2018 May 31;8(6):98 [FREE Full text] [doi: 10.3390/brainsci8060098] [Medline: 29857495]
17. Crowe J. Reform, revolution and disruption in mental health care: a consumer's perspective. Public Health Res Pract 2017 Apr 27;27(2) [FREE Full text] [doi: 10.17061/phrp2721711] [Medline: 28474048]
18. Piette JD, Krein SL, Striplin D, Marinec N, Kerns RD, Farris KB, et al. Patient-Centered Pain Care Using Artificial Intelligence and Mobile Health Tools: Protocol for a Randomized Study Funded by the US Department of Veterans Affairs Health Services Research and Development Program. JMIR Res Protoc 2016 Apr 07;5(2):e53 [FREE Full text] [doi: 10.2196/resprot.4995] [Medline: 27056770]
19. Spanakis EG, Santana S, Tsiknakis M, Marias K, Sakkalis V, Teixeira A, et al. Technology-Based Innovations to Foster Personalized Healthy Lifestyles and Well-Being: A Targeted Review. J Med Internet Res 2016 Jun 24;18(6):e128 [FREE Full text] [doi: 10.2196/jmir.4863] [Medline: 27342137]

20. Horgan D, Romao M, Morré SA, Kalra D. Artificial Intelligence: Power for Civilisation - and for Better Healthcare. Public Health Genomics 2019;22(5-6):145-161 [FREE Full text] [doi: 10.1159/000504785] [Medline: 31838476]
21. Atkins DC, Steyvers M, Imel ZE, Smyth P. Scaling up the evaluation of psychotherapy: evaluating motivational interviewing fidelity via statistical text classification. Implement Sci 2014 Apr 24;9:49 [FREE Full text] [doi: 10.1186/1748-5908-9-49] [Medline: 24758152]
22. Xiao B, Imel ZE, Georgiou PG, Atkins DC, Narayanan SS. "Rate My Therapist": Automated Detection of Empathy in Drug and Alcohol Counseling via Speech and Language Processing. PLoS One 2015;10(12):e0143055 [FREE Full text] [doi: 10.1371/journal.pone.0143055] [Medline: 26630392]
23. Imel ZE, Steyvers M, Atkins DC. Computational psychotherapy research: scaling up the evaluation of patient-provider interactions. Psychotherapy (Chic) 2015 Mar;52(1):19-30 [FREE Full text] [doi: 10.1037/a0036841] [Medline: 24866972]
24. Gyani A, Shafran R, Layard R, Clark DM. Enhancing recovery rates: lessons from year one of IAPT. Behav Res Ther 2013 Sep;51(9):597-606 [FREE Full text] [doi: 10.1016/j.brat.2013.06.004] [Medline: 23872702]
25. Clark DM, Fairburn CG, Wessely S. Psychological treatment outcomes in routine NHS services: a commentary on Stiles et al. (2007). Psychol Med 2008 May;38(5):629-634 [FREE Full text] [doi: 10.1017/S0033291707001869] [Medline: 17922943]
26. Macaskill ND. Improving clinical outcomes in REBT/CBT: The therapeutic uses of tape-recording. J Rational-Emot Cognitive-Behav Ther 1996 Sep;14(3):199-207. [doi: 10.1007/bf02238271]
27. Barnett JE. Utilizing technological innovations to enhance psychotherapy supervision, training, and outcomes. Psychotherapy (Chic) 2011 Jun;48(2):103-108. [doi: 10.1037/a0023381] [Medline: 21639653]
28. Haggerty G, Hilsenroth M. The use of video in psychotherapy supervision. British Journal of Psychotherapy. 2011/05/ 2011;27(2):01-210. [doi: 10.1111/j.1752-0118.2011.01232.x]
29. Shafranske E, Falender C. Clinical supervision. APA handbook of clinical psychologyducation and profession, Vol 5. American Psychological Association. APA handbooks in psychology® 2016:175-196. [doi: 10.1037/14774-012]
30. Shepherd L, Salkovskis PM, Morris M. Recording therapy sessions: an evaluation of patient and therapist reported behaviours, attitudes and preferences. Behav Cogn Psychother 2009 Mar;37(2):141-150. [doi: 10.1017/S1352465809005190] [Medline: 19364415]
31. Briggie A, Hilsenroth M, Conway F, Muran J, Jackson J. Patient comfort with audio or video recording of their psychotherapy sessions: Relation to symptomatology, treatment refusal, duration, and outcome. Professional Psychology: Research and Practice 2016;47(1):66-76. [doi: 10.1037/a0040063]
32. Singla DR, Hollon SD. The importance of implementation science in scaling up psychological treatments. Behav Res Ther 2020 Jul;130:103614. [doi: 10.1016/j.brat.2020.103614] [Medline: 32299610]
33. Kalmady SV, Greiner R, Agrawal R, Shivakumar V, Narayanaswamy JC, Brown MRG, et al. Towards artificial intelligence in mental health by improving schizophrenia prediction with multiple brain parcellation ensemble-learning. NPJ Schizophr 2019 Jan 18;5(1):2 [FREE Full text] [doi: 10.1038/s41537-018-0070-8] [Medline: 30659193]
34. Hill CE, Thompson BJ, Ladany N. Therapist use of silence in therapy: a survey. J Clin Psychol 2003 Apr;59(4):513-524. [doi: 10.1002/jclp.10155] [Medline: 12652641]
35. Sexton H, Littauer H, Sexton A, Tømmerås E. Building an alliance: Early therapy process and the client–therapist connection. Psychotherapy Research 2005 Jan;15(1-2):103-116. [doi: 10.1080/10503300512331327083]
36. Hollon S. Developing clinical practice guidelines to enhance clinical decision making. In: Magnavita JJ, editor. Clinical decision making in mental health practice. Arlington, VA: American Psychological Association; 2016:125-146.
37. McCabe R, Howes C, Purver M. Linguistic indicators of severity and profess in online text-based therapy for depression. 2014 Jun Presented at: Proceedings of the Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality; 2014; Baltimore, Maryland, USA p. 7-16 URL: https://www.aclweb.org/anthology/W14-3202 [doi: 10.3115/v1/w14-3202]
38. Funk B, Sadeh-Sharvit S, Fitzsimmons-Craft EE, Trockel MT, Monterubio GE, Goel NJ, et al. A Framework for Applying Natural Language Processing in Digital Health Interventions. J Med Internet Res 2020 Feb 19;22(2):e13855 [FREE Full text] [doi: 10.2196/13855] [Medline: 32130118]
39. Plateau C, Brookes F, Pugh M. Guided Recovery: An Interpretative Phenomenological Analysis of Service Users' Experiences of Guided Self-Help for Bulimic and Binge Eating Disorders. Cognitive and Behavioral Practice 2018 May;25(2):310-318 [FREE Full text] [doi: 10.1016/j.cbpra.2017.08.004]
40. Garland AF, Kruse M, Aarons GA. Clinicians and outcome measurement: what's the use? J Behav Health Serv Res 2003;30(4):393-405. [doi: 10.1007/BF02287427] [Medline: 14593663]
41. Fernandes J, Brunton I, Strudwick G, Banik S, Strauss J. Physician experience with speech recognition software in psychiatry: usage and perspective. BMC Res Notes 2018 Oct 01;11(1):1-5 [FREE Full text] [doi: 10.1186/s13104-018-3790-y] [Medline: 30285818]
42. Hollon S, DeRubeis R. Outcome studies in cognitive therapy. In: Leahy RL, editor. Science and Practice in Cognitive Therapy: Foundations, Mechanisms, and Applications. New York, NY: Guilford Press; 2018:48-61.
43. Clark GI, Egan SJ. The Socratic Method in Cognitive Behavioural Therapy: A Narrative Review. Cognitive Therapy and Research 2015 Jul 25;39(6):863-879. [doi: 10.1007/s10608-015-9707-3]

44. Can D, Marín RA, Georgiou PG, Imel ZE, Atkins DC, Narayanan SS. "It sounds like...": A natural language processing approach to detecting counselor reflections in motivational interviewing. J Couns Psychol 2016 Apr;63(3):343-350 [FREE Full text] [doi: 10.1037/cou0000111] [Medline: 26784286]
45. Thwaites R, Bennett-Levy J. Conceptualizing Empathy in Cognitive Behaviour Therapy: Making the Implicit Explicit. Behav. Cognit. Psychother 2007 Jul 6;35(05). [doi: 10.1017/s1352465807003785]
46. Firth J, Torous J, Carney R, Newby J, Cosco TD, Christensen H, et al. Digital Technologies in the Treatment of Anxiety: Recent Innovations and Future Directions. Curr Psychiatry Rep 2018 May 19;20(6):44 [FREE Full text] [doi: 10.1007/s11920-018-0910-2] [Medline: 29779065]
47. Lattie EG, Nicholas J, Knapp AA, Skerl JJ, Kaiser SM, Mohr DC. Opportunities for and Tensions Surrounding the Use of Technology-Enabled Mental Health Services in Community Mental Health Care. Adm Policy Ment Health 2020 Jan;47(1):138-149. [doi: 10.1007/s10488-019-00979-2] [Medline: 31535235]
48. Delgadillo J, Gonzalez Salas Duhne P. Targeted prescription of cognitive-behavioral therapy versus person-centered counseling for depression using a machine learning approach. J Consult Clin Psychol 2020 Jan;88(1):14-24. [doi: 10.1037/ccp0000476] [Medline: 31841021]
49. Sadeh-Sharvit S, Fitzsimmons-Craft EE, Taylor CB, Yom-Tov E. Predicting eating disorders from Internet activity. Int J Eat Disord 2020 Sep;53(9):1526-1533. [doi: 10.1002/eat.23338] [Medline: 32706444]
50. Chien I, Enrique A, Palacios J, Regan T, Keegan D, Carter D, et al. A Machine Learning Approach to Understanding Patterns of Engagement With Internet-Delivered Mental Health Interventions. JAMA Netw Open 2020 Jul 01;3(7):e2010791 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.10791] [Medline: 32678450]
51. Colombo D, Suso-Ribera C, Fernandez-Álvarez J. Exploring Affect Recall Bias and the Impact of Mild Depressive Symptoms: An Ecological Momentary Study. In: Cipresso P, Serino S, Villani D, editors. Pervasive Computing Paradigms for Mental Health. Cham: Springer; 2019:A.
52. Jacobson NC, Chung YJ. Passive Sensing of Prediction of Moment-To-Moment Depressed Mood among Undergraduates with Clinical Levels of Depression Sample Using Smartphones. Sensors (Basel) 2020 Jun 24;20(12) [FREE Full text] [doi: 10.3390/s20123572] [Medline: 32599801]
53. Fuller-Tyszkiewicz M, Richardson B, Klein B, Skouteris H, Christensen H, Austin D, et al. A Mobile App-Based Intervention for Depression: End-User and Expert Usability Testing Study. JMIR Ment Health 2018 Aug 23;5(3):e54 [FREE Full text] [doi: 10.2196/mental.9445] [Medline: 30139722]

Abbreviations

AI: artificial intelligence
GAD-7: Generalized Anxiety Disorder-7
MBC: measurement-based care
ML: machine learning
NLP: natural language processing
PHQ-9: Patient Health Questionnaire-9

Edited by G Eysenbach; submitted 24.05.20; peer-reviewed by K McKay, C Kaylor-Hughes, A Baumel; comments to author 15.06.20; revised version received 24.07.20; accepted 28.10.20; published 26.11.20.

Please cite as:
Sadeh-Sharvit S, Hollon SD
Leveraging the Power of Nondisruptive Technologies to Optimize Mental Health Treatment: Case Study
JMIR Ment Health 2020;7(11):e20646
URL: http://mental.jmir.org/2020/11/e20646/
doi:10.2196/20646
PMID:33242025

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org