JMIR Mental Health

Contents

Original Paper

Patient and Therapist Expectations for a Blended Cognitive Behavioral Therapy Program for Depression: Qualitative Exploratory Study

Ece Atik[1], MSc; Magnus Schückes[2], MSc; Jennifer Apolinário-Hagen[3], PhD

[1]Ruhr University Bochum, Bochum, Germany

[2]Institute for SME Research and Entrepreneurship, University of Mannheim, Mannheim, Germany

[3]Institute of Occupational, Social and Environmental Medicine, Faculty of Medicine, Centre for Health and Society, Heinrich Heine University Dusseldorf, Dusseldorf, Germany

Corresponding Author:
Magnus Schückes, MSc
Institute for SME Research and Entrepreneurship
University of Mannheim
L 9, 1–2
Mannheim, 68161
Germany
Phone: 49 621 181 2273
Email: schueckes@bwl.uni-mannheim.de

Abstract

Background: Blended cognitive behavioral therapy (bCBT)—the combination of digital elements and face-to-face psychotherapy—has been proposed to alleviate challenges that patients and therapists face in conventional cognitive behavioral therapy. There is growing evidence that adding digital elements to face-to-face psychotherapy can contribute to better treatment outcomes. However, bCBT programs still show considerable shortcomings, and knowledge on how to improve digital apps using a bCBT protocol is limited.

Objective: This study aimed to inductively identify functions and qualities that are expected from a bCBT treatment for depression in the eyes of patients and psychotherapists who were not currently receiving or practicing bCBT treatment.

Methods: We used a qualitative exploratory study design and conducted 3 focus group interviews (n=6 in each) and 5 semistructured in-depth interviews with therapists as well as 11 individual interviews with patients with a primary diagnosis of depression and currently undergoing cognitive behavioral therapy treatment in Germany. Themes and categories were established inductively from transcribed interview records based on a rigorous coding method.

Results: Both therapists and patients expected a digital app to provide patients with the opportunity to track their mood, work on therapeutic homework activities, easily access an intervention set for harder moments, and efficiently facilitate administrative tasks. The desire to be able to customize bCBT protocols to individual patient circumstances was evident in both patient and therapist interviews. Patients differed with respect to what content and the amount of material the app should focus on as well as the method of recording experiences. Therapists viewed digital apps as potentially aiding in their documentation work outside of sessions. Different attitudes surfaced on the topic of data security, with patients not as concerned as therapists.

Conclusions: Both patients and therapists had substantially positive attitudes toward the option of an integrated bCBT treatment. Our study presents novel findings on the expectations and attitudes of patients and therapists.

(JMIR Ment Health 2022;9(12):e36806) doi:10.2196/36806

XSL•FO
RenderX

Introduction

Background

Depression is one of the most common mental health–related diagnoses worldwide [1]. It causes debilitating psychological pain, problems in one's social life, and loss of productivity to individuals who have it. Furthermore, the hopelessness that many patients with depression feel is linked to suicidal behavior [2]. For many individuals, depression has a recurring trajectory over the life span, and each recurrent episode increases the likelihood of development of later episodes [3]. Beyond the hardships that individuals go through, depression also imposes a considerable burden on health care systems worldwide as the psychological and pharmaceutical treatments for depression are costly [4].

Cognitive behavioral therapy (CBT) is an effective and well-researched psychotherapy approach with extensive evidence on its effectiveness in the treatment of a wide range of mental health disorders, including major depression and dysthymia [5-7]. Especially for the treatment of depression, CBT has been shown to provide superior treatment outcomes compared with alternative psychotherapy methods [8]. In CBT, a high degree of adherence to the therapy is key to a successful treatment process and long-term outcomes. Insufficient engagement with therapeutic resources (especially between sessions) and dropouts from psychotherapy because of poor engagement are potential hazards for patients that may impair its effectiveness [9]. Research shows that patients' adherence to assigned homework within their treatment program—especially behaviorally demanding assignments—is considerably low [10,11]. According to therapists, more than half of patients face difficulties in completing their homework as prescribed, which in some cases leads to considerable negative consequences for therapy outcomes [12]. In contrast, there is a wealth of evidence that homework compliance is a significant predictor of better treatment outcomes [13-16]. Given these circumstances, it is not surprising that relapse and recurrence rates are high in CBT treatments. For example, a meta-analysis addressing relapse and recurrence rates following CBT programs for depression reported that relapse and recurrence rates can be as high as 46.5% throughout the treatment and highly predicted by residual depressive symptoms after the completion of therapy [17].

To alleviate these problems and support patients in CBT adherence, scholars have proposed to leverage the advantages of internet-based CBT (iCBT) in conventional face-to-face CBT treatment plans [18]. iCBT treatment options consist of web-based platforms offered to patients for use through smartphones or computers that provide them with psychoeducational content and homework tasks digitally [19]. By integrating an iCBT app and face-to-face CBT treatment into the same treatment protocol, a blended CBT (bCBT) modality offers a new treatment method that aims to incorporate the merits of both [20]. Although it promotes patient autonomy and makes psychotherapy easily accessible outside the therapy sessions, bCBT still preserves personal contact with a therapist and recognizes the therapeutic relationship as an important factor in treatment success [21]. bCBT programs have been suggested to intensify the therapy process by making it easier to record a patient's activities and thoughts, review the content of the therapy between sessions, and organize homework. Organizing homework in particular leads patients to focus on completion and enhances learning as a result [22]. bCBT programs also promise to reduce the total therapy duration [23,24].

Although there is no consensus in the literature on a single taxonomy for bCBT programs [18,25,26], bCBT treatments have primarily been categorized as either *sequential* or *simultaneous* according to how the face-to-face and digital components of blended therapy are introduced and scheduled [18]. In sequential bCBT treatments, patients' individual work with a digital mental health app either precedes or follows the face-to-face treatment with a psychotherapist [18]. Stepped care approaches could exemplify sequential blended treatments [27,28]. Simultaneous bCBT programs offer face-to-face and digital elements within the same period [18], oftentimes in the form of working on therapeutic homework through the digital app between face-to-face sessions with a therapist [29,30]. Our focus in this study is on simultaneous bCBT treatment as we believe this approach has the greatest potential to efficiently achieve positive long-term therapy goals.

Several studies have confirmed the additional benefits of simultaneous bCBT compared with face-to-face therapy in alleviating depression symptoms and mental health–related burdens on patients [18,25,29,31-33], whereas there are also studies where patient mental health outcomes in bCBT conditions were found to be similar to usual treatment [21,34-36]. Important is that none of those studies found that bCBT resulted in worse mental health outcomes. Similarly, some studies find bCBT to be a more cost-effective alternative to standard CBT [24,34,37], whereas others find longer total treatment time and higher costs in bCBT conditions when the amount of therapist contact remains constant [21,38]. More research is needed to shed light on these various outcome findings. The bCBT programs in these studies face shortcomings as they are rarely developed as a unique synthesis of the digital app and face-to-face therapy but simply use them simultaneously as distinct or *adjunct* therapy elements [29,39]. As an alternative to adjunct programs, more personalized and interconnected digital and face-to-face elements would offer an *integrated* bCBT program [25,30].

Objectives

To overcome the obstacles associated with conventional CBT programs and leverage the advantages that bCBT programs promise, understanding the demands and requirements of an integrated bCBT program is essential. Accordingly, this study aimed to discover patients' and therapists' attitudes and expectations from digital mental health apps to develop better-integrated bCBT protocols. Existing knowledge on how to design an efficient digital mental health app for a bCBT protocol is heavily based on qualitative data collected during pilot studies of newly developed apps [30,40,41]. Unlike those studies, this research aimed to collect input from patients and therapists who were naive in bCBT programs, which means

that they were not currently receiving a bCBT treatment and were not introduced to any bCBT app for the goals of the research. Principally, we also assumed that they had no previous experience with a bCBT program as integration of digital health apps in the psychological health care system is new in Germany and there were no bCBT options at the time of data collection or earlier [42,43]. As a result, we were able to identify the needs and expectations of patients and therapists from a perspective of inexperience with bCBT programs. Such an approach would also allow us to discover if patients and therapists have different user needs in the digital working spaces within a mental health app. Developers of new apps in this field would thereby have a general perspective on the needs of users rather than user reflections guided (and directly influenced) by previous interactions with already extant platforms.

We designed this explorative study to this end. The following research questions were addressed: (1) What are therapists' and patients' attitudes toward integrating digital interventions into conventional face-to-face CBT? What are the expected benefits? What are their concerns? (2) Do patients and therapists have different needs or attitudes regarding digital apps? (3) How do these attitudes and needs (including differences) translate into features of a digital mental health or software app?

Methods

Setting

We used a qualitative research design for this study. Qualitative methods are especially suitable for studies based on explorative research as these methods enable deriving direct insights from this study without presuppositions from earlier research [44,45]. We conducted semistructured interviews with both CBT-oriented psychotherapists and patients who were in psychotherapeutic treatment at the time of the study to unravel the perspectives of both sides of a therapeutic relationship. We deliberately targeted a patient sample with depression as the advantages of bCBT over face-to-face therapy appear most evident in this patient population [25,29,33,34]. Patients and therapists were given an explanatory introduction to the concept of bCBT programs to ensure a common understanding of the target topic of inquiry before the interviews took place. In these interactions, we made clear that our focus was on simultaneous bCBT protocols that present digital apps within the same period as face-to-face treatment [18]. We communicated that our goal would ultimately be to develop a digital app based on and incorporating the participants' feedback.

Participants and Recruitment

Patients

We conducted one-to-one interviews with 11 patients who were at the time of the interviews undergoing an individual standard face-to-face psychotherapy program (ie, no group therapy) for the treatment of unipolar depression at practices associated with the training institute Academy of Behavioral Therapy in Cologne, Germany. In total, 82% (9/11) female patients and 18% (2/11) male patients (mean age 33.27, SD 11.78; range 22-57 years) participated in the study. We approached patients via email based on referrals from participating therapists from

the same training institute. Inclusion criteria for patients were being between the ages of 18 and 65 years, having been diagnosed with mild to moderate depression by their therapists and assessed as not currently suicidal, having sufficient German language skills, and owning a smartphone with the iOS or Android operating system with internet access. Patients were excluded if they were experiencing or had experienced psychotic symptoms or substance dependence in the past or were currently using a digital mental health app for the treatment of a mental health diagnosis. All participating patients gave informed consent to take part in this study and provide researchers with personal data.

Psychotherapists

We conducted 3 focus group interview sessions with a total of 18 therapists and 5 individual therapist interviews. Focus groups and individual interviews were combined to obtain practitioners' collective opinions within the context of interaction with their colleagues as well as individually, where topics were addressed with greater specificity. Overall, we had 23 psychotherapist participants in our study (n=19, 83% female and n=4, 17% male; mean age 31.91, SD 8.11; range 25-58 years). All therapists participating in the study were either certified practitioners (7/8, 88% female and 1/8, 12% male; mean age 39, SD 10.38; range 28-58 years) or psychotherapists in their last year of training (12/15, 80% female and 3/15, 20% male; mean age 28.13, SD 2.33; range 25-33 years), and all were practicing CBT. Except for 1 certified therapist in the focus groups interviews who had a PhD in clinical psychology, all therapists had a Master of Science. Participating therapists were recruited from multiple practices in Germany and via the training institute Academy of Behavioral Therapy, Cologne, Germany, and referrals made by this training institute. Potential participants were approached directly through email or telephone. Participating therapists gave informed consent to take part in this study and provide researchers with personal data.

For both samples, potential participants were approached directly through email or telephone. As we recruited participants through referrals, we did not record demographic information of those who declined participation in the study or did not respond to the invitation.

Ethics Approval

The study was approved by the Ethical Board of the University of Mannheim (EK Mannheim 38/2020). Interviews were conducted by coauthor MS, who is experienced as an interviewer and has conducted several qualitative studies in the past focused on digitalization and product development. There was no relationship between the interviewer and participants before this study. Interviews with participants were conducted between February 2021 and May 2021. Participants did not receive any monetary compensation for taking part.

Data Collection

Interviews With Patients

Semistructured interviews were used to explore patients' attitudes, expectations, and concerns regarding digital mental health apps. The interviews with patients were conducted on

the web via tele- and videoconferencing and lasted an average of 35 (SD 9) minutes. The interviews were audio recorded with the consent of the patients and transcribed within 24 hours of the completion of the interviews.

The following questions were the main points addressed and directed to patients: (1) Can you imagine a digital mental health app offering assistance to your ongoing psychotherapy? How would such an option help? Which functions would be helpful for you? How much time would you dedicate to working with such an app? (2) Which functions can you imagine within a digital mental health app? Which would you consider unhelpful or disturbing? (3) What would you think about having a digital mental health app recommended to you by your therapist? Would you feel forced to use it? (4) What would be your attitude toward sharing your therapy-related data with your therapist through the app? What would be your attitude if the app used your data to offer a better version of the program? (5) Do you have any concerns regarding receiving treatment using a bCBT program? What are they? Do you have concerns regarding data security?

Focus Groups and Subsequent Interviews With Psychotherapists

We first conducted focus groups to explore psychotherapists' attitudes, expectations, and concerns regarding digital mental health apps. During the focus groups, web-based and digital tools (ie, Miro web-based whiteboard) were used to document the discussion and results, and meeting notes were taken. Focus group interviews were conducted on the web via videoconferencing, moderated by MS, and lasted approximately 3 hours on average. An assistant was present to take notes and introduce the web-based tool (ie, Miro). The focus groups were not audio recorded (notes were taken instead). The following questions guided the focus group interviews: (1) What do you expect from a digital health app? What are the pros and cons? (2) What does a day in the life of a psychotherapist look like? How can a digital health app support you? (3) How are the usual treatments of depression designed? What is special about these patients? How can a digital health app support usual treatment for depression?

Subsequent individual therapist interviews were conducted to explore specific issues that came up during the focus group sessions in more detail. Individual interviews were conducted on the web via tele- and videoconferencing and lasted an average of 39 (SD 3) minutes. The interviews were audio recorded with the consent of the psychotherapists and transcribed within 24 hours of the completion of the interviews.

Data Analysis

The transcripts and meeting notes from the interviews were coded according to grounded theory, a systematic data analysis methodology that focuses on inductively developing abstract theoretical conceptions from empirical data [46,47]. For the coding of transcriptions, Dedoose software (version 8.0.35; SocioCultural Research Consultants) for qualitative research was used. Patient and therapist interviews were coded separately and each interview independently. First, patient interviews were coded. Then, therapist interviews were coded. As a result of qualitative analysis of patient interviews, 4 main categories (expected features of bCBT apps, expected qualities of bCBT apps, qualities to avoid in bCBT apps, and concerns regarding bCBT apps) and several subcategories within these main categories were formed by inductive category formation based on the content of the interviews and codes assigned to text passages within them. Similarly, based on therapist interviews, 5 main categories (day of a therapist, attitudes toward digital apps, expected features of bCBT apps, expected qualities of bCBT apps, and concerns regarding app integration into therapy) and several subcategories within these main categories were formed inductively. We iterated between the developing model and the data until a viable set of first-order codes, second-order themes, and aggregate dimensions was identified, stopping when we reached "theoretical saturation" [48]. A joint agreement was reached among all authors to stop recruitment when saturation was achieved. The coding procedure was conducted by multiple researchers to ensure intercoder reliability—EA completed the first round of coding the data, and MS reviewed the coding scheme. A second coding round was then performed by EA and MS. Subsequently, the coding scheme was reviewed again by EA, MS, and JAH, and some adaptations were discussed. Finally, JAH approved the categories. As there were only minor adaptations that emerged from cross-checking, 2 coding rounds were considered sufficient. The completed COREQ (Consolidated Criteria for Reporting Qualitative Research) [49] checklist can be found in Multimedia Appendix 1. The interview quotes included in this paper were translated by a professional translator from German into English.

Results

Patient Interviews

Overview

During the interviews, patients shared their conception of a digital app intended to deliver an integrated bCBT program. In general, patients had very positive attitudes toward bCBT programs. They made specific references to expected features and qualities they would like to see in the apps and those they would like to avoid. Participants also addressed their concerns regarding bCBT (ie, under which therapeutic circumstances or conditions they would not feel comfortable using a digital app). The categories, themes, and dimensions emerged from the coding of patient interviews are illustrated in Table 1.

Table 1. Results of the interviews with patients.

First-order categories	Second-order themes	Third-order aggregate dimensions
• Keep track of details and timing of bad days • Reminders on incomplete mood and symptom entries • Overview of mood and symptoms through time	Mood and symptom tracking	Expected features of bCBT[a] apps
• Communication, calendar, and appointment management • Easier management of questionnaires and diagnostics • Help structure patient everyday activities • Provide a timeline of activity history (for therapist and patient)	Organizational functions	Expected features of bCBT apps
• Provide psychoeducation • Engaging with therapy activities • Collection of emergency therapeutic tools	Therapeutic content and resources	Expected features of bCBT apps
• Need to regard the app as useful, relevant, and integrated • Pleasant and easy to understand • Option to use it from a computer • Different media to convey content (visual, text, and auditory)	Usability	Expected qualities of bCBT apps
• Personalized reminder schedule for different users • Flexible tools to express oneself	Individualized features	Expected qualities of bCBT apps
• Feeling of success through use • Marking certain activities as "done" and seeing personal history in the form of completion steps • Setting notifications when required or by choice	Gamification	Expected qualities of bCBT apps
• Engagement with the app should not replace a face-to-face session • Introduce app only with accompanying therapist supervision • App should not encourage frequent and spontaneous contact with the therapist • App should not create a potential for dependency	Priority and balance with face-to-face treatment	Expected qualities of bCBT apps
• Dispersed use throughout the day is preferred • Need to be able to work on the app in a quiet and personal space • A total weekly amount of 1-2 hours of app use is foreseen	Use time fitted to patients' needs	Expected qualities of bCBT apps
• People with disabilities and who speak minority languages should be able to use the app • Therapy app helps destigmatize psychotherapy	Inclusion	Expected qualities of bCBT apps
• Mood and symptom tracking helps more efficient and realistic communication of one's current situation to the therapist in sessions • Open-ended thought record can help patients come to the sessions better equipped (specific triggers and questions recorded in a timely fashion)	Transfer to in-person sessions	Expected qualities of bCBT apps
• Educating language and lifestyle tips • Heavy numerical demonstration of one's progress • Advertisements and user cookies	Feeling distant and technical complexity	Qualities to avoid in bCBT apps
• Too much content • Overwhelming sound, color, or text • Too many push notifications • Excessive focus on one's depressive thoughts and experiences	Input bombardment	Qualities to avoid in bCBT apps
• Presentation of therapy activities as "must-do" • Constantly aiming to advance and be happy	Pressuring patients	Qualities to avoid in bCBT apps
• Recommendation by therapist is acceptable • Alternative description to "prescription" wording should be used as it is too much for some patients • Therapist prescription is positive if therapist explains how and why the app helps	Framing of a bCBT plan	Concerns regarding bCBT apps

XSL•FO
RenderX

First-order categories	Second-order themes	Third-order aggregate dimensions
• App's use of data to provide better service is acceptable • Sharing relevant therapy data with therapist is meaningful and acceptable • Data security is not a major theme of concern	Sharing of therapy-related data	Concerns regarding bCBT apps

[a]bCBT: blended cognitive behavioral therapy.

Expected Features of bCBT Apps

Mood and Symptom Tracking

A common desire of patients was a function through which they could keep regular track of their mood and symptoms regarding their mental health issues. In addition to recording details such as specific times when moods would occur, interviews also revealed a common need for the option to review one's mood and symptom history whenever they either wanted or needed to look back on their treatment path:

> *Sometimes, for me, depressive episodes have been so often that I forget that I was doing fine two weeks ago. Getting reminded that there have been times or there are things which can make me believe I can cope with this would be helpful.* [Patient 4]

Patients noted that, to achieve regular mood and symptom tracking, the digital app could ask questions that are delivered to their smartphone regularly via push notifications. Some patients also noted that push notifications could indicate incomplete entries, thus serving as an additional reminder mechanism to increase program adherence. In addition to tracking their mood, patients also indicated a desire to record more detailed notes and how they coped with the situation. Hence, we would infer that adding such notes provides a richer and qualitative character to the mood-recording function.

Organizational Functions

Patients identified numerous possible app functions, such as a calendar, appointment management, communication with the therapist via a messaging platform, and practical help in dealing with questionnaires used for diagnostics:

> *Adding functions such as making appointments, canceling appointments, and rescheduling appointments would be helpful. And this also touches on the topic of achieving goals, in the sense of, "Where do I actually stand now in my therapy?" and "How many sessions does my health insurance still cover?"* [Patient 8]

Interviewees envisioned a digital app that offered additional help to organize their daily life, such as a personal checklist of activities and options to schedule (and receive reminders for) selected activities:

> *If it (app) says you've had a glass of water, check; you've brushed your teeth today, check. You've managed to walk 10 minutes today, check. Then those would be things for me which are somehow feasible, which help me in a depressive phase to take care of myself, but which don't put pressure on me that I have to do them.* [Patient 3]

Apart from assistance in organizing one's daily life, an app could also broadly support a patient's psychotherapy progress by recording important tasks and activities when they were completed and providing an accurate timeline that could be reviewed:

> *It would be cool if one has something like a book, I mean of course digital, where it (the app) says, yes, you've done this or that. Where you can kind of enter milestones or something, because I don't even remember a lot of things.* [Patient 2]

A patient highlighted that such a timeline would benefit not only patients but also their therapists in overviewing patient history. In this way, functions within the app (such as a digital journal) could offer another channel of communication between therapists and patients that takes place outside of sessions.

Therapeutic Content and Resources

Patients expressed their willingness to use a digital app to maintain the connection with their therapist via therapy-related material outside of regularly scheduled meetings:

> *I would find that perfect if it (the app) would provide clarification about how psychotherapy works and which specific form of therapy is the one that I have been working out with my therapist.* [Patient 8]

A common point addressed by several patients concerned "emergency" situations—specific instances in daily life where the patients felt a notable sudden deterioration in their feelings, symptoms, or thought patterns. Patients drew attention to the difficulty of remembering possible intervention techniques to make themselves feel better in such moments. A collection of a toolkit for harder moments offered through the digital app was seen as a very useful support function during such episodes:

> *When you're suddenly in a situation where you'd like to see the therapist right away but she's not there...having some tools in the app that can calm you down and keep yourself under control for that moment [would be helpful].* [Patient 10]

As a result, app developers should consider including options where access to such interventions is at the topmost layer such that patients can easily access such information as needed in a crisis.

Expected Qualities of bCBT Apps

Usability

Patients expressed a common need to see the digital app as relevant to their treatment goals and easy to integrate into their daily routine:

XSL•FO

RenderX

I would use it frequently if I can somehow integrate it into my everyday life. If I have the need to look into it in the morning and if it becomes part of my everyday life, like part of my routine, then I would use it frequently. [Patient 6]

Patients would like an app that is easy to understand and looks esthetically pleasing. The app should contain various kinds of media (text, pictures, graphics, or audio) to present content to create a pleasant user experience:

I'm a purely visual person. I totally struggle when something contains only textual processing. I'm a picture person. So if it's intuitive and presented with pictures, I find it easier to use. [Patient 5]

A specific reference to the option of using the app in a desktop format was made by 18% (2/11) of the participants (because of the possibility of typing faster on a computer).

Individualized Features

The need for an individualized program was the most commonly raised issue in interviews among all patients. Interviewees reported that they would like to engage with a digital app that can be tailored to the needs of individual users:

It would definitely be good to be able to set it up in such a way that you can have it individualized, that it is related to your therapy progress, that you get a list of what's coming up or what you need to do at a certain time, depending on how you feel or how receptive you are. [Patient 1]

Patients also advocated for flexibility in how they expressed themselves. They recommended open-ended questions or customizable settings instead of only multiple-choice questions or standard options such as emojis to record their mood and feelings:

How many smileys would I have to choose from, six, seven, eight? What if the worst smiley was a crying smiley but I don't cry on a bad depression day. For me a day where I cry is when I can admit to myself that things aren't going the way I want them to but when I'm much more open about it and also communicate with people about it. That means a crying smiley would not be the definition of disaster for me. [Patient 4]

Gamification

The need to include gamification elements in a digital app emerged in the interviews. To maintain continuous use of the app over a longer time—which is very relevant for psychotherapy—patients wanted to experience a feeling of success. Such a moment could be ticking some items on a checklist as "done," with goals broken down into smaller components to be accomplished in a certain order. As a patient put it, "I'm just super gamified in my consumption, and I also just jump at reward systems...I think it's very pleasant in the design when you see that you've done something when you complete something" (Patient 3).

Priority and Balance With Face-to-face Treatment

Every patient expressed the need to see the digital app as well integrated with face-to-face psychotherapy treatment. They highlighted that good integration and engagement with the app should never lead to replacing the face-to-face sessions with a therapist. They were unanimous that such an app should not be offered without accompanying therapist supervision:

I really think the interpersonal contact with the therapist is the key to success in the whole story...The combination of real sessions and the app is totally important because otherwise it's basically a lifestyle product for me. It's not a clinical aid if I had to use it all by myself. I think I just wouldn't use it then. [Patient 5]

At the same time, a design that enables frequent spontaneous contact with the therapist was seen as something to be avoided. A patient stated the following:

I think direct contact with therapists on the app is a bit too much to ask of the therapist. I would rather not do that. [Patient 7]

Digital apps should be developed in a way that considers the risk to therapists regarding the loss of personal boundaries.

Use Time Fitted to Patient Needs

A total weekly amount of approximately 1 to 2 hours spent on a digital app was considered the right amount of time. Most of the participants (9/11, 82%) preferred a dispersed app use time throughout the day over a single continuous interaction. Those who preferred dispersed use patterns noted the possibility of taking advantage of the time spent on the bus or train. However, patients differed in their preferences in this respect, with a patient highlighting their need to work on the app in a quiet and personal space:

I couldn't work on that in a noisy environment. I would really have to have my peace of mind somehow to be able to carry out such exercises. [Patient 9]

Some patients noted the need for the app to accommodate potential variability in use patterns across time and that there should not be the expectation that patients sustain the same amount and frequency of use of the app at all times:

Especially at the beginning [of treatment], I can imagine that in the first few weeks you kind of spend significantly more time...When it later shifts to exercises and making appointments and reading, it could be less and then become more again. [Patient 8]

Inclusion

The chance to contribute to ending the stigma regarding mental health issues also emerged as an expected benefit. Some patients considered the addition of a digital app to one's daily life as a good opportunity to also integrate organic conversations on receiving psychotherapy treatment in discussions with their loved ones:

If such an app leads mental illness to become a bit more socially acceptable so that people can also talk

about it—that there are offers of help and also modern means of communication between therapist and patient—I think that would be a good thing. [Patient 5]

Another expectation with respect to inclusion concerned the importance of developing a digital app suitable for people with disabilities, such as visual or auditory impairments, as well as for people who speak different native languages:

It would be so supportive to have functions for visually and hearing-impaired people, and also linguistically. In Germany in particular, making the app available in Arabic or Turkish would be great. These things can really help to promote communication. [Patient 3]

Transfer to In-Person Sessions

There emerged an expectation in patients that having an integrated digital mental health app should also positively translate to face-to-face sessions. Easy access to mood and symptom-tracking functions would enable patients to talk about their symptoms more realistically. Noting which events or thoughts triggered specific moods or symptoms as well as questions that occurred to them would help patients come to sessions more aware and, therefore, use the time more efficiently:

You would then (with the app) actually be able to use the therapy hours more efficiently because reviewing your timeline (in the app) would already have gotten a bit involved in your thoughts about your progress. [Patient 5]

Qualities to Avoid in bCBT Apps

Feeling Distant and Technical Complexity

Patients converged on the point that they did not want to engage with a digital app that used an overly educational or distant voice with a "know-it-all" tone. Stereotypical advice such as simple lifestyle tips were identified as potential examples of this:

If you always have something appearing on your screen saying "go out in the sun" and "breathe in some fresh air" or such—I would drop something like that. [Patient 2]

Content that is difficult to understand by lay audiences, the overuse of numbers, and demonstrating patient progress through the excessive use of statistics or visuals would be too demanding for patients. Even though patients acknowledge that graphics and numerically charting progress could potentially be very helpful, they expressed concern about such materials not being easy to understand and use by the audience. Patients also reported working with a program that did not display advertisements or use cookies to be desirable.

Input Bombardment

Most patients (7/11, 64%) expressed concerns about being overwhelmed by a digital app. Too much content; design aspects involving too much sound, color, or text; and frequent notifications sent to the patient's phone were some examples of excessive input that they noted. Some interviewees underlined that frequent questions on mood and symptoms might lead patients to focus excessively on their depressive feelings and experiences and be drawn into a depressive spiral. This risk should also be kept in mind when deciding how frequently a digital app asks patients to reflect on their depression.

Pressuring Patients

Therapy activities involving the digital app should not be presented as "must" assignments (as such framing would be either harmful or counterproductive) but rather as suggestions. This approach would fit better with the patients' vision of the digital app filling the role of a "supportive companion":

In my case, in a burnout case, it would have been counterproductive if it (the app) had told me, "but now you have to do at least five exercises until the next session."...If I said that voluntarily to myself, it's okay, but if it had been associated with pressure like that, it would not be good. I would probably have dropped out. [Patient 5]

Patients emphasized that it would be a substantial drawback to using a digital app if it constantly pressured them to be happy. Questions prompting patients to record their moods and overview history should not give the impression that the patient can only record good feelings as allowing for negativity and embracing it is an essential part of the therapy:

When depressed it might be nice sometimes to see, oh I've actually got out of bed three times now. It's great that I get this push notification telling me that. But I believe that it could also be very brutal if things are conveyed too positively. I am not a therapist but I know that my therapist has often told me that I get up too often and don't allow myself to say I'm sick today. [Patient 4]

Concerns Regarding bCBT Apps

Framing of a bCBT Plan

In the interviews, we asked patients what they would think about having a bCBT program recommended by their psychotherapist. For most patients (6/11, 55%), such a recommendation would be viewed positively, both because of natural curiosity regarding a new therapeutic medium and trust in their therapist. A few patients (5/11, 45%) balked at the idea of their therapist "prescribing" a bCBT program, considering it unhelpful and "too much" in the context of psychotherapy:

Prescriptions always have something brutal about them. I mean, if the therapist were to say to me, "The app that we both will have is really helpful and we need to use it." That would make it sound a thousand times better to me than saying, "I'll prescribe it for you." Prescription always has something very, very stringent, like it is unavoidable, in my opinion. I would have a problem with my therapist saying that I have to use this app. [Patient 8]

For those who were concerned about the idea of having a digital program prescribed to them, simply having their therapist explain to them why they needed to use it and how it could help

was sufficient to resolve their resistance. Patients' doubts were instead centered on how the "prescription" was presented to them.

Sharing of Therapy-Related Data

The possibility of sharing data with their therapist through the app and with app developers to enhance programs were specific issues that interviewers posed explicit questions about to patients. Surprisingly, the fact that patients would have to share a lot of personal data with a digital app was not a major point of concern for them. Hence, data security concerns appeared to be an unlikely reason for the target audience to walk away from a digital app. Overall, patients were very willing to share personal data this way, with an offhand remark by a patient—"If the app can do that, why not?" (Patient 1)—being representative of the attitudes of the patients as a whole. It is noteworthy that sharing data with the app providers was an acceptable option only if their data were anonymous, secured, and not used for any other purposes.

Therapist Interviews

Overview

During the interviews, therapists' daily routines and their organization and preparation workload outside of sessions were first addressed. We explored their usual work habits to better understand the issues that may benefit from additional support from a bCBT plan. Interviewed therapists were open to and curious about digital smartphone app use in the context of psychotherapy. They identified potential beneficial app features and qualities that could enhance the quality of their workday. Participants also voiced their concerns regarding bCBT (ie, points that should be critically and carefully evaluated before deciding to work with a bCBT plan). The categories, themes, and dimensions emerged from the coding of therapist interviews are illustrated in Table 2.

Table 2. Results of the interviews with therapists.

First-order categories	Second-order themes	Third-order aggregate dimensions
• Planning sessions and becoming familiar with patient history • Note taking after the session • Analog versus digital organization of therapy documents	Nature of organizational work	Day of a therapist
• Appointment management is stressful • Therapy worksheets get lost or are difficult to organize • Insufficient current digital organization tools • Paper-based documentation makes planning difficult (especially outside the office) • Session planning and after-session notes are sometimes lengthy • Guidebooks for planning sessions are too rigid and difficult to adapt to individual patients	Problems that emerge during the workday	Day of a therapist
• Practicality • Open and curious toward apps	Attitudes toward app use	Attitudes toward digital apps
• Positive attitudes toward app use in psychotherapy • Therapy app could potentially enhance therapeutic relationship • Prescription of therapy app fosters adoption • Unsatisfactory current therapy apps	Attitudes toward app use in psychotherapy	Attitudes toward digital apps
• Documentation and writing session protocols • Support session preparation via patient history and therapeutic resources • Appointment management • Monitoring of patient progress	Benefits of bCBT[a] apps for therapists	Expected features of bCBT apps
• Mood and symptom tracking • Collecting therapeutic toolkit for harder moments • Better questionnaire and diagnostic management • Offer and remind regarding therapy resources and interventions • Encourage independence of patients	Benefits of bCBT apps for patients	Expected features of bCBT apps
• Ease the organizational workload and protect in-session time • Contribute to the connection between patient and therapist • Extra help in structuring therapy • Visualization of achieved milestones and successes	General benefits of bCBT apps	Expected features of bCBT apps
• Personalization of intervention content • Dynamic versus monotonous experience to support changes in use over time • Easy and intuitive to use • No advertisements	High usability	Expected qualities of bCBT apps
• App should not distort therapist work-life balance • App should not replace in-person sessions but complement them • App should preserve seriousness of therapy	Concerns regarding therapeutic qualities of a bCBT app	Concerns regarding app integration into therapy
• Data security • Wasting time with unhelpful or addictive features	General concerns regarding smartphone apps	Concerns regarding app integration into therapy

[a]bCBT: blended cognitive behavioral therapy.

Day of a Therapist

Nature of Organizational Work

The work outside of sessions with patients consists mainly of preparation for the upcoming sessions and taking notes about the session after patients leave. Therapists also have to note and manage appointment schedules and share documents with the patient's insurance company.

For the organization and collection of therapy-related files (session protocols, questionnaire results of patients, homework papers, and insurance papers), some therapists used a paper-based file collection, whereas others used digital software that is designed to organize this work. Therapists similarly differed in their preference for use of paper-and-pen calendars versus digital systems for appointment management. Interestingly, many therapists noted that they used their existing personal calendars for this purpose and not a distinct work calendar.

Problems That Emerge During the Workday

A common problem for both therapists and patients was that homework sheets were difficult to manage in their paper-based format. A therapist spoke for many when she shared that "these worksheets somehow get lost or the patients don't bring them to session. And that means they can't do them" (Therapist 4).

A similar problem with paper-based organization systems is the case therapy documentation that therapists have to manage. Documents are typically stored in huge folders in the office such that it is difficult to manage them and they are impossible to work with outside the office:

> *I find it really annoying if it's only on paper because then I can't look at them again at home. I'm not in the office every day and sometimes I think of the patients and want to check something, but then I would have to go to the practice first to check or ask a colleague to look at the file.* [Therapist 5]

However, those therapists who used digital software for documentation purposes were also rarely satisfied with the programs they had been using, with a therapist describing theirs as a "mechanical and dull program and also not that easy to work with...There are tens of spaces to click and hundreds of rows in front of you" (a therapist in the focus group discussion).

Finally, therapists noted that session planning before and note taking after sessions sometimes takes longer than planned, and the nondigital tools used to facilitate them (eg, guidebooks for planning sessions) are not helpful as they are not flexible (eg, session guidebooks follow a rigid suggestion format that is difficult to adapt to individual patients).

Attitudes Toward Digital Apps

Attitudes Toward App Use

The therapists interviewed were open and curious about digital smartphone apps in general and thought that digitalization brings about new and creative practical solutions to everyday matters. In this respect, they did not view psychotherapy as an exception and saw using a bCBT program in mental health treatment as a viable option that both therapists and patients would benefit from.

Attitudes Toward App Use in Psychotherapy

Therapists had diverging opinions about whether a digital ingredient within the therapeutic relationship would potentially enhance the relationship between the therapist and the patient. Some therapists stated that a digital app would contribute to improved therapeutic relationships, whereas others thought that it would not. One noted that a digital app could give patients a sense of continuing care and attention from their therapists outside of sessions:

> *I think it's quite nice for some patients if they think to themselves: Oh, my therapist has already looked in the app about what I've been doing this week. So I think that can be beneficial for their relationship.* [Therapist 1]

Others expressed skepticism regarding the positive impact of a digital app on a therapeutic relationship built on face-to-face interactions within the context of a therapy session:

> *I don't exactly know how an app is supposed to bond me and my patient. My experience is that the relationship simply grows from face-to-face contact and from the feelings that you feel for each other, and not from a worksheet or a digital platform.* [Therapist 3]

Importantly, no interviewed therapist believed that a digital app would damage the therapeutic relationship.

Therapists responded positively to the idea of prescribing a bCBT program, noting that it "does more justice" to the reality of the current generation of patients. They thought that a prescription could foster adoption by patients who already invest "two to three hours a day" in screen time and would prescribe a bCBT plan if they were convinced that the digital app supported the face-to-face therapy process "not only to deal with the therapy in the session but to accelerate this everyday transfer to promote and make it easier" (Therapist 3).

None of the interviewed therapists had suggested a psychotherapy-related app to a patient before, let alone prescribed one. The apps that the therapists were familiar with were deemed not satisfactory enough to recommend their use by patients.

Expected Features of bCBT Apps

Benefits of bCBT Apps for Therapists

Therapists remarked that digital documentation of many different kinds of therapy-related material, whether therapy activities, information on patients, or session protocols, would be helpful. A therapist specifically noted that having their session protocols digitally available would allow them to "search for keywords if you know that a topic has come up before or something" (Therapist 3). Better documentation functionality would also lead to easier planning for sessions, where therapists usually need to look at their notes and session protocols from previous meetings.

Another potentially useful contribution of a digital app would be providing therapeutic information that therapists could consult when needed:

> *So if I am focused on a topic, then I read up in my books, or I would look again in my personal notes. I google general things, for example grief counseling or things like that. Then I write down things that I find interesting and take them with me to the session. So yeah, I would use the app if I could access that information through it.* [Therapist 4]

Moreover, therapists expressed the desire to see basic information about a patient and where they stand in their therapy progress in a brief and easy-to-understand visual format:

> *It would be really helpful to have such an overview in which relevant diagnostics and symptom history for the patient can also be seen. And their status, information on the next session and perhaps also*

something like, "where does the patient stand right now in the therapy?" [Therapist 2]

Finally, many therapists (3/5, 60%) cited setting and canceling appointments as well as seeing their scheduled appointments in calendar form as helpful functions in an app.

Benefits of bCBT Apps for Patients

Within the context of CBT treatment, receiving help from a digital app by working with therapy resources and interventions would be one of the main possible benefits. Therapists stated that the app could offer intervention activities and therapeutic content as well as reminders for patients. Although it would support patients' between-session progress, the app would also ease administrative aspects of the work therapists do:

> *I have the feeling that maybe the homework would be done more often [with the app] and I wouldn't always have to print everything out. Sometimes there's information that I don't feel like explaining in the session; maybe depending on the app patients could read about it.* [Therapist 5]

Many other functions that therapists wanted were also in line with patients' expectations. Similar to patients, therapists named mood and symptom tracking, easier management and overview of diagnostics questionnaires, and collection of a toolkit that includes specific therapeutic interventions to use in harder moments as potentially helpful for patients. The ability to recall past data was prominently cited by many therapists:

> *If the app saves the data, you could also look again specifically at how it was three months ago because I've often had the experience that patients can't remember.* [Therapist 5]

An opinion shared by a few participants (2/5, 40%) was that the opportunity to engage with therapy material more intensively during a patient's personal time through the app would encourage the independence of patients:

> *It could contribute to the agency of my patients, where I wouldn't be the sole communicator of therapeutic knowledge and advice and so on. When they read some of the stuff on the app, and do homework more often than before with the app, it could give a sense of increased control and autonomy.* [A therapist in the focus group discussion]

General Benefits of bCBT Apps

Therapists thought that both they and their patients would take advantage of functions that ease the organizational workload and, therefore, preserve limited in-session time for more efficient use. Both parties would benefit from routinely available extra help to structure the therapy. Similar to patients, therapists also brought up the potential for a sense of continued togetherness between the therapist and patient and an enhanced therapeutic relationship because of a bCBT app. In addition, the potential to visualize achieved milestones and successes would make them easier to understand and aim for by patients and therapists.

Expected Qualities of bCBT Apps

Unsurprisingly, therapists were interested in a smooth, easy, and intuitive experience for both themselves and their patients when using the digital app. Therapists particularly noted that, as a course of treatment for many patients typically takes months, the digital app should offer a dynamic and engaging experience instead of a potentially monotonous one. The app needs to support patients as their functions change over the course of treatment through effective notification strategies. Personalization of intervention content for each patient would also lead to patients following the bCBT protocol over a longer time. The therapists stressed that the app should not incorporate advertisements (especially in the context of a prescription app).

Concerns Regarding App Integration Into Therapy

Concerns Regarding the Therapeutic Qualities of a bCBT App

Some of the therapists interviewed (2/5, 40%) were concerned that introducing a digital medium to the psychotherapy work might give the impression of appearing overly available to patients. They especially pointed to the potential ability of patients to send them messages and commented that such a function would distort their work-life balance. A therapist suggested that the app could keep a record of such queries and that the therapist could access them to "get an idea of what has preoccupied the patient" without feeling the obligation to respond.

Another major concern was defining the role of a digital app within the therapeutic relationship. A digital app should not change the frame of in-person sessions so that its role replaces a session. It should be clear to patients that the digital program has a complementary function supporting sessions with the therapist. A therapist mentioned the concern that the digital app might potentially look too game-like, which could be counterproductive for therapy purposes. Although it should offer a user-friendly interface, it should also preserve the seriousness of therapy.

General Concerns Regarding Smartphone Apps

Unlike patients in our interviews, all therapists were concerned with the fact that an app they would prescribe to their patients should ensure patients' data security. The difference between the responses of patients and therapists in this respect might be tied to the responsibility that therapists feel in the context of suggesting or even prescribing the app:

> *I think it's really important for the patients that I can tell them credibly it's safe and that the data will not be misused for any other purposes—that it won't be the case that insurance will not pay for treatment anymore depending on what they write on there or their insurance premiums will go up.* [Therapist 3]

Therapists highlighted that it is also crucial to ensure that, for a digital mental health app, the known danger of addiction when using digital apps (such as social media) is not present. Patients should not waste their time with unhelpful functions or (even worse) become dependent on the app in ways that contribute to the mental health issues the patient is grappling with.

Discussion

Principal Findings

Our study explored patients' and psychotherapists' attitudes toward the introduction of a digital mental health app into their ongoing psychotherapy process, including their expected features and concerns regarding bCBT protocols. We gained valuable insights into the perspectives of both potential user groups. Both patients and therapists had positive attitudes toward bCBT and digital health apps in general. The results provide not only a rich frame of reference regarding what functions are expected from an integrated digital app but also a guide on what to avoid.

We suggest that this qualitative study is an important contribution to the relevant literature as scientific inquiry into how to construct a digital app for bCBT protocols has not been extensive enough to create a solid framework [26]. Current knowledge is mostly based on interview data with patients or therapists regarding their personal experiences with existing bCBT programs combined with initial evaluations of programs that have been developed [30,40,41]. Our study alternatively offers fruitful data on patient and therapist experiences and expectancies that were not guided by any particular digital program. As patients and therapists in our study were not actively using a bCBT program and were assumed to lack experience with such programs given that these were not available earlier [43] (and no sample intervention was provided), they were not oriented toward a specific existing digital health program while voicing their views on potentially helpful and unhelpful digital program features. In conducting such an exploration without focusing on a particular digital health app, our study resonates with the work by van der Vaart et al [50], although their study used survey questions and qualitative data in parallel. A similar methodological approach to ours was used in the qualitative investigation by Cerga-Pashoja et al [51], but they conducted their exploration only with therapists. With this study, those who wish to develop integrated blended solutions for existing problems in psychotherapy can find the voices of both patients and therapists regarding potential problems and practical solutions. To our knowledge, the explicit distinction between *adjunct* and *integrated* digital app designs in bCBT programs has not been made earlier in the literature. On the basis of studies on existing digital apps calling themselves as such [29,30], we pointed out this distinction and showed in this study that an integrated bCBT design is more desirable to potential users.

In this study, patients and therapists mostly agreed on which functions would be expected in a digital app, and the differences were few but notable. Data security proved to be one of the fundamental issues that therapists needed assurance about, whereas patients were not as concerned. Conversely, patients saw potential negative connotations attached to the prescription of a digital mental health program, whereas therapists did not. These findings inform researchers and developers working in the field of digital mental health apps on how different functional aspects were viewed positively or negatively and

underline how psychotherapy is experienced differently by patients and therapists.

Our findings are in line with the existing consensus in the bCBT literature on the need for customizability of digital apps to an individual patient's needs in terms of both content and use [26,41,50]. Both therapists and patients view the rigidity and nonexclusiveness of existing digital tools as the reasons for their current limited use [52]. Our study supports these findings. Similarly, references to a desire for high usability in a digital program were made by both therapists and patients, supporting an earlier study with therapists pointing to "intuitive usability and logical structure of online platform" as a facilitator of bCBT use [41]. Attempts to develop new digital mental health programs or modify existing ones should address these concerns.

Aside from simply replicating the findings on the need for personalization, we argued that there could be 2 different interpretations of what personalization of a digital program means. According to one interpretation of the data, different patients should not be given identical psychological intervention content or feedback, which is consistent with previous research [26,41,50,53,54]. A second interpretation suggests that some patients view any restricted means for personal expression as a feature to avoid (as in the example of smiley faces or multiple-choice questions). Patients reported that these methods were inadequate to their communication needs. To our knowledge, this call for more open-ended and expressive means of recording one's mood and symptom history as well as personal thoughts is a new interpretation of customization in the digital mental health domain.

Previous findings on the needs of patients and therapists within the context of bCBT programs agreed that digital apps should avoid overwhelming patients [40,51]. For example, the volume of homework exercises delivered through the digital app [40] and the reminders and notifications provided [51] should not be excessive. Our findings endorse these earlier studies. Similarly, our findings supported the findings of van der Vaart et al [50] that therapists are concerned about preserving boundaries and keeping in-person therapy sessions central. Our study suggests that face-to-face and web-based elements should focus on distinct goals within a bCBT protocol, with the latter concentrating on the practical aspects of the therapy (such as homework assignments, diaries, and psychoeducation), whereas process-related matters and evaluation of the patient's condition are reserved for face-to-face meetings with the therapist [50]. Therapists had differing views on the potential effect of digital mental health apps on the therapeutic relationship between a patient and a therapist, with some suggesting that treatment through bCBT protocols may foster this relationship, whereas others expressed skepticism. Remarkably, none of the therapists thought that digital mental health apps would hinder the relationship between patient and therapist. These findings are important considering that the therapeutic relationship is known to be one of the strongest predictors of treatment success [55-57].

Receiving extra assistance in a bCBT treatment was put forth as an important potential benefit by both patients and therapists in our interviews. Given that both patients and therapists are

eager to try a digital app to promote their therapy progress, digital apps appear to offer a practical solution to widespread adherence issues regarding patients completing assigned homework [10,12].

Limitations

As our patient sample included only those who were diagnosed with depression and was small (n=11), the needs and expectations of this patient group might not be generalizable to all patients in psychotherapy. Patients with different mental health diagnoses might have different needs and express them differently than the population we studied. Moreover, our study included a selective sample of adult patients who were already undergoing CBT-oriented therapy and were seemingly motivated and held positive attitudes regarding bCBT programs. Patients who are not yet in treatment for their symptoms, such as those on waiting lists, or those who are not receiving a first-line treatment such as CBT might hold different views from those of our sample. Similarly, we cannot make claims about the potential of bCBT programs in children and teenagers with our findings as we deliberately targeted only adult patients. Further research is needed to understand if and in what conditions nonadult users can benefit from bCBT programs. For some patients, careful evaluation of their suitability for a bCBT program might be necessary. For instance, based on our finding that therapists believe it is crucial to ensure that digital app use does not lead to excessive use of digital devices and potential addiction as a result, patients who are prone to technology addiction may need higher caution before a bCBT program is introduced to them.

A limitation of our study is that we did not ask patients about their current or previous involvement with stand-alone iCBT programs or other digital apps. Differential familiarity with such programs might potentially influence user views. Furthermore, we did not record how long participating patients had been in psychotherapy. Different experiences in terms of completed therapy duration might have led to differing opinions of patients. In contrast, examining patient views independent of the duration

they had spent in therapy could also be considered a strength because, as we did not set any patient duration in therapy as an inclusion criterion, our study potentially covers a higher variety of patients. Finally, the explicit distinction we made between *adjunct* and *integrated* bCBT programs was mainly based on our observations in the literature. More research is needed to support the validity of such a clear conceptual distinction.

Our therapist sample consisted of only those working with a CBT orientation and was small (n=23). It should be noted that previous research has showed that therapists who work with a CBT orientation have more positive attitudes toward the idea of using digital and blended interventions in the context of psychotherapy compared with therapists using other psychotherapy orientations [58,59]. Thus, the results of this study might not be generalizable to all psychotherapists regardless of their therapy orientations.

All patients and therapists participating in this study resided in or close to Cologne, Germany. Hence, one cannot directly generalize findings to other cultures or other regions of Germany.

Further research is needed to investigate expectations and opinions of different patient samples as well as comparisons of participant expectations of various psychotherapy approaches.

Conclusions

Our qualitative study found that patients and therapists who had never participated in a bCBT program could imagine themselves using one. They were willing to incorporate a digital app into their current face-to-face therapy practices. Both parties in the therapy relationship were open and curious about the possibilities and opportunities that technology could bring to their treatment journey. Participants provided a detailed framework for what features should be included in digital apps, independent of any particular mental health app introduced to them before the study. Therefore, researchers and developers of new technological solutions can use our findings as an independent guide when constructing new digital apps.

Conflicts of Interest

EA worked as a part-time researcher for Elona Health, which aims to develop digital health apps in the field of mental health care, during the time when this study was conducted and reported. MS is a shareholder of Elona Health. JAH declares no conflict of interest.

Multimedia Appendix 1
COREQ (Consolidated Criteria for Reporting Qualitative Research) checklist.
[PDF File (Adobe PDF File), 1032 KB - mental_v9i12e36806_app1.pdf]

References

1. Richards D. Prevalence and clinical course of depression: a review. Clin Psychol Rev 2011 Nov;31(7):1117-1125. [doi: 10.1016/j.cpr.2011.07.004] [Medline: 21820991]

2.   Gillham JE, Shatté AJ, Freres DR. Preventing depression: a review of cognitive-behavioral and family interventions. Appl Prev Psychol 2000 Mar;9(2):63-88 [FREE Full text] [doi: 10.1016/s0962-1849(00)80007-4]

3.   Judd LL. The clinical course of unipolar major depressive disorders. Arch Gen Psychiatry 1997 Nov;54(11):989-991. [doi: 10.1001/archpsyc.1997.01830230015002] [Medline: 9366654]

4.   König H, König HH, Konnopka A. The excess costs of depression: a systematic review and meta-analysis. Epidemiol Psychiatr Sci 2019 Apr 05;29:e30 [FREE Full text] [doi: 10.1017/S2045796019000180] [Medline: 30947759]

5.   Cuijpers P, Cristea IA, Karyotaki E, Reijnders M, Huibers MJ. How effective are cognitive behavior therapies for major depression and anxiety disorders? A meta-analytic update of the evidence. World Psychiatry 2016 Oct;15(3):245-258 [FREE Full text] [doi: 10.1002/wps.20346] [Medline: 27717254]

6.   Haby MM, Donnelly M, Corry J, Vos T. Cognitive behavioural therapy for depression, panic disorder and generalized anxiety disorder: a meta-regression of factors that may predict outcome. Aust N Z J Psychiatry 2006 Jan;40(1):9-19. [doi: 10.1080/j.1440-1614.2006.01736.x] [Medline: 16403033]

7.   Hofmann SG, Asnaani A, Vonk IJ, Sawyer AT, Fang A. Erratum to: the efficacy of cognitive behavioral therapy: a review of meta-analyses. Cogn Ther Res 2014 Jan 8;38(3):368 [FREE Full text] [doi: 10.1007/s10608-013-9595-3]

8.   Tolin DF. Is cognitive-behavioral therapy more effective than other therapies? A meta-analytic review. Clin Psychol Rev 2010 Aug;30(6):710-720. [doi: 10.1016/j.cpr.2010.05.003] [Medline: 20547435]

9.   Burns DD, Nolen-Hoeksema S. Coping styles, homework compliance, and the effectiveness of cognitive-behavioral therapy. J Consult Clin Psychol 1991 Apr;59(2):305-311. [doi: 10.1037//0022-006x.59.2.305] [Medline: 2030192]

10. Haynes RB, McDonald HP, Garg AX. Helping patients follow prescribed treatment: clinical applications. JAMA 2002 Dec 11;288(22):2880-2883. [doi: 10.1001/jama.288.22.2880] [Medline: 12472330]

11. Kazantzis N, Lampropoulos GK. Reflecting on homework in psychotherapy: what can we conclude from research and experience? J Clin Psychol 2002 May;58(5):577-585. [doi: 10.1002/jclp.10034] [Medline: 11967882]

12. Helbig S, Fehm L. Problems with homework in CBT: rare exception or rather frequent? Behav Cogn Psychother 2004 Jun 14;32(3):291-301 [FREE Full text] [doi: 10.1017/s1352465804001365]

13. Burns DD, Spangler DL. Does psychotherapy homework lead to improvements in depression in cognitive-behavioral therapy or does improvement lead to increased homework compliance? J Consult Clin Psychol 2000 Feb;68(1):46-56. [doi: 10.1037//0022-006x.68.1.46] [Medline: 10710839]

14. Kazantzis N, Deane FP, Ronan KR. Homework assignments in cognitive and behavioral therapy: a meta-analysis. Clin Psychol Sci Pract 2000 Jun;7(2):189-202 [FREE Full text] [doi: 10.1093/clipsy.7.2.189]

15. Kazantzis N, Deane FP, Ronan KR. Assessing compliance with homework assignments: review and recommendations for clinical practice. J Clin Psychol 2004 Jun;60(6):627-641. [doi: 10.1002/jclp.10239] [Medline: 15141396]

16. Kazantzis N, Whittington C, Zelencich L, Kyrios M, Norton PJ, Hofmann SG. Quantity and quality of homework compliance: a meta-analysis of relations with outcome in cognitive behavior therapy. Behav Ther 2016 Sep;47(5):755-772. [doi: 10.1016/j.beth.2016.05.002] [Medline: 27816086]

17. Wojnarowski C, Firth N, Finegan M, Delgadillo J. Predictors of depression relapse and recurrence after cognitive behavioural therapy: a systematic review and meta-analysis. Behav Cogn Psychother 2019 Sep;47(5):514-529. [doi: 10.1017/S1352465819000080] [Medline: 30894231]

18. Erbe D, Eichert H, Riper H, Ebert DD. Blending face-to-face and internet-based interventions for the treatment of mental disorders in adults: systematic review. J Med Internet Res 2017 Sep 15;19(9):e306 [FREE Full text] [doi: 10.2196/jmir.6588] [Medline: 28916506]

19. Wright JH, Mishkind M, Eells TD, Chan SR. Computer-assisted cognitive-behavior therapy and mobile apps for depression and anxiety. Curr Psychiatry Rep 2019 Jun 27;21(7):62. [doi: 10.1007/s11920-019-1031-2] [Medline: 31250242]

20. Cuijpers P, Kleiboer A, Karyotaki E, Riper H. Internet and mobile interventions for depression: opportunities and challenges. Depress Anxiety 2017 Jul;34(7):596-602. [doi: 10.1002/da.22641] [Medline: 28471479]

21. 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 [FREE Full text] [doi: 10.1016/j.invent.2015.01.001]

22. Tang W, Kreindler D. Supporting homework compliance in cognitive behavioural therapy: essential features of mobile apps. JMIR Ment Health 2017 Jun 08;4(2):e20 [FREE Full text] [doi: 10.2196/mental.5283] [Medline: 28596145]

23. Jacmon J, Malouff JM, Taylor N. Treatment of major depression: effectiveness of cognitive-behavioural therapy with an internet course as a central component. E J Appl Psychol 2009 Jan 03;5(2):1-8 [FREE Full text] [doi: 10.7790/ejap.v5i2.153]

24. Kooistra LC, Wiersma JE, Ruwaard J, Neijenhuijs K, Lokkerbol J, van Oppen P, et al. Cost and effectiveness of blended versus standard cognitive behavioral therapy for outpatients with depression in routine specialized mental health care: pilot randomized controlled trial. J Med Internet Res 2019 Oct 29;21(10):e14261 [FREE Full text] [doi: 10.2196/14261] [Medline: 31663855]

25. Schuster R, Laireiter AR, Berger T, Moritz S, Meyer B, Hohagen F, et al. Immediate and long-term effectiveness of adding an Internet intervention for depression to routine outpatient psychotherapy: subgroup analysis of the EVIDENT trial. J Affect Disord 2020 Sep 01;274:643-651. [doi: 10.1016/j.jad.2020.05.122] [Medline: 32663998]

26. 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 09;3(1):e9 [FREE Full text] [doi: 10.2196/mental.4534] [Medline: 26860537]

27. Haug T, Nordgreen T, Öst LG, Kvale G, Tangen T, Andersson G, et al. Stepped care versus face-to-face cognitive behavior therapy for panic disorder and social anxiety disorder: predictors and moderators of outcome. Behav Res Ther 2015 Aug;71:76-89. [doi: 10.1016/j.brat.2015.06.002] [Medline: 26081010]

28. Klein AA, Slaymaker VJ, Dugosh KL, McKay JR. Computerized continuing care support for alcohol and drug dependence: a preliminary analysis of usage and outcomes. J Subst Abuse Treat 2012 Jan;42(1):25-34. [doi: 10.1016/j.jsat.2011.07.002] [Medline: 21862275]

29. Berger T, Krieger T, Sude K, Meyer B, Maercker A. Evaluating an e-mental health program ("deprexis") as adjunctive treatment tool in psychotherapy for depression: results of a pragmatic randomized controlled trial. J Affect Disord 2018 Feb;227:455-462. [doi: 10.1016/j.jad.2017.11.021] [Medline: 29154168]

30. Månsson KN, Skagius Ruiz E, Gervind E, Dahlin M, Andersson G. Development and initial evaluation of an Internet-based support system for face-to-face cognitive behavior therapy: a proof of concept study. J Med Internet Res 2013 Dec 10;15(12):e280 [FREE Full text] [doi: 10.2196/jmir.3031] [Medline: 24326065]

31. Baumeister H, Paganini S, Sander LB, Lin J, Schlicker S, Terhorst Y, et al. Effectiveness of a guided Internet- and mobile-based intervention for patients with chronic back pain and depression (WARD-BP): a multicenter, pragmatic randomized controlled trial. Psychother Psychosom 2021;90(4):255-268. [doi: 10.1159/000511881] [Medline: 33321501]

32. Klein JP, Berger T, Schröder J, Späth C, Meyer B, Caspar F, et al. Effects of a psychological internet intervention in the treatment of mild to moderate depressive symptoms: results of the EVIDENT study, a randomized controlled trial. Psychother Psychosom 2016;85(4):218-228 [FREE Full text] [doi: 10.1159/000445355] [Medline: 27230863]

33. Thase ME, Wright JH, Eells TD, Barrett MS, Wisniewski SR, Balasubramani GK, et al. Improving the efficiency of psychotherapy for depression: computer-assisted versus standard CBT. Am J Psychiatry 2018 Mar 01;175(3):242-250 [FREE Full text] [doi: 10.1176/appi.ajp.2017.17010089] [Medline: 28969439]

34. Ly KH, Topooco N, Cederlund H, Wallin A, Bergström J, Molander O, et al. Smartphone-supported versus full behavioural activation for depression: a randomised controlled trial. PLoS One 2015 May 26;10(5):e0126559 [FREE Full text] [doi: 10.1371/journal.pone.0126559] [Medline: 26010890]

35. Pérez JC, Fernández O, Cáceres C, Carrasco Á, Moessner M, Bauer S, et al. An adjunctive internet-based intervention to enhance treatment for depression in adults: randomized controlled trial. JMIR Ment Health 2021 Dec 16;8(12):e26814 [FREE Full text] [doi: 10.2196/26814] [Medline: 34927594]

36. Rasing SP, Stikkelbroek YA, den Hollander W, Riper H, Deković M, Nauta MH, et al. Pragmatic quasi-experimental controlled trial evaluating the outcomes of blended CBT compared to face-to-face CBT and treatment as usual for adolescents with depressive disorders. Int J Environ Res Public Health 2021 Mar 17;18(6):3102 [FREE Full text] [doi: 10.3390/ijerph18063102] [Medline: 33802913]

37. Thase ME, McCrone P, Barrett MS, Eells TD, Wisniewski SR, Balasubramani GK, et al. Improving cost-effectiveness and access to cognitive behavior therapy for depression: providing remote-ready, computer-assisted psychotherapy in times of crisis and beyond. Psychother Psychosom 2020;89(5):307-313 [FREE Full text] [doi: 10.1159/000508143] [Medline: 32396917]

38. Langergaard A, Mathiasen K, Søndergaard J, Sørensen SS, Laursen SL, Xylander AA, et al. Economic evaluation alongside a randomized controlled trial of blended cognitive-behavioral therapy for patients suffering from major depressive disorder. Internet Interv 2022 Apr;28:100513 [FREE Full text] [doi: 10.1016/j.invent.2022.100513] [Medline: 35242594]

39. Etzelmueller A, Radkovsky A, Hannig W, Berking M, Ebert DD. Patient's experience with blended video- and internet based cognitive behavioural therapy service in routine care. Internet Interv 2018 Jun;12:165-175 [FREE Full text] [doi: 10.1016/j.invent.2018.01.003] [Medline: 30135780]

40. Kooistra LC, Ruwaard J, Wiersma JE, van Oppen P, van der Vaart R, van Gemert-Pijnen JE, et al. Development and initial evaluation of blended cognitive behavioural treatment for major depression in routine specialized mental health care. Internet Interv 2016 May;4:61-71 [FREE Full text] [doi: 10.1016/j.invent.2016.01.003] [Medline: 30135791]

41. 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]

42. Gumz A, Boettcher J. Digitale interventionen in der Psychotherapie. Psychotherapeut 2021 Aug 30;66(5):369-371 [FREE Full text] [doi: 10.1007/s00278-021-00531-4]

43. Stoff-Ahnis S. Digitale gesundheitsanwendungen – das erste jahr aus sicht der gesetzlichen krankenversicherung. MedR 2022 Apr 12;40(4):285-290 [FREE Full text] [doi: 10.1007/s00350-022-6161-6]

44. Kitzinger J. Qualitative research. Introducing focus groups. BMJ 1995 Jul 29;311(7000):299-302 [FREE Full text] [doi: 10.1136/bmj.311.7000.299] [Medline: 7633241]

45. Malagon-Maldonado G. Qualitative research in health design. HERD 2014;7(4):120-134. [doi: 10.1177/193758671400700411] [Medline: 25303432]

46.   Chun Tie Y, Birks M, Francis K. Grounded theory research: a design framework for novice researchers. SAGE Open Med 2019 Jan 2;7:2050312118822927 [FREE Full text] [doi: 10.1177/2050312118822927] [Medline: 30637106]

47.   Corbin J, Strauss A. Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory. 4th edition. Thousand Oaks, CA, USA: Sage Publications; 2014.

48.   Glaser BG, Strauss AL. Discovery of Grounded Theory: Strategies for Qualitative Research. Chicago, IL, USA: Aldine Publishing Company; 1967.

49.   Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]

50.   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]

51.   Cerga-Pashoja A, Doukani A, Gega L, Walke J, Araya R. Added value or added burden? A qualitative investigation of blending internet self-help with face-to-face cognitive behaviour therapy for depression. Psychother Res 2020 Nov;30(8):998-1010. [doi: 10.1080/10503307.2020.1720932] [Medline: 32024447]

52.   Schuster R, Topooco N, Keller A, Radvogin E, Laireiter AR. Advantages and disadvantages of online and blended therapy: replication and extension of findings on psychotherapists' appraisals. Internet Interv 2020 Sep;21:100326 [FREE Full text] [doi: 10.1016/j.invent.2020.100326] [Medline: 32477885]

53.   Beattie A, Shaw A, Kaur S, Kessler D. Primary-care patients' expectations and experiences of online cognitive behavioural therapy for depression: a qualitative study. Health Expect 2009 Mar;12(1):45-59 [FREE Full text] [doi: 10.1111/j.1369-7625.2008.00531.x] [Medline: 19250152]

54.   Bendelin N, Hesser H, Dahl J, Carlbring P, Nelson KZ, Andersson G. Experiences of guided internet-based cognitive-behavioural treatment for depression: a qualitative study. BMC Psychiatry 2011 Jun 30;11:107 [FREE Full text] [doi: 10.1186/1471-244X-11-107] [Medline: 21718523]

55.   Lambert MJ, Barley DE. Research summary on the therapeutic relationship and psychotherapy outcome. Psychother Theory Res Pract Train 2001;38(4):357-361 [FREE Full text] [doi: 10.1037/0033-3204.38.4.357]

56.   Leach MJ. Rapport: a key to treatment success. Complement Ther Clin Pract 2005 Nov;11(4):262-265. [doi: 10.1016/j.ctcp.2005.05.005] [Medline: 16290897]

57.   Marziali E, Alexander L. The power of the therapeutic relationship. Am J Orthopsychiatry 1991 Jul;61(3):383-391. [doi: 10.1037/h0079268] [Medline: 1951645]

58.   Schröder J, Berger T, Meyer B, Lutz W, Hautzinger M, Späth C, et al. Attitudes towards internet interventions among psychotherapists and individuals with mild to moderate depression symptoms. Cogn Ther Res 2017 Apr 22;41(5):745-756 [FREE Full text] [doi: 10.1007/s10608-017-9850-0]

59.   Schuster R, Pokorny R, Berger T, Topooco N, Laireiter AR. The advantages and disadvantages of online and blended therapy: survey study amongst licensed psychotherapists in Austria. J Med Internet Res 2018 Dec 18;20(12):e11007 [FREE Full text] [doi: 10.2196/11007] [Medline: 30563817]

Abbreviations

bCBT: blended cognitive behavioral therapy
CBT: cognitive behavioral therapy
COREQ: Consolidated Criteria for Reporting Qualitative Research
iCBT: internet-based cognitive behavioral therapy

Use of an Ingestible, Sensor-Based Digital Adherence System to Strengthen the Therapeutic Relationship in Serious Mental Illness

Anabel G Richey[1,2*], MSc; Ildiko Kovacs[3], MD; Sara Browne[2*], MD

[1]Department of Psychiatry, Oxford University, Oxford, United Kingdom

[2]Department of Medicine, University of California, San Diego, La jolla, CA, United States

[3]Department of Psychiatry, University of California, San Diego, CA, United States

[*]these authors contributed equally

Corresponding Author:
Sara Browne, MD
Department of Medicine
University of California, San Diego
9452 Medical Center Dr
La jolla, CA, 92037
United States
Phone: 1 858 822 6563
Email: shbrowne@health.ucsd.edu

Abstract

Serious mental illness is a chronic condition that requires long-term pharmacological treatment. Adherence to oral antipsychotic medication has specific nuances that affects patients and physicians alike. For patients with serious mental illness, nonadherence increases their risk of hospitalization and relapse. Nonadherence is a formidable barrier for physicians in accurately assessing medication efficacy and helping patients achieve their fullest potential. A digital adherence system approved by the Food and Drug Administration can provide near–real time aripiprazole ingestion information. The system records ingestions through an embedded ingestible sensor in oral aripiprazole, which sends a transient local signal to a patch worn on the patient's torso that is then stored on a paired smartphone app. With patient permission, these data can be viewed remotely by their physician, along with a patient's mood, activity, and time spent resting. Such data are able to do the following: reveal broad patterns of medication adherence behavior to the patient as well as their physician; help physicians and patients understand and create more realistic expectations for adherence; promote discussion of treatment options; and minimize therapeutic appointment time devoted to determining actual adherence, thereby maximizing the time available to address each patient's distinctive reasons for their adherence pattern. Crucially, extra time created during appointments can be used to strengthen the therapeutic relationship, which may translate into both improvements in adherence and patient attitude toward their medication. Future investigations are needed to examine how this technology impacts the development of training and best practice guidelines for its use. Otherwise, the potential benefits of this technology may be lost, or worse, inadequate and inappropriate use may harm the therapeutic relationship.

Introduction

Overview

The art has three factors, the disease, the patient, the physician. The physician is the servant of the art. The patient must cooperate with the physician in combatting the disease. [Hippocrates]

Heralded as the father of western medicine, Hippocrates' contributions to medicine remain foundational to current perceptions of disease and the body [1]. He used multiple strategies to treat ailments, including herbal remedies in a manner similar to modern prescribed pharmaceuticals [2,3]. Hippocrates' understood that the medicinal art form requires a compassionate, holistic, yet highly individualized series of investigations into how the patient and disease intertwine. He understood there must be trust and honesty between both parties, and a doctor's help can only go as far as patients are willing to cooperate. Every patient has the power to render any oral medication 100% ineffective, no matter how much money, time,

or human resources the pharmaceutical industry infuses into its creation; a drug is only effective if it is taken [4].

Medication adherence is especially important in cases of serious mental illness (SMI), a term referring to bipolar disorder, schizoaffective disorder, and schizophrenia [5]. Reported adherence rates to antipsychotics vary, ranging from 24% to 90%, with a mean of 58% [6]. Nonadherent patients with schizophrenia have 55% higher odds of being admitted to hospital; adequate adherence is imperative to ensure remission and prevent SMI relapse [7].

The first ingestible-sensor–based digital adherence system (iDAS) to follow and support oral medication adherence was approved by the Food and Drug Administration in 2014 [8]. The major advancement associated with this medical device is capture of real-time data on actual drug ingestion along with capture of simultaneous physiological data [9]. Specific approval of the ingestible sensor system with the atypical antipsychotic aripiprazole was obtained in 2017 [10,11]. Aripiprazole, originally approved for use in schizophrenia [12], is also currently approved for use as an adjunct or monotherapy in bipolar I mania [13-17] and major depressive disorder [18]. Approval of digital aripiprazole included the use of an app tailored specifically for patients with serious mental illness under the name Abilify Mycite System [19].

This viewpoint discusses how iDAS adherence data may be used to benefit psychiatric consultations and positively impact the physician-patient relationship. We focus on how iDAS may be used to serve and strengthen the therapeutic relationship (TR), foundational to any beneficial outcome of any technology application [20]. Could iDAS represent a new era of physician-patient interactions, a technological turning point that Hippocrates could not have foreseen?

For the sake of clarity, at the outset, our discussion assumes patients give informed consent before and throughout their use of iDAS. We focus on how iDAS may be used to serve and strengthen the TR, which is foundational to any beneficial outcome of any technological application [20]. The ethics of informed consent is of critical importance when using this technology (see analysis and discussion by Beriain and Gonzalez [21] and Torous and Roberts [20]). As the cornerstone of any TR is trust, and a large portion of such trust comes from continued consent over time, we subsequently discuss dynamic consent during iDAS use.

Background

A Food and Drug Administration–approved iDAS, comprising the ingestible sensor and a generic app [8], has been used with multiple oral medications [22-24]. The iDAS for aripiprazole, developed in collaboration with Otsuka America Pharmaceutical, Inc [25], was approved with an app developed specifically for the use of patients with SMI and released under the name Abilify Mycite System [26]. The iDAS for aripiprazole has 3 components, which are aripiprazole with embedded ingestible sensor, a patch, and a paired smartphone with installed iDAS app (Figure 1). When ingested and in contact with gastric fluids, the sensor sends an electrochemical signal unique to aripiprazole that is recognized and stored by a patch worn on the user's torso. The patch then communicates via Bluetooth to the patient's paired smartphone with installed app, which stores the time and date of successful medication ingestion [27]. Over 90% of pill ingestions are detected by the patch in 3 minutes, while the patch can take up to 2 hours to sync ingestion data with the app [28,29]. The app has several features, including tracking mood, rest, and activity.

Figure 1. Depiction of the components of aripiprazole ingestible-sensor–based digital adherence system (iDAS). (1) Aripiprazole with an embedded ingestible sensor, a patch, and paired smartphone with an installed iDAS app. (2) When ingested and in contact with gastric fluids, the sensor sends an electrochemical signal unique to aripiprazole that is recognized and stored by a patch worn on the user's torso. (3) The patch then communicates via Bluetooth to the patient's paired smartphone with an installed app that stores the time and date of successful medication ingestion. (4-5) The patient allows the prescribing provider to view ingestion data via a website that is accessed through a secure server.
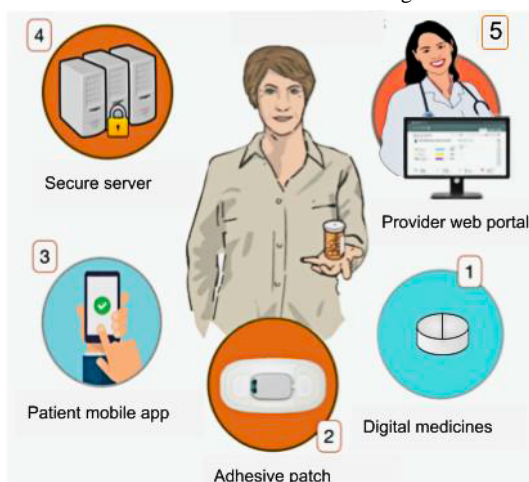

Figure 2 shows the displays of each feature in addition to noting if the feature's data are manually entered. The patient may give permission for their physician (and other trusted persons) to access a web-based iDAS portal where their medication adherence data can be view remotely. The app currently makes

clear attempts to gather a holistic data set, including information on mood and activity from the patient. It uses friendly interface displays such as "You're all good!" once a pill is ingested. The app has a *casual* appearance and does not, for example, display, "You're all good! Aripiprazole ingested: 11:24 AM. Ingestion
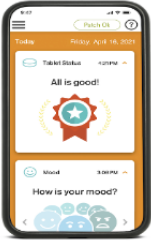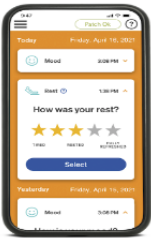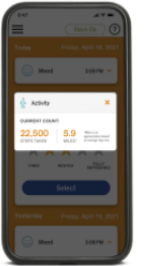
data sent to Dr. Argus Panoptes at 12:24 AM." The *feel-good* design of iDAS chooses not to highlight physician monitoring, for better or for worse.

The type of data collection by iDAS contributes to the underappreciated "art of observation of the individual in its entirety" [30], providing objective inside data about a patient's habits, thus falling under the umbrella of devices that contribute to digital phenotyping, as defined by Torous et al [31]. Several other tools have been developed to integrate digital phenotyping data into patient care, including The SilverCloud platform using digital phenotyping data to delineate different subtypes of internet-based cognitive behavioral therapy [32] and the

MindLAMP platform, aimed at preventing relapse in people with schizophrenia spectrum disorders [33,34].

Abilify iDAS is intended to augment treatment as usual by a physician, a model incorporated into the MindLAMP [34] and Horyzons platforms [35]. Abilify iDAS and these platforms have consistently demonstrated feasibility of use and efficacy with treatment as usual in psychiatric patients with SMI [34,36-41]. These reports substantiate the proposal that Abilify iDAS could support the TR. By and large, more data on patients are not inherently *beneficial* nor *harmful*; rather, there are needs for careful consideration of what information the iDAS can provide and how this may be used to build trust and mutual understanding in psychiatric consultations.

Figure 2. Features of the aripiprazole iDAS app. All pill ingestion information automatically enters the app and must be shared with a health care provider with patient's permission. Other features, including mood, rest, and activity, do not have to be shared with a physician. All features of the app can be shared with 5 other people including friends or family, or another physician. iDAS: ingestible-sensor–based digital adherence system. (Images adapted and used with permission from MyCite Menu Overview, 2021, Otsuka America Pharmaceutical, Inc).

App Feature	App feature and the associated display	Manual entry or automatically sensed by DMS?	Data from this feature required to be shared with a physician? (yes=✓, no=✗)
Abilify ingestion tracking		**Sensed Automatically by patch** Information is sent to a secure cloud where both the patient and their physician can view it on the iDAS portal	✓
Rest		**Sensed automatically by patch** Does not measure sleep-- only rest time Patient can manually rate rest time from 1-5 stars	✗
Mood		**Entered Manually** Patient chooses from 7 general emojis, labelled: - Excited - Okay - Good - Happy - Angry - Depressed - Sad	✗
Activity level via step count		**Sensed automatically by patch**	✗

What Near–Real Time Adherence Data May Bring to Psychiatric Consultations

Confirmation of a Patient's Adherence

The iDAS technology allows for an accurate summary of medication adherence available for medical visits. Through no fault of their own, patients may demonstrate recall bias at appointments by erroneously reporting medication adherence or experience [42]. The physician and patient can look back at the adherence data together and discuss how medication adherence is going.

Inform Dosage Adjustment

The features of the iDAS may inform physicians tailoring the aripiprazole dosage. Two common aripiprazole side effects include restlessness and akathisia as well as somnolence [43]. iDAS ingestion data, in conjunction with rest and activity data, may be helpful in gauging a patient's experience with these side effects. In theory, a doctor could adjust the aripiprazole dose in a patient who complains of being unable to sit still, frequently rates their rest at 1 star, and has relatively good medication adherence. It is important to note, however, that Abilify iDAS is not licensed to inform the modification of aripiprazole dosage [44]. Nonetheless, real-time rest information in a patient complaining of somnolence would be difficult to ignore. In situations where the patient's app data and verbal recollection concur, a physician may feel more assured in changing a patient's dose.

Temporal Tracking of SMI Relapse Relative to Medication Adherence

Another benefit of the aripiprazole iDAS is the potential to track whether a symptomatic relapse was caused by nonadherence or a medication's lack of efficacy. As Weiden [45] describes, relapse caused by inefficacy of a medication may begin prior to nonadherence, and without clear data, relapses can be misattributed to behavioral nonadherence. To illustrate this, a clinician may assume that their patient ceased medication and relapsed, leading them to encourage their patient to resume their medication. However, using the aripiprazole iDAS, a clinician could review ingestion data and build a more accurate timeline of medication cessation in relation to their patient's relapse. If the physician concludes that the relapse began before medication cessation, the physician can adjust the treatment regimen. Similarly, physicians can also see when, despite maintaining good adherence throughout, a patient relapsed and clearly requires medication dosage or regimen changes.

Improve Patient's and Physician's Understanding Patterns of Adherence

The duration and reliability of iDAS data allow patients and physicians to isolate trends in adherence behavior. Much of the discussion of the commercial Abilify Mycite product focuses on its advantage to physicians, who are required to prescribe it. Crucially, aripiprazole iDAS data may also aid a patient in learning about themselves. Seemingly random instances of nonadherence may be associated with broader patterns. These patterns may be illuminated by simultaneously considering the rest, activity, and mood recordings with ingestion data over time. Browne et al [9] produced visualizations coupling iDAS physiological data with adherence data in patients with diabetes, which allowed easily accessible interpretation of individual behavior patterns over time. iDAS data can provide the opportunity for patients to use ingestion data as a lens to reveal patterns in their own life, rhetorically asking, "What does nonadherence tell me about my own life?" Physicians may use adherence data in an opposite fashion: life patterns can be used as a lens to reveal adherence data. A physician may rhetorically ask, "What does a patient's life tell me about their nonadherence?"

Thus, the iDAS system can be used as a recordkeeper to help understand patterns of adherence in patients, in ways that directly benefit the psychiatric consultation. After all, Hippocrates states that physicians are servants "to the art of medicine" rather than patients being servants to the physician. All art forms involve flexibility coupled with creativity, but this art form, as every physician knows, requires the most precious commodity of all in medical practice, which is *time*.

Time Opened up During Appointments May Support the Therapeutic Relationship

Easily accessible and highly accurate adherence data provided by iDAS shorten time spent on adherence *detective work* on the part of the physician during consultations. This opens up appointment time, which may be used to start healthy conversations around adherence patterns. In a study of conversations surrounding antipsychotic nonadherence during psychiatric consultations, patients often tried to minimize the risk of a "disciplinary" reply from their physician when disclosing nonadherence [46]. Patients are frequently ashamed of their noncompliance, especially when they are attached to their treating physician. To illustrate this, 1 patient began the disclosure of nonadherence by saying "My mood's been fine but you're not gonna be very happy [to know] …" [46]. This study of nonadherence disclosures highlights how difficult they can be for the patient. The authors compare patients' methods of nonadherence disclosure to methods used to deliver *bad news*. Patients should not worry or feel embarrassed by medication nonadherence as there can be many reasons for it. While the physician needs to spend time, interpreting patterns provided by iDAS, this can be conducted with the patient, spending more time exploring reasons for their behavior patterns.

Closing the Loop: Strengthening the Therapeutic Relationship as a Means to Antipsychotic Treatment Adherence

The value of adherence data cannot be overstated, even in situations where patients seem to be making little progress in becoming adherent to their medication regime. How can iDAS be valuable when the patient is firmly nonadherent, despite a kind and honest effort from their physician? In fact, these situations of firm nonadherence may create tension between the physician and patient, creating the sense that both parties are at a *standstill*. Nevertheless, true medication adherence cannot

exist without a patient's notion of trust and confidence in their physician. The availability of iDAS data means time—which would otherwise have been spent trying to investigate whether a patient is adherent to their regime at all—can be spent to improve this trust during consultations.

In their book on therapeutic alliance, Leslie Greenberg and Adam Horvath explicitly state that the core of therapeutic alliance relies on a sense of collaboration [47]. One cannot create a genuine collaboration with someone they do not trust. Therapeutic alliance is contingent on factors beyond agreement of treatment-related goals, with one of the most important factors being the personal bond between patient and physician [47,48]. Forming such a bond takes time. Extra appointment time can facilitate this bond, particularly when there may not be agreement between a patient and physician.

The bond contributing to therapeutic alliance requires mutual trust and regard [49]. Extra appointment time facilitated by iDAS use may be spent strengthening the therapeutic alliance by openly and nonjudgmentally discussing possible reasons for nonadherence, trying to find the specific reasons why the individual patient is not taking the medication as prescribed [50]. Most patients are not aware of the multiple layers of attitudinal and behavioral factors leading to therapeutic adherence; when it is discussed in a cooperative, nonthreatening, and constructive way, patients are usually very interested in understanding their own behaviors [51]. Physicians need to provide encouragement and positive feedback for patients' effort [52].

Patients with schizophrenia who established a good therapeutic alliance during the first 6 months of treatment had a significantly increased likelihood of adhering to medication regimes [53]. This initial alliance was also significantly associated with improved treatment outcomes using less medication when compared to patients without good therapeutic alliance [54]. Subsequent studies on TR and antipsychotic adherence in patients with schizophrenia demonstrated that TR is associated with both adherence and attitude toward treatment with antipsychotics [54]. The use of the Helping Alliance Scale [54,55] to measure TR strength in clinicians and their patients (higher scores signify a better TR) showed each unit increase in the clinician's TR rating was accompanied by a 65.9% increase in the odds ratio of their patient having good medication adherence. For each patient-rated TR unit increase, there was a 20.8% increase in the odds ratio of good medication adherence [54]. Consequently, extra time spent strengthening the physician-patient bond could lead to better adherence in patients taking antipsychotics.

Dynamic Consent With Third Party Availability During iDAS Use May Minimize Risks to the Therapeutic Relationship

The bond that forms the cornerstone of the TR is based on mutual trust between physician and patient, which not only builds over time, but may also change over time. To support this bond over time iDAS used with dynamic consent may increase trust. Prictor et al [56] defined dynamic consent as "an approach to consent that enables people, through an interactive digital interface, to make granular decisions about their ongoing participation." The iDAS MyCite app contains inclusion of a digital interface allowing patients at any time to withdraw consent to all or certain types of data sharing with some (or all) of previously approved persons, including their physician [57]. Of course, it would be ideal if a patient felt comfortable telling their doctor that they no longer want certain portions of their data shared with their health care team. However, some patients may prefer to withdraw consent on a digital platform by clicking a "stop sharing data" button. It may also be important to consider having access to a third party, such as a nurse or a *consent representative*, to allow the patient to discuss their ongoing iDAS consent and relationship with the doctor. This would allow open conversations about withdrawing consent to be made outside of the doctor-patient space and create a space of open conversations about data sharing without a *direct* disruption in the doctor-patient relationship. The availability of a third party, such as a nurse, can directly address questions of "Who do you want to allow to see your data?", "What data are they allowed to see?" and "How do you want these data to be used?" This would allow a patient to discuss their patient-physician relationship, patient-data relationship, and physician-data relationship with someone in health care who is not their doctor. The third-party patient advocate role would mitigate patient's concerns about withdrawing consent and would also allow a patient to feel that their feelings and concerns are their health care team's *priority*. Dynamic consent, in addition to the availability of a third party, facilitates methods for key patient data control proposed by Vayena and Blasimme [58]—control over who accesses one's health care data and how those data are used. Placing the patient at the center of the consent process over time would likely optimize patient autonomy and physician-patient confidence that iDAS was supporting their relationship.

Conclusion

Assuming a patient agrees to use it over time, the aripiprazole iDAS can provide information on an individual patient's adherence, which may help to inform dosage and trace SMI relapse relative to medication adherence. iDAS data can be an *icebreaker* in conversations surrounding medication adherence or how patients feel about their current medication regime. Through iDAS use, patients may learn more about their relationship with adherence, and physicians may gain a more accurate perspective on individual patients, improving both the patient and physician's understanding of adherence patterns. Ultimately, this tool only reports behavioral data, which does not change a patient's attitude toward antipsychotics. It is up to the physician to practice their art in a way that compassionately understands and encourages patients to reach their full potential. This can only be done through time spent developing a strong and healthy therapeutic relationship, which itself has repeatedly been shown to influence patient treatment adherence. As data accumulate from patients and their physicians choosing to use the aripiprazole iDAS, a more complete picture will emerge of opportunities to tailor and provide targeted use needed for different SMI diagnoses and individual patient characteristics, such as gender. It is critical, however, that future investigations examine how this technology

impacts the physician-patient relationship to develop training and provide best-practice guidelines on how to use it to strengthen the therapeutic relationship. Otherwise, the potential benefits of this technological advancement may be lost, or worse still, inadequate and inappropriate use may result in harm to the therapeutic relationship.

Acknowledgments

Conflicts of Interest

None declared.

References

1. Yapijakis C. Hippocrates of Kos, the father of clinical medicine, and Asclepiades of Bithynia, the father of molecular medicine. Review. In Vivo 2009;23(4):507-514 [FREE Full text] [Medline: 19567383]
2. Orfanos C. From Hippocrates to modern medicine. J Eur Acad Dermatol Venereol 2007 Jul;21(6):852-858. [doi: 10.1111/j.1468-3083.2007.02273.x] [Medline: 17567335]
3. Mavrogenis AF, Saranteas T, Markatos K, Kotsiou A, Tesseromatis C. Pharmacies for pain and trauma in ancient Greece. Int Orthop 2019 Jun 9;43(6):1529-1536. [doi: 10.1007/s00264-018-4219-x] [Medline: 30413852]
4. Lindenfeld J, Jessup M. 'Drugs don't work in patients who don't take them' (C. Everett Koop, MD, US Surgeon General, 1985). Eur J Heart Fail 2017 Nov 10;19(11):1412-1413 [FREE Full text] [doi: 10.1002/ejhf.920] [Medline: 28891126]
5. Chang C, Hayes RD, Perera G, Broadbent MTM, Fernandes AC, Lee WE, et al. Life expectancy at birth for people with serious mental illness and other major disorders from a secondary mental health care case register in London. PLoS One 2011;6(5):e19590 [FREE Full text] [doi: 10.1371/journal.pone.0019590] [Medline: 21611123]
6. Cramer JA, Rosenheck R. Compliance with medication regimens for mental and physical disorders. Psychiatr Serv 1998 Feb;49(2):196-201. [doi: 10.1176/ps.49.2.196] [Medline: 9575004]
7. Mahabaleshwarkar R, Lin D, Fishman J, Blair T, Hetherington T, Palmer P, et al. The Impact of Once-Monthly Paliperidone Palmitate on Healthcare Utilization Among Patients With Schizophrenia Treated in an Integrated Healthcare System: A Retrospective Mirror-Image Study. Adv Ther 2021 Apr 11;38(4):1958-1974 [FREE Full text] [doi: 10.1007/s12325-021-01626-9] [Medline: 33704681]
8. Ingestible Event Marker approval, February 7, 2014. US Food and Drug Administration, Center for Drug evaluation and Research. URL: https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpmn/pmn.cfm?ID=K133263 [accessed 2022-04-01]
9. Browne SH, Behzadi Y, Littlewort G. Let Visuals Tell the Story: Medication Adherence in Patients with Type II Diabetes Captured by a Novel Ingestion Sensor Platform. JMIR Mhealth Uhealth 2015 Dec 31;3(4):e108 [FREE Full text] [doi: 10.2196/mhealth.4292] [Medline: 26721413]
10. Drug approval package: Abilify MyCite (aripiprazole). US Food and Drug Administration. URL: https://www.accessdata.fda.gov/drugsatfda_docs/nda/2017/207202Orig1s000TOC.cfm [accessed 2022-03-21]
11. FDA Newsroom. US Food and Drug Administration. URL: https://www.fda.gov/NewsEvents/Newsroom [accessed 2022-08-09]
12. Grady MA, Gasperoni TL, Kirkpatrick P. Aripiprazole. Nat Rev Drug Discov 2003 Jun;2(6):427-428. [doi: 10.1038/nrd1114] [Medline: 12790153]
13. Keck PE, Marcus R, Tourkodimitris S, Ali M, Liebeskind A, Saha A, Aripiprazole Study Group. A placebo-controlled, double-blind study of the efficacy and safety of aripiprazole in patients with acute bipolar mania. Am J Psychiatry 2003 Sep;160(9):1651-1658. [doi: 10.1176/appi.ajp.160.9.1651] [Medline: 12944341]
14. Sachs G, Sanchez R, Marcus R, Stock E, McQuade R, Carson W, Aripiprazole Study Group. Aripiprazole in the treatment of acute manic or mixed episodes in patients with bipolar I disorder: a 3-week placebo-controlled study. J Psychopharmacol 2006 Jul 09;20(4):536-546. [doi: 10.1177/0269881106059693] [Medline: 16401666]
15. Fagiolini A. Practical guidance for prescribing with aripiprazole in bipolar disorder. Curr Med Res Opin 2008 Sep 13;24(9):2691-2702. [doi: 10.1185/03007990802341426] [Medline: 18702850]
16. Vieta E, T'joen C, McQuade RD, Carson WH, Marcus RN, Sanchez R, et al. Efficacy of adjunctive aripiprazole to either valproate or lithium in bipolar mania patients partially nonresponsive to valproate/lithium monotherapy: a placebo-controlled study. Am J Psychiatry 2008 Oct;165(10):1316-1325. [doi: 10.1176/appi.ajp.2008.07101560] [Medline: 18381903]
17. Young AH, Oren DA, Lowy A, McQuade RD, Marcus RN, Carson WH, et al. Aripiprazole monotherapy in acute mania: 12-week randomised placebo- and haloperidol-controlled study. Br J Psychiatry 2009 Jan 02;194(1):40-48. [doi: 10.1192/bjp.bp.108.049965] [Medline: 19118324]

18.   Pae C, Forbes A, Patkar AA. Aripiprazole as adjunctive therapy for patients with major depressive disorder: overview and implications of clinical trial data. CNS Drugs 2011 Feb;25(2):109-127. [doi: 10.2165/11538980-000000000-00000] [Medline: 21254788]

19.   Abilify MyCite summary review, October 20, 2017. US Food and Drug Administration, Center for Drug evaluation and research. URL: https://www.accessdata.fda.gov/drugsatfda_docs/nda/2017/207202Orig1s000SumR.pdf [accessed 2022-04-01]

20.   Torous J, Roberts LW. The Ethical Use of Mobile Health Technology in Clinical Psychiatry. J Nerv Ment Dis 2017 Jan;205(1):4-8. [doi: 10.1097/NMD.0000000000000596] [Medline: 28005647]

21.   de Miguel Beriain I, Morla González M. 'Digital pills' for mental diseases: an ethical and social analysis of the issues behind the concept. J Law Biosci 2020;7(1):lsaa040 [FREE Full text] [doi: 10.1093/jlb/lsaa040] [Medline: 34221420]

22.   Sulkowski M, Luetkemeyer AF, Wyles DL, Martorell C, Muir A, Weisberg I, et al. Impact of a digital medicine programme on hepatitis C treatment adherence and efficacy in adults at high risk for non-adherence. Aliment Pharmacol Ther 2020 Jun 30;51(12):1384-1396. [doi: 10.1111/apt.15707] [Medline: 32352586]

23.   Browne SH, Umlauf A, Tucker AJ, Low J, Moser K, Gonzalez Garcia J, et al. Wirelessly observed therapy compared to directly observed therapy to confirm and support tuberculosis treatment adherence: A randomized controlled trial. PLoS Med 2019 Oct 4;16(10):e1002891 [FREE Full text] [doi: 10.1371/journal.pmed.1002891] [Medline: 31584944]

24.   Browne SH, Vaida F, Umlauf A, Tucker AJ, Blaschke TF, Benson CA. Supporting the art: Medication adherence patterns in persons prescribed ingestible sensor-enabled oral pre-exposure prophylaxis to prevent HIV infection. Clinical Infectious Diseases. In Press 2022. [doi: 10.1093/cid/ciac280]

25.   Otsuka and Proteus announce the first U.S. FDA approval of a digital medicine system: ABILIFY MYCITE (aripiprazole tablets with sensor). Otsuka. 2017. URL: https://www.otsuka-us.com/discover/articles-1075 [accessed 2019-10-01]

26.   Abilify MyCite approval letter, November 17, 2017. US Food and Drug Administration, Center for Drug evaluation and Research. URL: https://www.accessdata.fda.gov/drugsatfda_docs/nda/2017/207202Orig1s000Approv.pdf [accessed 2022-04-01]

27.   Papola D, Gastaldon C, Ostuzzi G. Can a digital medicine system improve adherence to antipsychotic treatment? Epidemiol Psychiatr Sci 2018 Mar 12;27(3):227-229. [doi: 10.1017/s2045796018000082]

28.   Goold I. Digital tracking medication: big promise or Big Brother? Law, Innovation and Technology 2019 Sep 12;11(2):203-230. [doi: 10.1080/17579961.2019.1665791]

29.   Abilify MyCite FAQs. AbilifyMyCite. URL: https://www.abilifymycite.com/faqs [accessed 2021-06-20]

30.   Prakash J, Chaudhury S, Chatterjee K. Digital phenotyping in psychiatry: When mental health goes binary. Ind Psychiatry J 2021;30(2):191-192 [FREE Full text] [doi: 10.4103/ipj.ipj_223_21] [Medline: 35017799]

31.   Torous J, Kiang MV, Lorme J, Onnela J. New Tools for New Research in Psychiatry: A Scalable and Customizable Platform to Empower Data Driven Smartphone Research. JMIR Ment Health 2016 May 05;3(2):e16 [FREE Full text] [doi: 10.2196/mental.5165] [Medline: 27150677]

32.   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]

33.   Vaidyam A, Halamka J, Torous J. Enabling Research and Clinical Use of Patient-Generated Health Data (the mindLAMP Platform): Digital Phenotyping Study. JMIR Mhealth Uhealth 2022 Jan 07;10(1):e30557 [FREE Full text] [doi: 10.2196/30557] [Medline: 34994710]

34.   Rodriguez-Villa E, Rozatkar AR, Kumar M, Patel V, Bondre A, Naik SS, et al. Cross cultural and global uses of a digital mental health app: results of focus groups with clinicians, patients and family members in India and the United States. Glob Ment Health (Camb) 2021 Aug 24;8:e30 [FREE Full text] [doi: 10.1017/gmh.2021.28] [Medline: 34512999]

35.   Alvarez-Jimenez M, Bendall S, Lederman R, Wadley G, Chinnery G, Vargas S, et al. On the HORYZON: moderated online social therapy for long-term recovery in first episode psychosis. Schizophr Res 2013 Jan;143(1):143-149. [doi: 10.1016/j.schres.2012.10.009] [Medline: 23146146]

36.   Peters-Strickland T, Pestreich L, Hatch A, Rohatagi S, Baker RA, Docherty J, et al. Usability of a novel digital medicine system in adults with schizophrenia treated with sensor-embedded tablets of aripiprazole. NDT 2016 Oct;Volume 12:2587-2594. [doi: 10.2147/ndt.s116029]

37.   Fowler JC, Cope N, Knights J, Fang H, Skubiak T, Shergill SS, et al. Hummingbird Study: Results from an Exploratory Trial Assessing the Performance and Acceptance of a Digital Medicine System in Adults with Schizophrenia, Schizoaffective Disorder, or First-Episode Psychosis. NDT 2021 Feb;Volume 17:483-492. [doi: 10.2147/ndt.s290793]

38.   Peters-Strickland T, Hatch A, Adenwala A, Atkinson K, Bartfeld B. Human factors evaluation of a novel digital medicine system in psychiatry. Neuropsychiatr Dis Treat 2018;14:553-565 [FREE Full text] [doi: 10.2147/NDT.S157102] [Medline: 29497299]

39.   Fowler JC, Cope N, Knights J, Phiri P, Makin A, Peters-Strickland T, et al. Hummingbird Study: a study protocol for a multicentre exploratory trial to assess the acceptance and performance of a digital medicine system in adults with schizophrenia, schizoaffective disorder or first-episode psychosis. BMJ Open 2019 Jun 27;9(6):e025952 [FREE Full text] [doi: 10.1136/bmjopen-2018-025952] [Medline: 31253613]

40.   Alipour A, Gabrielson S, Patel PB. Ingestible Sensors and Medication Adherence: Focus on Use in Serious Mental Illness. Pharmacy (Basel) 2020 Jun 16;8(2):103 [FREE Full text] [doi: 10.3390/pharmacy8020103] [Medline: 32560178]

41.   Alvarez-Jimenez M, Koval P, Schmaal L, Bendall S, O'Sullivan S, Cagliarini D, et al. The Horyzons project: a randomized controlled trial of a novel online social therapy to maintain treatment effects from specialist first-episode psychosis services. World Psychiatry 2021 Jun;20(2):233-243 [FREE Full text] [doi: 10.1002/wps.20858] [Medline: 34002511]

42.   Fleischhacker WW. Aripiprazole. Expert Opin Pharmacother 2005 Oct 03;6(12):2091-2101. [doi: 10.1517/14656566.6.12.2091] [Medline: 16197361]

43.   Highlights of prescribing information for ABILIFY MYCITE. US Food and Drug Administration. URL: https://www.accessdata.fda.gov/drugsatfda_docs/label/2017/207202lbl.pdf [accessed 2022-10-18]

44.   Your guide to starting the ABILIFY MYCITE® System. AbilifyMyCite. URL: https://www.abilifymycite.com/starting-mycite [accessed 2022-10-18]

45.   Weiden P. Understanding and addressing adherence issues in schizophrenia: from theory to practice. J Clin Psychiatry 2007;68 Suppl 14:14-19. [Medline: 18284273]

46.   Quirk A, Chaplin R, Hamilton S, Lelliott P, Seale C. Communication about adherence to long-term antipsychotic prescribing: an observational study of psychiatric practice. Soc Psychiatry Psychiatr Epidemiol 2013 Apr 25;48(4):639-647. [doi: 10.1007/s00127-012-0581-y] [Medline: 23007293]

47.   Horvath AO, Greenberg LS, editors. The Working Alliance: Theory, Research, and Practice. New York, US: John Wiley & Sons; 1994.

48.   Harbishettar V, Krishna K, Srinivasa P, Gowda M. The enigma of doctor-patient relationship. Indian J Psychiatry 2019;61(10):776. [doi: 10.4103/psychiatry.indianjpsychiatry_96_19]

49.   Chipidza FE, Wallwork RS, Stern TA. Impact of the Doctor-Patient Relationship. Prim. Care Companion CNS Disord 2015 Oct 22:13-37. [doi: 10.4088/pcc.15m01840]

50.   Brown MT, Bussell JK. Medication adherence: WHO cares? Mayo Clin Proc 2011 Apr;86(4):304-314 [FREE Full text] [doi: 10.4065/mcp.2010.0575] [Medline: 21389250]

51.   Steiner JF, Earnest MA. The language of medication-taking. Ann Intern Med 2000 Jun 06;132(11):926-930. [doi: 10.7326/0003-4819-132-11-200006060-00026] [Medline: 10836931]

52.   Osterberg L, Blaschke T. Adherence to medication. N Engl J Med 2005 Aug 04;353(5):487-497. [doi: 10.1056/NEJMra050100] [Medline: 16079372]

53.   Frank AF, Gunderson JG. The role of the therapeutic alliance in the treatment of schizophrenia. Relationship to course and outcome. Arch Gen Psychiatry 1990 Mar 01;47(3):228-236. [doi: 10.1001/archpsyc.1990.01810150028006] [Medline: 1968329]

54.   McCabe R, Bullenkamp J, Hansson L, Lauber C, Martinez-Leal R, Rössler W, et al. The therapeutic relationship and adherence to antipsychotic medication in schizophrenia. PLoS One 2012 Apr 27;7(4):e36080 [FREE Full text] [doi: 10.1371/journal.pone.0036080] [Medline: 22558336]

55.   Priebe S, Gruyters T. The role of the helping alliance in psychiatric community care. A prospective study. J Nerv Ment Dis 1993 Sep;181(9):552-557. [doi: 10.1097/00005053-199309000-00004] [Medline: 8245923]

56.   Prictor M, Lewis MA, Newson AJ, Haas M, Baba S, Kim H, et al. Dynamic Consent: An Evaluation and Reporting Framework. J Empir Res Hum Res Ethics 2020 Jul 15;15(3):175-186. [doi: 10.1177/1556264619887073] [Medline: 31729900]

57.   AbilifyMyCite. URL: https://www.otsuka-us.com/products-solutions-and-patient-support/abilify-mycite-information [accessed 2022-08-07]

58.   Vayena E, Blasimme A. Biomedical Big Data: New Models of Control Over Access, Use and Governance. J Bioeth Inq 2017 Dec 5;14(4):501-513 [FREE Full text] [doi: 10.1007/s11673-017-9809-6] [Medline: 28983835]

Abbreviations

iDAS: ingestible-sensor–based digital adherence system
SMI: serious mental illness
TR: therapeutic relationship

XSL•FO
RenderX

Original Paper

Cross-Platform Detection of Psychiatric Hospitalization via Social Media Data: Comparison Study

Viet Cuong Nguyen[1], BSc; Nathaniel Lu[2,3], MA; John M Kane[2,3,4], MD; Michael L Birnbaum[2,3,4], MD; Munmun De Choudhury[1], PhD

[1]School of Interactive Computing, Georgia Institute of Technology, Atlanta, GA, United States

[2]Department of Psychiatry, The Zucker Hillside Hospital, Northwell Health, Glen Oaks, NY, United States

[3]The Feinstein Institute for Medical Research, Northwell Health, Manhasset, NY, United States

[4]The Donald and Barbara Zucker School of Medicine at Hofstra/Northwell, Hempstead, NY, United States

Corresponding Author:
Viet Cuong Nguyen, BSc
School of Interactive Computing
Georgia Institute of Technology
756 W Peachtree St NW
Atlanta, GA, 30318
United States
Phone: 1 404 279 2941
Email: johnny.nguyen@gatech.edu

Abstract

Background: Previous research has shown the feasibility of using machine learning models trained on social media data from a single platform (eg, Facebook or Twitter) to distinguish individuals either with a diagnosis of mental illness or experiencing an adverse outcome from healthy controls. However, the performance of such models on data from novel social media platforms unseen in the training data (eg, Instagram and TikTok) has not been investigated in previous literature.

Objective: Our study examined the feasibility of building machine learning classifiers that can effectively predict an upcoming psychiatric hospitalization given social media data from platforms unseen in the classifiers' training data despite the preliminary evidence on identity fragmentation on the investigated social media platforms.

Methods: Windowed timeline data of patients with a diagnosis of schizophrenia spectrum disorder before a known hospitalization event and healthy controls were gathered from 3 platforms: Facebook (254/268, 94.8% of participants), Twitter (51/268, 19% of participants), and Instagram (134/268, 50% of participants). We then used a 3×3 combinatorial binary classification design to train machine learning classifiers and evaluate their performance on testing data from all available platforms. We further compared results from models in intraplatform experiments (ie, training and testing data belonging to the same platform) to those from models in interplatform experiments (ie, training and testing data belonging to different platforms). Finally, we used Shapley Additive Explanation values to extract the top predictive features to explain and compare the underlying constructs that predict hospitalization on each platform.

Results: We found that models in intraplatform experiments on average achieved an F_1-score of 0.72 (SD 0.07) in predicting a psychiatric hospitalization because of schizophrenia spectrum disorder, which is 68% higher than the average of models in interplatform experiments at an F_1-score of 0.428 (SD 0.11). When investigating the key drivers for divergence in construct validities between models, an analysis of top features for the intraplatform models showed both low predictive feature overlap between the platforms and low pairwise rank correlation (<0.1) between the platforms' top feature rankings. Furthermore, low average cosine similarity of data between platforms within participants in comparison with the same measurement on data within platforms between participants points to evidence of identity fragmentation of participants between platforms.

Conclusions: We demonstrated that models built on one platform's data to predict critical mental health treatment outcomes such as hospitalization do not generalize to another platform. In our case, this is because different social media platforms consistently reflect different segments of participants' identities. With the changing ecosystem of social media use among different demographic groups and as web-based identities continue to become fragmented across platforms, further research on holistic approaches to harnessing these diverse data sources is required.

Introduction

Background

Despite its relatively low prevalence compared with other mental health disorders, the burden of schizophrenia spectrum disorder (SSD) on patients, families, and society is substantial [1]. To mitigate the burden of SSD, early diagnosis and treatment are crucial. However, psychotic disorders, including SSD, often receive delayed attention and care, resulting in worse health outcomes [2,3]. At the same time, the use of social media is high among patients with serious psychotic disorders such as SSD, especially among adolescents and young adults, when SSD typically emerges [4,5]. For instance, Birnbaum et al [4] studied social media use among adolescents and young adults with psychotic and mood disorders and found that 97.5% of participants (mean age 18.3 years) regularly used social media, spending approximately 2.6 (SD 2.5) hours per day on the web. Similarly, Miller et al [5] studied the use of digital technologies among patients diagnosed with SSD and found that, among participants with access to the internet, 98% reported using at least one social media service and 57% used social media daily.

Given this information, there has been an established body of research on using social media data to identify and predict psychiatric outcomes of social media users with SSD using machine learning classifiers [6-8]. The most robust data sources available to train these classifiers consist of textual content posted on the web. Prior work in speech and text analysis among patients with SSD has identified reliable linguistic markers associated with SSD, which have been successfully used as features for the aforementioned classifiers [7,9,10]. These include certain word frequencies, word categories, and self-referential pronouns [11,12]. Given that the use of image- and video-based social media platforms such as Instagram, Snapchat, and TikTok is associated with youths, there has also been prior work in the analysis of images comparing between patients with SSD and healthy controls [13,14]. Hänsel et al [14] identified additional image markers associated with SSD, such as the image's colorfulness and saturation and the average number of faces per image. By exploiting these markers, previous research conducted by Birnbaum et al [15] and Ernala et al [8] built classifiers to distinguish between users with a confirmed diagnosis of SSD and healthy controls on Facebook and Twitter with area under the receiver operating characteristic curve (AUROC) scores of 0.75 and 0.82, respectively.

Although such results demonstrate the potential of automated techniques in predicting the mental health outcomes of individuals with SSD via social media data, many research gaps remain that need to be addressed before psychiatrists can reliably deploy such techniques for clinical purposes. Most prior work in this area primarily focused on a single source of social media data, either exclusively from Twitter or Facebook, for downstream classification and analysis tasks [16]. However, previous research has also shown that many social media users, especially youths, use different social media platforms for different purposes because of their variety in affordances and culture. Among youths, Facebook use is associated with keeping up with close and distant friends, whereas Instagram and Snapchat use is associated with self-expression and gratification [17,18]. In addition, researchers have argued that social media users have fragmented identities across platforms [19,20]. Therefore, using a single source of social media data to build psychiatric hospitalization prediction models may potentially lead to low-sensitivity prediction models, making them unsuitable for clinical purposes. However, few studies have quantified the extent to which classifiers trained on data from one social media platform are generalizable to other platforms. To this end, our study aimed to measure the generalizability of social media–based classifiers aimed at predicting upcoming psychiatric hospitalizations to data from unseen social media platforms. In addition, we aimed to surface any evidence of the differing fragmented identities that are reflected on 3 popular social media platforms—Twitter, Facebook, and Instagram—that might affect the models' generalizability.

Objectives

The research question we attempted to answer was as follows: given the preliminary evidence of fragmented identities that are reflected on the investigated social media platforms, can we build classifiers that can effectively detect users at risk of an upcoming psychiatric hospitalization using social media data from platforms unseen in the training data?

To answer our research question, we collated textual and image content (if available) from consenting participants' social media data from Facebook, Twitter, and Instagram. We then trained platform-specific classifiers to distinguish between social media data from healthy controls and data from patients with SSD with an upcoming psychiatric hospitalization. We compared the performance of classifiers on testing data between seen and unseen social media platforms from the training data. We also compared and analyzed the top predictive features and the feature importance distributions between the 3 platform-specific classifiers, with a view toward finding potential empirical evidence for fragmented identities between the various social media platforms.

Methods

Recruitment

We recruited participants clinically diagnosed with SSD and clinically verified healthy controls aged between 15 and 35 years. These data were collected as part of a broader research initiative involving the authors of this paper to identify technology-based health information to provide early identification, intervention, and treatment for young adults with SSD [6].
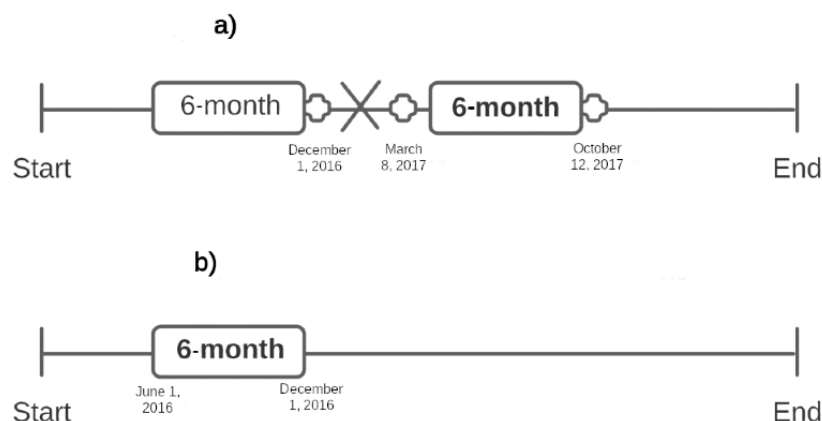
For participants with SSD aged between 15 and 35 years (141/268, 52.6%), diagnoses were based on clinical assessment of the most recent episode and were extracted from participants' medical records at the time of their consent. Participants in this group were recruited from the Northwell Health Zucker Hillside Hospital and collaborating institutions located in East Lansing, Michigan. Participants were excluded if they had an IQ of <70 (per clinical assessment), autism spectrum disorder, or substance-induced psychotic disorder.

In addition, healthy volunteers aged between 15 and 35 years (127/268, 47.4%) were approached and recruited from an existing database of eligible individuals who had already been screened for previous research projects at Zucker Hillside Hospital and had agreed to be recontacted for additional research opportunities. Healthy status was determined by either the Structured Clinical Interview for the Diagnostic and Statistical Manual of Mental Disorders conducted within the past 2 years or the Psychiatric Diagnostic Screening Questionnaire [21,22]. Participants were excluded if clinically significant psychiatric symptoms were identified during the screening process. Additional healthy volunteers were recruited from a southeastern university via a web-based student community research recruitment site. Finally, healthy volunteers were also recruited from the collaborating institutions located in East Lansing, Michigan.

Data Collection

All consenting participants were asked to download and share their Facebook, Twitter, and Instagram data archives. We collected all linguistic content from participants' Facebook and Twitter archives (ie, status updates and comments on Facebook and posts shared on Twitter). In addition, we collected image content from participants' Facebook and Instagram archives, including profile pictures and story photos.

Next, we also collected the medical history of each participant (following consent and adoption of Health Insurance Portability and Accountability Act–compliant policies). This included primary and secondary diagnosis codes, the total number of hospitalizations, and admission and discharge dates for each hospitalization event. Hospitalization data were collected from the medical records at the time of consent. As all consented patient participants in the study had also received care at the Zucker Hillside Hospital, the medical records at the hospital were accurate and up to date to the best of the hospital's efforts. We only counted psychiatric hospitalizations (not hospitalizations for other nonpsychiatric reasons). Thereafter, the study team accessed the corresponding consented patients' medical records to extract all their recorded hospitalization events in a similar manner to previous studies using this source of data [6,23].

Finally, we collected social media data from all available platforms for each participant with at least one known hospitalization event within a 6-month window before the latest hospitalization event, ensuring that there were no hospitalization events within these 6 months. This was done to ensure that the data gathered were representative of the participants' healthy mental status before symptomatic exacerbation and subsequent hospitalization. A 6-month period, which we refer to as the *windowed data*, was selected as it represents an interval of time long enough to identify changes signaling symptomatic exacerbation while also containing sufficient data required to train machine learning models. For healthy control participants without any hospitalizations, we randomly sampled a nonempty 6-month window of social media data for each available social media platform (nonempty meaning that there was at least some social media activity). Figure 1 provides a visual description of the windowing process.

Figure 1. Diagram representing the windowing process used to gather participants' social media data before hospitalization events. Bold text represents the selected data windows. Crosses represent hospitalization events. The X represents invalid data windows. A: Windowing—with hospitalizations; B: Windowing—without hospitalizations.

Feature Engineering

To encode participants' social media data for the downstream classification and analysis tasks outlined in our research objectives, we identified and extracted the following categories of features from these data for all 3 investigated social media platforms: (1) n-gram language features (n=500), (2) Linguistic Inquiry and Word Count (n=78), (3) lexico-semantic features

(n=3), (4) activity features (n=9), and (5) image features (n=23; Instagram and Facebook only).

The specific feature categories were chosen based on relevant previous literature, particularly relating to the use of social media data to infer mental health attributes and psychiatric outcomes [7,8]. Note that all features were computed at the

individual participant level. More details about this process can be found in Multimedia Appendix 1 [7,12,14,24-29].

Feature Selection

Using the aforementioned features, for each of the 3 examined social media platforms, we encoded available participants' textual and image data on Facebook and Instagram into 613-dimensional feature vectors and textual data on Twitter into 590-dimensional feature vectors. This yielded a Facebook data set of dimension 254×613, a Twitter data set of dimension 51×590, and an Instagram data set of dimension 134×613. We shall refer to these data sets as F, T, and I for Facebook, Twitter, and Instagram, respectively.

As the feature set might contain features that are noisy and irrelevant, the classification models may be unstable and produce suboptimal results [30]. To maximize the predictive power of the models while also reducing the redundancy and computational resources needed to train them, feature selection methods were used [30]. More specifically, we adopted the ANOVA F test to rank the features based on their F statistic under the test, which has been shown to produce optimal feature sets in previous research on the classification of social media data belonging to patients with SSD [8,11].

We trained a random forest model, with 5-fold stratified cross-validation to fine-tune hyperparameters, on data sets F, T, and I with an 80:20 train-test split, using only the top k percent of features based on the ranking given by the ANOVA F test on the classification, where k is between 10 and 100 in increments of 10. Via an examination of the evaluation metrics on the test sets (described in the Classification Algorithms and Metrics section), we determined that using only the top 20% of the features (based on their F statistic under the ANOVA F test) yielded the best results on unseen data across all 3 platforms. We will be using this subset of features moving forward.

Combinatorial Classification Methods

To answer the research question laid out in the Introduction section, we adopted a 3×3 combinatorial classification design, where we trained and tested machine learning models on the psychiatric hospitalization prediction task using all possible pairs of training and testing data sets. Figure 2 provides a visual description of our experimental design. For intraplatform experiments (where the training and testing data came from the same platform; eg, training and testing on Facebook data), we trained and tested the models on an 80 to 20 train-test label-stratified split based on the Scikit-learn *train_test_split()* function (version 0.24.1) [31]. For interplatform experiments (where the training and testing data came from different platforms; eg, training on Facebook data and testing on Instagram data), we trained the model on the entirety of the training data set and evaluated it on the entirety of the testing data set.

Figure 2. Diagram representing the classification experiments performed and their nature within the 3×3 combinatorial design.
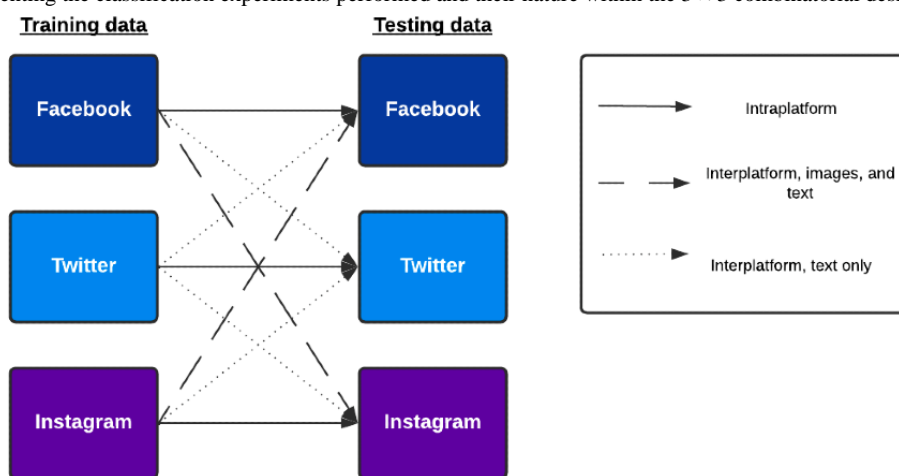
Classification Algorithms and Metrics

For both intra- and interplatform experiments, training data represented by the top 20% of features (as described in the Feature Selection section) were fed into a model to learn the classification task. We tried training the model over several algorithms, including random forest, logistic regression, support vector machine, and multilayer perceptron [32]. We selected these algorithms as they represented a variety of different types of learning algorithms [32]. This ensured that our analysis of performance differences between intra- and interplatform experiments would hold irrespective of the learning algorithm selection. We used the Scikit-learn implementation (version 0.24.1) for all the aforementioned algorithms [31]. For each algorithm, we fine-tuned its hyperparameters using 5-fold stratified cross-validation via the Scikit-learn *GridSearchCV()* pipeline, retaining the best hyperparameters per algorithm for analysis [31]. The chosen hyperparameters for each classification algorithm are provided in Textbox 1 (all other hyperparameters were left as default according to the Scikit-learn specification).

We measured the performance of the models using the metrics outlined in Textbox 2, all of which are commonly used in binary classification models. In this case, we abbreviated the number of true positives, true negatives, false positives, and false negatives as TP, TN, FP, and FN, respectively [33].

Textbox 1. Hyperparameters chosen for each classification algorithm.

Random forest

- max_depth: 15

- n_estimators: 100

- max_features: none

Logistic regression

- Penalty: l2

- C: 0.1

Support vector machine

- Kernel: rbf

- C: 0.01

- Gamma: scale

Multilayer perceptron

- Alpha: 0.0001

- Hidden_layer_sizes: (512, 256, 128)

Textbox 2. Metrics used to measure model performance.

Accuracy

- Also known as Rand accuracy, the ratio of correct predictions to all predictions

-

Precision

- The ratio of correct positive predictions to the total number of positive predictions

-

Recall

- The ratio of correct positive predictions to the total number of true positive instances

-

F_1-**score**

- The harmonic mean between precision and recall

-

Area under the receiver operating characteristic curve (AUROC)

- The AUROC, which plots the false positive rate against the true positive rate and, in practice, is often estimated using the trapezoidal rule with the following formula:

-

Feature Importance Selection

We used Shapley Additive Explanations (SHAP) to examine how certain features affected our model's decision to predict users with potential psychiatric hospitalization because of SSD given their social media data from the 3 inspected social media platforms. Our decision to use SHAP rather than other explainability methods stems from the fact that SHAP is not

only model-agnostic but also the most theoretically sound explainability framework among the available options. This is because SHAP feature scores can be calculated for localized samples and for the entire global data set [34]. SHAP is based on Shapley values, a game-theoretical concept that intuitively describes each feature's contribution to the outcome after considering all possible combinations of features [35].

For each of the intraplatform experiments within the 3×3 combinatorial design and each machine learning model, we calculated the average SHAP values for each of the features (ie, their importance to the prediction) across all instances within the testing set. We then recorded the list of features sorted in descending order according to the average SHAP values measured by each model. In the case of models with native support for feature importance extraction, including random forest (Gini importance) and logistic regression (feature coefficients), we also calculated and recorded them in an equivalent manner to SHAP values.

Robustness Checks

To ensure that our findings regarding differences in model performance between models and between intra- and interplatform experiments still held when certain aspects of the training and testing data sets were made more ideal, we performed several robustness checks, which are described in Multimedia Appendix 1.

Ethics Approval

The study was approved by the institutional review board of Northwell Health (the coordinating institution) and the institutional review board of the participating partners (Georgia Tech approval H21403). Participants were recruited from June 23, 2016, to December 4, 2020. Written informed consent was obtained from adult participants and legal guardians of participants aged <18 years. Assent was obtained from participating minors.

Results

Data Characteristics

In total, 268 participants (mean age 24.73, SD 5.64 years; male: 127/268, 47.4%; SSD: 141/268, 52.6%) with nonempty windowed data for at least one platform were included. Of these 268 participants, 254 (94.8%; SSD: 133/254, 52.4%) had valid windowed Facebook data, 51 (19%; SSD: 7/51, 13.7%) had valid windowed Twitter data, and 134 (50%; SSD: 42/134, 31.3%) had valid windowed Instagram data. Among participants with valid data for more than one platform, 17.5% (47/268; SSD: 5/47, 10.6%) had valid data for both Facebook and Twitter, 14.2% (38/268; SSD: 4/38, 10.5%) had valid data for both Twitter and Instagram, and 44.4% (119/268; SSD: 34/119, 28.6%) had valid data for both Facebook and Instagram. Finally, 14.2% (38/268; SSD: 4/38, 10.5%) of participants had valid data for all 3 platforms. Table 1 shows the demographic and clinical characteristics of these 268 participants. Table 2 describes the summary statistics, including mean and median, for these windowed data for each of the 3 social media platforms grouped by clinical status (SSD vs control). Figure 3 shows the distribution of available posts for participants in each of the 3 investigated platforms.

Table 1. Demographic and clinical characteristics of the participants (N=268).

| Characteristic	SSD[a] (n=141)	Control (n=127)	Full sample
Age (years), mean (SD)	24.86 (5.49)	24.57 (5.82)	24.73 (5.64)
Sex, n (%)			
Male	89 (63.1)	38 (29.9)	127 (47.4)
Female	52 (36.9)	89 (70.1)	141 (52.6)
Race or ethnicity, n (%)			
African American or Black	64 (45.4)	19 (15)	83 (31)
Asian	20 (14.2)	23 (18.1)	43 (16)
White	37 (26.2)	75 (59.1)	112 (41.8)
Mixed race or other	15 (10.6)	5 (3.9)	20 (7.5)
Hispanic	5 (3.5)	4 (3.1)	9 (3.4)
Pacific Islander	0 (0)	1 (0.8)	1 (0.4)
Primary diagnosis, n (%)			
Schizophrenia	67 (47.5)	N/A[b]	67 (25)
Schizophreniform	26 (18.4)	N/A	26 (9.7)
Schizoaffective	25 (17.7)	N/A	25 (9.3)
Unspecified SSDs	23 (16.3)	N/A	23 (8.6)
No diagnosis	N/A	127 (100)	127 (47.4)

[a]SSD: schizophrenia spectrum disorder.

[b]N/A: not applicable.

XSL·FO

RenderX

Table 2. Summary statistics for windowed data for both the control class and the schizophrenia spectrum disorder (SSD) class (ie, participants hospitalized with SSD). In this table, we consider data from Facebook, Twitter, and Instagram, as mentioned previously.

	Facebook (user: n=254; post: n=169,425)		Twitter (user: n=51; post: n=23,777)		Instagram (user: n=134; post: n=23,551)	
	SSD class	Control class	SSD class	Control class	SSD class	Control class
Total users, n (%)	133 (52)	121 (48)	7 (14)	44 (86)	42 (31)	92 (69)
Total posts, n (%)	114,793 (68)	54,632 (32)	991 (4)	22,786 (96)	7111 (30)	16,440 (70)
Posts, mean (SD)	863.1 (2365.1)	451.5 (818.87)	141.6 (255)	519.9 (1166.9)	169.3 (445.4)	178.7 (234.6)
Posts, median	260	184	37	138	54.5	103
Posts, range	2-23,589	1-4852	1-758	1-7056	1-2909	1-1328

Figure 3. Cumulative distribution function (CDF) curves of users and their number of posts for the schizophrenia spectrum disorder and control classes per data set: (A) Facebook (left), (B) Twitter (center), and (C) Instagram (right).
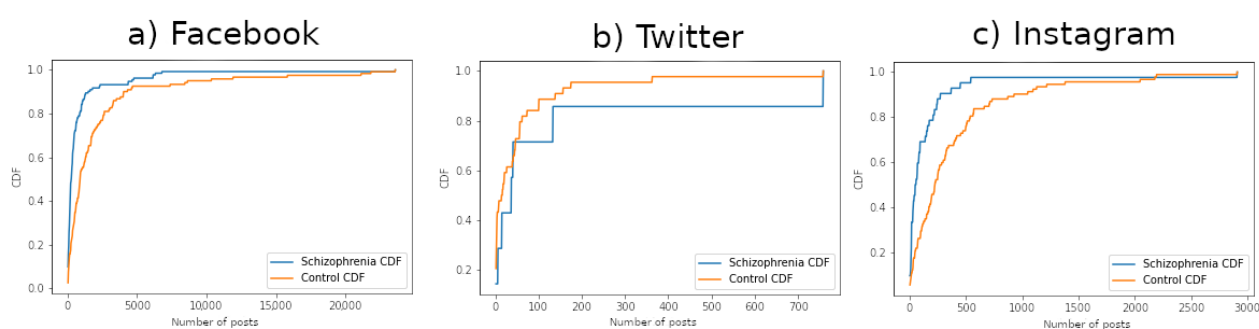
Results of Combinatorial Classification

We report the full results of the intraplatform experiments in Table 3. We also report the full results of the interplatform experiments in Tables 4 to 6. Finally, we report the receiver operating characteristic curves for the best-performing logistic regression model for the experiments from Tables 3 to 6 in Figure 4.

Elaborating on the results from Table 3, we found that, among the 4 classification algorithms that we used, the logistic regression model performed the best across the 3 intraplatform experiments, with the best performances for all of them. More elaborately, for the intraplatform experiments, performance reached its peak with the logistic regression model with an average F_1-score of 0.72 (SD 0.07), accuracy of 0.81 (SD 0.08), and AUROC of 0.749 (SD 0.06). In contrast, the worst-performing model (in this case, multilayer perceptron) achieved an average F_1-score of 0.521 (SD 0.19), accuracy of 0.714 (SD 0.19), and AUROC of 0.623 (SD 0.16) for the intraplatform experiments. Thus, we will be using the logistic regression model for further analysis regarding feature importance between platforms. These results align with previous

research and, thus, could be considered a soft replication of those findings [8,15].

By contrast, by aggregating the metrics for the interplatform experiments presented in Tables 4 to 6, the average F_1-score decreased to 0.428 (SD 0.11), accuracy decreased to 0.559 (SD 0.06), and AUROC decreased to 0.533 (SD 0.03) for the logistic regression model. This constitutes, on average, a drop of 40%, 31.4%, and 28.8% in F_1-score, accuracy, and AUROC score, respectively, from the intraplatform experiments. As just demonstrated, when comparing the effectiveness of models between intraplatform and interplatform experiments, we found a consistent drop in performance for all the investigated social media platforms. The drop in test F_1-score, given the best-performing logistic regression model, was the most drastic for Facebook at 0.364 (46%) and least drastic for Twitter at 0.08 (14%), averaging a drop of 0.285 (40%, SD 0.13) going from 0.713 for intraplatform experiments to 0.428 for interplatform experiments. Such trends hold even when disparities in data set size and dual-platform data availability (as described in the Methods section under Robustness Checks) are applied to the training and testing data (Multimedia Appendix 1).

Table 3. Classification results for all intraplatform classification experiments. In this table, for instance, Facebook indicates the Facebook-Facebook experiment.

Model	Facebook					Twitter					Instagram				
	Acc[a]	P[b]	R[c]	F_1	AUROC[d]	Acc	P	R	F_1	AU-ROC	Acc	P	R	F_1	AU-ROC
Random forest	0.739	0.739	0.738	0.738	0.709	0.745	0.150	0.116	0.116	0.494	0.7	0.648	0.637	0.637	0.681
SVM[e]	0.722	0.747	0.692	0.715	0.723	0.854	0.541	0.45	0.463	0.697	0.740	0.737	0.757	0.743	0.805
MLP[f]	0.506	0.406	0.507	0.367	0.516	0.845	0.458	0.45	0.426	0.692	0.792	0.771	0.794	0.77	0.840
Logistic regression	0.759	0.767	0.758	0.756	0.727	0.881	0.742	0.6	0.63	0.772	0.792	0.771	0.801	0.773	0.848

[a]Acc: accuracy.

[b]P: precision.

[c]R: recall.

[d]AUROC: area under the receiver operating characteristic curve.

[e]SVM: support vector machine.

[f]MLP: multilayer perceptron.

Table 4. Classification results for the interplatform classification experiments for Facebook training data.

Model	Twitter					Instagram				
	Acc[a]	P[b]	R[c]	F_1	AUROC[d]	Acc	P	R	F_1	AUROC
Random forest	0.392	0.221	0.88	0.354	0.579	0.379	0.328	0.952	0.488	0.537
SVM[e]	0.545	0.253	0.72	0.373	0.612	0.432	0.337	0.860	0.483	0.550
MLP[f]	0.587	0.240	0.55	0.334	0.573	0.435	0.332	0.812	0.471	0.539
Logistic regression	0.628	0.246	0.47	0.323	0.567	0.472	0.344	0.775	0.476	0.555

[a]Acc: accuracy.

[b]P: precision.

[c]R: recall.

[d]AUROC: area under the receiver operating characteristic curve.

[e]SVM: support vector machine.

[f]MLP: multilayer perceptron.

Table 5. Classification results for the interplatform classification experiments for Twitter training data.

Model	Facebook					Instagram				
	Acc[a]	P[b]	R[c]	F_1	AUROC[d]	Acc	P	R	F_1	AUROC
Random forest	0.531	0.569	0.378	0.452	0.536	0.628	0.331	0.207	0.252	0.512
SVM[e]	0.514	0.53	0.537	0.530	0.513	0.563	0.340	0.42	0.373	0.523
MLP[f]	0.533	0.561	0.440	0.492	0.536	0.557	0.325	0.395	0.356	0.512
Logistic regression	0.534	0.552	0.522	0.535	0.535	0.578	0.362	0.47	0.408	0.548

[a]Acc: accuracy.

[b]P: precision.

[c]R: recall.

[d]AUROC: area under the receiver operating characteristic curve.

[e]SVM: support vector machine.

[f]MLP: multilayer perceptron.

XSL•FO

RenderX

Table 6. Classification results for the interplatform classification experiments for Instagram training data.

Model	Facebook					Twitter				
	Acc[a]	P[b]	R[c]	F_1	AUROC[d]	Acc	P	R	F_1	AUROC
Random forest	0.51	0.523	0.612	0.563	0.507	0.751	0.369	0.42	0.386	0.624
SVM[e]	0.524	0.544	0.51	0.524	0.525	0.691	0.213	0.25	0.229	0.521
MLP[f]	0.554	0.584	0.48	0.526	0.557	0.683	0.201	0.23	0.214	0.51
Logistic regression	0.516	0.524	0.689	0.595	0.51	0.628	0.256	0.52	0.342	0.587

[a]Acc: accuracy.

[b]P: precision.

[c]R: recall.

[d]AUROC: area under the receiver operating characteristic curve.

[e]SVM: support vector machine.

[f]MLP: multilayer perceptron.

Figure 4. Receiver operating characteristic (ROC) curves for the classification experiments given the best logistic regression model. (A), (B), and (C) are curves for the Facebook, Twitter, and Instagram intraplatform results, respectively, from Table 3. (D) and (E) are the ROC curves for the interplatform experiments from Table 4, where Facebook was used as the training data.
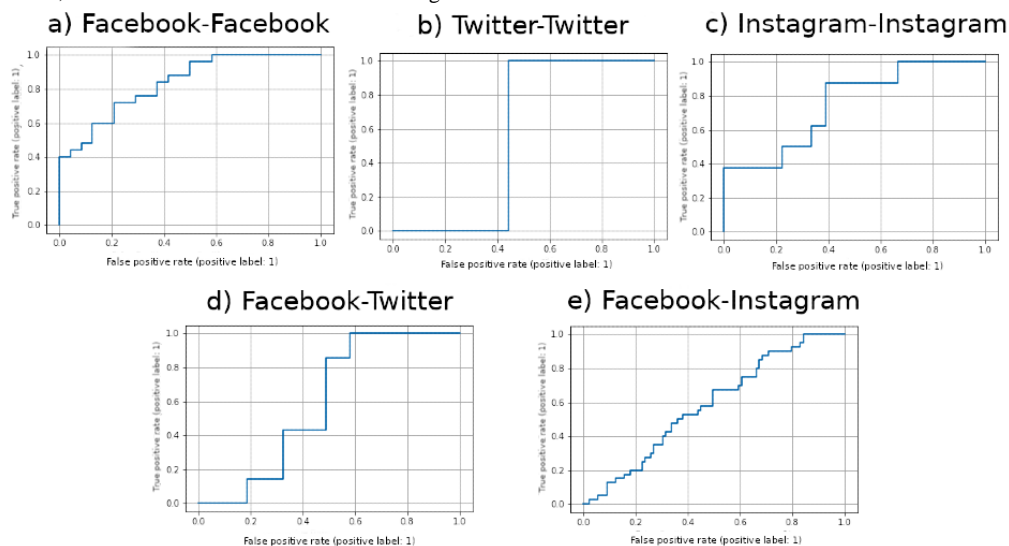
Feature Importance Analysis

We hypothesized that the decrease in performance from intraplatform experiments to interplatform experiments, as presented previously, was driven by differences in feature importance learned by models when trained on data from different social media platforms (even when they shared the same feature set). By extracting the list of SHAP features from the models per the method described previously, we found support for this hypothesis. Specifically, we observed little overlap between them across platforms among the top 25 features for each model and platform (when holding the model constant). On average, there were only 4.66 overlapping features for the same logistic regression classification model across platforms (the best-performing model based on the previous discussions). In addition, we found that the lists of feature importance for each of the platforms, based on the logistic regression model, had very weak rank correlation pairwise. Fully elaborating on the statistical results for the Kendall rank correlation coefficient, we found very weak rank correlations between the ranked lists of feature importance for Facebook and Twitter (τ_b=0.081; P=.003), Facebook and Instagram (τ_b=0.041; P=.01), and Twitter and Instagram (τ_b=0.055; P=.05). We report the average SHAP values and logistic regression coefficients of the top 10 features based on their SHAP values, along with their average value in the SSD class and the control class, in Table 7.

Table 7. Top 10 features for the logistic regression (LR) model for each of the platforms (Linguistic Inquiry and Word Count features are italicized) based on their Shapley Additive Explanations (SHAP) values.

Platform and feature acronym	Feature description	SHAP value	LR coefficient	SSD[a] group average (SD)	Control group average (SD)
Facebook					
Avg_post_readability	Average post readability, as measured using the SMOG[b] index	0.761	−0.268	5.6341 (2.74)	6.8048 (1.92)
Quant	Ratio of words within the "quantifiers" category	0.4195	−0.189	0.0012 (0.0012)	0.0016 (0.0012)
Negemo	Ratio of words within the "negative emotions" category	0.0953	0.244	0.0043 (0.0035)	0.0031 (0.0022)
Money	Ratio of words within the "money" category	0.0739	−0.216	0.0007 (0.001)	0.0011 (0.002)
Swear	Ratio of words within the "swear" category	0.0628	0.236	0.0017 (0.0025)	0.0007 (0.001)
Ratio_octile8	Ratio of activities from 9 PM to midnight	0.0443	0.077	0.1443 (0.149)	0.1241 (0.158)
Ratio_octile7	Ratio of activities from 6 PM to 9 PM	0.0409	0.177	0.1561 (0.1745)	0.1054 (0.125)
Anger	Ratio of words within the "anger" category	0.0095	0.191	0.0018 (0.002)	0.0009 (0.001)
Dream	Ratio of "dream" within the overall bag of words	0.0077	0.224	0.2028 (0.468)	0.0746 (0.24)
Fun	Ratio of "fun" within the overall bag of words	0.0043	−0.209	0.5722 (1.19)	1.1315 (1.76)
Twitter					
Conj	Ratio of words within the "conjunctions" category	0.2319	−0.063	0.0001 (0.0002)	0.0003 (0.0004)
Adj	Ratio of words within the "adjectives" category	0.1825	−0.05	0.0057 (0.004)	0.0080 (0.005)
Avg_post_negativity	Average post negativity, as calculated using the VADER[c] library	0.1509	0.082	0.071 (0.042)	0.0519 (0.036)
Male	Ratio of words within the "male" category	0.1355	0.039	0.0011 (0.0013)	0.0007 (0.001)
Ratio_octile_8	Ratio of activities from 9 PM to midnight	0.1265	0.045	0.0231 (0.356)	0.1227 (0.188)
Ingest	Ratio of words within the "ingest" category	0.0627	−0.056	0.0003 (0.0007)	0.0014 (0.0018)
Insight	Ratio of words within the "insight" category	0.0516	0.053	0.0044 (0.004)	0.0035 (0.003)
Power	Ratio of words within the "power" category	0.0308	−0.058	0.0024 (0.0026)	0.0042 (0.0036)
We	Ratio of words within the "we" category	0.0196	−0.056	0.0001 (0.0002)	0.0002 (0.0004)
Prep	Ratio of words within the "prepositions" category	0.0117	0.063	0.0028 (0.0026)	0.0017 (0.0017)
Instagram					
Avg_post_readability	Average post readability, as measured using the SMOG index	0.761	−0.203	5.1018 (1.15)	6.2564 (1.638)

XSL•FO
RenderX

Platform and feature acronym	Feature description	SHAP value	LR coefficient	SSD[a] group average (SD)	Control group average (SD)
Space	Ratio of words within the "space" category	0.733	−0.147	0.0031 (0.0025)	0.0042 (0.0025)
Affiliation	Ratio of words within the "affiliation" category	0.6839	−0.181	0.0032 (0.0027)	0.0056 (0.0034)
Friend	Ratio of words within the "friend" category	0.5336	−0.159	0.0009 (0.0027)	0.0018 (0.0034)
Female	Ratio of words within the "female" category	0.4576	−0.168	0.0008 (0.001)	0.0019 (0.0023)
Sad	Ratio of words within the "sad" category	0.4554	0.113	0.0011 (0.0008)	0.0007 (0.0012)
Quant	Ratio of words within the "quantifier" category	0.4195	−0.118	0.0012 (0.0013)	0.0019 (0.0016)
Away	Ratio of "away" within the overall bag of words	0.4064	−0.105	0.0768 (0.276)	0.2505 (0.5)
Assent	Ratio of words within the "assent" category	0.3913	−0.102	0.0008 (0.0012)	0.0013 (0.0014)
Next	Ratio of "next" within the overall bag of words	0.3854	−0.12	0.0957 (0.267)	0.6466 (1.236)

[a]SSD: schizophrenia spectrum disorder.

[b]SMOG: Simple Measure of Gobbledygook.

Attributing Divergent Construct Validity of Models to Divergent Identities on the Web

What could explain the observed differences in construct validities of the intraplatform models? Early in this paper, we posited that these differences might stem from people's identities being fragmented across different platforms. To situate that these divergent identities are indeed the drivers behind differential cross-platform model construct validities and, by extension, performance, we adopted a strategy to measure the differences within the extracted feature space between the investigated platforms for a given participant. As social media data for participants on all platforms are encoded via feature vectors in this study, we calculated the pairwise similarity between platform-specific data using cosine similarity [36]. More specifically, we calculated the average cosine similarity within participants between platforms and compared it with the average cosine similarity between participants within platforms for participants with SSD with data on all 3 platforms. Given that, even within the same social media platform, different people can have unique modes of expressing their identities, we used the latter as a baseline for assessing whether fragments of identities representing an individual across platforms diverge more or less than the divergence of identities between individuals.

We found that the average between-platform, within-participant cosine similarity was 0.3093 for Facebook-Twitter, 0.2304 for Facebook-Instagram, and 0.3905 for Twitter-Instagram. This was either lower than or similar to the average within-platform, between-participant cosine similarity for the investigated platforms: 0.5072 for Facebook, 0.5427 for Twitter, and 0.373 for Instagram. The same trend holds even when calculating the averages using data from both participants with SSD and healthy controls with data from all 3 platforms.

Discussion

Principal Findings

Our study aimed to measure the ability (or inability) of mental health classifiers to generalize across platforms and surface evidence of fragmented identities on social media among patients with SSD. Overall, we found that, across the board, models trained on data from social media platforms have poor generalizability when evaluated on data from other social media platforms even when holding the feature set constant across training and testing data. This trend holds true even in the 2 robustness tests, where the same participants and data set size were used in the training and testing data (as described in the Methods section). This trend is also true even when the training data come from a platform with high data availability and the testing data come from a platform with low data availability. For instance, the best F_1-score of the intraplatform models for Twitter (0.63) was 0.257 (69%) higher compared with the best F_1-score of the interplatform models for Twitter, where the training data came from Facebook (0.373).

Next, we discuss the findings regarding feature importance in more detail. First, looking at the theoretical validity of the top 10 features per platform and interpretation of the sign of the features' logistic regression coefficient, we found alignment with previous literature and evidence of clinical meaningfulness [7,8,11]. For instance, given the positive coefficient from the trained logistic regression model presented in Table 7, higher levels of use of lexicon indicative of negative emotions are highly predictive of SSD for Facebook (see the example post in Textbox 3 highlighting words such as "fear," "fail," and "hurts"). This confirms literature noting that a reduced ability to feel or express pleasure (anhedonia) is common in patients with SSD [37]. Similarly, previous research has found

anger-related terms commonly appearing in social media posts before the onset of early psychosis as well as preceding a psychiatric hospitalization [38]. This may explain why higher levels of use of lexicon indicative of the Linguistic Inquiry and Word Count category *Anger* are also highly predictive of SSD for Facebook (example post in Textbox 3 containing *Anger* words such as "shit" and "fucking"). Finally, words and phrases such as those in the Linguistic Inquiry and Word Count *Sad* category (eg, "useless," "sorry," and "sob") point to typical negative symptoms of SSD [39]. They can be indicative of a decreased sense of purpose and a seeming lack of interest in the world [39]. Models trained on Instagram successfully picked up such cues from the posts, where higher use of such vocabulary was indicative of an impending psychiatric hospitalization because of SSD.

That said, each model corresponding to each platform seemed to pick up contrasting signals from its respective training data, which is why we note the low overlap in the aforementioned top SHAP features. Among the few that overlap in the top 10 features reported previously, we found "avg_post_readability" to be picked up as a highly predictive feature by both Facebook and Instagram models, whereas "ratio_octile8" was selected by both Facebook and Twitter models. In our case, "avg_post_readability" is calculated using the Simple Measure of Gobbledygook index, which approximates the years of education needed to fully comprehend a piece of written text. The negative logistic regression coefficient and the averages of the SSD and control groups for this feature suggest that texts written by patients with SSD are simpler in nature, which is indicative of language dysfunction. This is a known negative symptom of schizophrenia and related psychotic disorders, as

observed in prior work [40]. In addition, higher levels of late-night activity such as web or social media use, captured in the "ratio_octile8" feature, have been known to be associated with deteriorated mental health [41]. Finally, we found significant divergence in the distribution of feature importance between the platforms, as indicated by the low pairwise Kendall τ (<0.1) for the platforms' feature importance rankings. These qualitative and quantitative results broadly imply that the models were being trained on considerably different data sources with differing content and contexts of use, which likely contributed to poor cross-platform model generalization.

At the crux of these differences, we found that the models had inherently different construct validity across platforms. Data on each platform reflect only a segment of an individual's identity—a segment that may be absent in another platform. The fragmentation of one's identity on social media can be most clearly seen among participants with data on all 3 platforms. In the analysis presented at the end of the Results section, we found low average pairwise cosine similarities within participants between platforms, especially when comparing with cosine similarities of different participants within the same platform. This indicates that, even within the same feature space for the same participant, social media data between platforms are likely to diverge into multiple distinct directions mapping to these fragments of identities. This divergence is at least equal to, if not even greater than, the divergence in identity presentation between different individuals within the same social media platform. Therefore, when models trained on data from one platform learn this specific fragment of identity, they are less effective on testing data that capture a different identity.

Textbox 3. Example (paraphrased and deidentified) posts representative of example top features to distinguish between schizophrenia spectrum disorder and control classes. Words indicative of the features are italicized.

NegEmo

- I *fear* to try and *fail*, because i don't want to be part of the STATISTIC of people that *failed*. It *hurts* when the opportunity passes by though.'

Swear

- Omfg the Damn *mf* #struggle to stay the *fking* sleep I'm like *wtf* this isn't fair I hate my Damn neck hurting like this *shit* isn't cool this pain waking me up every Damn hr

Sad

- Im a *useless sorry sob*

Anger

- Yo stay *tf* out my room unless we *fucking* cause I'm tired too tired for this *shit* and all my *shit* better be where i left it

---

Comparison With Prior Work

Our findings provide replicative validity to several threads in previous research. Specifically, we found that the performance of models trained on social media data with clinically verified labels (ie, SSD or control) is consistent with similar models presented in previous research, including those trained on similar patient populations and clinical sites [6,8]. Furthermore, linguistic differences reflecting serious mental health conditions between social media platforms found in our work have also

been elucidated upon in previous work. For instance, Guntuku et al [42] found that there is little overlap between words indicative of stress on Twitter and Facebook. In addition, our findings regarding the low performance of models for interplatform tasks compared with intraplatform tasks follow a similar vein to those of the study by Ernala et al [8]. In their study, they found that, despite the overwhelming advantage in data availability, models trained on social media data with self-reported labels significantly underperformed models trained on social media data with clinically verified labels when

evaluated on clinical testing data [8]. Similar to our experiments, such a difference in performance in the study by Ernala et al [8] was also noted to be caused by a mismatch in important features learned by the different models to differentiate between language and activity patterns deployed by patients with SSD and healthy controls. Overall, our analysis combined with previous results suggests that construct validities of predictive models trained on data from different social media platforms are dissimilar, reinforcing the need for continued exploration of novel social media–based early identification strategies with a special emphasis on uniting distinct fragments of identities for accurate identification and intervention.

Clinical Implications

Our findings have important implications for mental health research and practice. Hospitalization prediction for psychiatric illnesses by harnessing digital trace data has been of significant interest in recent years. These previous studies have explored the utility of smartphone sensor data (ie, geolocation, physical activity, phone use, and speech), wearables, and social media activity to predict symptom fluctuations as well as understand the diagnostic process and hospitalization identification [6,43-46]. Our work extends this body of research by critically examining how machine learning efforts that harness data from single sources may not be readily applicable to support hospitalization prediction in contexts where the same source of data is not present. For these models to be usable in the real world, we advocate for a comprehensive approach in which clinicians look to patterns gleaned through the integration of different data sources while augmenting their decision-making with objective measures derived from digital trace data. Social media data are also increasingly becoming a part of consultations [47,48]. Therefore, we suggest that clinicians consider both acknowledging and incorporating collateral information spanning multiple platforms into the way they monitor symptomatic exacerbation in their patients and modify treatment to prevent further hospitalizations.

Finally, digital interventions that are touted to be powered by social media data should consider the significant aspect of fragmented web-based identities of patients [49,50]. To intervene at the right time, at the right place, and for the right person, a comprehensive approach to understanding a patient's context for hospitalization prediction would be beneficial. However, we recognize that, in a domain as sensitive as mental health, combining data sources may further complicate the privacy and ethical risks to those who contribute their data—research has shown that information integration can enable the discovery of otherwise latent attributes, some of which may present grave feelings of discomfort and violation in individuals [51,52]. Therefore, we urge caution and call for new standards to protect the confidentiality and rights of this sensitive population and ensure that the enabled technologies are used in the service of positive outcomes for the patients.

Limitations and Future Work

Our work has some limitations that could be addressed in future research. First, despite the use of data augmentation techniques to rebalance the ratio between SSD data and control data for each data set and make the data set sizes of the 3 examined platforms (ie, Instagram, Twitter, and Facebook) comparable with each other, we acknowledge that a limited quantity of available data may have affected the observed classification performance. Although it is widely recognized that patient social media data are challenging to collect, as was the case in this study, future research may consider the potential of creating large benchmarked data sets that may support better reproducible research in this field [53]. Second, we acknowledge the demographic dissimilarity between participants with SSD and healthy controls, which may be a confounding factor in our study design. Furthermore, our methods did not examine or extract any features concerning video data, which are available on Facebook and especially Instagram. Given that youths nowadays are increasingly expressing themselves on social media via videos (especially on video-centric platforms such as TikTok), future research should aim to fill these gaps so that we can ensure the completeness of one's mental health records expressed on social media and other forms of networked communication. Along these lines, future research may also consider data from additional novel social media platforms that are increasingly being used by youths for their social goals, such as Snapchat and TikTok. Finally, it would be worthwhile to examine additional clinical questions such as suicidal risk to explore the extent to which identity fragmentation across social media platforms may affect the quality of inferences made from these data.

Conclusions

In this study, we showed that it is challenging to build effective models for predicting future psychiatric hospitalizations of patients with SSD on new social media data from platforms previously unseen in the models' training data. Specifically, we demonstrated that models built on one platform's data do not generalize to another as each platform consistently reflects different segments of participants' identities. This fragmentation of identity is empirically backed up by both significant differences in the construct validity of intraplatform classifiers and divergent feature vectors within participants between the 3 investigated social media platforms. To ensure the effective incorporation of digital technology into early psychosis intervention, especially in the prevention of relapse hospitalizations, further research must explore precisely how symptoms of mental illness manifest on the web through changing patterns of language and activity on various platforms as well as how comprehensive, ethical, and effective treatment and engagement strategies should be devised that function seamlessly across patients' fragmented web-based identities.

the early phases of patient data collection. The authors also thank members of the Social Dynamics and Wellbeing Lab at Georgia Tech for their valuable feedback during the various phases of the study.

Conflicts of Interest

Multimedia Appendix 1
Additional information on the feature selection process and robustness checks.
[DOCX File , 9 KB - mental_v9i12e39747_app1.docx]

References

1. Wolthaus JE, Dingemans PM, Schene AH, Linszen DH, Wiersma D, Van Den Bosch RJ, et al. Caregiver burden in recent-onset schizophrenia and spectrum disorders: the influence of symptoms and personality traits. J Nerv Ment Dis 2002 Apr;190(4):241-247. [doi: 10.1097/00005053-200204000-00005] [Medline: 11960085]
2. Birchwood M, Macmillan F. Early intervention in schizophrenia. Aust N Z J Psychiatry 1993 Sep;27(3):374-378. [doi: 10.3109/00048679309075792] [Medline: 8250779]
3. Lieberman JA, Fenton WS. Delayed detection of psychosis: causes, consequences, and effect on public health. Am J Psychiatry 2000 Nov;157(11):1727-1730. [doi: 10.1176/appi.ajp.157.11.1727] [Medline: 11058464]
4. Birnbaum ML, Rizvi AF, Confino J, Correll CU, Kane JM. Role of social media and the internet in pathways to care for adolescents and young adults with psychotic disorders and non-psychotic mood disorders. Early Interv Psychiatry 2017 Aug 23;11(4):290-295 [FREE Full text] [doi: 10.1111/eip.12237] [Medline: 25808317]
5. Miller BJ, Stewart A, Schrimsher J, Peeples D, Buckley PF. How connected are people with schizophrenia? Cell phone, computer, email, and social media use. Psychiatry Res 2015 Feb 28;225(3):458-463. [doi: 10.1016/j.psychres.2014.11.067] [Medline: 25563669]
6. Birnbaum ML, Ernala SK, Rizvi AF, Arenare E, R Van Meter A, De Choudhury M, et al. Detecting relapse in youth with psychotic disorders utilizing patient-generated and patient-contributed digital data from Facebook. NPJ Schizophr 2019 Oct 07;5(1):17 [FREE Full text] [doi: 10.1038/s41537-019-0085-9] [Medline: 31591400]
7. Mitchell M, Hollingshead K, Coppersmith G. Quantifying the language of schizophrenia in social media. In: Proceedings of the 2nd Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality. 2015 Presented at: 2nd Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality; Jun 5, 2015; Denver, Colorado. [doi: 10.3115/v1/w15-1202]
8. Ernala S, Birnbaum M, Candan K, Rizvi A, Sterling W, Kane J, et al. Methodological gaps in predicting mental health states from social media: triangulating diagnostic signals. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. 2019 Presented at: CHI '19: CHI Conference on Human Factors in Computing Systems; May 4 - 9, 2019; Glasgow Scotland Uk. [doi: 10.1145/3290605.3300364]
9. Rekhi G, Ang MS, Lee J. Clinical determinants of social media use in individuals with schizophrenia. PLoS One 2019 Nov 20;14(11):e0225370 [FREE Full text] [doi: 10.1371/journal.pone.0225370] [Medline: 31747434]
10. Zomick J, Levitan S, Serper M. Linguistic analysis of schizophrenia in Reddit posts. In: Proceedings of the Sixth Workshop on Computational Linguistics and Clinical Psychology. 2019 Presented at: Sixth Workshop on Computational Linguistics and Clinical Psychology; Jun, 2019; Minneapolis, Minnesota. [doi: 10.18653/v1/w19-3009]
11. Birnbaum ML, Ernala SK, Rizvi AF, De Choudhury M, Kane JM. A collaborative approach to identifying social media markers of schizophrenia by employing machine learning and clinical appraisals. J Med Internet Res 2017 Aug 14;19(8):e289 [FREE Full text] [doi: 10.2196/jmir.7956] [Medline: 28807891]
12. Ernala SK, Rizvi AF, Birnbaum ML, Kane JM, De Choudhury M. Linguistic markers indicating therapeutic outcomes of social media disclosures of schizophrenia. Proc ACM Human Comput Interact 2017 Dec 06;1(CSCW):1-27. [doi: 10.1145/3134678]
13. Auxier B, Anderson M. Social media use in 2021. Pew Research Center. 2021 Apr 7. URL: https://www.pewresearch.org/internet/2021/04/07/social-media-use-in-2021/ [accessed 2022-12-06]
14. Hänsel K, Lin IW, Sobolev M, Muscat W, Yum-Chan S, De Choudhury M, et al. Utilizing Instagram data to identify usage patterns associated with schizophrenia spectrum disorders. Front Psychiatry 2021 Aug 16;12:691327 [FREE Full text] [doi: 10.3389/fpsyt.2021.691327] [Medline: 34483987]

XSL•FO
RenderX

15. Birnbaum ML, Norel R, Van Meter A, Ali AF, Arenare E, Eyigoz E, et al. Identifying signals associated with psychiatric illness utilizing language and images posted to Facebook. NPJ Schizophr 2020 Dec 03;6(1):38 [FREE Full text] [doi: 10.1038/s41537-020-00125-0] [Medline: 33273468]

16. Chancellor S, De Choudhury M. Methods in predictive techniques for mental health status on social media: a critical review. NPJ Digit Med 2020;3:43 [FREE Full text] [doi: 10.1038/s41746-020-0233-7] [Medline: 32219184]

17. Kircaburun K, Griffiths MD. Instagram addiction and the big five of personality: the mediating role of self-liking. J Behav Addict 2018 Mar 01;7(1):158-170 [FREE Full text] [doi: 10.1556/2006.7.2018.15] [Medline: 29461086]

18. Bayer JB, Ellison NB, Schoenebeck SY, Falk EB. Sharing the small moments: ephemeral social interaction on Snapchat. Inform Commun Soc 2015 Sep 18;19(7):956-977. [doi: 10.1080/1369118x.2015.1084349]

19. Purwaningtyas MP, Alicya DA. The fragmented self: having multiple accounts in Instagram usage practice among Indonesian youth. J Media dan Komunikasi Indonesia 2020 Sep 24;1(2):171. [doi: 10.22146/jmki.58459]

20. Gündüz U. The effect of social media on identity construction. Mediterranean J Social Sci 2017;8(5):85-92.

21. First M, Spitzer RL, Gibbon M, Williams J. Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version. New York: Biometrics Research, New York State Psychiatric Institute; 2002.

22. Zimmerman M, Mattia JI. A self-report scale to help make psychiatric diagnoses: the psychiatric diagnostic screening questionnaire. Arch Gen Psychiatry 2001 Aug 01;58(8):787-794. [doi: 10.1001/archpsyc.58.8.787] [Medline: 11483146]

23. Ernala SK, Kashiparekh KH, Bolous A, Ali A, Kane J, Birnbaum ML, et al. A social media study on mental health status transitions surrounding psychiatric hospitalizations. Proc ACM Hum Comput Interact 2021 Apr 13;5(CSCW1):1-32 [FREE Full text] [doi: 10.1145/3449229] [Medline: 36267476]

24. Aizawa A. An information-theoretic perspective of tf–idf measures. Inform Process Manag 2003 Jan;39(1):45-65. [doi: 10.1016/s0306-4573(02)00021-3]

25. Tausczik YR, Pennebaker JW. The psychological meaning of words: LIWC and computerized text analysis methods. J Language Social Psychol 2009 Dec 08;29(1):24-54. [doi: 10.1177/0261927x09351676]

26. Mclaughlin G. SMOG grading - a new readability formula. J Reading 1969;12(8):639-646 [FREE Full text]

27. Hutto C, Gilbert E. VADER: a parsimonious rule-based model for sentiment analysis of social media text. In: Proceedings of the International AAAI Conference on Web and Social Media. 2014 Presented at: International AAAI Conference on Web and Social Media; Jun 1–4, 2014; Ann Arbor, Michigan, USA.

28. Garimella V, Alfayad A, Weber I. Social media image analysis for public health. In: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems. 2016 Presented at: CHI'16: CHI Conference on Human Factors in Computing Systems; May 7 - 12, 2016; San Jose California USA. [doi: 10.1145/2858036.2858234]

29. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: synthetic minority over-sampling technique. J Artif Intell Res 2002 Jun 01;16:321-357. [doi: 10.1613/jair.953]

30. Guyon I, Elisseeff A. An introduction to variable and feature selection. J Mach Learn Res 2003;3:1157-1182.

31. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: machine learning in Python. JMLR 2011;12(85):2825-2830.

32. Hastie T, Tibshirani R, Friedman J. The Elements of Statistical Learning Data Mining, Inference, and Prediction. Cham, Switzerland: Springer; 2009.

33. Powers DM. Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation. Int J Mach Learn Technol 2020:37-63.

34. Lundberg S, Lee S. A unified approach to interpreting model predictions. arXiv 2017.

35. Handbook of Game Theory with Economic Applications Volume 2. Amsterdam, Netherlands: Elsevier; 1992.

36. Jurafsky D, Martin JH. Speech and Language Processing An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition. Hoboken, New Jersey, United States: Pearson Prentice Hall; 2000.

37. Kwapil TR. Social anhedonia as a predictor of the development of schizophrenia-spectrum disorders. J Abnormal Psychol 1998 Nov;107(4):558-565. [doi: 10.1037/0021-843X.107.4.558]

38. Ringer JM, Lysaker PH. Anger expression styles in schizophrenia spectrum disorders: associations with anxiety, paranoia, emotion recognition, and trauma history. J Nerv Ment Dis 2014 Dec;202(12):853-858. [doi: 10.1097/NMD.0000000000000212] [Medline: 25386763]

39. Liu J, Chua JJ, Chong SA, Subramaniam M, Mahendran R. The impact of emotion dysregulation on positive and negative symptoms in schizophrenia spectrum disorders: a systematic review. J Clin Psychol 2020 Apr;76(4):612-624. [doi: 10.1002/jclp.22915] [Medline: 31909833]

40. Kuperberg G, Caplan D. Language dysfunction in schizophrenia. In: Neuropsychiatry. Philadelphia: Lippincott Williams and Wilkins; 2003.

41. Palmese LB, DeGeorge PC, Ratliff JC, Srihari VH, Wexler BE, Krystal AD, et al. Insomnia is frequent in schizophrenia and associated with night eating and obesity. Schizophr Res 2011 Dec;133(1-3):238-243 [FREE Full text] [doi: 10.1016/j.schres.2011.07.030] [Medline: 21856129]

42. Chandra Guntuku S, Buffone A, Jaidka K, Eichstaedt JC, Ungar LH. Understanding and measuring psychological stress using social media. In: Proceedings of the International AAAI Conference on Web and Social Media. 2019 Presented at:

International AAAI Conference on Web and Social Media; Jun 11-14, 2019; Munich, Germany. [doi: 10.1609/icwsm.v13i01.3223]

43. Ben-Zeev D, Scherer EA, Wang R, Xie H, Campbell AT. Next-generation psychiatric assessment: using smartphone sensors to monitor behavior and mental health. Psychiatr Rehabil J 2015 Sep;38(3):218-226 [FREE Full text] [doi: 10.1037/prj0000130] [Medline: 25844912]

44. Birnbaum ML, Kulkarni P, Van Meter A, Chen V, Rizvi AF, Arenare E, et al. Utilizing machine learning on internet search activity to support the diagnostic process and relapse detection in young individuals with early psychosis: feasibility study. JMIR Ment Health 2020 Sep 01;7(9):e19348 [FREE Full text] [doi: 10.2196/19348] [Medline: 32870161]

45. Eisner E, Bucci S, Berry N, Emsley R, Barrowclough C, Drake RJ. Feasibility of using a smartphone app to assess early signs, basic symptoms and psychotic symptoms over six months: a preliminary report. Schizophr Res 2019 Jun;208:105-113 [FREE Full text] [doi: 10.1016/j.schres.2019.04.003] [Medline: 30979665]

46. Zulueta J, Piscitello A, Rasic M, Easter R, Babu P, Langenecker SA, et al. Predicting mood disturbance severity with mobile phone keystroke metadata: a BiAffect digital phenotyping study. J Med Internet Res 2018 Jul 20;20(7):e241 [FREE Full text] [doi: 10.2196/jmir.9775] [Medline: 30030209]

47. Fisher CE, Appelbaum PS. Beyond googling: the ethics of using patients' electronic footprints in psychiatric practice. Harv Rev Psychiatry 2017;25(4):170-179. [doi: 10.1097/hrp.0000000000000145]

48. Rieger A, Gaines A, Barnett I, Baldassano CF, Connolly Gibbons MB, Crits-Christoph P. Psychiatry outpatients' willingness to share social media posts and smartphone data for research and clinical purposes: survey study. JMIR Form Res 2019 Aug 29;3(3):e14329 [FREE Full text] [doi: 10.2196/14329] [Medline: 31493326]

49. Yoo D, Birnbaum M, Van Meter A, Ali A, Arenare E, Abowd G, et al. Designing a clinician-facing tool for using insights from patients' social media activity: iterative co-design approach. JMIR Ment Health 2020 Aug 12;7(8):e16969. [doi: 10.2196/16969] [Medline: 32784180]

50. Yoo D, Ernala S, Saket B, Weir D, Arenare E, Ali A, et al. Clinician perspectives on using computational mental health insights from patients' social media activities: design and qualitative evaluation of a prototype. JMIR Ment Health 2021 Nov 16;8(11):e25455. [doi: 10.2196/25455] [Medline: 34783667]

51. Terrasse M, Gorin M, Sisti D. Social media, e‐health, and medical ethics. Hastings Center Report 2019 Feb 21;49(1):24-33. [doi: 10.1002/hast.975] [Medline: 30790306]

52. Thieme A, Belgrave D, Sano A, Doherty G. Machine learning applications. Interactions. 2020. URL: https://www.scss.tcd.ie/gavin.doherty/papers/Interactions-MLinMH.pdf [accessed 2022-11-23]

53. Househ M, Grainger R, Petersen C, Bamidis P, Merolli M. Balancing between privacy and patient needs for health information in the age of participatory health and social media: a scoping review. Yearb Med Inform 2018 Aug;27(1):29-36. [doi: 10.1055/s-0038-1641197] [Medline: 29681040]

Abbreviations

AUROC: area under the receiver operating characteristic curve
SHAP: Shapley Additive Explanations
SSD: schizophrenia spectrum disorder

Review

The Effectiveness of Internet-Guided Self-help Interventions to Promote Physical Activity Among Individuals With Depression: Systematic Review

Yiling Tang[1], MSc; Madelaine Gierc[1], PhD; Raymond W Lam[2], MD; Sam Liu[3], PhD; Guy Faulkner[1], PhD

[1]School of Kinesiology, University of British Columbia, Vancouver, BC, Canada

[2]Department of Psychiatry, University of British Columbia, Vancouver, BC, Canada

[3]School of Exercise Science, Physical & Health Education, University of Victoria, Vancouver, BC, Canada

Corresponding Author:
Yiling Tang, MSc
School of Kinesiology
University of British Columbia
6081 University Blvd
Vancouver, BC, V6T 1Z1
Canada
Phone: 1 6048222211
Email: yilingtt@student.ubc.ca

Abstract

Background: Depression is a prevalent and debilitating mental disorder and a leading cause of disability worldwide. Physical activity (PA) interventions have been shown to alleviate depressive symptoms. However, not all patients have access to PA programing tailored for depression. Internet-guided self-help (IGSH) interventions may be an effective option for increasing PA among people with depression who cannot or prefer not to access supervised exercise treatment.

Objective: We aimed to evaluate the effectiveness of IGSH interventions in increasing PA and alleviating depressive symptoms in people with depression.

Methods: A systematic literature search was conducted for randomized controlled trials and quasiexperimental studies using 9 electronic databases. The review was registered in PROSPERO (2020 CRD42020221713).

Results: A total of 4 randomized controlled trials (430 participants) met the inclusion criteria. Of these, 3 were web-based and 1 was app-based. Three studies found IGSH interventions to have medium to large effects on decreasing depressive symptoms but not on increasing PA compared with waitlist or usual care. One study showed increased self-reported PA but no significant difference in depressive symptoms in the intervention group compared with the control group. Goal setting was the most common behavior change technique used in the interventions. Dropout rates within the intervention groups were relatively low (0%-19%).

Conclusions: Our findings suggested that IGSH PA interventions are feasible and have the potential to reduce depressive symptoms in people with depression. More well-designed and tailored interventions with different combinations of behavior change techniques, particularly those targeting the emotion domain, are needed to assess the overall effectiveness and feasibility of using IGSH interventions to increase PA among people with depression.

Trial Registration: PROSPERO CRD42020221713; https://tinyurl.com/ysaua5bu

KEYWORDS

physical activity; eHealth; mobile health; mHealth; depression; systematic review; internet; mobile phone

Introduction

Background

Depression is a chronic mental health condition characterized by sadness; anhedonia; and secondary physical, cognitive, and emotional symptoms. People with depression often experience lower quality of life [1-3], increased risk of suicide [4,5], and increased risk of having comorbid chronic diseases (eg, diabetes, asthma, chronic lung disease, and coronary heart disease), which can result in premature death [6,7]. Historically, it was estimated

that >300 million people have experienced depression worldwide [8], and most people have never been diagnosed [6,9,10]. Surveillance studies have observed a marked increase in the prevalence and severity of depressive symptoms since the beginning of the COVID-19 pandemic. For instance, Bueno-Notivol et al [11] observed that the prevalence of self-reported depression worldwide was 7 times higher in 2020 (25%) than in 2017 (3.44%). Similarly, a meta-analysis by Nochaiwong et al [12] estimated that more than a quarter of people globally (28%) self-reported depressive symptoms during the COVID-19 outbreak in 2020.

Physical activity (PA)—"any bodily movement produced by skeletal muscles that results in energy expenditure [13]"—has long been recognized as an important health-promoting behavior. In recent years, it has also been shown to be beneficial in the prevention and treatment of depressive disorders [14,15]. Exercise, a subset of PA that is typically planned and structured with the goal of increasing or maintaining fitness [13], is now recommended as a monotherapy for mild to moderate depression in Canada [16]. PA presents several advantages over conventional treatments for depression (ie, psychotherapy and antidepressant medication). Among them, PA has minimal negative side effects, is affordable, and is potentially more accessible [17]. Rebar and Taylor [18] suggested that PA could be a cost-effective method for treating depression worldwide. As an additional benefit, a large body of literature confirmed the positive side effects of engaging in regular PA, including heightened health-related quality of life (ie, physical and mental well-being) [19], chronic disease prevention (including obesity, type 2 diabetes, coronary heart diseases, and several cancers), and reduced risk of premature mortality [20,21].

Globally, a large minority (31.1%) of adults do not meet the minimum recommended levels of PA [22]. Data from the 2016 to 2017 Canadian Health Measures Survey suggested that only 16% of Canadian adults meet the current PA recommendation of 150 minutes of moderate to vigorous PA (MVPA) per week [23]. People with depression are more likely to experience lower PA levels than those without depression [24,25], in part because of symptoms such as pain and discomfort, insomnia, cognitive difficulties, fatigue, and anhedonia [17,25,26]. Tailored approaches to help people with depression that initiate and maintain PA are needed [27].

Internet-Guided Self-help Interventions

Internet-guided self-help (IGSH) interventions, a form of eHealth intervention, could be 1 mechanism for supporting PA behavior changes among people with depression. eHealth interventions are defined as those that use information and communication technologies to enable health care, including supporting health behavior engagement [28,29]. As such, eHealth interventions are quite broad and include just-in-time adaptive interventions, wearable technology, telehealth, and social media. Research has demonstrated that most people with mental disorders have an interest in trying eHealth interventions (such as smartphone apps) to monitor and manage their health concerns [30,31].

IGSH interventions are characterized by web-based or app-based programs that are primarily self-guided. Some will offer limited

support from a professional or paraprofessional [32]. Compared with synchronous eHealth interventions that likely require costlier direct consultation (eg, telehealth, live internet-based therapy, and live-streamed exercise sessions) [33], IGSH programs delivered via the web or mobile devices have the potential for broad reach and scalability at a relatively low cost. They also allow participants to access the intervention content at their own pace [34-36]. These features may be particularly beneficial for people with depression who experience symptoms, such as fatigue and disrupted sleep patterns. In addition, many people with depression experience stigma that negatively affects treatment-seeking [37] and have preferences for managing symptoms on their own [38]. Offering self-help interventions could be potentially useful in increasing help-seeking rates by mitigating stigma [39].

Systematic reviews and meta-analyses have found that eHealth interventions are effective in increasing short-term PA participation in nonclinical populations, including young people [40,41], adults [42], and older adults [43,44]. Other research on eHealth interventions suggests that interventions that adopt theory or incorporate evidence-based behavior change techniques (BCTs) are generally associated with greater effects and adherence [45,46]. A systematic review of meta-analyses on the effectiveness of self-help and internet-guided interventions has suggested that these programs are effective in treating depression [36,47]. In fact, computerized cognitive behavioral therapy (CBT) has been recommended as a treatment for subthreshold or mild to moderate depression in the United Kingdom [48]. Similarly, a systematic review by Andersson and Cuijpers [49] suggested that guided internet-based psychological interventions were more effective than unguided interventions for depression among adults.

Less is known about whether IGSH PA interventions for people with depression are effective at (1) increasing PA engagement and (2) reducing depressive symptoms. To date, 2 reviews have investigated eHealth PA interventions for individuals with mental illnesses [50,51], and 1 review [52] specifically examined web-based interventions. However, these reviews investigated eHealth interventions for mental illness generally, rather than IGSH programs for depression specifically. These are noteworthy distinctions, as the strongest research evidence of the benefits of PA-based treatments is for depression [16]; to our knowledge, no clinical guidelines exist for PA-based treatment for anxiety, schizophrenia, or other mental health conditions. In addition, the reviews included both experimental and observational studies [50]. Given the rapid pace of technological development and growing concerns of physical inactivity among people with depression, it is likely that this research field has expanded in recent years. There is a need for an updated systematic review of high-quality studies specific to depression. Thus, the primary objective of this systematic review was to assess the effectiveness of IGSH interventions in promoting PA and alleviating depressive symptoms in people with depression. The secondary objective of this study was to understand study characteristics, such as attrition rates and intervention design, to explore factors associated with successful interventions and areas for future growth.

Methods

This systematic review was guided by the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) approach [53,54]. The protocol was registered in PROSPERO (CRD42020221713). The PRISMA checklist is provided in Multimedia Appendix 1.

Searches

A total of 2 rounds of search were conducted in December 2020 and November 2021. The search strategy was developed in consultation with a university rehabilitation sciences librarian. In all, 7 electronic databases were searched: MEDLINE (via Ovid), PsycINFO (EBSCOhost), CINAHL (EBSCOhost), Web of Science, Cochrane Central Register of Controlled Trials (CENTRAL; via Ovid), Embase (via Ovid), and SportDiscus (EBSCOhost). In addition, OpenGrey and ProQuest Dissertations were searched to identify gray literature that matched the inclusion criteria. Finally, reference lists of the included studies were searched to identify additional eligible studies.

The December 2020 search used a combination of controlled vocabulary (eg, Medical Subject Headings) and keywords related to "physical activity," "depression," and "eHealth." For comprehensiveness, the search included keywords for all types of eHealth interventions. The search strategies for all databases are presented in Multimedia Appendix 2. The selected keywords were obtained from previous systematic reviews and protocols in relevant areas.

A revised search was conducted in November 2021, with two changes: (1) the results were limited to the period from December 1, 2020, to November 5, 2021, and (2) the search strategy was updated to better reflect the inclusion criteria and reduce irrelevant records informed by the experience of the December 2020 search. For example, the keywords "text messaging" and "video conferencing" were deleted. The updated search strategies for all included databases are presented in Multimedia Appendix 3.

Eligibility Criteria

The study selection criteria were based on the Population, Intervention, Comparison, Outcome, Study (PICOS) design framework [55,56]. Studies that met the following criteria were included in this review.

Participants

Individuals with a clinical diagnosis of depression (eg, the Diagnostic and Statistical Manual of Mental Disorders criteria) or individuals with clinically significant depressive symptoms based on a validated self-report tool (eg, Patient Health Questionnaire 9-Item or Beck Depression Inventory-II [BDI-II] scores) were considered eligible for inclusion. Studies that used nonvalidated items (eg, "Are you depressed? Yes/No") or those who did not use recognized diagnostic criteria were excluded. There were no restrictions based on age, gender, nationality, or ethnicity.

Interventions

Interventions were considered eligible if they were delivered via web-based platforms or mobile apps (eg, via smartphones and tablets) and included content designed to increase PA (eg, asynchronous PA programing, education, and PA goal setting). Studies were included even if promoting PA was the secondary objective, so long as changes in PA were measured and reported. By definition, interventions were primarily self-guided or automatic; however, interventions that offered degrees of human support (eg, providing personalized feedback via telephone or email) were permitted. In contrast, interventions primarily based on in-person support (eg, telephone counseling and SMS text messaging) were excluded. Similarly, live-streamed interventions (eg, live Zoom yoga classes) were excluded, as these are synchronous events organized in a live internet-based space. There were no restrictions based on the type of PA, PA intensity, bout length, program frequency or duration, or follow-up period.

Comparison

Any comparators were considered for inclusion. For instance, a PA intervention delivered in person, an alternative eHealth intervention, or a waitlist control group. No restrictions were placed on the nature of the comparison group.

Outcomes

The primary outcome was a change in PA levels. Both device-based (eg, accelerometer and pedometer) and self-report (eg, PA diary and questionnaires) measures were considered. Secondary outcomes included changes in the severity of depression, assessed by validated measures, and the acceptability of treatment assessed by (1) reported indicators of intervention engagement [57] and (2) dropout rates at postintervention. Engagement was defined as the extent to which the participants undertook the intervention. Dropout rate, also called attrition rate, was defined as the percentage of participants who were randomized to a group but failed to complete it.

Study Design

Randomized controlled trials (RCTs) and quasiexperimental studies with control groups and pre- and posttest measurements were considered. Studies were eligible if they were a published original manuscript or thesis or dissertation. Papers were excluded if they did not contain original findings (eg, editorials or reviews), if they were research protocols or proposals, or if they were conference abstracts.

Other

Papers were considered eligible for inclusion if they were published in English. There were no limitations on the date of publication.

Study Selection

A total of 2 authors (first round: YT and JL; second round: YT and MG) independently screened the articles for inclusion using the agreed-upon study eligibility or ineligibility criteria. Screening occurred in 3 phases, facilitated by the Covidence systematic review software (Veritas Health Innovation). First, duplicate records were removed automatically using Covidence.

Second, titles and abstracts were independently screened for relevance by 2 reviewers. In case of disagreement, the reviewers met to discuss and reach a consensus. Finally, full-text copies of the selected studies were screened (YT, JL, DD, and MG) to confirm their eligibility. Disagreements between the 2 researchers were resolved through discussion. A third reviewer (GF) was consulted if consensus could not be reached. Following selection, the reference list for each included study was searched to identify additional eligible studies. The PRISMA flow diagrams are presented in Figure 1 and Figure 2.

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram of the first round of searches. PA: physical activity; RCT: randomized controlled trial.
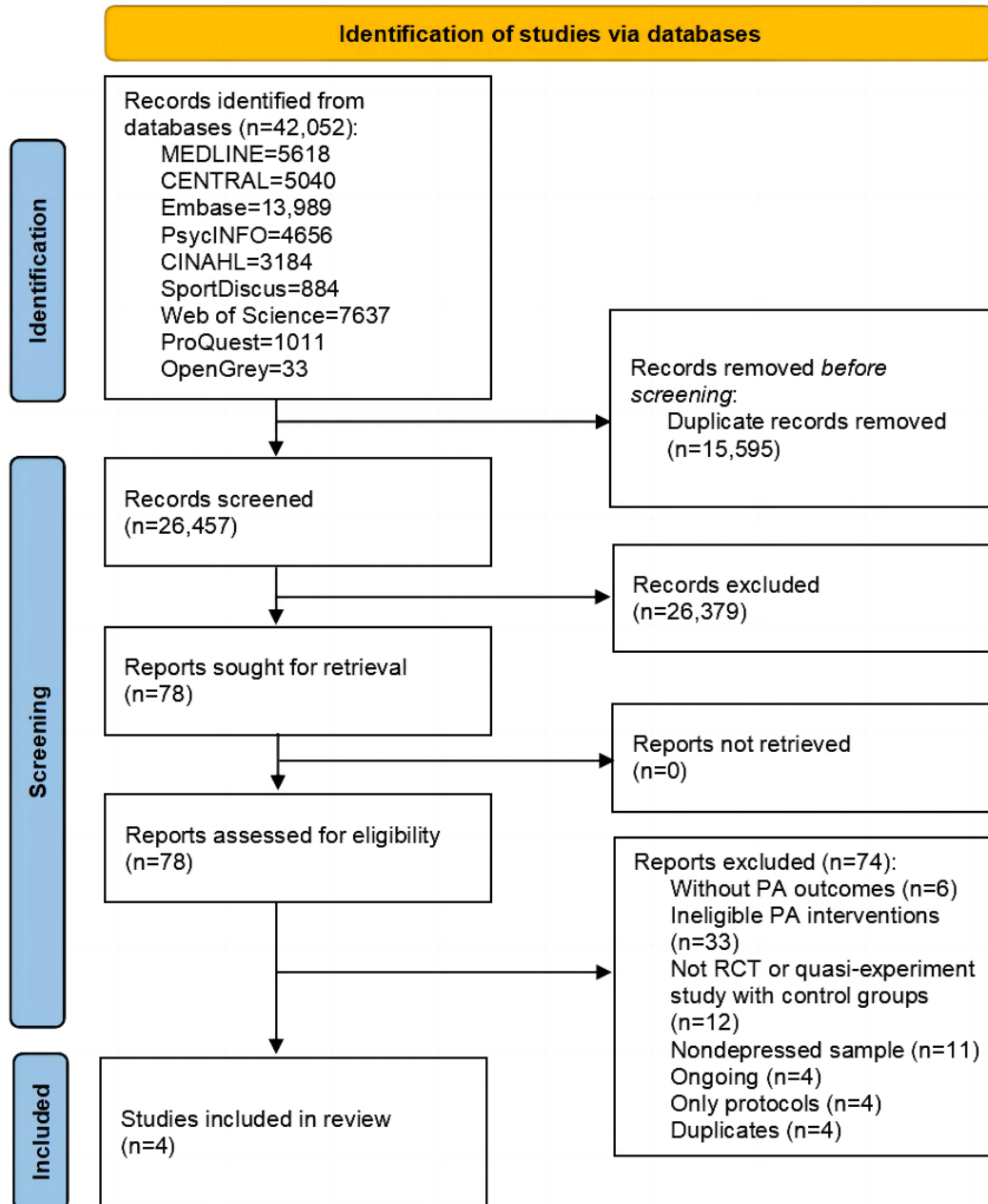
Figure 2. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram of the second round of searches. PA: physical activity; RCT: randomized controlled trial. *This study was excluded as being a duplicate of first round studies.
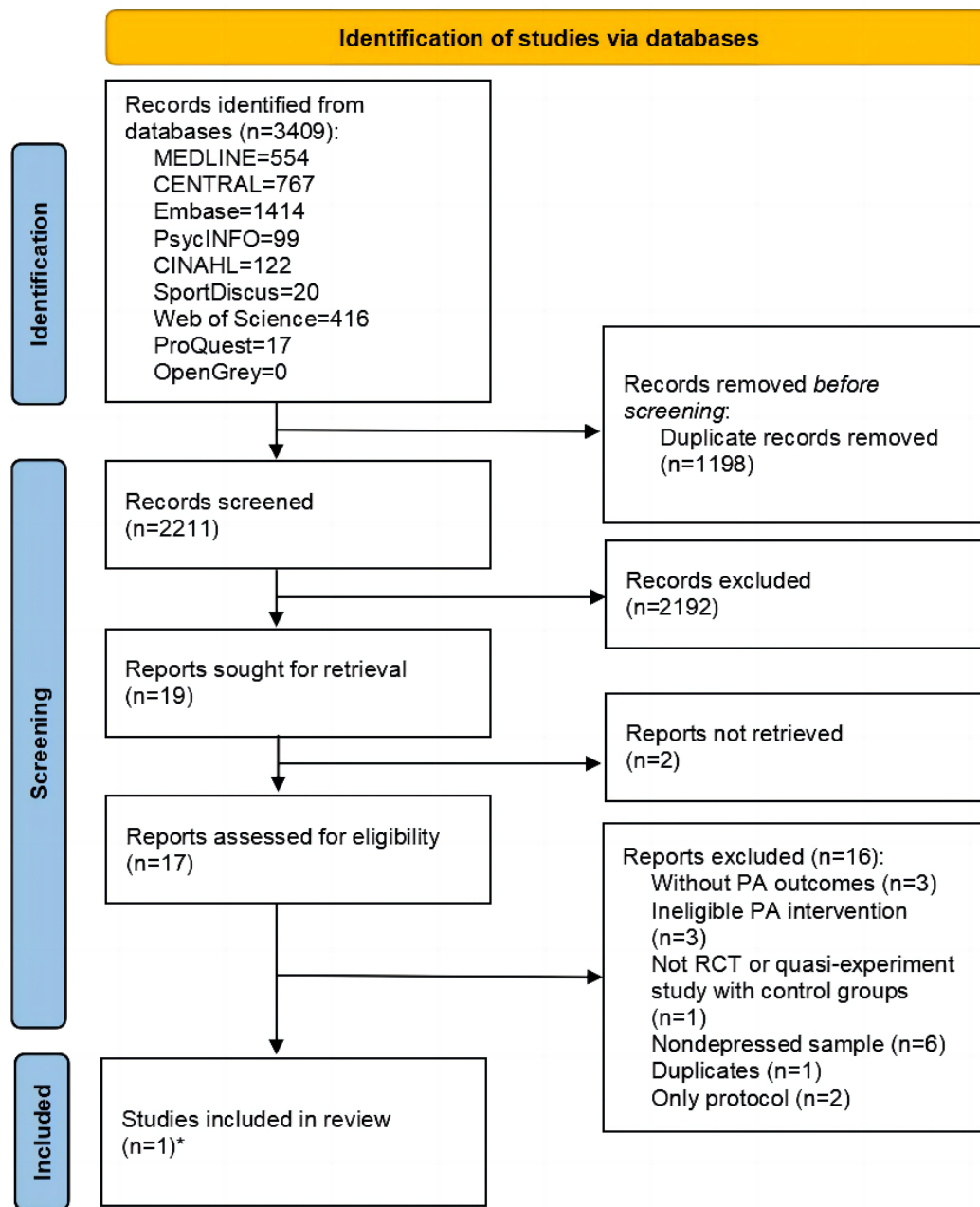
Data Extraction

The following items were extracted by 1 author (YT) and reviewed for accuracy by another reviewer (first round: DD and second round: MG): (1) author and year; (2) study design; (3) country; (4) sample size; (5) participant characteristics, such as mean age and gender; (6) inclusion and exclusion criteria; (7) intervention details, including the level of contact; (8) comparator; (9) main outcomes (measurement tools); (10) additional outcomes (measurement tools); (11) main findings with regard to PA and depression symptomatology; (12) follow-up period; (13) BCTs; and (14) funding sources. A third author (GF) was consulted to resolve discrepancies.

Level of contact was categorized as high, moderate, and low, following procedures described by Ma et al [47]. If the support was directly provided by a qualified therapist (eg, exercise trainers providing personalized feedback), the level of contact was high. If the support was provided by members of the research team or trained students, it was considered to be moderate. If there was no direct contact with an interventionist (eg, only automated reminder emails were sent out), the contact was considered to be low.

Risk of Bias

The Cochrane risk of bias tool (ROB 2; version 2) [58,59] was used to assess the risk of bias for both PA and depression outcomes. This tool includes 5 domains of risk of bias: randomization process, deviations from intended interventions, missing outcome data, measurement of the outcome, and selection of the reported result. The risk of bias for each outcome in each study was judged as high, some concerns, or low. Risk of bias was assessed by 2 independent reviewers (YT and MG). In case of disagreement, the reviewers met to discuss and reach

a consensus. The remaining disagreements were resolved by consulting a third author (GF).

Data Synthesis

Meta-analysis was not feasible because the included studies were too heterogeneous in their design and reporting of results. The study interventions and outcome characteristics were summarized using narrative synthesis and descriptive statistics. When the effect size was reported as Cohen d (standardized mean difference), it was interpreted as small (Cohen d=0.2), moderate (Cohen d=0.5), or large (Cohen d=0.8) [60].

Confidence in the Cumulative Evidence

The Grading of Recommendation, Assessment, Development and Evaluation (GRADE) [61] criteria were used to rate the certainty of the cumulative evidence based on the risk of bias, imprecision, inconsistency, indirectness, and publication bias. The quality of evidence was categorized as high, moderate, low, or very low.

Results

Overview

A total of 45,461 records were identified: 42,052 in December 2020 and an additional 3409 in November 2021. After removing duplicates, 28,668 articles underwent title or abstract screening and 95 underwent full-text review. The first round of searches (December 2020) produced 4 articles. The second round of searches (November 2021) produced 1 additional article, which was subsequently deleted because it was a duplicate. A total of 4 eligible studies were included in this systematic review. The PRISMA flow diagrams are presented in Figure 1 and Figure 2.

Characteristics of Included Trials

Study Design

Tables 1 and 2 summarize the characteristics of the 4 included studies. All 4 included studies were dual-arm RCTs. None of the studies used an active control group (eg, sham intervention) or treatment comparison group (eg, synchronous in-person therapy). Rather, the studies used either a waitlist or a treatment-as-usual group.

Table 1. Participant characteristics.

Study	Target population	Sample size	Age (years), mean (SD)	Sex (female or male)	Preexisting psychiatric treatment		
					Therapy	Medication	Combined
Guo et al [62], 2020	HIV and major depressive disorder	300; IG[a]: 150; CG[b]: 150	IG: 28 (5.8); CG: 28.6 (5.9)	IG: 8/142; CG: 15/135	0[c]	0[c]	0[c]
Haller et al [63], 2018	Adults (20-65 years old) with major depression	20; IG: 14; CG: 6	IG: 43 (14); CG: 51 (12)	IG: 10/4; CG: 3/3	IG: 2; CG: 0	IG: 7; CG: 3	IG: 3; CG: 3
Lambert et al [64], 2018	Adults with at least moderate depressive symptoms	62; IG: 32; CG: 30	IG: 39.3 (12.0); CG: 36.9 (12.6)	IG: 26/6; CG: 26/4	IG: 1; CG: 7	IG: 18; CG: 18	IG: N/A[d]; CG: N/A
Ström et al [65], 2013	Mild to moderate major depression diagnosis and sedentary lifestyle	48; IG: 24; CG: 24	IG: 48.8 (12.7); CG: 49.6 (8.7)	IG: 20/4; CG: 20/4	IG: 0; CG: 0	IG: 3; CG: 4	IG: N/A; CG: N/A

[a]IG: intervention group.

[b]CG: control group

[c]Potential participants excluded if they were currently on psychiatric treatment.

[d]N/A: not applicable.

Table 2. Characteristics of the included trials.

Study	Intervention	Duration of intervention (week)	Control	Attrition rate	Outcome measurements	Effect size
Guo et al [62], 2020	Cognitive behavioral stress management course+PA[a] promotion via WeChat app	12	Waitlist control+usual care of HIV	8.7%; IG[b]: 11 cases; CG[c]: 15 cases	PA: METs[d] calculated from Chinese version of the GPAQ[e]; depression: CES-D[f] (main), PHQ-9[g] (secondary)	PA outcome: nonsignificant between-group differences (3 mo MET: −1898 (−4285 to 489); P=.12); depressive outcome: between-group mean difference −5.77 (95% CI −7.82 to −3.71), Cohen d=0.66; P<.001
Haller et al [63], 2018	Web-based platform with weekly exercise schedules and motivational feedback, as well as an additional biweekly group training session	8	Treatment as usual	15%; IG: 3 cases; CG: 0 cases	PA: Baecke questionnaire; depression: QIDS-SR[h] and QIDS-C[i]	Total HPA[j] outcome: eta²=0.36; P=.007; Depressive outcome: after 6-12 days: nonsignificant between-group differences (QIDS-SR: P=.06; eta²=0.2) posttreatment: nonsignificant between-group differences
Lambert et al [64], 2018	Web-based modular-fashioned course with evidence-based treatment based on behavioral activation and PA promotion	8	Waitlist control+treatment as usual	19%; IG: 7 cases; CG: 5 cases	Device-based PA: min per week of objective MVPA[k] in 10-min bouts; self-reported PA: IPAQ-SF[l]; depression: PHQ-8[m]	PA outcome: nonsignificant between-group differences—between-group adjusted mean difference: device-based PA: 16.4 (−43.7 to 76.5); self-reported PA: 0.2 (−8.7 to 9.2); depressive outcome: between-group adjusted mean difference −3.6 (95% CI −6.1 to −1.1); Cohen d=0.93
Ström et al [65], 2013	Web-based guided self-help PA program with 9 text modules, written feedback, and home assignments, from therapists	9	Waitlist control	0%; IG: 0 cases; CG: 0 cases	PA: IPAQ[n]; depression: BDI-II[o] (main); and MADRS-S[p] (secondary)	PA outcome: nonsignificant between-group differences (Cohen d=0.20); depressive outcome: between-group—BDI-II: Cohen d=0.67; MADRS-S: Cohen d=0.62

[a]PA: physical activity.

[b]IG: intervention group.

[c]CG: control group.

[d]METs: metabolic equivalents.

[e]GPAQ: Global Physical Activity Questionnaire.

[f]CES-D: Centre for Epidemiological Studies Depression.

[g]PHQ-9: Patient Health Questionnaire 9-item.

[h]QIDS-SR: Quick Inventory of Depressive Symptomatology—self-report.

[i]QIDS-C: Quick Inventory of Depressive Symptomatology—clinician rating.

[j]HPA: habitual physical activity.

[k]MVPA: moderate to vigorous physical activity.

[l]IPAQ-SF: International Physical Activity Questionnaire—Short Form.

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

[n]IPAQ: International Physical Activity Questionnaire.

[o]BDI-II: Beck Depression Inventory—second version.

[p]MADRS-S: Montgomery-Åsberg Depression Rating Scale: Self-rated version.

XSL·FO

RenderX

Sample Characteristics

Sample Characteristics

The studies were conducted in 4 different countries: China [62], Germany [63], the United Kingdom [64], and Sweden [65]. Together, the 4 studies included 430 people with depression. The sample size varied from 20 [63] to 300 [62]. Guo et al [62] included adults with a clinical diagnosis of HIV, with the majority (92.3%) of participants being men. The remaining 3 (75%) studies included higher rates of female participation (65% [63] to 86.7% [64]). The mean age ranged from 28.3 (SD 5.8) [62] to 49.2 (SD 10.07) years [65]. The 3 (75%) studies [63-65] included participants receiving psychiatric treatment, including psychotherapy or medication. Although the studies varied in their reporting of sociodemographic characteristics, it was generally observed that the samples were well-educated (ie, the majority of participants had completed high school or more) and currently employed. In addition, the 3 (75%) studies [62,64,65] reported participants' baseline PA levels, with different proportions of them being physically active (13% [64] to 60.5% [65]).

Measures

Of the 4 studies, 3 (75%) used self-report methods to measure PA: metabolic equivalents calculated from the Global Physical Activity Questionnaire [62], habitual PA scores calculated from the Baecke Physical Activity Questionnaire [63], and the International Physical Activity Questionnaire [65]. In contrast, Lambert et al [64] used a combination of self-report (International Physical Activity Questionnaire–Short Form) and device-based measures (accelerometry, with MVPA reported as minutes per week in 10-minute bouts).

Intervention Characteristics

Interventions typically provide discrete "modules" of information with distinct learning objectives. One study by Guo et al [62] delivered the intervention via a preexisting commercial app (WeChat). The other 3 (75%) trials used web-based PA interventions. The intervention length ranged from 8 weeks [63,64] to 3 months [62]. In addition, 2 (50%) studies [63,64] reported short-term postintervention effects, whereas Guo et al [62] and Ström et al [65] assessed longer-term maintenance (6- and 9-month and 6-month maintenance, respectively). All 4 studies incorporated weekly contact. Moreover, 2 (50%) studies [63,64] reported a moderate level of contact. Haller et al [63] and Ström et al [65] included a high level of contact with participants receiving feedback from a therapist via phone or an encrypted web-based platform, respectively. In addition, Haller et al [63] incorporated an optional biweekly face-to-face training session led by a sports therapist.

Attrition and Engagement

Dropout rates were relatively low, ranging from 0% [65] to 19% [64] across the 4 studies (Table 2). In terms of intervention engagement, participants in the study by Guo et al [62] completed 55% of the program (ie, 9 content modules and 3 review modules). In the study by Haller et al [63], participants were recommended to complete a maximum of 3 endurance sessions and 2 strength training sessions per week for 8 weeks. Participants completed 84% (16 [IQR 9-19] of 19 [IQR 15-21]) of recommended endurance sessions and 90% (9 [IQR 4-12] of

the 10 [IQR 8-13]) of recommended strength training sessions. In the study by Lambert et al [64], only 1 participant used all 13 modules: the introduction module, 8 weekly modules, 1 generic problem-solving module, and 3 unlockable modules. A total of 53% (17/32) of the participants completed at least the introduction module and the first 2 weekly modules, and 25% (8/32) participants completed at least 4 weekly modules. In the study Ström et al [65], 58% (14/24) of the intervention participants completed all the 9 modules.

Effectiveness of Interventions

All 4 studies [62-65] reported a change in depressive symptoms as the main outcome and change in PA as a secondary outcome. All studies observed a positive effect of the intervention on depression severity between baseline and posttreatment, 3 (75%) of which [62,64,65] reported significant and moderate or large between-group differences.

The app-based intervention by Guo et al [62] found a moderate to large effect on depressive symptoms in the intervention group versus the waitlist control group (mean difference −5.77; 95% CI −7.8 to −3.71; Cohen d=0.66; $P<.001$), which was sustained at 6 months (Cohen d=0.63; $P<.001$) and 9 months (Cohen d=0.51; $P<.001$) after intervention. However, there were no significant changes in PA levels (metabolic equivalents) from baseline to follow-up in either group.

Haller et al [63], Lambert et al [64], and Ström et al [65] all examined web-based PA interventions. Similar to Guo et al [62], Haller et al [63] reported significantly reduced depressive symptoms on the Quick Inventory of Depressive Symptomatology Clinician Rating (P=.02) and Self-Report Quick Inventory of Depressive Symptomatology (P=.001) from before to after the intervention. However, there were no significant differences between the intervention and control groups in reducing depressive symptoms after 6 to 12 days (η^2=0.2; P=.06) and after treatment. In terms of PA improvement, Haller et al [63] reported significantly increased total habitual PA (η^2=0.36; P=.007) in the intervention group after the 8-week intervention. In the study by Ström et al [65], the web-based intervention was effective in reducing depressive symptoms, reflected both in Montgomery-Åsberg Depression Rating Scale: Self-rated version and BDI-II. For example, the BDI-II results showed a moderate between-group effect size (Cohen d=0.67; 95% CI 0.09-1.25). Ström et al [65] also reported increased PA levels in both the intervention group and the control group. However, this change did not significantly differ between groups. In Lambert et al [64], the results showed greater changes in depressive symptoms (Patient Health Questionnaire 8-item scores: adjusted mean difference −3.6; 95% CI −6.1 to −1.1) at 2 months in the intervention group. The intervention group also reported a higher median of minutes of device-measured MVPA in 10-minute bouts (97.6 [IQR 49.7-166.3]) than the control group (13.0 [IQR 0.0-131.4]), although this difference was not statistically significant.

Theory and BCTs

A large body of evidence illustrates that interventions are more effective when they are informed by theory and integrated with evidence-based BCTs. The studies by Lambert et al [64,66]

were the only study that explicitly describes the adoption of a behavior change or knowledge translation framework, the Centre for eHealth Research and Disease Management road map, to guide intervention design and evaluation.

Three (75%) studies were informed by a therapeutic approach: Guo et al [62] adopted cognitive behavioral and stress management principles; Lambert et al [64] was informed by behavior activation; and Ström et al [65] incorporated aspects of acceptance and commitment therapy (ACT) and motivational interviewing. With regard to behavior change theories, Lambert et al [64] applied Self-Determination Theory (SDT) to behavioral activation as the underlying theory of PA behavior change and Ström et al [65] was informed by the SDT and the Transtheoretical (Stages of Change) Model. The remaining studies [62,63] described an atheoretical or eclectic approach that integrated a few common BCTs for PA promotion.

BCTs for increasing PA were coded according to Michie et al [67] BCT Taxonomy. In the intervention groups, "goal setting (behavior)" (n=4), "problem solving" (n=3), "review behavior goals" (n=3), "feedback on behavior" (n=3), "self-monitoring of behavior" (n=3), and "information about health consequences" (n=3) were the 6 most commonly used BCTs. The highest number of BCTs were identified in both Ström et al [65] (n=15) and Lambert et al [64] (n=15). Table 3 provides an overview of BCT use across the studies.

Table 3. Intervention behavior change techniques for physical activity promotion (N=4).

BCT[a]	Guo et al [62], 2020	Haller et al [63], 2018	Lambert et al [64], 2018	Ström et al [65], 2013	Total (n/N)
Number of BCTs, n	7	8	15	15	__[b]
1.1 Goal setting (behavior)	✓[c]	✓	✓	✓	4/4
1.2 Problem solving	✓	—	✓	✓	3/4
1.4 Action planning	—	—	✓	✓	2/4
1.5 Review behavior goals	—	✓	✓	✓	3/4
1.7 Review outcome goals	—	—	—	✓	1/4
1.8 Behavior contract	—	✓	—	—	1/4
2.2 Feedback on behavior	✓	✓	—	✓	3/4
2.3 Self-monitoring of behavior	✓	—	✓	✓	3/4
2.6 Biofeedback	—	✓	—	—	1/4
3.3 Social support (emotional)	—	✓	—	—	1/4
4.1 Instructions on how to perform the behavior	✓	✓	✓	—	3/4
4.4 Behavior experiments	—	—	✓	—	1/4
5.1 Information about health consequences	✓	—	✓	✓	3/4
5.4 Monitoring of emotional consequences	—	—	✓	—	1/4
5.6 Information about emotional consequences	—	—	✓	✓	2/4
6.1 Demonstration of the behavior	—	✓	✓	—	2/4
7.1 Prompts or cues	—	—	✓	—	1/4
8.3 Habit formation	—	—	—	✓	1/4
8.7 Graded tasks	—	—	✓	—	1/4
9.1 Credible source	—	—	✓	—	1/4
9.2 Pros and cons	—	—	—	✓	1/4
9.3 Comparative imagining of future outcomes	—	—	—	✓	1/4
10.4 Social reward	—	—	✓	—	1/4
10.9 Self-reward	—	—	—	✓	1/4
11.2 Reduce negative emotions	✓	—	—	—	1/4
15.3 Focus on past success	—	—	—	✓	1/4
16.2 Imaginary reward	—	—	—	✓	1/4

[a]BCT: behavior change technique.

[b]Not available.

[c]Indicates that a behavior change technique was used.

Risk of Bias Within Studies

The ROB 2 [58,59] tool was used to assess the risk of bias for each outcome (PA and depression) of each included RCT (Figure 2). Overall, the "randomization process" was properly used in all studies. In addition, there were no concerns related to "deviations from the intended interventions" and "selection of the reported result." However, in terms of PA, 2 (50%) studies [63,64] were judged as having a high risk of overall bias owing to "missing outcome data." Three (75%) studies [62,63,65] showed some concerns owing to the subjective measurement

of PA outcomes. In terms of depression, 3 (75%) studies [62-64] presented concerns owing to missing outcome data or outcome measurements.

Quality of Cumulative Evidence

The Grading of Recommendation, Assessment, Development and Evaluation assessments are presented in Table 4. The 4 selected RCTs started as high-quality evidence but were subsequently rated down owing to inconsistency and imprecision.

Table 4. Quality of evidence of outcomes of interest.

Outcomes	Risk of bias	Publication bias	Inconsistency	Indirectness	Imprecision	Quality of evidence
Physical activity	Some concerns to high[a]	Not suspected	Serious	Not serious	Serious	❌ Very low
Depression	Low to some concerns[a]	Not suspected	Not serious	Not serious	Serious	❌ Moderate

[a]Detailed information is presented in Figure 3.

Figure 3. Risk of bias assessments of physical activity and depression of included studies.
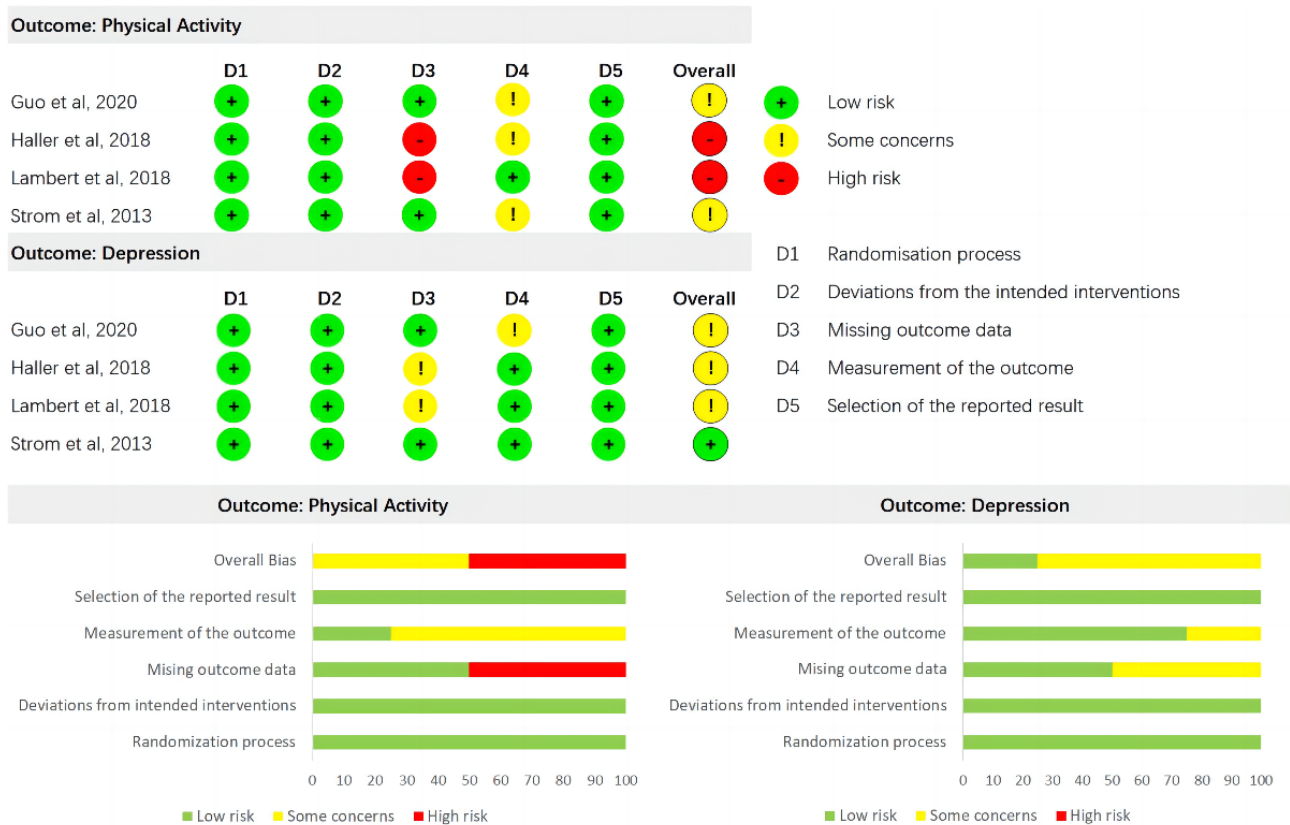
Discussion

Principal Findings

To our knowledge, this is the first systematic review of IGSH interventions to increase PA in people with depression. A total of 4 studies met our inclusion criteria: 3 (75%) web-based RCTs and 1 (25%) app-based RCT. The 3 (75%) studies [62,64,65] reported that PA levels increased at postintervention in both

treatment and control groups. Overall, between-group differences in PA were small and generally not significant. In contrast, 3 (75%) studies found significant medium to large reductions in depressive symptoms; the final study reported a significant reduction in both the intervention and control groups. Collectively, these results suggested that IGSH interventions may be helpful for managing clinical depression.

Why did these interventions fail to increase PA levels? Although explanations may vary across studies, several common factors were noted. In 2 studies [62,64], PA was presented as part of a larger depression management strategy, in conjunction with other components. For example, the Run4Love intervention by Guo et al [62] featured sessions on cognitive behavioral and stress management strategies. As PA was not the primary focus of these interventions, the nonsignificant results were not altogether surprising. More-intensive interventions, that is, those with a central focus on PA behavior change, may be required to produce clinically meaningful changes in PA behavior. From a methodological perspective, the 3 (75%) studies had small sample sizes, which is associated with an increased risk of type 2 error [68].

An additional explanation for the nonsignificant differences in PA is low engagement. Low engagement has previously been identified as a major challenge to eHealth interventions and is likely to negatively impact the effectiveness of interventions [69-71]. In our review, 3 of the 4 (75%) studies were marked by low participant engagement despite different tactics to enhance adherence (eg, reminders via phone calls or emails and tailored feedback). In addition, Lambert et al [64] included some level of tailored content by incorporating graded tasks, where participants could select their preferred physical activities (easier to harder) for the following week from week 3 onward. However, 47% (15/32) of the participants did not reach week 3. The only study with a positive effect on PA change (Haller et al [63]) reported a remarkably higher adherence rate, with participants completing 84% of the recommended endurance sessions and 90% of recommended strength sessions. This study not only tailored the exercise duration or intensity for participants from the beginning of the intervention but also provided weekly personalized feedback. This indicated that personalized IGSH interventions could potentially be more effective than standardized interventions in increasing PA—something that is well-established in the broader literature [72]. Similarly, the importance of personalized components in engagement with technology-based interventions has been emphasized by both people with depression and practitioners [73-75]. Additional supportive strategies, such as providing weekly individualized feedback to enhance adherence, should be incorporated throughout the intervention. Future IGSH interventions should consider combining tailored content based on participants' preferences and conditions (physical or mental) and weekly personalized feedback for potentially improving engagement with the intervention content.

In addition, the types of measurements used may have affected the results. Only Lambert et al [64] incorporated both subjective and device-based PA measurements. Even then, less than half of the participants in both the intervention and control groups provided valid device-based PA data, which may have significantly affected PA outcomes. The other 3 (75%) studies [62,63,65] relied solely on self-reported PA measures, which were subject to recall bias. It is worth noting that none of these self-reported PA measures were developed specifically for people with mental health disorders and may fail to accurately assess PA in the selected studies [76]. Moreover, 3 (75%) studies [62,64,65] incorporated a "self-monitoring of behavior" BCT

component. For example, Ström et al [65] provided the intervention group with a pedometer. Consequently, we would expect the treatment and control groups to differ in their awareness and ability to recall PA. Finally, approximately half of the participants were physically active before enrolling in the study [62,64]. In short, PA promotion effects may be influenced by ceiling effects.

Depressive Symptoms

In contrast to PA outcomes, more consistent evidence was found for depressive symptom reduction. Three of the included studies [62,64,65] reported a medium to large effect on decreasing depression symptoms relative to controls. The final study [63] reported a significant reduction in both the intervention and control groups. These results were consistent with a previous meta-analysis that found IGSH interventions to be effective in reducing depression among college students [47]. As PA is generally unaffected by these interventions, the improvements observed across studies are likely attributable to the various therapeutic components used. Within the field of clinical psychology, evidence-based therapies are regularly informed by theory; for example, CBT, behavioral activation, ACT, and interpersonal therapy all possess distinct theoretical foundations that uniquely guide case conceptualization and treatment activities. The 3 studies incorporated ≥1 therapeutic modalities: cognitive behavioral and stress management approaches [62], behavioral activation [64], and ACT and motivational interviewing [65]. This is in line with an earlier study suggesting that therapist-guided web-based CBT has a large effect on depression outcomes [77], and other studies suggesting behavioral activation strategies, delivered either in person or over the internet, can be as effective as CBT for depression management [78-80].

One of the 4 studies [63] did not observe a significant between-group decrease in depressive symptoms, although early antidepressive findings (6-12 days) were noted on the threshold of significance ($P=.06$). There were 2 possible explanations. First, the sample size was small (n=20) and the groups were unequal (intervention group=14; control group=6). Therefore, it is likely that this study was underpowered [81]. Second, the authors observed that 1 of the 6 participants in the control group reported a full remission from severe depression after treatment. Although not impossible, this is an unusual occurrence and likely resulted in an inflated index of depressive symptom reduction in the control group. Therefore, the between-group difference was considered less reliable. In contrast, the significant and meaningful reduction in depression in the IG group was believed to be attributed to PA participation. In short, all studies reported that IGSH interventions are feasible and effective in treating depression.

Intervention Characteristics and BCTs

In general, research suggests that theory-based interventions are more effective than interventions without theory [45,46,82]. In addition to providing a framework for selecting evidence-based BCTs, theories also aid in understanding factors that mediate behavior changes and the reasons for intervention success or failure [83,84]. Only 1 study in this review explicitly identified the empirical or theoretical basis for BCT selection

for PA promotion. Specifically, Lambert et al [64,66] reported that SDT was adopted to inform their eMotion intervention. Similarly, Ström et al [65] were inspired by an earlier study that used SDT as the theoretical framework. Apart from theories, the importance of intervention development frameworks for guiding the development process of behavior change interventions has also been highlighted to reduce the risk of research waste and increase the effectiveness and sustainability of interventions [85,86]. Only Lambert et al [64,66] identified a framework: in their case, the Center for eHealth Research and Disease Management road map.

In contrast to theory, all 4 interventions used multiple BCTs. The most common BCTs identified were goal setting, problem solving, feedback, reviewing behavioral goals, providing information about health consequences, and instructions on how to perform the behavior. As noted by Bohlen et al [82], these BCTs are frequently used in interventions for the general population. Within the literature, there is growing interest in how to optimally combine BCTs to support behavior change, with several studies suggesting that BCT combinations might differ across populations and behavioral targets [87-90]. A previous factorial trial of an internet-based intervention [91] found that combining action planning, coping planning, and self-monitoring induced and amplifies the effect of increasing MVPA in healthy adults. Interestingly, both Lambert et al [64] and Ström et al [65], who included this combination of BCTs, reported no change in PA behavior compared with the control groups. Although methodological limitations must be emphasized, this observed inconsistency raised the question of tailoring BCTs to the population. We cannot assume that strategies that are effective for the general population are a good fit for people with depression. Rather, different or additional BCTs may be required to support PA promotion.

A scoping review of the barriers to and facilitators of exercise in people with depression suggested that interventions should incorporate BCTs that target the emotion domain (eg, to address low mood) [92]. Cane et al [93] recommended 4 BCTs for the emotion domain: "social support (emotional)," "information about emotional consequences," "self-assessment of affective consequences," and "reduce negative emotions." In our review, all 4 studies [62-65] addressed aspects of emotion (eg, both Lambert et al [64] and Ström et al [65] touched upon the emotional consequences of PA), although none of the studies used all 4 BCTs to target this domain. The lack of population-tailored BCTs might partly explain the inconsistent results regarding PA outcomes. Future IGSH interventions for PA behavior change among people with depression should test different combinations of BCTs and consider integrating more BCTs that focus on the emotion domain.

Attrition and Engagement

The included IGSH interventions showed relatively low dropout rates (0%-19%). In contrast, Meyerowitz-Katz et al [94] reported a dropout rate of 40% (95% CI 16%-63%) among RCTs on mobile health interventions. Josephine et al [95] also found a mean intervention dropout rate of 37% within internet-based and mobile-based interventions for people with depression. There have been some studies suggesting that higher levels of

therapy contact have a positive impact on acceptance of the intervention [96,97]. All the included studies incorporated at least a moderate level of participant contact, including weekly contact with study-affiliated personnel. Three studies [62,63,65] provided individualized support; these reported lower attrition and higher study engagement compared with the study by Lambert et al [64], who provided weekly automated reminder emails. This is consistent with previous studies that suggested regular motivational feedback may enhance the adherence to internet-based intervention [98,99]. Similarly, our findings suggest that regular individualized feedback may contribute to reduced dropout rates. Further research is needed to explore the effects of different levels of interventionist contact on IGSH interventions.

Strengths and Limitations

This review provides an updated summary of IGSH interventions for PA promotion in people with depression. Two rounds of searches were conducted to ensure inclusion of all available evidence. A notable strength of this review was the application of the BCT taxonomy to identify potentially useful ingredients for informing future intervention development. We addressed the gaps observed in previous reviews by including only RCTs and specifically focusing on people with depression.

Similar to all studies, this review was not without limitations. First, both the limited number of studies in this review and their heterogeneity in measurement precluded the meta-analysis and limited our ability to draw clear conclusions. Encouragingly, this is a rapidly evolving field, and several promising protocols were identified during article screening (eg, study by Sylvia et al [100]). We are hopeful that more evidence will become available in the next several years. Second, none of the studies had a primary outcome of PA, and because of the measures used, there is a very low level of certainty of PA evidence included. Third, only studies published in English were searched and screened. Fourth, this study used an a priori definition of IGSH interventions, which conceptualized these interventions as technology-facilitated and primarily self-guided, with the option of in-person support. Although having a clear definition is a strength, we did not explicitly designate the degree of in-person support that would render the studies ineligible. Study selection was determined by researchers through a process of discussion. As such, some studies were excluded from this study, which other research groups may have considered eligible or ineligible. Finally, we coded the BCTs using a dichotomous system (ie, yes or no). This review was unable to speak to the quality and rigor of BCT administration within the included studies.

Future Directions

Despite increasing interest in eHealth and IGSH interventions, this review identified only 4 studies that (1) included a PA intervention component, (2) assessed changes in PA, and (3) specifically targeted people with depression. However, there were notable gaps in the research design. None of the 4 identified studies identified PA behavior change as their primary objective; interventions contained relatively little content focused on PA promotion; studies generally featured small samples; and half of the included studies did not incorporate

behavior change theory for BCT selection or follow a systematic framework of behavioral intervention development. In contrast, not only is PA now recommended as a first-line monotherapy for mild to moderate depression but also there is a large and robust body of evidence regarding the different theories, BCT, and characteristics that can support successful eHealth or IGSH intervention. In summary, there appears to be a large gap between general research and its specific application in IGSH interventions to promote PA in people with depression.

Therefore, there is a great need for high-quality and thoughtful research on IGSH interventions for people with depression. On the basis of the results of this review, future research should be characterized by the following: (1) a specific focus on PA promotion within people with depression, including diverse populations of people with depression (eg, older adults, racialized communities, and individuals in rural or remote communities); (2) a priori consideration of theory, including using theory to guide BCT selection; (3) a rigorous development process following a systematic intervention development framework; (4) defining specific intervention targets (eg, meeting Canadian Network for Mood and Anxiety Treatments Guidelines [16] of 3×30-minute bouts of MVPA per week); (5)

exploration of the mediators and moderators of behavior change, as defined by theory; and (6) using validated tools to assess before and after changes in depressive symptoms and PA, with inclusion of objective measures when possible. Further questions for examination included exploring the optimum amount and modality of guided support, personalization and tailoring considerations, PA programing (eg, frequency, intensity, type, and time), the moderating effect of patient characteristics (eg, baseline fitness and symptom severity), and knowledge translation and intervention scaling.

Conclusions

An emerging body of evidence suggests that IGSH PA interventions are feasible and have the potential to reduce depressive symptoms in people with depression. More well-designed and tailored interventions are needed to assess the overall efficacy and feasibility of using IGSH interventions to help people with depression increase PA. Future research on such interventions should be theoretically informed in its development and implementation and test different combinations of BCTs, particularly those targeting the emotion domain, to verify their efficacy in increasing PA among people with depression.

Authors' Contributions

YT developed the search strategies and conducted 2 rounds of study search. YL and GF conducted the first round of the screening and data extraction. YL, MG, and GF conducted the second round of screening and reviewed the included studies. YL interpreted the findings and drafted the manuscript. RL, SL, GF, and MG critically reviewed and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 checklist.
[DOCX File , 29 KB - mental_v9i12e38049_app1.docx]

Multimedia Appendix 2
Search strategies for the first round of searches.
[DOCX File , 35 KB - mental_v9i12e38049_app2.docx]

Multimedia Appendix 3
Search strategies for the second round of searches.
[DOCX File , 33 KB - mental_v9i12e38049_app3.docx]

References

1. Angermeyer MC, Holzinger A, Matschinger H, Stengler-Wenzke K. Depression and quality of life: results of a follow-up study. Int J Soc Psychiatry 2002 Sep;48(3):189-199. [doi: 10.1177/002076402128783235] [Medline: 12413247]
2. Hansson L. Quality of life in depression and anxiety. Int Rev Psychiatry 2002;14(3):185-189. [doi: 10.1080/09540260220144966]

3. Sivertsen H, Bjørkløf GH, Engedal K, Selbæk G, Helvik AS. Depression and quality of life in older persons: a review. Dement Geriatr Cogn Disord 2015;40(5-6):311-339 [FREE Full text] [doi: 10.1159/000437299] [Medline: 26360014]

4. Cavanagh JT, Carson AJ, Sharpe M, Lawrie SM. Psychological autopsy studies of suicide: a systematic review. Psychol Med 2003 Apr;33(3):395-405. [doi: 10.1017/s0033291702006943] [Medline: 12701661]

5. Ferrari AJ, Charlson FJ, Norman RE, Patten SB, Freedman G, Murray CJ, 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]

6. Lotfaliany M, Bowe SJ, Kowal P, Orellana L, Berk M, Mohebbi M. Depression and chronic diseases: co-occurrence and communality of risk factors. J Affect Disord 2018 Dec 01;241:461-468. [doi: 10.1016/j.jad.2018.08.011] [Medline: 30149333]

7. The WHO special initiative for mental health (2019-2023): universal health coverage for mental health. World Health Organization. 2019. URL: https://www.who.int/publications/i/item/special-initiative-for-mental-health-(2019-2023) [accessed 2022-02-15]

8. 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 [accessed 2022-02-15]

9. Arokiasamy P, Uttamacharya, Kowal P, Capistrant BD, Gildner TE, Thiele E, et al. Chronic noncommunicable diseases in 6 low- and middle-income countries: findings from wave 1 of the World Health Organization's study on global ageing and adult health (SAGE). Am J Epidemiol 2017 Mar 15;185(6):414-428 [FREE Full text] [doi: 10.1093/aje/kww125] [Medline: 28399566]

10. Lecrubier Y. Widespread underrecognition and undertreatment of anxiety and mood disorders: results from 3 European studies. J Clin Psychiatry 2007;68 Suppl 2:36-41. [Medline: 17288506]

11. Bueno-Notivol J, Gracia-García P, Olaya B, Lasheras I, López-Antón R, Santabárbara J. Prevalence of depression during the COVID-19 outbreak: a meta-analysis of community-based studies. Int J Clin Health Psychol 2021;21(1):100196 [FREE Full text] [doi: 10.1016/j.ijchp.2020.07.007] [Medline: 32904715]

12. Nochaiwong S, Ruengorn C, Thavorn K, Hutton B, Awiphan R, Phosuya C, et al. Global prevalence of mental health issues among the general population during the coronavirus disease-2019 pandemic: a systematic review and meta-analysis. Sci Rep 2021 May 13;11(1):10173 [FREE Full text] [doi: 10.1038/s41598-021-89700-8] [Medline: 33986414]

13. Caspersen CJ, Powell KE, Christenson GM. Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. Public Health Rep 1985;100(2):126-131 [FREE Full text] [Medline: 3920711]

14. Kandola A, Ashdown-Franks G, Hendrikse J, Sabiston CM, Stubbs B. Physical activity and depression: towards understanding the antidepressant mechanisms of physical activity. Neurosci Biobehav Rev 2019 Dec;107:525-539. [doi: 10.1016/j.neubiorev.2019.09.040] [Medline: 31586447]

15. Teychenne M, White RL, Richards J, Schuch FB, Rosenbaum S, Bennie JA. Do we need physical activity guidelines for mental health: what does the evidence tell us? Ment Health Phys Act 2020 Mar;18:100315 [FREE Full text] [doi: 10.1016/j.mhpa.2019.100315]

16. Ravindran AV, Balneaves LG, Faulkner G, Ortiz A, McIntosh D, Morehouse RL, CANMAT Depression Work Group. Canadian Network for Mood and Anxiety Treatments (CANMAT) 2016 clinical guidelines for the management of adults with major depressive disorder: section 5. Complementary and alternative medicine treatments. Can J Psychiatry 2016 Sep;61(9):576-587 [FREE Full text] [doi: 10.1177/0706743716660290] [Medline: 27486153]

17. Blumenthal JA, Smith PJ, Hoffman BM. Is exercise a viable treatment for depression? ACSMs Health Fit J 2012 Jul;16(4):14-21 [FREE Full text] [doi: 10.1249/01.FIT.0000416000.09526.eb] [Medline: 23750100]

18. Rebar AL, Taylor A. Physical activity and mental health; it is more than just a prescription. Ment Health Phys Act 2017 Oct;13:77-82. [doi: 10.1016/j.mhpa.2017.10.004]

19. Anokye NK, Trueman P, Green C, Pavey TG, Taylor RS. Physical activity and health related quality of life. BMC Public Health 2012 Aug 07;12:624 [FREE Full text] [doi: 10.1186/1471-2458-12-624] [Medline: 22871153]

20. Carlson SA, Adams EK, Yang Z, Fulton JE. Percentage of deaths associated with inadequate physical activity in the United States. Prev Chronic Dis 2018 Mar 29;15:E38 [FREE Full text] [doi: 10.5888/pcd18.170354] [Medline: 29602315]

21. Kraus WE, Powell KE, Haskell WL, Janz KF, Campbell WW, Jakicic JM, 2018 PHYSICAL ACTIVITY GUIDELINES ADVISORY COMMITTEE*. Physical activity, all-cause and cardiovascular mortality, and cardiovascular disease. Med Sci Sports Exerc 2019 Jun;51(6):1270-1281 [FREE Full text] [doi: 10.1249/MSS.0000000000001939] [Medline: 31095084]

22. Hallal PC, Andersen LB, Bull FC, Guthold R, Haskell W, Ekelund U, Lancet Physical Activity Series Working Group. Global physical activity levels: surveillance progress, pitfalls, and prospects. Lancet 2012 Jul 21;380(9838):247-257. [doi: 10.1016/S0140-6736(12)60646-1] [Medline: 22818937]

23. Physical Activity, Sedentary Behaviour and Sleep (PASS) Indicators Data Tool. Center for Surveillance and Applied Research, Public Health Agency of Canada. Ottawa, ON, Canada: Public Health Agency of Canada; 2021. URL: https://health-infobase.canada.ca/pass/ [accessed 2022-02-15]

24. Song MR, Lee YS, Baek JD, Miller M. Physical activity status in adults with depression in the National Health and Nutrition Examination Survey, 2005-2006. Public Health Nurs 2012;29(3):208-217. [doi: 10.1111/j.1525-1446.2011.00986.x] [Medline: 22512422]

25. Stubbs B, Koyanagi A, Schuch FB, Firth J, Rosenbaum S, Veronese N, et al. Physical activity and depression: a large cross-sectional, population-based study across 36 low- and middle-income countries. Acta Psychiatr Scand 2016 Dec;134(6):546-556. [doi: 10.1111/acps.12654] [Medline: 27704532]

26. Vancampfort D, Firth J, Schuch FB, Rosenbaum S, Mugisha J, Hallgren M, et al. Sedentary behavior and physical activity levels in people with schizophrenia, bipolar disorder and major depressive disorder: a global systematic review and meta-analysis. World Psychiatry 2017 Oct;16(3):308-315 [FREE Full text] [doi: 10.1002/wps.20458] [Medline: 28941119]

27. Gerber M, Holsboer-Trachsler E, Pühse U, Brand S. Exercise is medicine for patients with major depressive disorders: but only if the "pill" is taken!. Neuropsychiatr Dis Treat 2016 Aug 5;12:1977-1981 [FREE Full text] [doi: 10.2147/NDT.S110656] [Medline: 27540294]

28. Ritterband LM, Andersson G, Christensen HM, Carlbring P, Cuijpers P. Directions for the international society for research on internet interventions (ISRII). J Med Internet Res 2006 Sep 29;8(3):e23 [FREE Full text] [doi: 10.2196/jmir.8.3.e23] [Medline: 17032639]

29. Strecher V. Internet methods for delivering behavioral and health-related interventions (eHealth). Annu Rev Clin Psychol 2007;3:53-76. [doi: 10.1146/annurev.clinpsy.3.022806.091428] [Medline: 17716048]

30. Proudfoot J, Parker G, Hadzi Pavlovic D, Manicavasagar V, Adler E, Whitton A. Community attitudes to the appropriation of mobile phones for monitoring and managing depression, anxiety, and stress. J Med Internet Res 2010 Dec 19;12(5):e64 [FREE Full text] [doi: 10.2196/jmir.1475] [Medline: 21169174]

31. Torous J, Friedman R, Keshavan M. Smartphone ownership and interest in mobile applications to monitor symptoms of mental health conditions. JMIR Mhealth Uhealth 2014 Jan 21;2(1):e2 [FREE Full text] [doi: 10.2196/mhealth.2994] [Medline: 25098314]

32. Borgueta AM, Purvis CK, Newman MG. Navigating the ethics of Internet-guided self-help interventions. Clin Psychol (New York) 2018 Jun;25(2):e12235 [FREE Full text] [doi: 10.1111/cpsp.12235] [Medline: 33664555]

33. Naslund JA, Marsch LA, McHugo GJ, Bartels SJ. Emerging mHealth and eHealth interventions for serious mental illness: a review of the literature. J Ment Health 2015;24(5):321-332 [FREE Full text] [doi: 10.3109/09638237.2015.1019054] [Medline: 26017625]

34. Barak A, Klein B, Proudfoot JG. Defining internet-supported therapeutic interventions. Ann Behav Med 2009 Aug;38(1):4-17. [doi: 10.1007/s12160-009-9130-7] [Medline: 19787305]

35. McCrabb S, Lane C, Hall A, Milat A, Bauman A, Sutherland R, et al. Scaling-up evidence-based obesity interventions: a systematic review assessing intervention adaptations and effectiveness and quantifying the scale-up penalty. Obes Rev 2019 Jul;20(7):964-982. [doi: 10.1111/obr.12845] [Medline: 30868745]

36. Van't Hof E, Cuijpers P, Stein DJ. Self-help and Internet-guided interventions in depression and anxiety disorders: a systematic review of meta-analyses. CNS Spectr 2009 Feb;14(2 Suppl 3):34-40. [doi: 10.1017/s1092852900027279] [Medline: 19238128]

37. Clement S, Schauman O, Graham T, Maggioni F, Evans-Lacko S, Bezborodovs N, et al. What is the impact of mental health-related stigma on help-seeking? A systematic review of quantitative and qualitative studies. Psychol Med 2015 Jan;45(1):11-27. [doi: 10.1017/S0033291714000129] [Medline: 24569086]

38. Jennings KS, Cheung JH, Britt TW, Goguen KN, Jeffirs SM, Peasley AL, et al. How are perceived stigma, self-stigma, and self-reliance related to treatment-seeking? A three-path model. Psychiatr Rehabil J 2015 Jun;38(2):109-116. [doi: 10.1037/prj0000138] [Medline: 25844914]

39. Levin ME, Krafft J, Levin C. Does self-help increase rates of help seeking for student mental health problems by minimizing stigma as a barrier? J Am Coll Health 2018;66(4):302-309. [doi: 10.1080/07448481.2018.1440580] [Medline: 29447600]

40. Lau PW, Lau EY, Wong DP, Ransdell L. A systematic review of information and communication technology-based interventions for promoting physical activity behavior change in children and adolescents. J Med Internet Res 2011 Jul 13;13(3):e48 [FREE Full text] [doi: 10.2196/jmir.1533] [Medline: 21749967]

41. McIntosh JR, Jay S, Hadden N, Whittaker PJ. Do E-health interventions improve physical activity in young people: a systematic review. Public Health 2017 Jul;148:140-148. [doi: 10.1016/j.puhe.2017.04.001] [Medline: 28482322]

42. Davies CA, Spence JC, Vandelanotte C, Caperchione CM, Mummery WK. Meta-analysis of Internet-delivered interventions to increase physical activity levels. Int J Behav Nutr Phys Act 2012 Apr 30;9:52 [FREE Full text] [doi: 10.1186/1479-5868-9-52] [Medline: 22546283]

43. Núñez de Arenas-Arroyo S, Cavero-Redondo I, Alvarez-Bueno C, Sequí-Domínguez I, Reina-Gutiérrez S, Martínez-Vizcaíno V. Effect of eHealth to increase physical activity in healthy adults over 55 years: a systematic review and meta-analysis. Scand J Med Sci Sports 2021 Apr;31(4):776-789. [doi: 10.1111/sms.13903] [Medline: 33280182]

44. Muellmann S, Forberger S, Möllers T, Bröring E, Zeeb H, Pischke CR. Effectiveness of eHealth interventions for the promotion of physical activity in older adults: a systematic review. Prev Med 2018 Mar;108:93-110. [doi: 10.1016/j.ypmed.2017.12.026] [Medline: 29289643]

45. Murray E. Web-based interventions for behavior change and self-management: potential, pitfalls, and progress. Med 2 0 2012 Aug 14;1(2):e3 [FREE Full text] [doi: 10.2196/med20.1741] [Medline: 25075231]

46.   Webb TL, Joseph J, Yardley L, Michie S. Using the internet to promote health behavior change: a systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy. J Med Internet Res 2010 Feb 17;12(1):e4 [FREE Full text] [doi: 10.2196/jmir.1376] [Medline: 20164043]

47.   Ma L, Huang C, Tao R, Cui Z, Schluter P. Meta-analytic review of online guided self-help interventions for depressive symptoms among college students. Internet Interv 2021 Sep;25:100427 [FREE Full text] [doi: 10.1016/j.invent.2021.100427] [Medline: 34401386]

48.   Depression in adults: recognition and management. National Institute for Health and Care Excellence. 2009 Oct 28. URL: https://www.nice.org.uk/guidance/cg90/ifp/chapter/Treatments-for-mild-to-moderate-depression [accessed 2022-02-20]

49.   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]

50.   Moran J, Kelly G, Haberlin C, Mockler D, Broderick J. The use of eHealth to promote physical activity in people with mental health conditions: a systematic review. HRB Open Res 2018 Aug 10;1:5 [FREE Full text] [doi: 10.12688/hrbopenres.12796.3]

51.   Rosenbaum S, Newby JM, Steel Z, Andrews G, Ward PB. Online physical activity interventions for mental disorders: a systematic review. Internet Interv 2015 May;2(2):214-220 [FREE Full text] [doi: 10.1016/j.invent.2015.04.001]

52.   Carneiro L, Rosenbaum S, Ward PB, Clemente FM, Ramirez-Campillo R, Monteiro-Júnior RS, et al. Web-based exercise interventions for patients with depressive and anxiety disorders: a systematic review of randomized controlled trials. Braz J Psychiatry 2022;44(3):331-341 [FREE Full text] [doi: 10.1590/1516-4446-2021-2026] [Medline: 34852034]

53.   Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. Syst Rev 2021 Mar 29;10(1):89 [FREE Full text] [doi: 10.1186/s13643-021-01626-4] [Medline: 33781348]

54.   Page MJ, Moher D, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. BMJ 2021 Mar 29;372:n160 [FREE Full text] [doi: 10.1136/bmj.n160] [Medline: 33781993]

55.   Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. J Clin Epidemiol 2009 Oct;62(10):e1-34 [FREE Full text] [doi: 10.1016/j.jclinepi.2009.06.006] [Medline: 19631507]

56.   Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Ann Intern Med 2009 Aug 18;151(4):264-W64 [FREE Full text] [doi: 10.7326/0003-4819-151-4-200908180-00135] [Medline: 19622511]

57.   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]

58.   Higgins JP, Savović J, Page MJ, Elbers RG, Sterne JA. Assessing risk of bias in a randomized trial. In: Higgins JP, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al, editors. Cochrane Handbook for Systematic Reviews of Interventions. 2nd edition. Hoboken, NJ, USA: Wiley; 2019:205-228.

59.   Sterne JA, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ 2019 Aug 28;366:l4898 [FREE Full text] [doi: 10.1136/bmj.l4898] [Medline: 31462531]

60.   Nakagawa S, Cuthill IC. Effect size, confidence interval and statistical significance: a practical guide for biologists. Biol Rev Camb Philos Soc 2007 Nov;82(4):591-605. [doi: 10.1111/j.1469-185X.2007.00027.x] [Medline: 17944619]

61.   Balshem H, Helfand M, Schünemann HJ, Oxman AD, Kunz R, Brozek J, et al. GRADE guidelines: 3. Rating the quality of evidence. J Clin Epidemiol 2011 Apr;64(4):401-406. [doi: 10.1016/j.jclinepi.2010.07.015] [Medline: 21208779]

62.   Guo Y, Hong YA, Cai W, Li L, Hao Y, Qiao J, et al. Effect of a WeChat-based intervention (Run4Love) on depressive symptoms among people living with HIV in China: a randomized controlled trial. J Med Internet Res 2020 Feb 09;22(2):e16715 [FREE Full text] [doi: 10.2196/16715] [Medline: 32044751]

63.   Haller N, Lorenz S, Pfirrmann D, Koch C, Lieb K, Dettweiler U, et al. Individualized Web-based exercise for the treatment of depression: randomized controlled trial. JMIR Ment Health 2018 Oct 12;5(4):e10698 [FREE Full text] [doi: 10.2196/10698] [Medline: 30314962]

64.   Lambert JD, Greaves CJ, Farrand P, Price L, Haase AM, Taylor AH. Web-based intervention using behavioral activation and physical activity for adults with depression (the eMotion study): pilot randomized controlled trial. J Med Internet Res 2018 Jul 16;20(7):e10112 [FREE Full text] [doi: 10.2196/10112] [Medline: 30012547]

65.   Ström M, Uckelstam C, Andersson G, Hassmén P, Umefjord G, Carlbring P. Internet-delivered therapist-guided physical activity for mild to moderate depression: a randomized controlled trial. PeerJ 2013 Oct 3;1:e178 [FREE Full text] [doi: 10.7717/peerj.178] [Medline: 24109561]

66.   Lambert JD, Greaves CJ, Farrand P, Haase AM, Taylor AH. Development of a web-based intervention (eMotion) based on behavioural activation to promote physical activity in people with depression. Ment Health Phys Act 2017 Oct;13:120-136. [doi: 10.1016/j.mhpa.2017.10.003]

67.   Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change

interventions. Ann Behav Med 2013 Aug;46(1):81-95 [FREE Full text] [doi: 10.1007/s12160-013-9486-6] [Medline: 23512568]

68. Banerjee A, Chitnis UB, Jadhav SL, Bhawalkar JS, Chaudhury S. Hypothesis testing, type I and type II errors. Ind Psychiatry J 2009 Jul;18(2):127-131 [FREE Full text] [doi: 10.4103/0972-6748.62274] [Medline: 21180491]

69. Andrews G, Basu A, Cuijpers P, Craske MG, McEvoy P, English CL, et al. Computer therapy for the anxiety and depression disorders is effective, acceptable and practical health care: an updated meta-analysis. J Anxiety Disord 2018 Apr;55:70-78 [FREE Full text] [doi: 10.1016/j.janxdis.2018.01.001] [Medline: 29422409]

70. Schubart JR, Stuckey HL, Ganeshamoorthy A, Sciamanna CN. Chronic health conditions and Internet behavioral interventions: a review of factors to enhance user engagement. Comput Inform Nurs 2011 Feb;29(2):81-92 [FREE Full text] [doi: 10.1097/NCN.0b013e3182065eed] [Medline: 21164337]

71. Schure MB, Lindow JC, Greist JH, Nakonezny PA, Bailey SJ, Bryan WL, et al. Use of a fully automated Internet-based cognitive behavior therapy intervention in a community population of adults with depression symptoms: randomized controlled trial. J Med Internet Res 2019 Nov 18;21(11):e14754 [FREE Full text] [doi: 10.2196/14754] [Medline: 31738173]

72. Ghanvatkar S, Kankanhalli A, Rajan V. User models for personalized physical activity interventions: scoping review. JMIR Mhealth Uhealth 2019 Jan 16;7(1):e11098 [FREE Full text] [doi: 10.2196/11098] [Medline: 30664474]

73. Patoz MC, Hidalgo-Mazzei D, Blanc O, Verdolini N, Pacchiarotti I, Murru A, et al. Patient and physician perspectives of a smartphone application for depression: a qualitative study. BMC Psychiatry 2021 Jan 29;21(1):65 [FREE Full text] [doi: 10.1186/s12888-021-03064-x] [Medline: 33514333]

74. Stiles-Shields C, Montague E, Lattie EG, Kwasny MJ, Mohr DC. What might get in the way: barriers to the use of apps for depression. Digit Health 2017 Jun 8;3:2055207617713827 [FREE Full text] [doi: 10.1177/2055207617713827] [Medline: 29942605]

75. Doherty G, Coyle D, Sharry J. Engagement with online mental health interventions: an exploratory clinical study of a treatment for depression. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2012 Presented at: CHI '12; May 5-10, 2012; Austin, TX, USA p. 1421-1430. [doi: 10.1145/2207676.2208602]

76. Rosenbaum S, Morell R, Abdel-Baki A, Ahmadpanah M, Anilkumar TV, Baie L, et al. Assessing physical activity in people with mental illness: 23-country reliability and validity of the simple physical activity questionnaire (SIMPAQ). BMC Psychiatry 2020 Mar 06;20(1):108 [FREE Full text] [doi: 10.1186/s12888-020-2473-0] [Medline: 32143714]

77. Ruwaard J, Schrieken B, Schrijver M, Broeksteeg J, Dekker J, Vermeulen H, et al. Standardized web-based cognitive behavioural therapy of mild to moderate depression: a randomized controlled trial with a long-term follow-up. Cogn Behav Ther 2009;38(4):206-221. [doi: 10.1080/16506070802408086] [Medline: 19221919]

78. Richards DA, Ekers D, McMillan D, Taylor RS, Byford S, Warren FC, et al. Cost and Outcome of Behavioural Activation versus Cognitive Behavioural Therapy for Depression (COBRA): a randomised, controlled, non-inferiority trial. Lancet 2016 Aug 27;388(10047):871-880 [FREE Full text] [doi: 10.1016/S0140-6736(16)31140-0] [Medline: 27461440]

79. Carlbring P, Hägglund M, Luthström A, Dahlin M, Kadowaki Å, Vernmark K, et al. Internet-based behavioral activation and acceptance-based treatment for depression: a randomized controlled trial. J Affect Disord 2013 Jun;148(2-3):331-337. [doi: 10.1016/j.jad.2012.12.020] [Medline: 23357657]

80. Dimidjian S, Barrera Jr 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]

81. Sullivan GM, Feinn R. Using effect size-or why the P value is not enough. J Grad Med Educ 2012 Sep;4(3):279-282 [FREE Full text] [doi: 10.4300/JGME-D-12-00156.1] [Medline: 23997866]

82. Bohlen LC, Michie S, de Bruin M, Rothman AJ, Kelly MP, Groarke HN, et al. Do combinations of behavior change techniques that occur frequently in interventions reflect underlying theory? Ann Behav Med 2020 Nov 01;54(11):827-842 [FREE Full text] [doi: 10.1093/abm/kaaa078] [Medline: 32959875]

83. Brawley LR. The practicality of using social psychological theories for exercise and health research and intervention. J Appl Sport Psychol 1993 Sep;5(2):99-115. [doi: 10.1080/10413209308411309]

84. Stacey FG, James EL, Chapman K, Courneya KS, Lubans DR. A systematic review and meta-analysis of social cognitive theory-based physical activity and/or nutrition behavior change interventions for cancer survivors. J Cancer Surviv 2015 Jun;9(2):305-338 [FREE Full text] [doi: 10.1007/s11764-014-0413-z] [Medline: 25432633]

85. O'Cathain A, Croot L, Sworn K, Duncan E, Rousseau N, Turner K, et al. Taxonomy of approaches to developing interventions to improve health: a systematic methods overview. Pilot Feasibility Stud 2019 Mar 12;5:41 [FREE Full text] [doi: 10.1186/s40814-019-0425-6] [Medline: 30923626]

86. Czajkowski SM, Hunter CM. From ideas to interventions: a review and comparison of frameworks used in early phase behavioral translation research. Health Psychol 2021 Dec;40(12):829-844. [doi: 10.1037/hea0001095] [Medline: 34990185]

87. French DP, Olander EK, Chisholm A, Mc Sharry J. Which behaviour change techniques are most effective at increasing older adults' self-efficacy and physical activity behaviour? A systematic review. Ann Behav Med 2014 Oct;48(2):225-234. [doi: 10.1007/s12160-014-9593-z] [Medline: 24648017]

XSL•FO

RenderX

88. McEwan D, Harden SM, Zumbo BD, Sylvester BD, Kaulius M, Ruissen GR, et al. The effectiveness of multi-component goal setting interventions for changing physical activity behaviour: a systematic review and meta-analysis. Health Psychol Rev 2016;10(1):67-88. [doi: 10.1080/17437199.2015.1104258] [Medline: 26445201]

89. Olander EK, Fletcher H, Williams S, Atkinson L, Turner A, French DP. What are the most effective techniques in changing obese individuals' physical activity self-efficacy and behaviour: a systematic review and meta-analysis. Int J Behav Nutr Phys Act 2013 Mar 03;10:29 [FREE Full text] [doi: 10.1186/1479-5868-10-29] [Medline: 23452345]

90. van Genugten L, Dusseldorp E, Webb TL, van Empelen P. Which combinations of techniques and modes of delivery in internet-based interventions effectively change health behavior? A meta-analysis. J Med Internet Res 2016 Jun 07;18(6):e155 [FREE Full text] [doi: 10.2196/jmir.4218] [Medline: 27268104]

91. Schroé H, Van Dyck D, De Paepe A, Poppe L, Loh WW, Verloigne M, et al. Which behaviour change techniques are effective to promote physical activity and reduce sedentary behaviour in adults: a factorial randomized trial of an e- and m-health intervention. Int J Behav Nutr Phys Act 2020 Oct 07;17(1):127 [FREE Full text] [doi: 10.1186/s12966-020-01001-x] [Medline: 33028335]

92. Glowacki K, Duncan MJ, Gainforth H, Faulkner G. Barriers and facilitators to physical activity and exercise among adults with depression: a scoping review. Ment Health Phys Act 2017 Oct;13:108-119. [doi: 10.1016/j.mhpa.2017.10.001]

93. Cane J, Richardson M, Johnston M, Ladha R, Michie S. From lists of behaviour change techniques (BCTs) to structured hierarchies: comparison of two methods of developing a hierarchy of BCTs. Br J Health Psychol 2015 Feb;20(1):130-150. [doi: 10.1111/bjhp.12102] [Medline: 24815766]

94. Meyerowitz-Katz G, Ravi S, Arnolda L, Feng X, Maberly G, Astell-Burt T. Rates of attrition and dropout in app-based interventions for chronic disease: systematic review and meta-analysis. J Med Internet Res 2020 Sep 29;22(9):e20283 [FREE Full text] [doi: 10.2196/20283] [Medline: 32990635]

95. Josephine K, Josefine L, Philipp D, David E, Harald B. Internet- and mobile-based depression interventions for people with diagnosed depression: a systematic review and meta-analysis. J Affect Disord 2017 Dec 01;223:28-40. [doi: 10.1016/j.jad.2017.07.021] [Medline: 28715726]

96. Bennett SD, Cuijpers P, Ebert DD, McKenzie Smith M, Coughtrey AE, Heyman I, et al. Practitioner Review: unguided and guided self-help interventions for common mental health disorders in children and adolescents: a systematic review and meta-analysis. J Child Psychol Psychiatry 2019 Aug;60(8):828-847. [doi: 10.1111/jcpp.13010] [Medline: 30775782]

97. O'Brien M, Daley D. Self-help parenting interventions for childhood behaviour disorders: a review of the evidence. Child Care Health Dev 2011 Sep;37(5):623-637. [doi: 10.1111/j.1365-2214.2011.01231.x] [Medline: 21585419]

98. Ströhle A. Physical activity, exercise, depression and anxiety disorders. J Neural Transm (Vienna) 2009 Jun;116(6):777-784. [doi: 10.1007/s00702-008-0092-x] [Medline: 18726137]

99. Andersson G, Bergström J, Holländare F, Carlbring P, Kaldo V, Ekselius L. Internet-based self-help for depression: randomised controlled trial. Br J Psychiatry 2005 Nov;187:456-461. [doi: 10.1192/bjp.187.5.456] [Medline: 16260822]

100. Sylvia LG, Faulkner M, Rakhilin M, Amado S, Gold AK, Albury EA, et al. An online intervention for increasing physical activity in individuals with mood disorders at risk for cardiovascular disease: design considerations. J Affect Disord 2021 Aug 01;291:102-109. [doi: 10.1016/j.jad.2021.04.094] [Medline: 34029880]

Abbreviations

ACT: acceptance and commitment therapy
BCT: behavior change technique
BDI-II: Beck Depression Inventory-second version
CBT: cognitive behavioral therapy
IGSH: internet-guided self-help
MVPA: moderate to vigorous physical activity
PA: physical activity
PICOS: Population, Intervention, Comparison, Outcome, Study
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RCT: randomized controlled trial
SDT: Self-Determination Theory

Review

Leveraging Mobile Health to Manage Mental Health/Behavioral Health Disorders: Systematic Literature Review

Clemens Scott Kruse[1*], PhD, MHA, MBA, MSIT; Jose A Betancourt[1*], DrPH; Matthew Gonzales[1*], MHA; Kennedy Dickerson[1*], MHA; Miah Neer[1*], MHA

School of Health Administration, Texas State University, San Marcos, TX, United States
[*]all authors contributed equally

Corresponding Author:
Clemens Scott Kruse, PhD, MHA, MBA, MSIT
School of Health Administration
Texas State University
601 University Drive
Encino, 250
San Marcos, TX, 78666
United States
Phone: 1 512 245 4462
Fax: 1 512 245 8712
Email: scottkruse@txstate.edu

Abstract

Background: Mental health is a complex condition, highly related to emotion. The COVID-19 pandemic caused a significant spike in depression (from isolation) and anxiety (event related). Mobile Health (mHealth) and telemedicine offer solutions to augment patient care, provide education, improve symptoms of depression, and assuage fears and anxiety.

Objective: This review aims to assess the effectiveness of mHealth to provide mental health care by analyzing articles published in the last year in peer-reviewed, academic journals using strong methodology (randomized controlled trial).

Methods: We queried 4 databases (PubMed, CINAHL [Cumulative Index to Nursing and Allied Health Literature], Web of Science, and ScienceDirect) using a standard Boolean search string. We conducted this systematic literature review in accordance with the Kruse protocol and reported it in accordance with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analysis) 2020 checklist (n=33).

Results: A total of 4 interventions (mostly mHealth) from 14 countries identified improvements in primary outcomes of depression and anxiety as well as in several secondary outcomes, namely, quality of life, mental well-being, cognitive flexibility, distress, sleep, self-efficacy, anger, decision conflict, decision regret, digestive disturbance, pain, and medication adherence.

Conclusions: mHealth interventions can provide education, treatment augmentation, and serve as the primary modality in mental health care. The mHealth modality should be carefully considered when evaluating modes of care.

Trial Registration: PROSPERO International Prospective Register of Systematic Reviews CRD42022343489; https://www.crd.york.ac.uk/PROSPERO/display_record.php?RecordID=343489

Introduction

Rationale

Mental health is a complex topic that is highly related to emotion, because emotional regulation is necessary for daily functioning [1]. Emotional regulation is necessary for friendships and intimate relationships [2]. The perception of mental health can be culturally based and highly related to self-actualization [3]. For the purposes of this manuscript, we define mental health as being able to work creatively and productively, to relate to others in a way that is mutually satisfying, and to feel comfortable when alone, usually developing a rich and fulfilling inner life [4]. The

psychopathology of mental illness is also complicated. Mental illness can manifest itself in terms of depression and anxiety, which are the major foci of this manuscript. Depression can be described as a mood (sadness or lack of enjoyment), a symptom, a syndrome, or a disorder [5]. During the COVID-19 pandemic, college students were particularly affected by depression due to the isolation created by the pandemic thrust upon them during a time in their lives when they expected to be highly socially active [6]. Anxiety often refers to multiple mental and physiological phenomena, including fear, distress, and a constant state of worry over events or actual situations [7]. The pandemic created COVID-19 anxiety, which has been added to a list of anxiety disorders [8].

Telemedicine and telehealth are defined as healing at a distance using information communication technologies to improve health outcomes [9]. The World Health Organization does not distinguish between these terms, so they will be used interchangeably in this manuscript. Telemedicine has existed for decades, but it became an essential modality of care during the COVID-19 pandemic. The highly contagious nature of COVID-19 prevented face-to-face appointments, and many providers were forced into this modality before they thought they were ready [10]. Despite this challenge, many providers discovered the effectiveness this modality can enable, including the use of mobile health (mHealth) and eHealth [11]. mHealth is a component of eHealth and telehealth that enables medical and public health practices through mobile devices, such as smartphones, patient monitoring devices, and other wireless devices [12]. Mobile devices have blurred the lines between tablets and computers, and the computing power of mobile devices enables the use of many apps formerly only available on a computer. This study focuses on the intersection between mHealth and mental health care.

Studies have shown that some college students are comfortable with mental health screening through mHealth modalities in the areas of performance expectancy and social influence [13]. Social influence and the stigma associated with mental health care greatly influence this age bracket (18-24 years) and their willingness to answer questions over their mobile devices. Telemedicine offers safety and efficiency, but several limitations prevent wide adoption such as technical problems, patient distraction, lack of confidentiality, compromised therapeutic alliance, and the management of unstable patients through a distance modality [6].

mHealth for an older population has not been entirely successful as well mostly due to the digital divide inherent to older populations and overall digital literacy [14]. Practitioners recognize this gap, and some have even tried to bridge the gap with intensive training to enable their patients to participate in this modality of care. Unfortunately, they found that a substantial number of older adults either do not have the technology or cannot negotiate the technology, despite this training. This population experiences high rates of depression, either due to illnesses, such as cancer (incidence as high as 58%) [15], or dementia (incidence as high as 34%) [16]. The overall prevalence of depression among older adults is estimated at 28.4% [17].

One systematic review from 2022 analyzed 26 articles over 10 years [18]. It found a major reduction in symptoms of depression among trials that included participants with moderate to severe depression, but not as strong as an effect among trials involving participants with mild to moderate depression. The reviewers concluded that app-based interventions have a moderate effect to reduce symptoms of depression.

Another systematic review from 2022 analyzed 21 articles about telepsychiatry over 15 years [19]. It identified familiar themes such as equivalence to in-person, convenience, overcoming remoteness, and timely access to treatment with specialists. The reviewers concluded that the video teleconferencing technology enables telepsychiatry in both the home and the emergency department on par with in-person reviews with the additional benefit of wider access and timeliness to treatment.

Objectives

The purpose of this review is to analyze studies published over the last year that examine mHealth as an intervention to both screen for and treat symptoms of mental health among adults aged over 18 years with strong methodological design (randomized controlled trial [RCT]). The intention of this review is to focus on mental health interventions during 1 year of the COVID-19 pandemic.

Methods

Protocol and Registration

This review is conducted in accordance with the Kruse protocol for writing a systematic review. It is reported in accordance with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analysis) checklist. This review is registered with PROSPERO (registration number CRD42022343489).

Eligibility Criteria

This review focused on studies with strong methodology including participants who were adults (>18 years of age) that were published in peer-reviewed, academic journals over the last year. Other systematic reviews were excluded to avoid confounding the results. All reporting is in accordance with the PRISMA 2020 standard [20].

Information Sources

A total of 4 research databases were queried: PubMed (MEDLINE), CINAHL (Cumulative Index to Nursing and Allied Health Literature), Web of Science, and ScienceDirect from Embase. They were queried on August 14, 2022. These databases were chosen due to their availability and comprehensive indexing of research. Their availability make it easier for others to duplicate our study. Their comprehensive indexing ensures we capture a majority of the literature in our search.

Search Strategy

Our initial process included a search on Google Scholar to better understand the topic and recent work published. We gathered information from Google Scholar into an MS Excel spreadsheet to enable the review team to read the articles so that we could see the key terms these studies used for indexing. Using the

Medical Subject Headings (MeSH) from the US National Library of Medicine, we created a Boolean search string to exhaustively query the databases without redundant terms. We used the same search strategy in all databases and similar filter strategies.

Selection Process

In accordance with the Kruse protocol [21], we searched key terms in all databases using a Boolean search string: (mhealth OR "mobile app" OR telemedicine) AND ("mental health" OR "behavioral health" OR "depression"). We removed duplicates, filtered the results, and screened abstracts for applicability to our objective statement [21]. We selected only studies with strong methods (eg, RCT). The RCT and true experiments were chosen due to their rigorous adherence to control groups and comparisons to the same.

Data Collection Process

The Kruse protocol standardized an MS Excel spreadsheet as a data extraction tool and as an analysis tool [21]. The spreadsheet's standardized fields allowed the collection of additional data at each step of the process, thus making analysis robust and useful to both clinicians and administrators. Through 3 consensus meetings we finalized the group of articles for analysis, identified themes, and finalized the additional analysis through a data synthesis method.

Data Items

In accordance with the Kruse protocol [21], we collected the following fields of data at each step: Google Scholar step (date of publication, authors, study title, journal, impact factor from Journal Citations Reports, study design, key terms, experimental intervention, results, and comments from each reviewer); filter articles step (the number of results before and after each filter was applied in all 4 databases); abstract screening step (database source, date of publication, authors, study title, journal, screening decision for each reviewer, notes about rejections, consensus meeting 1 determination of screening decision, and a set of rejection criteria); analysis step (database source, date of publication, authors, study title, participants, experimental intervention, results compared with a control group, medical outcomes, study design, sample size, bias effect size, country of origin, statistics used, patient satisfaction, facilitators to adoption, barriers to adoption, and the strength and quality of evidence).

Study Risk and Reporting of Bias Assessment

Risk of bias was determined through multiple means. Reviewers used the John's Hopkins Nursing Evidence-Based Practice (JHNEBP) tool assessment of the strength and quality of evidence [22]. Strength of evidence is based on the study design (eg, RCT, quasi-experimental, qualitative), and the quality of evidence is based on adequate sample size, adequate control, and consistency of results. Reviewers also made a note of other observations of bias such as selection, sample, design, or publication bias. These observations and the JHNEBP assessment of the strength and quality of evidence were used for interpreting the results, because bias can limit the internal and external validity of the results [23].

Effect Measures

The preferred methodologies for this review were the RCT and true experiments because these are the strongest group of methodologies in the JHNEBP tool. The preferred measures of effect were the Cohen d, but other measures of effect were also collected, such as the odds ratio. All measures of effect were tabulated for those studies in which they were reported.

Synthesis Methods

We performed a thematic analysis of the data extracted [24]. This thematic analysis helped us make sense of the data by grouping same or similar observations into themes. Themes and individual observations were then tabulated for both reporting and to enable inferences.

Additional Analyses and Certainty Assessment

Sensitivity, specificity, and effect size were tabulated and included in the data extraction. Combined with the narrative analysis, this provided us with certainty assessments. The frequency of themes does not imply importance, but it does provide confidence in the data analyzed.

Ethics Approval

No human subjects were used in this research. It is therefore IRB exempt.

Results

Overview

Figure 1 illustrates the study selection process. The query of 4 databases resulted in 23,713 results; however, 21,426 of these results were duplicates. After filtering and screening, reviewers were left with 67 articles eligible for review. The reviewers chose to only analyze the RCTs, because this is the highest strength of evidence in the JHNEBP. This can be seen in Figure 1 under "methods not strong," meaning we excluded all but RCTs and true experiments. The remaining group of studies for analysis was 33.

Figure 1. Article selection process. CINAHL: Cumulative Index to Nursing and Allied Health Literature; WoS: Web of Science.
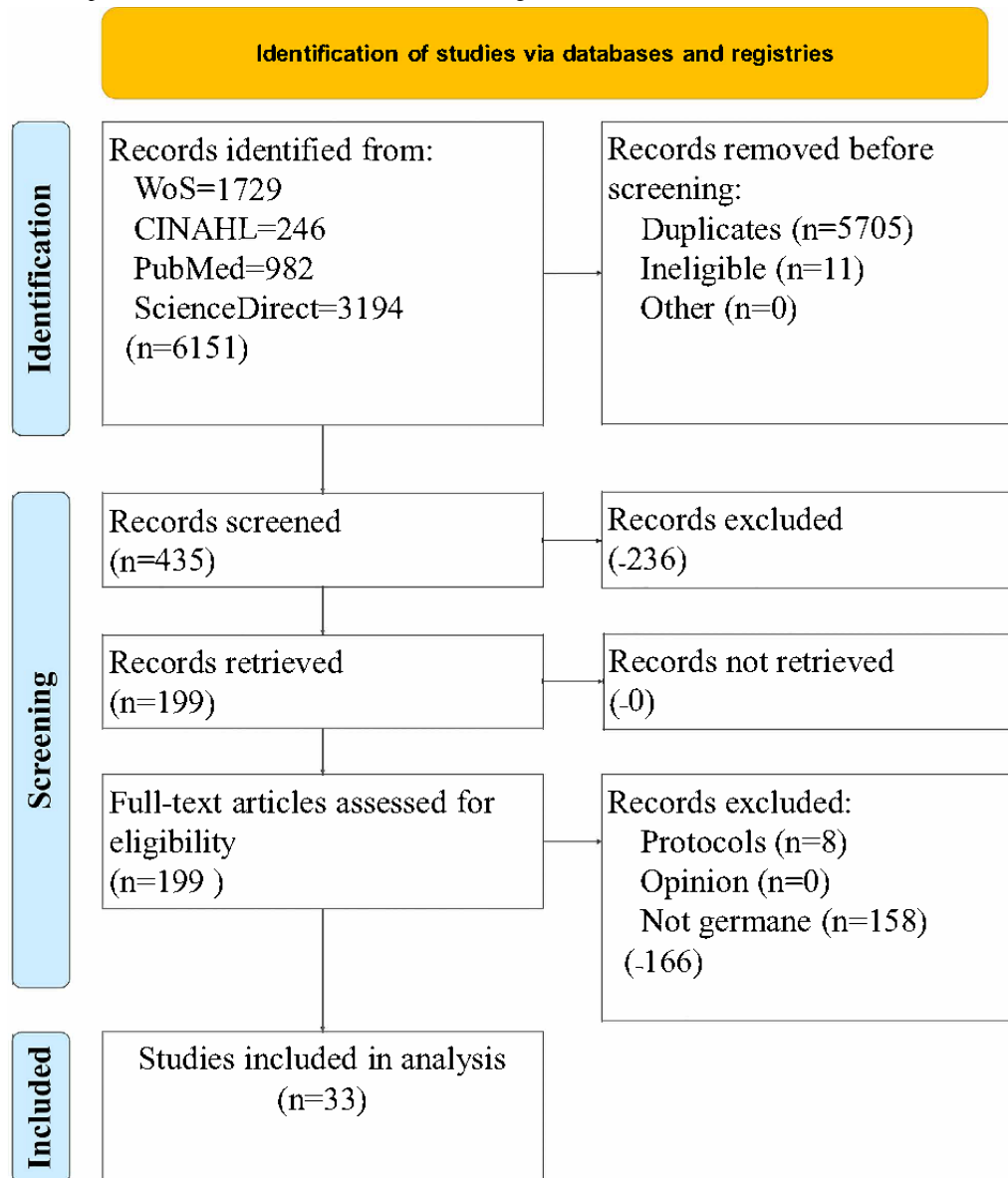
Study Selection

Study Characteristics

Following the PRISMA checklist [20] and the Kruse protocol [21], we extracted the following data fields: participants, intervention, comparison with the control or other group, medical outcomes, and study design (PICOS; Table 1) as a way to summarize study characteristics, as required by PRISMA. Of the 33 studies analyzed over the 1-year period, 17 were from 2021 [25-41] and 16 were from 2022 [42-57]. All studies were RCTs involving adults, and 48% (16/33) involved an average age of 50 years or above. About 64% (21/33) used an mHealth or eHealth app as the intervention, 21% (7/33) used telemedicine over computer or mobile device, 12% (4/33) used telephone, and 3% (1/33) used telemonitoring. Studies originated from 14 countries, but 42% (14/33) originated from the United States and 9% (3/33) from China, which accounted for over half of the studies.

Table 1. Summary of study characteristics (PICOS[a]).

Authors	Participants	Experimental intervention	Results (compared with the control group)	Medical outcomes reported	Study design
Acierno et al [25]	Adult females with PTSD[b] from sexual trauma, average age 43.4 years, 64% African American	Telemedicine	Reduced depression, but there were no differences in dose received or PTSD symptom reduction	Decrease in depression, but not statistically significant	RCT[c]
Baek et al [26]	Adults, average age 41.9 years, 67% female	mHealth[d] app (MibyeongBogam [MBBG])	The intervention group showed a decrease in depression (P=.003), anxiety (P=.01), sleep disturbance (P=.02), anger (P=.003), and pain (P=.02) greater than the control group; also fatigue (P=.6) and digestive disturbance (P=.76) were not statistically significant	Decreased depression, anxiety, sleep disturbance, anger, pain, fatigue, digestive disturbance	RCT
Colomina et al [27]	Adults, average age 73 years, 66% female	mHealth self-management app (CONNECARE)	Decreased anxiety more than the control group, but there were no differences in depression symptom reduction	Decreased anxiety, but not statistically significant over traditional care	RCT
Dobkin et al [28]	Adult veterans, average age 67.8 years, 100% male, 92% White	Video-to-home cognitive behavioral therapy	Intervention outperformed treatment as usual across all 3 measures of depression (P=.001), decreased anxiety, but not statistically significant	Decreased depression and anxiety	RCT
Domogalla et al [29]	Adults with psoriasis, average age 49 years, 60% male	mHealth study and disease management app	Significant reduction in HADS[e], HADS-D[f] (P=.04), and HADS-A[g] (P=.05) more than those in the control group	Decreased anxiety and depression	RCT
Fang et al [30]	Adults 100% female	mHealth app (Pink Journey)	Decreased anxiety, depression, decision conflict, and decision regret more than the control group, but not statistically significant from control; decreased body image distress (P=.027)	Decreased anxiety, depression, decision conflict, decision regret, and body image distress	RCT
Fortney et al [31]	Adults average age 39 years, 70% female, 66% White	Telepsychiatry	Decreased depression and anxiety, but with small effect	Decreased depression and anxiety	RCT
Huberty et al [32]	Adults average age 44.2 years, 78% female, 56% White	mHealth app	Decreased anxiety (P<.001) and depression (P<.001) more than the control group	Decreased anxiety and depression	RCT
Jones et al [33]	Adults average age 53.4 years	mHealth app (WRAP)	Decreased HADS more than the control group	Decreased anxiety and depression	RCT
Krzyzanowska et al [34]	Adults <40<75 years, median age 55 years	Telephone	No effect on anxiety, depression, or self-efficacy	No effect on anxiety, depression, or self-efficacy	RCT
Moskowitz et al [35]	Adults average age 37.95 years, 74% female	eHealth	Decreased depression (P<.06) more than the control group	Decreased depression	RCT
Pakrad et al [36]	Adults average age 62.7 years, 82% male	mHealth app	Decreased anxiety (P<.028), stress (P<.022), and quality of life (P<.001) more than the control. Decreased depression more than the control group, but not statistically significant (P<.063)	Decreased depression, anxiety, and stress, and increased quality of life	RCT
Rollman et al [37]	Adults average age 63.9 years, 56% male, 74% White	Telephone	Decreased depression more than the control group	Decreased depression	RCT
Romijn et al [38]	Adults, average age 36.25 years, 75% White	eHealth cognitive behavioral therapy (inference-based cognitive behavioral therapy)	Decreased anxiety more than the control group	Decreased anxiety	RCT

Authors	Participants	Experimental intervention	Results (compared with the control group)	Medical outcomes reported	Study design
Su and Yu [39]	Adults, average age 55.75 years, 85% male, 100% Chinese	eHealth	Decreased anxiety more than the control group, no effect on depression	Decreased anxiety	RCT
Taguchi et al [40]	Adults, average age 50 years, 67% female	Video-based cognitive behavioral therapy	Decreased depression and anxiety (not statistically significant from the control)	Decreased depression and anxiety	RCT
Wong et al [41]	Adults >60 years, average age 72 years, 82% female	mHealth app	Decreased depression (not statistically significant over the control group), increased medication adherence (P<.001), self-efficacy (P<.16), and quality of life (P<.04)	Decreased depression, improved medication adherence, self-efficacy, and quality of life	RCT
Aikens et al [42]	Adults, average age 48.6 years, 81% female, 74% White	Telephone (automated interactive voice response)	Decreased depression more than the control group, with medium effect; increased self-efficacy	Decreased depression, increased self-efficacy	RCT
Akin-Sari et al [43]	Adults, average age 23 years, 78% female	mHealth app	Decreased depression and COVID-19 distress more than the control group	Decreased depression, decreased COVID-19 distress	RCT
Bathgate et al [44]	Adults, average age 32.4 years, 81% female, 97% White	Telemedicine	Decreased depression more than the control (P=.78), anxiety but not more than the control (P=.6), increased coping self-efficacy but not more than the control (P=.93), increased quality of life (physical functioning, social functioning, and vitality)	Decreased depression and anxiety, increased coping self-efficacy and quality of life	RCT
Catuara-Solarz et al [45]	Adults, average age 40 years, 54% female	mHealth app	Decreased anxiety (P=.04), increase in resilience (P=.001), sleep (P=.01), and mental well-being (P=.02) more than the control group	Decreased anxiety, increase in resilience, sleep, and mental well-being	RCT
Deady et al [46]	Adults, average age 40 years, 74% male	mHealth app (HeadGear)	Improved depression, anxiety, resilience, and well-being more than the control group (P=.0031)	Improved depression, anxiety, resilience, and well-being	RCT
Drew et al [47]	Adults, average age 48.4 years, 100% male	eHealth app (SHED-IT)	Improved depression, sleep, cognitive flexibility more than the control	Improved depression, sleep, cognitive flexibility	RCT
Guo et al [48]	Adults, average age 28.3 years, 95% male, 100% Chinese	mHealth, social media (Run4Love)	Improved depression more than the control	Decreased depression	RCT
Gustafson et al [49]	Adults, >65 years, average age 76.5 years, 74% female, 89% White	eHealth app (ElderTree)	Improved depression (OR[h] –0.20, P=.034) and overall mental health quality of life (OR 0.32, P=.007) more than the control group	Decreased depression, increased mental health, increased quality of life	RCT
Kuhn et al [50]	Adults, average age 44.5 years, mostly male	mHealth app	Decreased depression (d=–0.8, P<.012) and sleep-related impairment (d=–0.6, P<.04) more than the control group	Decreased depression and sleep-related impairment	RCT
Lopez et al [51]	Adults, average age 44 years, 100% female	Telemedicine	Reduced depression, but there were no differences in dose received or PTSD symptom reduction	Decrease in depression, but not statistically significant	RCT
Mitchell et al [52]	Adults, average age 51 years, 60% female	Telemedicine cognitive behavioral therapy (RED-D)	Decreased depression and readmission (P<.012) more than the control	Decreased depression	RCT
Nardi et al [53]	Adults, average age 42.9 years, 93% female	mHealth app (unwinding anxiety)	Decreased anxiety (P=.005) and worry (P=.01) more than the control	Decreased anxiety and worry	RCT
Orman et al [54]	Adults, average age 68.4 years, 64% male	Telephone	Decreased depression and anxiety greater than usual care, short-term positive effect on quality of life	Decreased anxiety and depression, and increased quality of life	RCT

Authors	Participants	Experimental intervention	Results (compared with the control group)	Medical outcomes reported	Study design
Sun et al [55]	Adults, 100% Chinese	mHealth app (mindfulness)	Decreased depression and anxiety (P=.024) greater than usual care, but depression was not statistically different	Decreased anxiety and depression	RCT
Volpato et al [56]	Adults, average age 76.2 years, 51% male	mHealth cognitive behavioral therapy	Decreased anxiety and depression, but not statistically significant than the control. Improved adherence to noninvasive ventilation (P<.001) and quality of life (P<.002)	Decreased anxiety and depression, improved quality of life, and noninvasive ventilation	RCT
Ware et al [57]	Adults, average age 59 years, 56% male	Telemonitoring	No effect on anxiety or depression. Improved self-care maintenance, management, confidence, and physical quality of life	Improved self-care maintenance, management, confidence, and physical quality of life	RCT

[a]PICOS: participants, intervention, comparison with the control or other group, medical outcomes, and study design.

[b]PTSD: posttraumatic stress disorder.

[c]RCT: randomized controlled trial.

[d]mHealth: mobile health.

[e]HADS: Hospital Anxiety and Depression Scale

[f]HADS-D: Hospital Anxiety and Depression Scale-Depression.

[g]HADS-A: Hospital Anxiety and Depression Scale-Anxiety.

[h]OR: odds ratio.

Risk of Bias in and Across Studies

The JHNEBP quality assessment tool identified the strength and quality of evidence [22]. Strength of evidence is defined by methodology: level I is defined as true experiments and RCTs; level II is defined as quasi-experiments; and level III is defined as nonexperimental, observational, and qualitative studies. Levels IV and V are defined as expert opinions and editorials. We only used RCTs in our systematic literature review, so the group for analysis was 100% (33/33) level I. The quality of evidence in the JHNEBP tool is defined by sample size, consistency of results (based on established measurement standards), control groups, conclusions, and literature reviews. Each level accepts a lower standard. Level A is defined by consistent results with sufficient sample sizes (based on power analysis), adequate control groups, definitive conclusions, and consistent recommendations based on extensive literature reviews. Level B is defined by reasonably consistent results, sufficient sample sizes, some control groups, fairly definitive conclusions, and reasonably consistent recommendations based on fairly comprehensive literature reviews. Level C is defined by little evidence with inconsistent results, insufficient sample sizes, and nondefinitive conclusions. In our group for analysis, only 1 RCT (1/33, 3%) was defined as level B, while the rest were defined as level A (32/33, 97%).

Reviewers also noted instances of bias, because bias can limit external and internal validity [23]. There were 30 instances of selection bias (affecting internal validity), and 28 instances of sample bias (affecting external validity). The latter were usually due to a high percentage of sex or race in the sample. The former were due to studies conducted in 1 region of 1 country.

Results of Individual Studies

Table 2 summarizes the results of individual studies through themes. Themes were identified when the same or similar observation occurred in the literature. An observation-to-theme match can be found in Multimedia Appendices 1 and 2. The other data items collected (sample size, bias, effect size, country of origin, statistics used, and JHNEBP strength and quality of evidence ratings) can be found in Multimedia Appendix 3. The average sample size was 331.

Table 2. Summary of themes, sorted chronologically by author.

Authors	Intervention themes	Results themes	Medical outcomes themes	Patient satisfaction themes	Effectiveness themes	Barrier themes
Acierno et al [25]	Telemedicine	• Reduced depression • No statistical significance for at least one condition	• Reduced depression	• Satisfied	• Reduced depression • Enabled preference for telemedicine	• May not be the preferred treatment method • Staff training • Low reimbursement
Baek et al [26]	mHealth[a]/eHealth app	• Reduced depression • Reduced anxiety • Increased sleep • Decreased anger • Decreased pain • Decreased digestive disturbance • No statistical significance for at least one condition	• Reduced depression • Reduced anxiety • Increased sleep • Decreased anger • Decreased pain • Decreased digestive disturbance	• Not reported	• Reduced depression • Reduced anxiety • Increased sleep • Decreased anger • Decreased pain • Decreased digestive disturbance	• May not be the preferred treatment method • Staff training • Low reimbursement
Colomina et al [27]	mHealth/eHealth app	• Reduced anxiety • No effect on depression • No statistical significance for at least one condition	• Reduced anxiety	• Satisfied	• Reduced health costs per patient • Reduced anxiety	• May not be the preferred treatment method • Staff training • Low reimbursement
Dobkin et al [28]	Telemedicine	• Reduced depression • Reduced anxiety • No statistical significance for at least one condition	• Reduced anxiety • Reduced depression	• Satisfied	• Reduced anxiety • Reduced depression • Extended care to rural patients	• May not be the preferred treatment method • Staff training • Low reimbursement
Domogalla et al [29]	mHealth/eHealth app	• Reduced anxiety • Reduced depression	• Reduced anxiety • Reduced depression	• Satisfied	• Reduced anxiety • Reduced depression	• May not be the preferred treatment method • Staff training
Fang et al [30]	mHealth/eHealth app	• Reduced anxiety • Reduced depression • Decreased decision conflict • Decreased decision regret • Decreased distress • No statistical significance for at least one condition	• Reduced anxiety • Reduced depression • Decreased decision conflict • Decreased decision regret • Decreased distress	• Satisfied	• Reduced anxiety • Reduced depression • Decreased decision conflict • Decreased decision regret • Decreased distress	• May not be the preferred treatment method • Staff training • Low reimbursement
Fortney et al [31]	Telemedicine	• Reduced anxiety • Reduced depression	• Reduced anxiety • Reduced depression	• Not reported	• Reduced anxiety • Reduced depression	• May not be the preferred treatment method • Staff training • Low reimbursement
Huberty et al [32]	mHealth/eHealth app	• Reduced anxiety • Reduced depression	• Reduced anxiety • Reduced depression	• Satisfied	• Reduced anxiety • Reduced depression	• May not be the preferred treatment method • Staff training
Jones et al [33]	mHealth/eHealth app	• Reduced anxiety • Reduced depression	• Reduced anxiety • Reduced depression	• Satisfied	• Reduced anxiety • Reduced depression	• May not be the preferred treatment method • Staff training

Authors	Intervention themes	Results themes	Medical outcomes themes	Patient satisfaction themes	Effectiveness themes	Barrier themes
Krzyzanowska et al [34]	Telephone	• No effect on anxiety • No effect on depression • No effect on self-efficacy	• None	• Not reported	• None	• May not be the preferred treatment method • Staff training
Moskowitz et al [35]	mHealth/eHealth app	• Reduced depression	• Reduced depression	• Satisfied	• Reduced depression	• May not be the preferred treatment method • Staff training • Low reimbursement
Pakrad et al [36]	mHealth/eHealth app	• Reduced anxiety • Decreased distress • Increased quality of life • Reduced depression • No statistical significance for at least one condition	• Reduced anxiety • Decreased distress • Increased quality of life • Reduced depression	• Satisfied	• Reduced anxiety • Decreased distress • Reduced depression • Increased quality of life	• May not be the preferred treatment method • Staff training • Low reimbursement
Rollman et al [37]	Telephone	• Reduced depression	• Reduced depression	• Not reported	• Reduced depression	• May not be the preferred treatment method • Staff training • Low reimbursement
Romijn et al [38]	mHealth/eHealth app	• Reduced anxiety	• Reduced anxiety	• Satisfied	• Reduced anxiety	• May not be the preferred treatment method • Staff training
Su and Yu [39]	mHealth/eHealth app	• Reduced anxiety • No effect on depression	• Reduced anxiety	• Not reported	• Reduced anxiety	• May not be the preferred treatment method • Staff training
Taguchi et al [40]	Telemedicine	• Reduced anxiety • Reduced depression • No statistical significance for at least one condition	• Reduced anxiety • Reduced depression	• Not reported	• Reduced anxiety • Reduced depression	• May not be the preferred treatment method • Staff training
Wong et al [41]	mHealth/eHealth app	• Reduced depression • No statistical significance for at least one condition • Increased medication adherence • Increased self-efficacy • Increased quality of life	• Reduced depression • Increased medication adherence • Increased self-efficacy • Increased quality of life	• Not reported	• Reduced depression • Increased medication adherence • Increased self-efficacy • Increased quality of life	• May not be the preferred treatment method • Staff training • Low reimbursement
Aikens et al [42]	Telephone	• Reduced depression • Increased self-efficacy	• Reduced depression • Increased self-efficacy	• Not reported	• Reduced depression • Increased self-efficacy	• May not be the preferred treatment method • Staff training • Low reimbursement

Authors	Intervention themes	Results themes	Medical outcomes themes	Patient satisfaction themes	Effectiveness themes	Barrier themes
Akin-Sari et al [43]	mHealth/eHealth app	• Reduced depression • Decreased distress	• Reduced depression • Decreased distress	• Not reported	• Reduced depression • Decreased distress	• May not be the preferred treatment method • Staff training • Low reimbursement
Bathgate et al [44]	Telemedicine	• Reduced depression • Reduced anxiety • Increased self-efficacy • Increased quality of life • No statistical significance for at least one condition	• Reduced depression • Reduced anxiety • Increased self-efficacy • Increased quality of life	• Satisfied	• Reduced depression • Reduced anxiety • Increased self-efficacy • Increased quality of life	• May not be the preferred treatment method • Staff training • Low reimbursement
Catuara-Solarz et al [45]	mHealth/eHealth app	• Reduced anxiety • Decreased fatigue/increased resilience • Increased sleep • Increased mental well-being/cognitive flexibility	• Reduced anxiety • Decreased fatigue/increased resilience • Increased sleep • Increased mental well-being/cognitive flexibility	• Satisfied	• Reduced anxiety • Decreased fatigue/increased resilience • Increased sleep • Increased mental well-being/cognitive flexibility	• May not be the preferred treatment method • Staff training
Deady et al [46]	mHealth/eHealth app	• Reduced depression • Reduced anxiety • Decreased fatigue/increased resilience • Increased mental well-being/cognitive flexibility	• Reduced depression • Reduced anxiety • Decreased fatigue/increased resilience • Increased mental well-being/cognitive flexibility	• Not reported	• Reduced depression • Reduced anxiety • Decreased fatigue/increased resilience • Increased mental well-being/cognitive flexibility	• May not be the preferred treatment method • Staff training
Drew et al [47]	mHealth/eHealth app	• Reduced depression • Increased sleep • Increased mental well-being/cognitive flexibility	• Reduced depression • Increased sleep • Increased mental well-being/cognitive flexibility	• Not reported	• Reduced depression • Increased sleep • Increased mental well-being/cognitive flexibility	• May not be the preferred treatment method • Staff training
Guo et al [48]	mHealth/eHealth app	• Reduced depression	• Reduced depression	• Not reported	• Reduced depression	• May not be the preferred treatment method • Staff training
Gustafson et al [49]	mHealth/eHealth app	• Reduced depression • Increased mental well-being/cognitive flexibility • Increased quality of life	• Reduced depression • Increased mental well-being/cognitive flexibility • Increased quality of life	• Not reported	• Reduced depression • Increased mental well-being/cognitive flexibility • Increased quality of life	• May not be the preferred treatment method • Staff training • Low reimbursement
Kuhn et al [50]	mHealth/eHealth app	• Reduced depression • Increased sleep	• Reduced depression • Increased sleep	• Not reported	• Reduced depression • Increased sleep	• May not be the preferred treatment method • Staff training • Low reimbursement
Lopez et al [51]	Telemedicine		• Reduced depression	• Not reported	• Reduced depression	

Authors	Intervention themes	Results themes	Medical outcomes themes	Patient satisfaction themes	Effectiveness themes	Barrier themes
		• Reduced depression • No statistical significance for at least one condition				• May not be the preferred treatment method • Staff training • Low reimbursement
Mitchell et al [52]	Telemedicine	• Reduced depression	• Reduced depression	• Satisfied	• Reduced depression • Decreased readmissions	• May not be the preferred treatment method • Staff training • Low reimbursement
Nardi et al [53]	mHealth/eHealth app	• Reduced anxiety • Decreased distress	• Reduced anxiety • Decreased distress	• Not reported	• Reduced anxiety • Decreased distress	• May not be the preferred treatment method • Staff training • Low reimbursement
Orman et al [54]	Telephone	• Reduced depression • Reduced anxiety • Increased quality of life	• Reduced depression • Reduced anxiety • Increased quality of life	• Not reported	• Reduced depression • Reduced anxiety • Increased quality of life	• May not be the preferred treatment method • Staff training
Sun et al [55]	mHealth/eHealth app	• Reduced depression • Reduced anxiety • No statistical significance for at least one condition	• Reduced depression • Reduced anxiety	• Not reported	• Reduced depression • Reduced anxiety	• May not be the preferred treatment method • Staff training
Volpato et al [56]	mHealth/eHealth app	• Reduced anxiety • Reduced depression • Increased quality of life	• Reduced anxiety • Reduced depression • Increased quality of life	• Not reported	• Reduced anxiety • Reduced depression • Increased quality of life	• May not be the preferred treatment method • Staff training • Low reimbursement
Ware et al [57]	Telemonitoring	• No effect on anxiety • No effect on depression • Increased self-efficacy • Increased mental well-being/cognitive flexibility • Increased quality of life	• No effect on anxiety • No effect on depression • Increased self-efficacy • Increased mental well-being/cognitive flexibility • Increased quality of life	• Not reported	• Increased self-efficacy • Increased mental well-being/cognitive flexibility • Increased quality of life	• May not be the preferred treatment method • Staff training

[a]mHealth: mobile health.

Results of Syntheses, Additional Analysis, and Certainty of Evidence

A thematic analysis helped makes sense of the data collected. Although thematic analyses are typically used for qualitative analysis, other systematic reviews have also used this technique to makes sense of all observations from the data extraction process, whether the studies were qualitative or quantitative [10,11]. The themes and observations are tabulated into affinity matrices for interpretation.

Patient Satisfaction

Patient satisfaction was not reported in all studies (20/33, 61%); however, 13/33 (39%) reported users were satisfied or highly satisfied. At the point where these data were collected, users were very pleased with the user interface and any further progress in the interventions improved their mental health.

Results of Interventions, Compared With the Control Groups

Table 3 summarizes the results of interventions compared with the control groups (treatment as usual, in-person). This section is designed for the scientist or researcher. A total of 11 themes

and 2 individual observations were identified by the reviewers for a total of 48 occurrences in the literature, whereas 11 themes and 7 observations were noted for a total of 95 observations. In the 33 studies analyzed, 26 (79%) showed an improvement in symptoms of depression [25,26,28-33,35-37,40-44,46-52,54-56], while 19 (58%) showed an improvement in anxiety [26-33,36,38-40,44-46,53-56]. Only 11/33 (33%) reported that at least one symptom was not statistically significant when compared with results from the control group, but the improvement was still noted [25-28,30,36,40,41,44,51,55]. As many as 7/33 (21%) showed an increase in quality of life [36,41,44,49,54,56,57], and 5 (15%) showed an increase in mental well-being, cognitive flexibility, or confidence

[45-47,49,57]. A total of 4 themes emerged that each appeared 4/33 times (12%): decreased distress (body image or COVID-19 distress) [30,36,43,53], increased sleep (less sleep disturbance or sleep-related impairment) [26,45,47,50], increased self-efficacy [41,42,44,57], and the intervention had no effect on depression [27,34,39,57]. A total of 2 themes each appeared 2/33 times (6%): Decreased fatigue (increased resilience) [45,46] and the intervention had no effect on anxiety [34,57]. The following observations only occurred once in the literature: decrease in anger [26], decrease in decision conflict [30], decrease in decision regret [30], decreased digestive disturbance [26], decreased pain [26], increased medication adherence [41], and the intervention had no effect on self-efficacy [34].

Table 3. Results compared with the control groups.

Results themes and observations	Frequency, n (n=95)
Reduced depression [25,26,28-33,35-37,40-44,46-52,54-56]	26
Reduced anxiety [26-33,36,38-40,44-46,53-56]	19
No statistical significance for at least one condition [25-28,30,36,40,41,44,51,55]	11
Increased quality of life [36,41,44,49,54,56,57]	7
Increased mental well-being/cognitive flexibility [45-47,49,57]	5
Decreased distress [30,36,43,53]	4
Increased sleep [26,45,47,50]	4
Increased self-efficacy [41,42,44,57]	4
No effect on depression [27,34,39,57]	4
Decreased fatigue/increased resilience [45,46]	2
No effect on anxiety [34,57]	2
Decreased anger [26]	1
Decreased decision conflict [30]	1
Decreased decision regret [30]	1
Decreased digestive disturbance [26]	1
Decreased pain [26]	1
Increased medication adherence [41]	1
No effect on self-efficacy [34]	1

Medical Outcomes Commensurate With the Use of mHealth

Table 4 summarizes the medical outcomes observed. A total of 8 themes and 6 individual observations were recorded commensurate with the adoption of mHealth for the management

of mental health for a total of 78 occurrences. The results (compared with the control groups) and medical outcomes were very similar, but they are focused on observations for the provider. The only difference was 1 study [34] reported no effect on depression or anxiety.

Table 4. Medical outcomes commensurate with the adoption of mobile health.

Medical outcomes themes and observations	Frequency, n (n=78)
Reduced depression [25,26,28-33,35-37,40-44,46-52,54-56]	26
Reduced anxiety [26-33,36,38-40,44-46,53-56]	19
Increased quality of life [36,41,44,49,54,56,57]	7
Increased mental well-being/cognitive flexibility [45-47,49,57]	5
Decreased distress [30,36,43,53]	4
Increased sleep [26,45,47,50]	4
Increased self-efficacy [41,42,44,57]	4
Decreased fatigue/increased resilience [45,46]	2
Decreased anger [26]	1
Decreased decision conflict [30]	1
Decreased decision regret [30]	1
Decreased digestive disturbance [26]	1
Decreased pain [26]	1
Increased medication adherence [41]	1
None [34]	1

Effectiveness of mHealth to Manage Mental Health

Table 5 summarizes the themes and observations related to the effectiveness of mHealth in managing mental health. These were highly similar to the results and medical outcomes, but they include observations for the health care administrator. A total of 8 themes and 11 observations were noted for 82 observations. Because of the similarities with the previous tables, only the differences will be reported. One study noted the reduction in readmission when mHealth was used as part of the follow-up to focus on mental health conditions [52]. One study highlighted that while the intervention did not result in a statistically significant reduction in depression and anxiety over normal care, it enabled a preference for telemedicine, if a patient prefers it [25]. One study highlighted how the intervention can extend care to rural patients [28]. Another study highlighted a reduction in cost of care per patient when using mHealth over traditional care [27]. Two studies reported no effects on mental health from the intervention [34,57].

Table 5. Clinical and administrative effectiveness of mobile health to manage mental health.

Effectiveness themes and observations	Frequency, n (n=165)
Reduced depression [25,26,28-33,35-37,40-44,46-52,54-56]	26
Reduced anxiety [26-33,36,38-40,44-46,53-56]	19
Increased quality of life [36,41,44,49,54,56,57]	7
Increased mental well-being/cognitive flexibility [45-47,49,57]	5
Decreased distress [30,36,43,53]	4
Increased self-efficacy [41,42,44,57]	4
Increased sleep [26,45,47,50]	4
Decreased fatigue/increased resilience [45,46]	2
Decreased anger [26]	1
None [34,57]	2
Decreased decision conflict [30]	1
Decreased decision regret [30]	1
Decreased digestive disturbance [26]	1
Decreased pain [26]	1
Decreased readmissions[a] [52]	1
Enabled preference for telemedicine [25]	1
Extended care to rural patients [28]	1
Increased medication adherence [41]	1
Reduced health costs per patient[a] [27]	1
Administrative observations[a]	82

[a]Collected data that show effectiveness.

Barriers to the Adoption of mHealth for Mental Health Care

Three barriers were identified in the literature for the adoption of mHealth for mental health care. mHealth and telemedicine may not be the preferred modality of treatment for some patients or providers. This, along with the requirement to train staff, was identified in the literature 33 times [25-57]. The other barrier was for countries that must receive reimbursement for care, and that telemedicine modalities are not often fully reimbursed to the point where they would pay for the intervention. This observation occurred 19 times [25-28,30,31,35-37,41-44,49-53,56].

Interactions Between Interventions

When mHealth was used as the intervention, a reduction in both depression and anxiety was reported [26,27,29,30,32,33,35,36,38,39,41,43,45-50,53,55,56]. The same can be said for telemedicine [25,28,31,40,44,51,52], and mostly for telephone intervention [34,37,42,54]. Only 1 of each telephone [34] intervention and telemonitoring [57] intervention had no effect on depression or anxiety.

Discussion

Summary of Evidence

This systematic literature review analyzed 33 RCTs from 14 countries published over a 1-year period in peer-reviewed, academic journals using adults as participants (half of which were older adults) to analyze the effectiveness of mHealth for mental health care. A total of 4 interventions were studied: mHealth or eHealth apps, telemedicine (delivered over either a computer or a mobile device), telephone, and telemonitoring. Strong study methodologies resulted in low bias within and across studies. Observations of both sample and selection bias were noted, but there was nothing significant to report from these sources of bias. Overall, the interventions resulted in 26 instances of reduced depression; 19 instances of reduced anxiety; 7 instances of increased quality of life; 5 instances of increased mental well-being; 4 instances of decreased stress, increased self-efficacy, and increased sleep; 2 instances of decreased fatigue or increased resilience; and 1 instance each of decreased anger, decision regret, decision conflict, digestive disturbance, pain, readmission, and health care costs per patient. Only 2 studies reported no effect on depression and anxiety.

Future research should focus on standardizing mHealth into clinical practice guidelines for the treatment of some depression and anxiety issues. mHealth interventions can be rapidly deployed to a wide range of patient for very little money [27].

All but 2 studies reported improvements in at least one area of care [34,57]. This shows great promise for this modality of care.

Results from this systematic review should empower providers to adopt some mHealth interventions to augment or supplant existing practices; however, a few barriers should be addressed. While mHealth interventions can be conveniently adopted by some providers, it may not be the preferred modality for some patients. Providers should be sensitive to patient preferences. As mHealth and telemedicine modalities are introduced to provider clinics, staff training will have to take place, but after initial training has occurred, small refresher training should be all that is necessary. Finally, while many countries introduced reimbursement mechanisms during the pandemic, many have expired or have not been renewed. This is an important policy point that this review documents. It is imperative that the efficacy of this modality be recognized as beneficial to patients, and as such they should be reimbursed appropriately.

Limitations

This study has several limitations. While confirmation bias can create problems for researchers, multiple reviewers were used to control for this bias. While selection bias can be a problem with internal validity, multiple research databases were used to control for this bias. Publication bias is one this study did not control for. Because we only used studies published in peer-reviewed academic journals, it is possible there are other studies without positive results that we failed to include in the analysis. Our review used a Boolean search string from MeSH to ensure the search was exhaustive, but this technique may have overlooked articles indexed with terms other than those in MeSH. The short time frame of acceptance criteria (1 year) may also have introduced a limitation because there may have been older studies with results worthy of analysis. Including these studies may have biased the results during the pandemic.

Conclusions

mHealth is an effective tool to augment or, in some cases, supplant certain treatments of mental health care. It has been shown to improve depression and anxiety (primary research objectives) and many other conditions such as distress, sleep disturbance, pain, digestive disturbance, anger, fatigue, decision regret, and self-efficacy. Although some studies reported results that were not statistically significant, all but 2 interventions showed improvement in at least one area of care. These results are promising for both patients and providers seeking additional methods of care.

Data Availability

Data from this study can be obtained by asking the lead author.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Observation-to-theme conversion (Intervention, Results, Medical Outcomes).
[DOCX File , 31 KB - mental_v9i12e42301_app1.docx]

Multimedia Appendix 2
Observation-to-theme conversion (Patient Satisfaction, Effectiveness, Barriers).
[DOCX File , 30 KB - mental_v9i12e42301_app2.docx]

Multimedia Appendix 3
Other observations incident to the data extraction process (sample size, country of origin, effect size, statistics used, JHNEBP strength and quality of evidence).
[DOCX File , 23 KB - mental_v9i12e42301_app3.docx]

References

1. Freud S. Civilization and its discontents. Guelph, ON, Canada: Broadview Press; Dec 16, 2015.
2. Gottman J. The roles of conflict engagement, escalation, and avoidance in marital interaction: A longitudinal view of five types of couples. In: Interpersonal Development (1st Edition). London, UK: Routledge; 2017:359-368.
3. Davison GC, Neale JM. Abnormal Psychology: An Experimental Clinical Approach. Hoboken, NJ: Wiley; 1974.
4. Gilmore S. The Counselor-in-Training. Hoboken, NJ: Prentice Hall (Appleton-Century-Crofts); 1973.
5. Muñoz R. Depression Prevention: Research Directions. London, UK: Hemisphere Publishing Corp; 1987.
6. Turner BJ, Siegel S. Telemedicine and mental health care of young adults during COVID-19: A qualitative study. Practice Innovations 2022 Mar;7(1):64-70. [doi: 10.1037/pri0000170]
7. Foa EB, Franklin M, McLean C, McNally RJ, Pine D, Jane Costello E, et al. Defining anxiety disorders. In: Evans D, editor. Treating and Preventing Adolescent Mental Health Disorders: What We Know and What We Don't Know (2nd Edition). Oxford, UK: Oxford Academic; Sep 1, 2017:189-212.

8. Silva WAD, de Sampaio Brito TR, Pereira CR. COVID-19 anxiety scale (CAS): Development and psychometric properties. Curr Psychol 2022;41(8):5693-5702 [FREE Full text] [doi: 10.1007/s12144-020-01195-0] [Medline: 33204058]

9. World Health Organization (WHO). Telemedicine: Opportunities and Developments in Member States: Report on the Second Global Survey on eHealth. Geneva, Switzerland: World Health Organization; 2010.

10. Kruse C, Heinemann K. Facilitators and Barriers to the Adoption of Telemedicine During the First Year of COVID-19: Systematic Review. J Med Internet Res 2022 Jan 04;24(1):e31752 [FREE Full text] [doi: 10.2196/31752] [Medline: 34854815]

11. Kruse CS, Beane A. Health Information Technology Continues to Show Positive Effect on Medical Outcomes: Systematic Review. J Med Internet Res 2018 Feb 05;20(2):e41 [FREE Full text] [doi: 10.2196/jmir.8793] [Medline: 29402759]

12. World Health Organization (WHO). mHealth: New horizons for health through mobile technologies. WHO. Geneva, Switzerland: World Health Organization; 2011. URL: https://www.afro.who.int/publications/mhealth-new-horizons-health-through-mobile-technologie#:~:text=mHealth%3A%20New%20horizons%20for%20health%20through%20mobile%20technologie,powerful%20combination%20of%20factors%20is%20driving%20this%20change [accessed 2022-10-14]

13. Mitchell KM, Holtz BE, McCarroll AM. Assessing College Students' Perceptions of and Intentions to Use a Mobile App for Mental Health. Telemed J E Health 2022 Apr;28(4):566-574. [doi: 10.1089/tmj.2021.0106] [Medline: 34166099]

14. Daly JR, Depp C, Graham SA, Jeste DV, Kim H, Lee EE, et al. Health Impacts of the Stay-at-Home Order on Community-Dwelling Older Adults and How Technologies May Help: Focus Group Study. JMIR Aging 2021 Mar 22;4(1):e25779 [FREE Full text] [doi: 10.2196/25779] [Medline: 33690146]

15. Massie MJ. Prevalence of depression in patients with cancer. J Natl Cancer Inst Monogr 2004(32):57-71. [doi: 10.1093/jncimonographs/lgh014] [Medline: 15263042]

16. Park S. Which of the Cornell Scale for Depression in Dementia or the Geriatric Depression Scale is more useful to screen for depression in older adults? Asian J Psychiatr 2022 Jun;72:103147. [doi: 10.1016/j.ajp.2022.103147] [Medline: 35533539]

17. Hu T, Zhao X, Wu M, Li Z, Luo L, Yang C, et al. Prevalence of depression in older adults: A systematic review and meta-analysis. Psychiatry Res 2022 May;311:114511. [doi: 10.1016/j.psychres.2022.114511] [Medline: 35316691]

18. Serrano-Ripoll MJ, Zamanillo-Campos R, Fiol-DeRoque MA, Castro A, Ricci-Cabello I. Impact of Smartphone App-Based Psychological Interventions for Reducing Depressive Symptoms in People With Depression: Systematic Literature Review and Meta-analysis of Randomized Controlled Trials. JMIR Mhealth Uhealth 2022 Jan 27;10(1):e29621 [FREE Full text] [doi: 10.2196/29621] [Medline: 35084346]

19. Noble D, Haveland S, Islam MS. Integrating Telepsychiatry Based Care in Rural Acute Community Mental Health Services? A systematic literature review. Asia Pacific Journal of Health Management 2022;17(2):i1105 [FREE Full text] [doi: 10.24083/apjhm.v17i2.1105]

20. Page M, McKenzie J, Bossuyt P, Boutron I, Hoffmann T, Mulrow C, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71 [FREE Full text] [doi: 10.1136/bmj.n71] [Medline: 33782057]

21. Kruse CS. Writing a Systematic Review for Publication in a Health-Related Degree Program. JMIR Res Protoc 2019 Oct 14;8(10):e15490 [FREE Full text] [doi: 10.2196/15490] [Medline: 31527018]

22. Newhouse R, Dearholt S, Poe S, Pugh L, White K. The Johns Hopkins nursing evidence-based practice rating scale. The Johns Hopkins Hospital. Baltimore, MD: The Johns Hopkins Hospital; 2005. URL: https://guides.upstate.edu/c.php?g=1023176&p=7411259 [accessed 2022-10-14]

23. Pannucci CJ, Wilkins EG. Identifying and avoiding bias in research. Plast Reconstr Surg 2010 Aug;126(2):619-625 [FREE Full text] [doi: 10.1097/PRS.0b013e3181de24bc] [Medline: 20679844]

24. Braun V, Clarke V. Using thematic analysis in psychology. Qualitative Research in Psychology 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]

25. Acierno R, Jaffe AE, Gilmore AK, Birks A, Denier C, Muzzy W, et al. A randomized clinical trial of in-person vs. home-based telemedicine delivery of Prolonged Exposure for PTSD in military sexual trauma survivors. J Anxiety Disord 2021 Oct;83:102461. [doi: 10.1016/j.janxdis.2021.102461] [Medline: 34391978]

26. Baek Y, Jeong K, Lee S, Kim H, Seo B, Jin H. Feasibility and Effectiveness of Assessing Subhealth Using a Mobile Health Management App (MibyeongBogam) in Early Middle-Aged Koreans: Randomized Controlled Trial. JMIR Mhealth Uhealth 2021 Aug 19;9(8):e27455 [FREE Full text] [doi: 10.2196/27455] [Medline: 34420922]

27. Colomina J, Drudis R, Torra M, Pallisó F, Massip M, Vargiu E, CONNECARE-Lleida Group. Implementing mHealth-Enabled Integrated Care for Complex Chronic Patients With Osteoarthritis Undergoing Primary Hip or Knee Arthroplasty: Prospective, Two-Arm, Parallel Trial. J Med Internet Res 2021 Sep 02;23(9):e28320 [FREE Full text] [doi: 10.2196/28320] [Medline: 34473068]

28. Dobkin RD, Mann SL, Weintraub D, Rodriguez KM, Miller RB, St Hill L, et al. Innovating Parkinson's Care: A Randomized Controlled Trial of Telemedicine Depression Treatment. Mov Disord 2021 Nov;36(11):2549-2558. [doi: 10.1002/mds.28548] [Medline: 33710659]

29.   Domogalla L, Beck A, Schulze-Hagen T, Herr R, Benecke J, Schmieder A. Impact of an eHealth Smartphone App on the Mental Health of Patients With Psoriasis: Prospective Randomized Controlled Intervention Study. JMIR Mhealth Uhealth 2021 Oct 25;9(10):e28149 [FREE Full text] [doi: 10.2196/28149] [Medline: 34431478]

30.   Fang S, Lin P, Kuo Y. Long-Term Effectiveness of a Decision Support App (Pink Journey) for Women Considering Breast Reconstruction Surgery: Pilot Randomized Controlled Trial. JMIR Mhealth Uhealth 2021 Dec 10;9(12):e31092 [FREE Full text] [doi: 10.2196/31092] [Medline: 34890354]

31.   Fortney JC, Bauer AM, Cerimele JM, Pyne JM, Pfeiffer P, Heagerty PJ, et al. Comparison of Teleintegrated Care and Telereferral Care for Treating Complex Psychiatric Disorders in Primary Care: A Pragmatic Randomized Comparative Effectiveness Trial. JAMA Psychiatry 2021 Nov 01;78(11):1189-1199 [FREE Full text] [doi: 10.1001/jamapsychiatry.2021.2318] [Medline: 34431972]

32.   Huberty J, Puzia ME, Green J, Vlisides-Henry RD, Larkey L, Irwin MR, et al. A mindfulness meditation mobile app improves depression and anxiety in adults with sleep disturbance: Analysis from a randomized controlled trial. Gen Hosp Psychiatry 2021;73:30-37 [FREE Full text] [doi: 10.1016/j.genhosppsych.2021.09.004] [Medline: 34537477]

33.   Jones RA, Mueller J, Sharp SJ, Vincent A, Duschinsky R, Griffin SJ, et al. The impact of participant mental health on attendance and engagement in a trial of behavioural weight management programmes: secondary analysis of the WRAP randomised controlled trial. Int J Behav Nutr Phys Act 2021 Nov 07;18(1):146 [FREE Full text] [doi: 10.1186/s12966-021-01216-6] [Medline: 34743721]

34.   Krzyzanowska MK, Julian JA, Gu C, Powis M, Li Q, Enright K, et al. Remote, proactive, telephone based management of toxicity in outpatients during adjuvant or neoadjuvant chemotherapy for early stage breast cancer: pragmatic, cluster randomised trial. BMJ 2021 Dec 08;375:e066588 [FREE Full text] [doi: 10.1136/bmj-2021-066588] [Medline: 34880055]

35.   Moskowitz JT, Addington EL, Shiu E, Bassett SM, Schuette S, Kwok I, et al. Facilitator Contact, Discussion Boards, and Virtual Badges as Adherence Enhancements to a Web-Based, Self-guided, Positive Psychological Intervention for Depression: Randomized Controlled Trial. J Med Internet Res 2021 Sep 22;23(9):e25922 [FREE Full text] [doi: 10.2196/25922] [Medline: 34550076]

36.   Pakrad F, Ahmadi F, Grace SL, Oshvandi K, Kazemnejad A. Traditional vs Extended Hybrid Cardiac Rehabilitation Based on the Continuous Care Model for Patients Who Have Undergone Coronary Artery Bypass Surgery in a Middle-Income Country: A Randomized Controlled Trial. Arch Phys Med Rehabil 2021 Nov;102(11):2091-2101.e3. [doi: 10.1016/j.apmr.2021.04.026] [Medline: 34175270]

37.   Rollman BL, Anderson AM, Rothenberger SD, Abebe KZ, Ramani R, Muldoon MF, et al. Efficacy of Blended Collaborative Care for Patients With Heart Failure and Comorbid Depression: A Randomized Clinical Trial. JAMA Intern Med 2021 Oct 01;181(10):1369-1380 [FREE Full text] [doi: 10.1001/jamainternmed.2021.4978] [Medline: 34459842]

38.   Romijn G, Batelaan N, Koning J, van Balkom A, de Leeuw A, Benning F, et al. Acceptability, effectiveness and cost-effectiveness of blended cognitive-behavioural therapy (bCBT) versus face-to-face CBT (ftfCBT) for anxiety disorders in specialised mental health care: A 15-week randomised controlled trial with 1-year follow-up. PLoS One 2021;16(11):e0259493 [FREE Full text] [doi: 10.1371/journal.pone.0259493] [Medline: 34767575]

39.   Su JJ, Yu DS. Effects of a nurse-led eHealth cardiac rehabilitation programme on health outcomes of patients with coronary heart disease: A randomised controlled trial. Int J Nurs Stud 2021 Oct;122:104040. [doi: 10.1016/j.ijnurstu.2021.104040] [Medline: 34333211]

40.   Taguchi K, Numata N, Takanashi R, Takemura R, Yoshida T, Kutsuzawa K, et al. Clinical Effectiveness and Cost-effectiveness of Videoconference-Based Integrated Cognitive Behavioral Therapy for Chronic Pain: Randomized Controlled Trial. J Med Internet Res 2021 Nov 22;23(11):e30690 [FREE Full text] [doi: 10.2196/30690] [Medline: 34813489]

41.   Wong AKC, Wong FKY, Chow KKS, Wong SM, Lee PH. Effect of a Telecare Case Management Program for Older Adults Who Are Homebound During the COVID-19 Pandemic: A Pilot Randomized Clinical Trial. JAMA Netw Open 2021 Sep 01;4(9):e2123453 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.23453] [Medline: 34499135]

42.   Aikens JE, Valenstein M, Plegue MA, Sen A, Marinec N, Achtyes E, et al. Technology-Facilitated Depression Self-Management Linked with Lay Supporters and Primary Care Clinics: Randomized Controlled Trial in a Low-Income Sample. Telemed J E Health 2022 Mar;28(3):399-406. [doi: 10.1089/tmj.2021.0042] [Medline: 34086485]

43.   Akin-Sari B, Inozu M, Haciomeroglu AB, Trak E, Tufan D, Doron G. Cognitive training using a mobile app as a coping tool against COVID-19 distress: A crossover randomized controlled trial. J Affect Disord 2022 Aug 15;311:604-613 [FREE Full text] [doi: 10.1016/j.jad.2022.05.118] [Medline: 35636516]

44.   Bathgate CJ, Kilbourn KM, Murphy NH, Wamboldt FS, Holm KE. Pilot RCT of a telehealth intervention to reduce symptoms of depression and anxiety in adults with cystic fibrosis. J Cyst Fibros 2022 Mar;21(2):332-338. [doi: 10.1016/j.jcf.2021.07.012] [Medline: 34366282]

45.   Catuara-Solarz S, Skorulski B, Estella-Aguerri I, Avella-Garcia CB, Shepherd S, Stott E, et al. The Efficacy of "Foundations," a Digital Mental Health App to Improve Mental Well-being During COVID-19: Proof-of-Principle Randomized Controlled Trial. JMIR Mhealth Uhealth 2022 Jul 01;10(7):e30976 [FREE Full text] [doi: 10.2196/30976] [Medline: 34978535]

46.   Deady M, Glozier N, Calvo R, Johnston D, Mackinnon A, Milne D, et al. Preventing depression using a smartphone app: a randomized controlled trial. Psychol Med 2022 Feb;52(3):457-466. [doi: 10.1017/S0033291720002081] [Medline: 32624013]

47. Drew RJ, Morgan PJ, Young MD. Mechanisms of an eHealth program targeting depression in men with overweight or obesity: A randomised trial. J Affect Disord 2022 Feb 15;299:309-317. [doi: 10.1016/j.jad.2021.12.001] [Medline: 34871640]

48. Guo Y, Li Y, Yu C, Xu H, Hong YA, Wang X, et al. Long-term Effects of a Social Media-Based Intervention (Run4Love) on Depressive Symptoms of People Living With HIV: 3-Year Follow-up of a Randomized Controlled Trial. J Med Internet Res 2022 Jun 28;24(6):e36809 [FREE Full text] [doi: 10.2196/36809] [Medline: 35763324]

49. Gustafson DH, Kornfield R, Mares M, Johnston DC, Cody OJ, Yang EF, et al. Effect of an eHealth intervention on older adults' quality of life and health-related outcomes: a randomized clinical trial. J Gen Intern Med 2022 Feb;37(3):521-530 [FREE Full text] [doi: 10.1007/s11606-021-06888-1] [Medline: 34100234]

50. Kuhn E, Miller KE, Puran D, Wielgosz J, YorkWilliams SL, Owen JE, et al. A Pilot Randomized Controlled Trial of the Insomnia Coach Mobile App to Assess Its Feasibility, Acceptability, and Potential Efficacy. Behav Ther 2022 May;53(3):440-457. [doi: 10.1016/j.beth.2021.11.003] [Medline: 35473648]

51. Lopez CM, Gilmore AK, Brown WJ, Hahn CK, Muzzy W, Grubaugh A, et al. Effects of Emotion Dysregulation on Post-treatment Post-traumatic Stress Disorder and Depressive Symptoms Among Women Veterans With Military Sexual Trauma. J Interpers Violence 2022 Aug;37(15-16):NP13143-NP13161. [doi: 10.1177/08862605211005134] [Medline: 33775153]

52. Mitchell SE, Reichert M, Howard JM, Krizman K, Bragg A, Huffaker M, et al. Reducing Readmission of Hospitalized Patients With Depressive Symptoms: A Randomized Trial. Ann Fam Med 2022;20(3):246-254 [FREE Full text] [doi: 10.1370/afm.2801] [Medline: 35606137]

53. Nardi W, Roy A, Dunsiger S, Brewer J. Analyzing the Impact of Mobile App Engagement on Mental Health Outcomes: Secondary Analysis of the Unwinding Anxiety Program. J Med Internet Res 2022 Aug 15;24(8):e33696 [FREE Full text] [doi: 10.2196/33696] [Medline: 35969440]

54. Orman Z, Thrift AG, Olaiya MT, Ung D, Cadilhac DA, Phan T, STANDFIRM (Shared Team Approach between NursesDoctors For Improved Risk factor Management) Investigators. Quality of life after stroke: a longitudinal analysis of a cluster randomized trial. Qual Life Res 2022 Aug;31(8):2445-2455. [doi: 10.1007/s11136-021-03066-y] [Medline: 35067819]

55. Sun S, Lin D, Goldberg S, Shen Z, Chen P, Qiao S, et al. A mindfulness-based mobile health (mHealth) intervention among psychologically distressed university students in quarantine during the COVID-19 pandemic: A randomized controlled trial. J Couns Psychol 2022 Mar;69(2):157-171. [doi: 10.1037/cou0000568] [Medline: 34264696]

56. Volpato E, Banfi P, Pagnini F. Promoting Acceptance and Adherence to Noninvasive Ventilation in Chronic Obstructive Pulmonary Disease: A Randomized Controlled Trial. Psychosom Med 2022 May 01;84(4):488-504 [FREE Full text] [doi: 10.1097/PSY.0000000000001053] [Medline: 35149638]

57. Ware P, Shah A, Ross HJ, Logan AG, Segal P, Cafazzo JA, et al. Challenges of Telemonitoring Programs for Complex Chronic Conditions: Randomized Controlled Trial With an Embedded Qualitative Study. J Med Internet Res 2022 Jan 26;24(1):e31754 [FREE Full text] [doi: 10.2196/31754] [Medline: 35080502]

Abbreviations

CINAHL: Cumulative Index to Nursing and Allied Health Literature
HADS: Hospital Anxiety and Depression Scale
HADS-A: Hospital Anxiety and Depression Scale-Anxiety
HADS-D: Hospital Anxiety and Depression Scale-Depression
JHNEBP: John's Hopkins Nursing Evidence-Based Practice
MeSH: Medical Subject Headings
mHealth: mobile health
OR: odds ratio
PICOS: participants, intervention, comparison with the control or other group, medical outcomes, and study design
PTSD: posttraumatic stress disorder
RCT: randomized controlled trial
WoS: Web of Science

XSL•FO

RenderX

Original Paper

Experiences of Patients and Therapists Testing a Virtual Reality Exposure App for Symptoms of Claustrophobia: Mixed Methods Study

Gwendolyn Mayer[1], Dr sc hum; Nadine Gronewold[1], MSc; Kirsten Polte[1], MSc; Svenja Hummel[1]; Joshua Barniske[1], BSc; Jakob J Korbel[2], MSc; Rüdiger Zarnekow[2], Prof Dr; Jobst-Hendrik Schultz[1], PD, MD

[1]Department of General Internal Medicine and Psychosomatics, Heidelberg University Hospital, Heidelberg, Germany
[2]Information and Communication Management, Technische Universität Berlin, Berlin, Germany

Corresponding Author:
Gwendolyn Mayer, Dr sc hum
Department of General Internal Medicine and Psychosomatics
Heidelberg University Hospital
Im Neuenheimer Feld, 130.3
Heidelberg, 69120
Germany
Phone: 49 62215635685
Email: gwendolyn.mayer@med.uni-heidelberg.de

Abstract

Background: The effectiveness of virtual reality exposure (VRE) in the treatment of anxiety disorders is well established. Several psychological mechanisms of VRE have been identified, whereby both emotional processing and the sense of presence play a key role. However, there are only few studies that contribute to our knowledge of examples of implementation in the case of VRE for claustrophobia based on patients' experiences and the perspective of therapists.

Objective: This study asks for key elements of a VRE app that are necessary for effective exposure for people with claustrophobic symptoms.

Methods: A mixed methods design was applied in which patients (n=15) and therapeutic experts (n=15) tested a VRE intervention of an elevator ride at 5 intensity levels. Intensity was varied by elevator size, duration of the elevator ride, and presence of virtual humans. Quantitative measures examined self-reported presence with the Igroup Presence Questionnaire (IPQ) ranging from 0 to 6 and 15 Likert-scaled evaluation items that had been developed for the purpose of this study, ranging from 1 to 5. In both measures, higher scores indicate higher levels of presence or agreement. Think-aloud protocols of the patients and semistructured interviews posttreatment of all participants were conducted to gain in-depth perspectives on emotional processes.

Results: The intervention induced a feeling of presence in patients and experts, posttreatment scores showed a high IPQ presence score (mean 3.84, SD 0.88), with its subscores IPQ spatial presence (mean 4.53, SD 1.06), IPQ involvement (mean 3.83, SD 1.22), and IPQ experienced realism (mean 2.75, SD 1.02). Patients preferred a setting in the presence of a therapist (mean 4.13, SD 0.83) more than the experts did (mean 3.33, SD 1.54). Think-aloud protocols of the patients revealed that presence and anxiety both were achieved. Qualitative interviews of patients and experts uncovered 8 topics: feelings and emotions, personal story, telepresence, potential therapeutic effects, barriers, conditions and requirements, future prospects, and realization. The intensity levels were felt to appropriately increase in challenge, with ambivalent results regarding the final level. Virtual humans contributed to feelings of fear.

Conclusions: Key elements of a VRE app for claustrophobic symptoms should include variation of intensity by adding challenging cues in order to evoke presence and anxiety. Virtual humans are a suitable possibility to make the intervention realistic and to provide a sense of closeness; however, some of the fears might then be related to symptoms of social phobia or agoraphobia. Patients may need the physical presence of a therapist, though not all of them share this view. A higher degree of sophistication in the intensity levels is needed to deliver targeted help for specific symptoms of anxiety.

Introduction

Virtual Reality Exposure in the Treatment of Anxiety Disorders

Anxiety disorders are the most prevalent mental disorders worldwide [1], with a considerable impact on the individual's quality of life [2] and on occupational outcomes [3]. A substantial share of patients with anxiety show comorbid depressive symptoms and remain chronically affected [4]. Treatment guidelines recommend cognitive behavioral therapy (CBT) as an effective treatment [5-7], of which exposure therapy (ET) is a core element of evidence-based approaches [8]. This specific therapy form builds upon the mechanisms of conditioning and learning. By exposing the patient to the feared situation or object, they may overcome the anxiety. Mechanisms behind this therapeutic effect have been identified as cognitive and emotional factors, such as inhibitory learning, emotional processing, and self-efficacy [9]. ET has been elaborated in various formats beyond the classical in vivo exposure, including the computer-based presentation of the feared stimulus. Digital technology was formerly used to create an exposure situation based on videos or images. More recent approaches use augmented reality (AR) and virtual reality (VR) apps with the use of mobile technology and head-mounted displays (HMDs), respectively [10]. Virtual exposure has the advantage of increasing the self-reflectiveness of patients [11] and provide a better level of standardization for clinicians [12].

The state of the art of virtual reality exposure (VRE) for anxiety disorders allows for multiple purposes in terms of objectives and scope of diagnoses. VRE is currently used for the assessment and treatment of anxiety and related disorders, such as obsessive-compulsive disorders and posttraumatic stress disorders [9,13], as well as a diagnostic tool for paranoid ideations [14]. Studies show promising results for the treatment of social anxiety disorders and fear of public speaking [15,16] by including virtual humans that may be controlled by a therapist [17]. A comparable effectiveness of VRE in the treatment of anxiety disorders compared to in vivo treatment has been successfully identified in the broad spectrum of anxiety-related disorders [18,19] and specific phobias [20,21]. Deterioration rates are likely to occur randomly and could not be traced back to the application of VRE therapy in a meta-analysis with several treatment approaches for anxiety [22].

Psychological Mechanisms of VRE in Specific Phobias

The effectiveness of ET for specific phobias in general can be explained by several factors. In early studies, the emotional processing theory by Foa and Kozak [23] served as an explanation [9]. Foa and Kozak [23] showed that in exposure, (1) the fear response has to be activated, then (2) a habituation (decrease) of fear will take place gradually, and (3) the initial reaction to the feared situation or object will decrease over several sessions. In line with these findings, the avoidance of relaxation during exposure has been detected to be an important

emotional prerequisite for a successful ET in several studies, as compiled by Böhnlein et al [10]. In their review, they found several other emotional, cognitive, and behavioral factors that explain the success of ET (eg, cognitive factors as self-efficacy before and the focus on changes in fear-relevant cognitions during the exposure), that is, by increasing awareness for cognitions before and after experiencing stress during ET. An important behavioral success factor is the variation of the context or stimulus variation [10]. With respect to the latter, different levels of stimuli variations seem to be beneficial for the reduction in phobic symptoms. For this reason, VRE apps usually offer different intensity levels by gradually adding more fear-evoking cues within the virtual environment (see, eg, Ref. [24] for social phobias and Ref. [25] for acrophobia).

When working with VR environments, the feared context or stimulus can only be felt by the user if there is a sense of presence, in other words a feeling of "being there," during the VR session [26]. Three aspects constitute the nature of presence, which are (1) spatial presence (ie, the degree of experiencing the environment as a 3D room to interact in), (2) involvement (ie, the strength of internal focus toward the VR environment and the degree of "forgetting" the real world), and (3) realness (ie, the degree of comparability of the VR environment with the real world) [27]. Study results show a positive relationship between the sense of presence and anxiety, although with a larger effect for fear of animals and fear of flying and nearly no effect for social phobias and claustrophobia [26]. Presence is closely related to phobic elements within the VR environment and therefore to a feeling of anxiety, but presence alone is not sufficient for a positive treatment outcome [28]. Presence is a concept that describes the perspective of the user, while the objective attribute of the technology enabling the user to feel presence is summarized by the term "immersion" [29]. In anxiety treatment, higher levels of immersion are associated with the correlation of presence and anxiety [26].

Nevertheless, presence seems to be no necessary requirement for inducing fears within VR, as shown in a study that investigated whether social anxiety could be induced more effectively by in vivo talking to somebody or during conversation with avatars. The authors emphasized that although virtual conversation was rated with lower presence, participants reported a higher degree of fear toward their virtual counterpart [30].

Relevance and Examples of the Treatment of Claustrophobia

Claustrophobia belongs to the category of several specific phobias and refers to an irrational fear and avoidance of enclosed places (elevators, tunnels, caves) and the inability to escape from them. The fear seems to be composed of a fear of suffocation and of confinement [31]. Epidemiological studies estimate a lifetime prevalence of 2.2%, with a treatment rate of 27.5% [32]. However, subclinical symptoms in daily life that are reported by the general population have been observed in

12.5% [33], and more low-threshold treatment offers might be necessary to meet the demands of these people. Specific phobias are associated with impairment and show high comorbidity with other mental health disorders [32]. Moreover, a recent population-based survey in 24 countries showed that patients suffering from specific phobias who were seeking professional help received helpful treatment only in 23% of cases from their first professional contact [34]. In the case of claustrophobia, the limitations due to this phobia become highly relevant for patients preparing for an MRI scan. Many radiographers report being confronted daily with claustrophobic patients [35]. VR technology is currently used for symptom reduction in MRI scans in educational programs [36,37] or for the purpose of distraction [38].

Only few studies describe VRE specifically aiming at the treatment of claustrophobia, although the first studies were conducted based on a case study as early as 1998 [39] and by the same author with 4 participants in 2000 [40]. As reported by Ling et al [26], the role of presence in VRE for claustrophobia was investigated only in 2 studies, which did not find a relationship between presence and the degree of anxiety felt. Following the above-mentioned review of Böhnlein et al [10], important behavioral factors might have been missing, such as sufficient variation of context, which is a necessary success factor of ET. Other studies make use of rather simple settings to induce claustrophobic fears (eg, a closed box [41] or a room with a sudden fire), which might trigger confounding fears, in addition to claustrophobia [42]. The experience of narrow space in reality is not only influenced by the size of a room but also by the presence of other people, such as in crowded places or supermarkets. This is especially the case in an elevator, where one has to share a room that is already narrow due to the presence of others. This might be one reason some authors consider claustrophobia a prodromal stage of agoraphobia [43].

The Therapeutic Experts' Perspective on VRE

Former study results show a generally high acceptance of VRE among therapists but at the same time a low familiarity [44]. Studies so far have analyzed the experiences and treatment effects on patients. However, we did not find any study that included the experiences made by therapists of different therapeutic backgrounds testing the same intervention as their potential patients, even in studies where the VR sessions were therapist-led by voiceover instructions and therapists thereby stayed in contact with their patients [45]. The perspective of therapeutic experts is essential to understand the experiences reported by patients during and after VRE and different views due to the variety of approaches (CBT or psychodynamic) might add valuable insights. For the implementation in anxiety treatment, it is also interesting to find out which factors might influence the therapists' attitudes and openness toward the use of VRE interventions in their therapies. As therapeutic experts will prescribe this kind of intervention in the future, understanding their perspective is crucial for future developments.

Objectives

The aim of this study was to test a VR app with different intensity levels for the future use of VRE for claustrophobic symptoms in order to find key elements that are necessary for exposure. To reach this goal, we asked whether our fully immersive intervention is sufficiently able to induce feelings of presence as a basic requirement to evoke anxiety. Furthermore, the following question was studied: Did intensity levels reflect an appropriate growth in challenge for future personalization of the app?

Additionally, we were also interested in the perspective of psychotherapeutic experts of different approaches regarding the content of the intensity levels and potential barriers and facilitators for implementation into their future clinical practice.

Methods

Study Design

This study was a nonrandomized feasibility study that analyzed the experiences and perspectives of patients and therapeutic experts regarding a VRE app for symptoms of claustrophobia.

The app was developed by clinicians and technology experts in the context of the nationally funded project SELFPASS (Self-administered Psycho-Therapy Systems) that has been described elsewhere [46,47]. The design and content of the intervention were closely oriented to psychotherapeutic manuals of CBT [48]. The intervention was designed as a fully immersive system with the use of HMD technology that has the advantage of shutting out physical reality, while providing a high degree of fidelity [49].

A convergent mixed methods approach was used, with quantitative measures pre- and posttreatment, think-aloud protocols during treatment, and qualitative interviews posttreatment [50]. This combination of methods allowed for the integration of generalizable aspects with quantitative scales, on the one hand, and a detailed description of individual experiences with qualitative material, on the other hand [51]. The latter had an impact on the sample size considerations. As we chose to carry out qualitative interviews and to analyze think-aloud protocols, we limited the group size to 15 participants in each group. We chose equal group sizes to be able to compare scales. However, as emphasized by Malterud et al [52], the more information power the sample holds, the lower the sample size that can be planned. One important factor that contributes to information power is the quality of dialogue that we considered as high due to the close contact to the participants during and after the intervention.

Recruitment and Procedures

Recruitment was performed by postings and mail distribution services to patients, employees, and experts within the Heidelberg University Hospital's Department of Internal Medicine and Psychosomatics in the Medical Faculty. The criteria for inclusion were a minimum age of 18 years and the capacity to provide consent. Patients were included when they reported suffering from symptoms of claustrophobia and when they had received a diagnosis of an anxiety disorder of any kind

by a physician at the Heidelberg University Hospital within the past year. A Structured Clinical Interview for DSM-5 (SCID) interview [53] was then conducted by author NG, who is a trained psychotherapist. This interview was conducted to exclude those patients in danger of suicidality or psychosis and to confirm the self-reported symptom profile. Experts were considered eligible if they held a university degree in medicine or psychology and were active in the field of psychotherapy, psychiatry, or psychosomatics. To obtain broad perspective of feedback, experts from various psychotherapeutic approaches were considered eligible, such as psychodynamic, psychoanalytic, and systemic approaches, as well as CBT.

The testing procedures were carried out in a separate room within the facilities of the hospital. For the time of the study, an area of at least 9 m^2 was reserved. The hardware consisted of an HTV VIVE Pro Eye headset, 2 base stations, and 2 controllers. The 2 base stations were set up in the corners of the room to serve as reference points for the headset and controllers. Participants' safety was guaranteed by the presence of at least 1 researcher, who took care for the position of the cable. Hygienic measures were applied before and after testing to reduce the risk of infection during the COVID-19 pandemic.

Ethical Considerations

Ethical approval was obtained from the Ethics Committee of the University of Heidelberg (S-746/2020), and recruitment started thereafter. All participants signed an informed consent form.

Measurements

Participants completed a short questionnaire with demographic details, the technology commitment scale (TCS) [54], and the State-Anxiety Scale of the State-Trait Anxiety Inventory (STAI-S) [55]. The concept of technology commitment refers to the individual readiness to use technology, which is based on 3 underlying concepts: technology acceptance, technology competence conviction, and technology control conviction. All 12 items range from 1 (fully disagree) to 5 (fully agree), with single items to be coded reversely. The instrument shows high correlation to the concept of self-efficacy; however, it is more closely related to the use of technology [56]. There are no cut-off values for the TCS; however, population studies have shown mean values of mean 3.73 (SD 0.62) in technology commitment, a mean of 3.27 (SD 0.94) for technology acceptance, a mean of 4.16 (SD 0.80) for technology competence conviction, and a mean of 3.75 (SD 0.74) for technology control conviction [56]. We expected no differences in technology commitment between patients and experts, as both groups should be willing to engage with the VRE. The STAI-S was applied pretreatment in order to validate the anxiety of patients before the intervention that should be higher than in the other group. A cut-off of 41 is

recommended in the literature to differentiate between healthy and clinical levels of anxiety [57]. We expected higher scores in the STAI-S for the patients.

Thereafter, they started the virtual tasks. After completion of at least 3 tasks, they filled out the Igroup Presence Questionnaire (IPQ) that measured the subscales spatial presence, involvement, experienced realism, and general presence on a 7-point Likert scale with 14 items, with higher results indicating a higher sense of presence [58]. Although the IPQ is a validated instrument, the authors present no norm values; however, studies by the authors with values ranging from 0 to 6 were interpreted as relevant when exceeding a mean of 3 during the intervention [59]. As the intervention was designed as a fully immersive app, we expected high levels of presence, as measured by the subscales of the IPQ in both groups.

A set of 15 evaluation items was developed for the purpose of this study. These items were Likert-scaled with a range from 1 (lowest) to 5 (highest). The first 15 items asked about the appropriateness of the design, its clarity, the length of the intervention, and the perceived usefulness. Two further items asked for the possibility to carry out the intervention in the future while being alone or whether there is a need of a therapist being present. We asked these questions to gain insights into the different perspectives; however, group comparisons were not relevant.

All participants were invited for 1 session of a duration of 60-90 minutes. They were asked to complete at least 3 tasks during the session. The experts were asked to complete tasks 1, 3, and 5. However, every participant could perform all tasks by choice.

The participants were invited to speak aloud during the intervention and to talk frankly about their observations, feelings, and emotions, if they wanted to do this. They were reminded to express their thoughts. However, some participants were too involved to make use of this opportunity and forgot to do so. This so-called think-aloud method was applied to obtain immediate insights into the experiences of the participants [60]. As our analysis focused on the experiences of patients carrying out VRE, we only reported on the results of the think-aloud protocols of the patients.

The semistructured interviews posttreatment included 6 questions that referred to an overall impression of the intervention, the perception of the intensity levels, the assessment of effects of the intervention for patients, and suggestions for improvement. Further questions relating to specific patient groups and implementation prerequisites were asked to the experts only. All questions are presented in Multimedia Appendix 1. The whole process is illustrated in Figure 1.

Figure 1. Process of measurements. IPQ: Igroup Presence Questionnaire; STAI: State-Trait Anxiety Inventory.
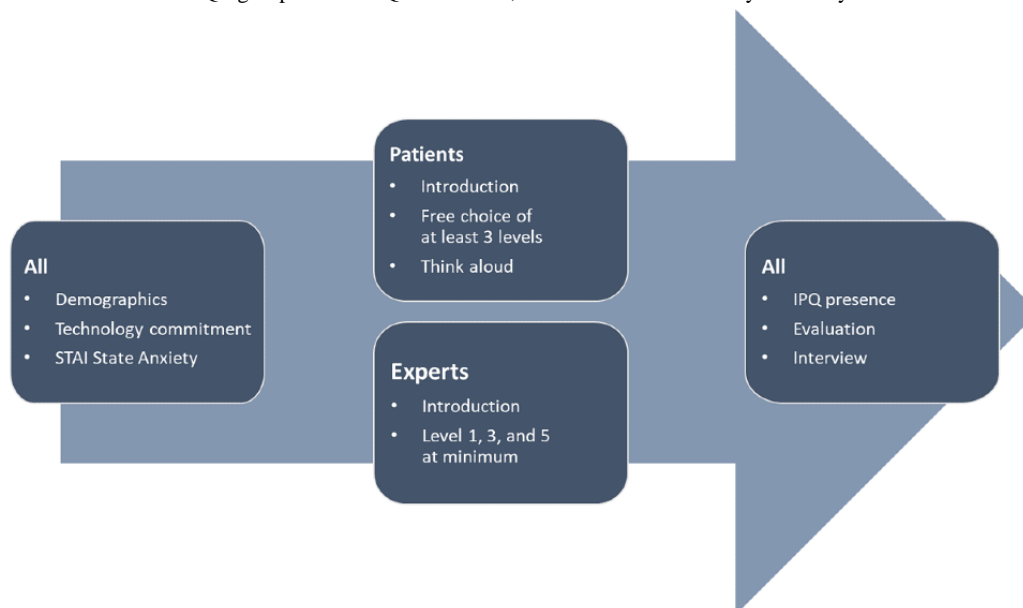
Virtual Environment

Before starting the virtual tasks, all participants were provided with an introduction by a female voice guiding them through the core functionalities of the program. This was performed using a scenario of the frame of an elevator placed on green grass, without a roof or walls, consisting only of a bottom panel, pillars, and an operating panel of the elevator in order to avoid feelings of claustrophobic anxiety at this stage (Figure 2). The participants were instructed to try out the operating panel and to understand 2 buttons: A yellow alarm clock could be pressed to receive the instruction again. A red "stop" button would immediately finish the level and lead to the main menu in the case of an emerging panic attack. The end of each session was marked by a gray cube that should be pressed downward (Figure 3). During the introduction as well as later during the tasks, the participants were able to see their own hands but not their body.

After completion of the virtual introduction, the participants reached the main menu, which was presented in a blue sky over green grass. Here, they were asked to choose 1 of 5 levels by pointing at and clicking a number with their controller. After that, 3 questions were visually and acoustically presented: "How high is your fear or unease when you think about this task?", "How high will your fear or unease be when you will have completed the task?", and "How high is your commitment to complete this task?" These questions were presented as 10-point scales from 0 (no fear) to 10 (extreme fear) and as percentages to assess the commitment. After the completion of each task, participants were asked to assess the amount of actual fear felt at that moment and to rate their willingness to perform this task again. After each successful task, the environment changed and trees grew on the green grass as an element of gamification (see Figure 4).

The virtual task was an elevator ride in a simulated office building. Five levels of difficulty with increasing "claustrophobic" challenges were created to graduate the intensity of exposure. Four factors were systematically manipulated to reach the increasing intensity: the size of the elevator, the duration of the ride, brightness, and the number of passengers. The size was varied with 3 sizes. The duration was varied by 3 variables: duration of opening and closing the door, duration to reach the next floor, and duration of stopping at each floor. The virtual passengers were able to act responsively when the participants sought eye contact by returning the gaze, but they were not able to talk. Examples are presented in Figures 5 and 6.

The first level was a simple task: reaching the second floor from the ground floor in a large empty elevator without any passengers (Figure 3). The second level was again a ride in a rather large elevator from the second to the sixth floor, with 1 female passenger already in the elevator. At the third level, 2 male passengers were in an elevator of medium size, leaving less space for the participant who had to go from the office on the 6th floor to the restaurant on the 12th (Figure 6). The fourth level included 3 passengers, among whom 1 female wore a COVID-19 mask and coughed (Figure 5). The size of the elevator was medium. The task was to get from the ninth floor down to the ground floor. At the final level, no passengers were in the elevator, but the task was to get from the parking area on the second floor to the ninth floor. At this stage, the elevator was as small as in none of the previous scenarios, with dark lighting and old elevator noises. With each level, the waiting time in front of the elevator increased, while the speed of the elevator during the ride reduced. In addition, the VR environment allowed participants to get out on another floor, but the level could only be completed in the respective level that had been communicated in the task.

Figure 2. Own hand pressing a button by using the controller during the introduction to the program.

Figure 3. View out of the elevator with the final cube (to be pressed down).

Figure 4. Start screen; after each level, more trees were "won." SELFPASS: Self-administered Psycho-Therapy Systems.

Figure 5. Female virtual passenger at level 4.

Figure 6. Two male virtual passengers at level 3.

Data Analysis

The quantitative data were first analyzed descriptively, and means, SDs, frequencies, and percentages are reported. The scores of the 3 scales STAI-S, TCS, and IPQ were calculated according to the manuals. We carried out a reliability analysis for the scales, the STAI-S reached α=.93, the TCS showed α=.87, and the IPQ's α was .81. Therefore, all scales reached a good level of reliability [61]. The IPQ items were recoded so that the original range of –3 to +3 was transferred to 0 to 6. We explored differences between the groups in the TCS and IPQ scores by *t* tests. In the case of a violation of the homogeneity assumption, we chose the Welch test results. In the case of the STAI, which is usually treated as an ordinal scale [62], we preferred a nonparametric test and performed a Mann-Whitney *U* test [63]. A significance level of *P*<.05 was considered statistically significant. Evaluation items were recoded, if necessary, and an evaluation score from 1 (low evaluation) to 5 (high evaluation) was calculated out of the first 13 items. Two further items referred to the need of assistance by a therapist and to the potential to carry out the intervention alone. These items were calculated separately. Statistical analysis was carried out using the Statistical Package for the Social Sciences (IBM SPSS Statistics ver. 24).

The think-aloud protocols of the patients and the semistructured interviews of all participants were transcribed and analyzed with the help of MAXQDA [64]. The think-aloud protocols were analyzed by 1 of the researchers (author JB) and supervised by author GM. Two coders (authors GM and SH) carried out the analysis of the interviews. The research team thoroughly

discussed the 2 coding systems, and disagreements were resolved. Following the rationale of analysis by an inductive approach, as suggested by Mayring [65], new codes were added when new aspects emerged from the data; however, after reaching half of the material, no new codes were accepted. In the qualitative part, we did not count codes, as the frequent occurrence of a topic might not reflect its importance but rather the willingness to talk longer on that topic than on another [66].

Results

Participants

A total of 30 participants took part in the study, with 15 (50%) patients and 15 (50%) experts. The mean age of all participants was 40.14 (SD 14.33) years. The patients' age ranged from 20 to 72 years and showed a mean of 46.07 (SD 17.48) years. The age of the experts ranged from 26 to 41 years, with a mean of 33.79 (SD 17.48) years. Demographic details are presented in Table 1.

Table 1. Demographic characteristics of the study sample.

Characteristics	Participants, n (%)		
	Patients (n=15)	Experts (n=15)	All (N=30)
Gender			
Male	8 (53.3)	9 (60.0)	17 (56.7)
Female	7 (46.7)	6 (40.0)	13 (43.3)
Profession			
Student	2 (13.3)	N/A[a]	N/A
Employee	3 (20.0)	N/A	N/A
Retired	4 (26.7)	N/A	N/A
Not employed	3 (20.0)	N/A	N/A
Other	3 (20.0)	N/A	N/A
Therapist background (multiple choice)			
Psychodynamic therapy	N/A	9 (60.0)	N/A
CBT[b]	N/A	5 (33.3)	N/A
Systemic family therapy	N/A	2 (13.3)	N/A
Other	N/A	1 (6.7)	N/A

[a]N/A: not applicable.

[b]CBT: cognitive behavioral therapy.

Pretreatment Scores

Patients showed the highest scores of state anxiety before starting the intervention. A Mann-Whitney U test indicated that the anxiety of the patients was higher than that of the experts: U=154, P=.03. Experts expressed higher scores in technology commitment than patients (see Table 2). The t tests showed significant differences in the total score between technology commitment (t_{28}=2.19, P=.04) and technology competence conviction (t_{28}=3.53, P=.002). Other subscores were not significantly different between the 2 groups (technology acceptance: t_{28}=0.90, P=.38; technology control conviction: t_{28}=1.01, P=.32).

Table 2. Scores of STAI-S[a] and TCS[b] pretreatment.

Scale	Patients (n=15), mean (SD)	Experts (n=15), mean (SD)	All (N=30), mean (SD)	P value
STAI-S (N=29)[c]	45.79 (10.46)	36.80 (8.29)	41.14 (10.30)	.03
TCS				
Total	3.74 (0.56)	4.23 (0.67)	3.99 (0.66)	.04
Technology acceptance	3.17 (0.87)	3.47 (0.96)	3.32 (0.91)	.38
Technology competence conviction	4.55 (0.91)	5.48 (0.47)	5.02 (0.86)	.002
Technology control conviction	3.50 (0.45)	3.75 (0.85)	3.63 (0.68)	.32

[a]STAI-S: State-Anxiety Scale of the State-Trait Anxiety Inventory.

[b]TCS: technology commitment scale.

[c]Data of 1 patient were missing for the STAI-S.

Posttreatment Scores

The VR intervention reached an IPQ presence score of mean 3.84 (SD 0.88), and its highest subscore was the IPQ spatial presence (mean 4.53, SD 1.06).

The t tests revealed significant differences in IPQ spatial presence (t_{28}=2.50, P=.02). There was no significant difference between the 2 groups regarding IPQ involvement (t_{28}=–0.30, P=.77), IPQ experienced realism (t_{28}=–0.71, P=.49), and IPQ total (t_{28}=0.48, P=.63).

For details of posttreatment IPQ scores, see Table 3.

Overall evaluation was high (mean 4.25, SD 0.32). Answers to the question "I think such interventions are better conducted in the presence of a therapist" had a mean score of 4.13 (SD 0.83) for patients and 3.33 (SD 1.54) for experts, while the question "I could imagine carrying out such interventions alone in the future" had a mean score of 4.20 (SD 1.01) by the patients and 4.00 (SD 1.07) by the experts. All details of further evaluation items are presented in Multimedia Appendix 2.

Table 3. Scores of IPQ[a] presence (range: 0=lowest, 6=highest).

Scale	Patients (n=15), mean (SD)	Experts (n=15), mean (SD)	All (N=30), mean (SD)	P value
IPQ (total)	3.77 (0.91)	3.92 (0.87)	3.84 (0.88)	.63
IPQ spatial presence	4.08 (1.12)	4.97 (0.82)	4.53 (1.06)	.02
IPQ involvement	3.90 (1.25)	3.77 (1.23)	3.83 (1.22)	.77
IPQ experienced realism	2.88 (1.22)	2.62 (0.80)	2.75 (1.02)	.49

[a]IPQ: Igroup Presence Questionnaire.

Qualitative Results I: Think-Aloud Method

Of the 15 patients, 12 (80%) decided to share their observations and experiences with the think-aloud technique during the intervention, while 3 (20%) made only short comments when technical problems occurred. All codes and subcodes are listed in Table 4. In the following section, the codes "feelings and emotions," "self-assessment," and "telepresence" are elaborated. The 3 codes "intensity levels," "own technical expertise," and "technical problems" are related to comments on the quality of different levels in comparison to other software and to problems during the handling of the program.

Table 4. Patients' expressions during the intervention (think-aloud method).

Codes	Subcodes
Feelings and emotions	fear, tension, relaxation, emotional coping, perception of the virtual humans, perception of the elevator
Self-assessment	anxiety, willingness, motivation, satisfaction
Telepresence	involvement, spatial presence, realism
Intensity levels	N/A[a]
Own technical expertise	N/A
Technical problems	control problems, software errors, interaction with program/supervisor

[a]N/A: not applicable.

Feelings and Emotions

A broad category of "feelings and emotions" described notions of fear, tension, and relaxation. Moreover, emotional coping was a feeling that was expressed as a positive emotional reaction to a threatening situation. Finally, perceptions of the virtual humans and of the elevator were expressed in a highly emotional way and therefore coded as feelings and emotions as well.

Many patients reported directly feeling claustrophobic symptoms, for example, in the following expression:

> *Oh God. Woah...why do you build something like that? I always feel like being buried alive.* [Patient, 48 years, male, fear/perception of the elevator]

The feelings and emotions that patients reported during the intervention revealed that their anxiety could not be summarized as claustrophobic symptoms alone. Due to the presence of 1-3 virtual humans from level 2 on, further symptoms were reported that might be explained by shame or a feeling of not being able to escape from the situation, as observed both in social phobic and in agoraphobic patients.

> *The two people, one person looked at me, I don't like being watched, it's sort of uncomfortable and then the noises of the person standing behind me,...exactly so the fear was then getting in and driving higher and when I'm out, I can't think of a better term for it, I felt relief there to get out.* [Patient, 31 years, male, fear/perception of the virtual humans]

Relaxation was expressed by patients after completion of a level:

> *I was happy to just be out and then the thing was, then I could literally leave it behind me.* [Patient, 31 years, male, relaxation]

Moreover, patients talked about applying *emotional coping* while feeling their symptoms of claustrophobia:

> *I would like to lean against the wall a bit, but that's not possible right now...Yeah. Wall in the back for sure, because nothing can come from behind...Yes, that gives security that nothing is coming from behind and since I also partly know, I can already tell that, ah ok (yells), I hate it when it counts down and I don't know what's coming out.* [Patient, 28 years, male, fear/emotional coping]

Self-Assessment

Before and after the intervention, the users were asked to assess the degree of anxiety that was expected in advance or experienced retrospectively, the degree of motivation, and the

willingness to repeat the intervention. Some patients took this task seriously and answered in a sophisticated way:

> *How willing are you to repeat this exercise? Since I take it as a learning success, I'll say 80%, 100% would be up here, because I'd like to escape from something like that. But since I just take it as a starting point to further work on my problems I'll just say 80%.* [Patient, 28 years, male, motivation/willingness]

Patients often stayed in contact with the psychologist during this part of the intervention. When they chose a level of anxiety felt with their controllers, it was coded as part of the self-assessment and not under the code "feelings and emotions":

> *Well, so also 5-6. I'm shaking already, see?* [Patient, 72 years, female, anxiety]

Telepresence

The patients reported feeling spatial presence, realism, and involvement during the VRE. The code *spatial presence* reflected the perception of the room, while *realism* was given when elements of the intervention tended to be confounded with reality.

> *To be honest, I'm a bit scared that I'll run into something, hehe.* [Patient, 29 years, male, spatial presence]

> *Whoa, isn't it normal? It's not normal that you're afraid of virtual people?* [Patient, 48 years, male, realism]

However, some of the patients expressed still being aware of the difference between the real and the virtual environment. This awareness provided a feeling of controllability, as shown in the following quote:

> *The good thing about it is that I know that nothing can happen to me in this case...With a normal elevator, it even starts when I go down a few centimeters...* [Patient, 58 years, male, realism]

Involvement, in turn, indicated that the participants were submerged in the experience.

> *But now I was so focused on it that I didn't understand which floor, haha. Believe 2nd, but there I am.* [Patient, 36 years, female, involvement]

Qualitative Results II: Semistructured Interviews

The analysis of the semistructured interviews posttreatment revealed 8 main categories with 34 subcategories in total (Table 5). In the following section, we report all coding in short and provide examples.

Table 5. Codes and subcodes of the qualitative analysis.

Codes	Subcodes
Feelings and emotions	fear, tension, controllability, distress, symptom improvement, positive experiences, perception of virtual humans
Personal story	history of anxiety, personal background
Telepresence	virtuality-reality discrepancy, immersion, realism, perception of the space
Potential therapeutic effects	needs of patients, experiences as a therapist, intensity levels, risks, therapeutic approach, comparison with in vivo, effect
Barriers	financial/technical effort, therapeutic approach
Conditions and requirements	costs, premises, therapeutic setting
Future prospects	therapeutic potential, target group, future challenges
Realization	gamification, hardware, software, suggestions for improvement, personalization

Feelings and Emotions

The participants talked about their experiences during the VR intervention in a detailed way and openly shared their personal feelings, such as fear, tension, or distress. *Fear* was observed in both groups. However, experts talked about their anxious feelings in another way (eg, using words such as "unpleasant" or "odd"). Two factors were often reported as triggers for the anxious feelings: the presence of the other virtual humans and the fear of getting stuck:

> I know my fears. I asked myself "what happens next?" and "does it get stuck?" That would have been the super disaster. Of course, then you can take off your glasses. However, I would not have done it, for whatever reason. [Patient, 53 years, male, fear/history of anxiety]

Moreover, most participants assessed the *perception of virtual humans* as strange and frightening. In particular, the male virtual humans, who were quite tall, caused negative feelings. One of the experts hinted at the possibility of potential risks for female patients with a sexual trauma. One patient even expressed feelings of being negatively concerned with fictitious expectations of the virtual humans:

> As I'm in a hotel or something like, um, waiting for an elevator…there is actually such a tension, um, that it could get uncomfortable…And this waiting and then the people in the elevator. You don't want to attract negative attention or anything. [Patient, 29 years, male, tension/perception of virtual humans/history of anxiety]

Some participants insisted on keeping control over the situation, as interpreted with *controllability*:

> Accordingly, uh, you always have it in your head: Yes, I can also cancel it. [Patient, 53 years, male, controllability/virtuality-reality discrepancy]

Many participants reported *positive experiences*. They assessed the VR intervention as appealing and useful. Specifically, patients talked about the individual *symptom improvement* that they observed within themselves after the intervention by passing through stepwise different levels. In this context, participants expressed the wish to have more than 5 intensity

levels, which in turn was an important factor for *realization* (see next).

Personal Story

As the participants were invited to talk frankly about personal feelings, some of them took the opportunity to talk about their own story, either their *personal background* or their *history of anxiety*. The latter included descriptions of patients that went beyond symptoms of claustrophobia, as most of them were burdened by complex psychosomatic symptom profiles. Some patients, for example, reported having difficulty breathing and were sometimes forced to take an elevator, even if the anxiety is ever-present. Other patients reported to have severe panic attacks that caused them to resort to psychiatric medication.

Some patients were able to develop own strategies of self-efficacy in their past:

> I had such terrible panic attacks. And it happened to me in the supermarket and I was so embarrassed, I could,…I never knew how to pay…So I had to face the situation every time. After two or three times I noticed: No, nothing happens in the supermarket. [Patient, 67 years, female, history of anxiety]

Telepresence

Nearly all participants talked about the degree of telepresence they experienced during the intervention. Their statements referred to the *discrepancy between virtuality and reality*, *immersion*, *realism of the environment*, and *perception of space*, of which all terms will be explained as follows: First, some participants reported feeling remaining doubts regarding the presence of the virtual world, which was termed as *virtuality-reality discrepancy*. The following example illustrates the ambivalence of this experience:

> Because there was in the back of your mind, it is now like this...it is not real in the sense that you normally have it. Where you then walk stairs." "Understand." "Before you get into the elevator." [Patient, 58 years, male, virtuality-reality discrepancy]

However, some participants said that only the cable reminded them of the existence of the real world, while 1 participant stated that she almost leaned against the elevator wall, as she usually does during an elevator ride. Both examples serve as an

illustration for *immersion* as an effectual prerequisite of the feeling of presence.

The *perception of space* was highly influenced by the presence of virtual humans, which seemed to contribute to the limited space in the elevator:

> The most difficult for me were the many people. That was the most difficult, yes. Yes, that makes the whole thing even tighter...Because of the tightness because of the people in such a small elevator. [Patient, 58 years, male, anxiety/perception of virtual humans/perception of the space]

As a final characteristic subcode within *telepresence*, *realism* described the sense that the design and realization of the intervention was conducted in a realistic way. This issue raised some critical remarks, as the graphical realization of the virtual humans was not assessed to be state-of-the-art technology, and therefore some quotes were categorized as well as *suggestions for improvement* within the main category "realization."

Potential Therapeutic Effects

The participants commented in detail on the perceived usefulness of the VR interventions for therapeutic purposes. *Patients' needs* were in the focus of strong concerns, as many participants observed that the VR intervention might not be suitable for every kind of patient. Elderly patients might not benefit from this kind of technology; moreover, some patients could need more individual cues to trigger their anxiety, and therefore, a broader range of intensity levels is needed. The latter directly refers to the subcode *intensity levels,* which were assessed as not fully building up on each other in terms of difficulty. The final intensity level took place at night and alone in a small elevator, whereas levels 3 and 4 included virtual humans. Many participants, but not all of them, found the sessions with the virtual humans more difficult than being completely alone at level 5.

The experts argued from their professional perspective when talking about their *experiences as therapists.* In doing so, they voiced the wish for controlling the situation and the time frame, as it might be necessary to stay for a long time in a situation until the anxiety decreases. The therapists appreciated the possibility of a smooth start for patients and assessed VRE as *comparable to in vivo* or as a starting point:

> I can also imagine that it will be well accepted, because you can't make a fool of yourself in that sense...It's just something different when you really go to the department store and do some exercise. [Expert, 40 years, male, therapeutic potential/comparison with in vivo]

Finally, *related risks* raised by the participants mainly referred to a potential scenario of doing the VRE alone without guidance. In this case, some participants identified an increased danger of being left alone with a panic attack.

Barriers

There were a few reasons participants would not use the VR intervention for therapeutic purposes. These reasons can be summarized by restrictions due to high *financial* or *technical*

efforts that were seen by those therapists who were talking about therapy in a private practice, where care is not provided by a clinic. Other barriers were reported by therapists who followed other *therapeutic approaches* than CBT. Especially psychodynamic therapists stated not to concentrate on a mere behavioral training when treating an anxiety disorder, as expressed, for example, in the following statement:

> And with whom I wouldn't do it? I don't think it will help those who—I think—organize or express their social needs by their fears. [Expert, 30 years, female]

Conditions and Requirements

The category "conditions and requirements" showed some overlapping meanings with "barriers"; however, additional aspects, such as requirements and practical considerations, were reported in this context. Subcodes included *cost*, *facilities*, and *therapeutic setting*. Many participants, especially experts, but also patients, considered the presence of a face-to-face therapist as important for the success of the VR intervention:

> In my eyes, the introduction...is actually very well suited to reduce this fear and to be able to get involved with this whole thing once. And therefore...it is definitely important to include other people to support, from a therapeutic perspective, definitely. [Patient, 31 years, male]

Nevertheless, some of the experts were convinced that patients should use the intervention as mere self-management training at home:

> So that you can say this has to be practiced now, another thing has to be practiced next week...of course, people must have a setup. Well, I don't feel like practicing it in a therapy session, I think so, really, but I would like to let people do it at home. [Expert, 41 years, male]

Experts who were in favor of a guided setting expressed a desire for the features to be directly controlled by the therapist during the session (eg, the degree of narrowness, the number of virtual humans). This aspect is strongly related to features of personalization, which were also included in "realization."

Future Prospects

Overall, the participants described some future prospects for this kind of VR intervention. They saw much *therapeutic potential* in the scenario for different ways of treatment (eg, by addressing more fears than just claustrophobia, by adding a glass floor for patients with acrophobia). Further, they defined *target groups* that were more or less suitable for comparable VR interventions. Future challenges were seen in the successful integration into clinical practice; however, many technological prospects might allow new possibilities of treatment, as 1 expert elaborated:

> I thought about whether you could still build in wearables, there are so many watches that measure the pulse, for example if you could somehow integrate it into the system that you have markers like biofeedback. I would think that's great because, I mean this, shall we say, physiological stress reaction

that you have that always stops. So if you are, let's say you are really claustrophobic and you stand in the elevator with all the people, then you have, you are in a panic state that lasts for a long time but at some point it stops and that's that what you show people over and over again with biofeedback. [Expert, 40 years, male]

Realization

Technical realization was assessed in a differentiated way by the participants. The *gamification* elements after each level were much appreciated. *Hardware* and *software* were seen as leaving room for future improvements, either by finding a solution without the annoying cable on the HMD or by delivering a more stable connection, as glitches and distortions were reported in some cases. However, none of the participants reported feeling motion sickness, which is a frequently felt consequence of VR interventions.

Suggestions for improvements were collected as well. They included ideas for more elaborated intensity levels that should be directly controllable by the therapist.

The participants recommended providing more opportunities for personalization features, not only delivered by the therapist, but also delivered for the patients themselves:

What I still find cool would-be customizability, no, that is, that you create a personal fear hierarchy and then maybe adjust the levels accordingly. [Expert, 30 years, female]

Discussion

Principal Findings

Although the effectiveness of VRE in the treatment of anxiety disorders is well studied and the comparability to in vivo trainings could be shown in numerous studies [9,14,18,67], to date, few studies have examined the experiences of patients in the case of VRE for claustrophobic symptoms. This study used a mixed methods approach to ask for the experiences and perspectives of patients and therapeutic experts testing a VRE app with different intensity levels for claustrophobia with the use of additional virtual humans within the environment. Our research intended to understand the inner processes of patients with anxiety during VRE sessions in order to define key elements necessary for an effective exposure setting. The perspective of therapeutic experts was added to understand potential facilitators and barriers, as well as professionals' view on the target group and treatment effects. In the long term, the results might serve to improve the design of comparable apps.

In our results, patients, who initially had the highest pretreatment anxiety, scored lower than the experts in the total presence score of the IPQ and in the subscore IPQ spatial presence but higher in IPQ involvement and IPQ experienced realism; however, the difference was only significant in the case of spatial presence. In addition, the qualitative interviews revealed that a considerable proportion of participants felt the discrepancy between reality and virtuality (eg, by feeling the cable in their back). Both groups gave high ratings in the evaluation of

feasibility and acceptability, and the scores were higher in the patient group.

The patients were asked to think aloud while carrying out the intervention, and their feelings of anxiety and tension support the assumption that the intervention successfully led to the desired effect. However, some symptoms were closely related to the presence of 1-3 virtual humans from the second intensity level on. These virtual humans were included to make the intervention more realistic and provide a further feeling of narrowness, but they may have evoked sociophobic or agoraphobic symptoms as well. Specific phobias and agoraphobia show a correlation of r=0.57 in the literature [68]; however, studies that report comorbidities specifically between claustrophobia and other anxiety-related disorders are missing. Further results could be derived from semistructured interviews with the participants. These results repeated the quantitative results, as feelings of presence and involvement were reported in both groups.

Recommendations for the methodology of VR clinical trials in health care were recently provided by an international working group. Birckhead et al [69] recommend 3 types of VR trials: VR1 studies for content development, VR2 studies for proof of concept, and VR3 studies for clinical evidence. Our study followed the rationale of VR2 studies. For those studies, the authors suggest investigating the following parameters: patient population, clinical setting, assessment of acceptability, feasibility and tolerability, and, finally, assessment of initial clinical efficacy. The latter was conceptualized as patient-reported outcomes, as objective clinical outcomes are recommended for randomized clinical trials (RCTs) only [69].

Following the criteria of Birckhead et al [69], the *patient population* was assessed by targeted recruiting, including SCID interviews, and supported by the results of the STAI-S, which exceeded the cut-off of 41 in the patient group [57]. Moreover, technology commitment was assessed in advance, and even though patients showed the lowest commitment, the 2 groups stayed comparable to the population sample of the scale developers [56]. Low levels of the total score of technology acceptance of patients might be explained by the lower score of technology competence conviction in comparison to the experts. This, in turn, might be due to the higher age of some of the patients and should be considered in future clinical use. The *clinical setting* was established as the intervention was tested within the facilities of the Heidelberg University Hospital and with the guidance of a trained psychologist. *Feasibility* and *acceptability* were assessed by evaluation items and by the semistructured interviews; the results indicate that both may be considered to have been reached. However, some experts remained skeptical, especially when they followed a psychodynamic therapy concept. Unlike cognitive behavioral therapists, psychodynamic approaches often aim at emotional experience rather than at habituating high levels of anxiety aroused by a situation [70]. *Tolerability* was assessed qualitatively, as the participants got an opportunity to report on their own experiences during the intervention. None of the participants had to stop the intervention due to feeling overwhelmed or motion sickness, which is a frequent limitation in the usage of VR apps [71]. The *initial clinical efficacy* was

supported by the qualitative statements and could be shown as patients reported personal success in their self-assessments. However, further sessions with time intervals between sessions would be necessary to prove efficacy.

Therapeutic Setting of VRE

An ambivalent topic was the question of whether the intervention is an exercise that can be conducted at home or whether the physical presence of a therapist is necessary. It has to be stated that therapeutic guidelines recommend that a therapist stay by the side of the patient [7]. On the contrary, there is growing evidence that a self-management app for VRE leads to symptom reduction and is at least not inferior to that with therapist guidance [45,72]. Hence, both kinds of apps are used in practice [73]. In Germany, there is a self-management app for anxiety with integrated VR training that is covered by health insurance [74]. However, therapists might miss the opportunity to interact with the patient during the intervention and control variables (eg choice of stimuli, duration, or intensity levels) [14].

In our results, the patients scored high on the question of whether the intervention should be accompanied by a therapist and at the same time they were also convinced that the intervention could be conducted alone at home. These ambivalent results may reflect the uncertainty toward the new technology. However, future studies are necessary to support our interpretation. With regard to in vivo treatment, we recommend that the first sessions be always conducted in the presence of a therapist who might then decide whether the patient is able to undergo the training alone at home.

In fact, the uptake of VR interventions in clinical practice remains hesitant [75]. Our results have shown that although most of our experts, regardless of their therapeutic background, showed positive attitudes toward the intervention, barriers might arise due ot costs and the need for a separate room. Another limiting factor seems to be a lack of evidence-based software that can be purchased and integrated into one's IT facilities. Either the software is available but not clinically tested, or the software is developed within a scientific infrastructure, well tested, but not commercially available [17].

Based on our results, we suggest the following key elements for a successful virtual exposure:

- Provide context variation by varying relevant factors regarding size, duration of a setting, and increasing darkness [10].
- Allow for systematic variation of the factors in order to provide an opportunity for individualized training.
- Add self-assessments within the treatment before and after a session regarding the respective anxiety level actually felt by the patient.
- Include virtual humans that should be personalized to the needs of the patients with respect to comorbidities (social phobia, traumatic experiences).
- Provide a safe therapeutic setting for exposure, as recommended by treatment guidelines [7].

Directions of Future Research

The potential of VRE might be enhanced by the possibility to add objective data to the personal feelings of the patient by integrating sensor data collected with a wearable wristband. These results can lead to personalized suggestions for the respective adequate intensity level. With the integration of real-time physiological data, a validation of the prerequisites of effective treatment might be possible, as former studies have shown that different scenarios in film, text, or VR induce different patterns of parasympathetic activation, with the lowest result in VR despite the highest self-reported presence [76]. This result is in line with the statement of Böhnlein et al [10], who found that the avoidance of relaxation is important for the success of VRE. Finally, future directions must meet the growing demands of personalized digital interventions. For example, the sex of the virtual humans as well as their age and culture-specific features should be tailored to the user [77].

Limitations

To the best of our knowledge, this is the first study to examine the experiences of patients with anxiety testing a VRE app for claustrophobia with virtual humans and made use of the think-aloud method with this target group. However, some limitations could not be avoided. First, claustrophobia is a seldom-given diagnosis, so we relied on self-reported symptoms of patients with a diagnosis of any anxiety disorder. Therefore, some experiences and feelings reported by the patients might be due to other fears and not just claustrophobia. As mentioned in the introduction, some authors already stated that claustrophobia can be understood as a prodromal stage of agoraphobia [43]. Second, we only measured state anxiety before starting the intervention, and therefore, an assessment of clinical efficacy was only conducted qualitatively and should be repeated with a longitudinal design, including more differentiated measures to assess the aspects of claustrophobia (eg, fear of suffocation and of confinement). Third, our anxiety assessment during the intervention was only implemented as a beta version. Future implementation will allow for data export for therapists to see the patients' progress. Fourth, although the think-aloud methodology is a renowned method to get in touch with the immediate feelings and emotions of a person, it might impede the process of immersion and presence. Future study designs need to find a way to resolve this paradox. Moreover, the fourth intensity level of our intervention included a virtual coughing woman wearing a protective mask. By this, we wanted to ensure a high degree of realism; however, it cannot be excluded that health-related fears may interfere with our results. Finally, the participants could choose an individual level of intensity. Future studies will investigate the exact effects of all levels with a higher number of participants in an experimental design with systematic variation.

Conclusion

A VRE app for claustrophobia with different intensity levels and with the presence of a gradually increasing number of virtual humans is feasible for inducing the desired degree of anxiety in patients in order to work with those fears during a therapy session. Key elements of a VRE app for claustrophobic symptoms should provide variation of intensity by adding

challenging cues in order to induce presence, which is a necessary state for inducing anxiety. Virtual humans can be included to make the intervention realistic and to provide a sense of closeness; however, some of the fears might then as well be caused by social phobia or agoraphobia. Patients may need the physical presence of a therapist, even if psychotherapists argue that the intervention might be conducted alone as well. More intensity levels are needed, with the option to adapt the intervention to a personalized symptom profile. By doing this, a more specific support might be provided.

Acknowledgments

Conflicts of Interest

None declared.

Multimedia Appendix 1
Interview questions.
[DOCX File , 19 KB - mental_v9i12e40056_app1.docx]

Multimedia Appendix 2
Results of the evaluation of patients and experts (means and SDs).
[DOCX File , 18 KB - mental_v9i12e40056_app2.docx]

References

1. Yang X, Fang Y, Chen H, Zhang T, Yin X, Man J, et al. Global, regional and national burden of anxiety disorders from 1990 to 2019: results from the Global Burden of Disease Study 2019. Epidemiol Psychiatr Sci 2021 May 06;30:e36-e36. [doi: 10.1017/s2045796021000275]

2. Wilner JG, Dufour S, Kennedy K, Sauer-Zavala S, Boettcher H, Farchione TJ, et al. Quality of life in heterogeneous anxiety disorders: changes across cognitive-behavioral treatments. Behav Modif 2020 May 07;44(3):343-364. [doi: 10.1177/0145445518815603] [Medline: 30525935]

3. Deady M, Collins D, Johnston D, Glozier N, Calvo R, Christensen H, et al. The impact of depression, anxiety and comorbidity on occupational outcomes. Occup Med (Lond) 2022 Jan 13;72(1):17-24. [doi: 10.1093/occmed/kqab142] [Medline: 34693972]

4. Solis EC, van Hemert AM, Carlier IV, Wardenaar KJ, Schoevers RA, Beekman AT, et al. The 9-year clinical course of depressive and anxiety disorders: new NESDA findings. J Affect Disord 2021 Dec 01;295:1269-1279 [FREE Full text] [doi: 10.1016/j.jad.2021.08.108] [Medline: 34706441]

5. National Institut for Health and Care Excellence (NICE). Generalised Anxiety Disorder and Panic Disorder in Adults: Management. 2011. URL: https://www.nice.org.uk/guidance/cg113 [accessed 2022-11-17]

6. Katzman M, Bleau P, Blier P, Chokka P, Kjernisted K, Van AM, et al. Canadian clinical practice guidelines for the management of anxiety, posttraumatic stress and obsessive-compulsive disorders. BMC Psychiatry 2014;14(Suppl 1):S1. [doi: 10.32920/14638581]

7. Bandelow B, Lichte T, Rudolf S, Wiltink J, Beutel M. S3-Leitlinie: Kurzfassung der Empfehlungen zur Behandlung von Angststörungen. In: S3-Leitlinie Angststörungen. Berlin: Springer; 2015:1-22.

8. Rosqvist J. Exposure Treatments for Anxiety Disorders : A Practitioner's Guide to Concepts, Methods, and Evidence-Based Practice. New York, NY: Routledge; 2005.

9. Meyerbröker K, Morina N. The use of virtual reality in assessment and treatment of anxiety and related disorders. Clin Psychol Psychother 2021 May 14;28(3):466-476 [FREE Full text] [doi: 10.1002/cpp.2623] [Medline: 34097318]

10. Böhnlein J, Altegoer L, Muck NK, Roesmann K, Redlich R, Dannlowski U, et al. Factors influencing the success of exposure therapy for specific phobia: a systematic review. Neurosci Biobehav Rev 2020 Jan;108:796-820. [doi: 10.1016/j.neubiorev.2019.12.009] [Medline: 31830494]

11. Riva G, Baños RM, Botella C, Mantovani F, Gaggioli A. Transforming experience: the potential of augmented reality and virtual reality for enhancing personal and clinical change. Front Psychiatry 2016 Sep 30;7:164 [FREE Full text] [doi: 10.3389/fpsyt.2016.00164] [Medline: 27746747]

12. Lognoul M, Nasello J, Triffaux JM. Virtual reality exposure therapy for post-traumatic stress disorders, obsessive-compulsive disorders and anxiety disorders: indications, added value and limitations. L'Encephale 2020;46(4):293-300.

XSL•FO
RenderX

13. van Loenen I, Scholten W, Muntingh A, Smit J, Batelaan N. The effectiveness of virtual reality exposure-based cognitive behavioral therapy for severe anxiety disorders, obsessive-compulsive disorder, and posttraumatic stress disorder: meta-analysis. J Med Internet Res 2022 Feb 10;24(2):e26736 [FREE Full text] [doi: 10.2196/26736] [Medline: 35142632]

14. Oing T, Prescott J. Implementations of virtual reality for anxiety-related disorders: systematic review. JMIR Serious Games 2018 Nov 07;6(4):e10965 [FREE Full text] [doi: 10.2196/10965] [Medline: 30404770]

15. Kahlon S, Lindner P, Nordgreen T. Virtual reality exposure therapy for adolescents with fear of public speaking: a non-randomized feasibility and pilot study. Child Adolesc Psychiatry Ment Health 2019 Dec 27;13(1):47 [FREE Full text] [doi: 10.1186/s13034-019-0307-y] [Medline: 31890004]

16. Premkumar P, Heym N, Brown DJ, Battersby S, Sumich A, Huntington B, et al. The effectiveness of self-guided virtual-reality exposure therapy for public-speaking anxiety. Front Psychiatry 2021 Aug 19;12:694610 [FREE Full text] [doi: 10.3389/fpsyt.2021.694610] [Medline: 34489755]

17. Bell IH, Nicholas J, Alvarez-Jimenez M, Thompson A, Valmaggia L. Virtual reality as a clinical tool in mental health research and practice. Dialogues Clin Neurosci 2022 Apr 01;22(2):169-177. [doi: 10.31887/dcns.2020.22.2/lvalmaggia]

18. Carl E, Stein AT, Levihn-Coon A, Pogue JR, Rothbaum B, Emmelkamp P, et al. Virtual reality exposure therapy for anxiety and related disorders: a meta-analysis of randomized controlled trials. J Anxiety Disord 2019 Jan;61:27-36. [doi: 10.1016/j.janxdis.2018.08.003] [Medline: 30287083]

19. Donnelly M, Reinberg R, Ito K, Saldana D, Neureither M, Schmiesing A, et al. Virtual reality for the treatment of anxiety disorders: a scoping review. Am J Occup Ther 2021 Nov 01;75(6):7506205040 [FREE Full text] [doi: 10.5014/ajot.2021.046169] [Medline: 34817595]

20. Wechsler TF, Kümpers F, Mühlberger A. Inferiority or even superiority of virtual reality exposure therapy in phobias: a systematic review and quantitative meta-analysis on randomized controlled trials specifically comparing the efficacy of virtual reality exposure to gold standard exposure in agoraphobia, specific phobia, and social phobia. Front Psychol 2019 Sep 10;10:1758 [FREE Full text] [doi: 10.3389/fpsyg.2019.01758] [Medline: 31551840]

21. Freitas JRS, Velosa VHS, Abreu LTN, Jardim RL, Santos JAV, Peres B, et al. Virtual reality exposure treatment in phobias: a systematic review. Psychiatr Q 2021 Dec 26;92(4):1685-1710. [doi: 10.1007/s11126-021-09935-6] [Medline: 34173160]

22. Fernández-Álvarez J, Rozental A, Carlbring P, Colombo D, Riva G, Anderson PL, et al. Deterioration rates in virtual reality therapy: an individual patient data level meta-analysis. J Anxiety Disord 2019 Jan;61:3-17. [doi: 10.1016/j.janxdis.2018.06.005] [Medline: 30057346]

23. Foa EB, Kozak MJ. Emotional processing of fear: exposure to corrective information. Psychol Bull 1986 Jan;99(1):20-35. [doi: 10.1037/0033-2909.99.1.20]

24. Hartanto D, Brinkman W, Kampmann I, Morina N, Emmelkamp P, Neerincx MA. Home-based virtual reality exposure therapy with virtual health agent support. In: Pervasive Computing Paradigms for Mental Health. Cham: Springer International; 2016.

25. Rothbaum BO, Hodges LF, Kooper R, Opdyke D, Williford JS, North M. Virtual reality graded exposure in the treatment of acrophobia: a case report. Behav Ther 1995;26(3):547-554. [doi: 10.1016/s0005-7894(05)80100-5]

26. Ling Y, Nefs HT, Morina N, Heynderickx I, Brinkman W. A meta-analysis on the relationship between self-reported presence and anxiety in virtual reality exposure therapy for anxiety disorders. PLoS One 2014 May 6;9(5):e96144 [FREE Full text] [doi: 10.1371/journal.pone.0096144] [Medline: 24801324]

27. Schubert TW. The sense of presence in virtual environments: a three-component scale measuring spatial presence, involvement, and realness. Zeitschrift für Medienpsychologie 2003 Apr;15(2):69-71. [doi: 10.1026//1617-6383.15.2.69]

28. Price M, Anderson P. The role of presence in virtual reality exposure therapy. J Anxiety Disord 2007 Jan;21(5):742-751 [FREE Full text] [doi: 10.1016/j.janxdis.2006.11.002] [Medline: 17145164]

29. Slater M, Sanchez-Vives MV. Enhancing our lives with immersive virtual reality. Front Robot AI 2016 Dec 19;3:74. [doi: 10.3389/frobt.2016.00074]

30. Powers MB, Briceno NF, Gresham R, Jouriles EN, Emmelkamp PM, Smits JA. Do conversations with virtual avatars increase feelings of social anxiety? J Anxiety Disord 2013 May;27(4):398-403. [doi: 10.1016/j.janxdis.2013.03.003] [Medline: 23746490]

31. Harris LM, Robinson J, Menzies RG. Evidence for fear of restriction and fear of suffocation as components of claustrophobia. Behav Res Ther 1999 Feb;37(2):155-159. [doi: 10.1016/s0005-7967(98)00110-7]

32. Wardenaar KJ, Lim CCW, Al-Hamzawi AO, Alonso J, Andrade LH, Benjet C, et al. The cross-national epidemiology of specific phobia in the World Mental Health Surveys. Psychol Med 2017 Feb 22;47(10):1744-1760. [doi: 10.1017/s0033291717000174]

33. Vadakkan C, Siddiqui W. Claustrophobia. Treasure Island, FL: StatPearls; 2021.

34. de Vries YA, Harris MG, Vigo D, Chiu WT, Sampson NA, Al-Hamzawi A, WHO World Mental Health Survey collaborators. Perceived helpfulness of treatment for specific phobia: findings from the World Mental Health Surveys. J Affect Disord 2021 Jun 01;288:199-209 [FREE Full text] [doi: 10.1016/j.jad.2021.04.001] [Medline: 33940429]

35. Hudson D, Heales C, Vine S. Radiographer Perspectives on current occurrence and management of claustrophobia in MRI. Radiography (Lond) 2022 Feb;28(1):154-161. [doi: 10.1016/j.radi.2021.09.008] [Medline: 34657799]

36.   Hudson D, Heales C, Vine S. Scoping review: ow is virtual reality being used as a tool to support the experience of undergoing magnetic resonance imaging? Radiography (Lond) 2022 Feb;28(1):199-207. [doi: 10.1016/j.radi.2021.07.008] [Medline: 34304960]

37.   Brown RK, Petty S, O'Malley S, Stojanovska J, Davenport MS, Kazerooni EA, et al. Virtual reality tool simulates MRI experience. Tomography 2018 Sep 01;4(3):95-98 [FREE Full text] [doi: 10.18383/j.tom.2018.00023] [Medline: 30320208]

38.   Kılıç A, Brown A, Aras I, Hui R, Hare J, Hughes L, et al. Using virtual technology for fear of medical procedures: a systematic review of the effectiveness of virtual reality-based interventions. Ann Behav Med 2021 Oct 27;55(11):1062-1079 [FREE Full text] [doi: 10.1093/abm/kaab016] [Medline: 33821879]

39.   Botella C, Baños R, Perpiñá C, Villa H, Alcañiz M, Rey A. Virtual reality treatment of claustrophobia: a case report. Behav Res Ther 1998 Feb;36(2):239-246. [doi: 10.1016/s0005-7967(97)10006-7]

40.   Botella C, Baños RM, Villa H, Perpiñá C, García-Palacios A. Virtual reality in the treatment of claustrophobic fear: a controlled, multiple-baseline design. Behav Ther 2000;31(3):583-595. [doi: 10.1016/s0005-7894(00)80032-5]

41.   Shiban Y, Peperkorn H, Alpers GW, Pauli P, Mühlberger A. Influence of perceptual cues and conceptual information on the activation and reduction of claustrophobic fear. J Behav Ther Exp Psychiatry 2016 Jun;51:19-26. [doi: 10.1016/j.jbtep.2015.11.002] [Medline: 26687921]

42.   Yeh S, Li Y, Zhou C, Chiu P, Chen J. Effects of virtual reality and augmented reality on induced anxiety. IEEE Trans Neural Syst Rehabil Eng 2018 Jul;26(7):1345-1352. [doi: 10.1109/tnsre.2018.2844083]

43.   Öst L, Alm T, Brandberg M, Breitholtz E. One vs five sessions of exposure and five sessions of cognitive therapy in the treatment of claustrophobia. Behav Res Ther 2001 Feb;39(2):167-183. [doi: 10.1016/s0005-7967(99)00176-x]

44.   Lindner P, Miloff A, Zetterlund E, Reuterskiöld L, Andersson G, Carlbring P. Attitudes toward and familiarity with virtual reality therapy among practicing cognitive behavior therapists: a cross-sectional survey study in the era of consumer VR platforms. Front Psychol 2019 Feb 8;10:176 [FREE Full text] [doi: 10.3389/fpsyg.2019.00176] [Medline: 30800086]

45.   Lindner P, Miloff A, Fagernäs S, Andersen J, Sigeman M, Andersson G, et al. Therapist-led and self-led one-session virtual reality exposure therapy for public speaking anxiety with consumer hardware and software: a randomized controlled trial. J Anxiety Disord 2019 Jan;61:45-54. [doi: 10.1016/j.janxdis.2018.07.003] [Medline: 30054173]

46.   Mayer G, Hummel S, Gronewold N, Oetjen N, Hilbel T, Schultz JH. Validity and Reliability of the Self-administered Psycho-TherApy-SystemS (SELFPASS) Item Pool for the Daily Mood Tracking of Depressive Symptoms: Cross-sectional Web-Based Survey. JMIR Ment Health 2021 Oct 18;8(10):e29615 [FREE Full text] [doi: 10.2196/29615] [Medline: 34661547]

47.   Mayer G, Hummel S, Oetjen N, Gronewold N, Bubolz S, Blankenhagel K, et al. User experience and acceptance of patients and healthy adults testing a personalized self-management app for depression: a non-randomized mixed-methods feasibility study. Digit Health 2022 Apr 07;8:20552076221091353 [FREE Full text] [doi: 10.1177/20552076221091353] [Medline: 35425641]

48.   Kossak HC. Verhaltenstherapie Methoden. In: Kognitiv-behaviorale Psychotherapie von Ängsten: Kurztherapie mit Hypnose - die Praxisanleitung. Berlin, Heidelberg: Springer; 2020:103-124.

49.   Makransky G, Lilleholt L. A structural equation modeling investigation of the emotional value of immersive virtual reality in education. Educ Tech Res Dev 2018 Feb 22;66(5):1141-1164. [doi: 10.1007/s11423-018-9581-2]

50.   Creswell J. A Concise Introduction to Mixed Methods Research. Thousand Oaks, CA: Sage; 2014.

51.   Fetters MD, Curry LA, Creswell JW. Achieving integration in mixed methods designs-principles and practices. Health Serv Res 2013 Dec 23;48(6 Pt 2):2134-2156 [FREE Full text] [doi: 10.1111/1475-6773.12117] [Medline: 24279835]

52.   Malterud K, Siersma VD, Guassora AD. Sample size in qualitative interview studies: guided by information power. Qual Health Res 2016 Nov 10;26(13):1753-1760. [doi: 10.1177/1049732315617444] [Medline: 26613970]

53.   Wittchen HU, Zaudig M, Fydrich T. Strukturiertes klinisches Interview für DSM-IV : SKID; eine deutschsprachige, erweiterte Bearbeitung der amerikanischen Originalversion des SCID. Göttingen: Hogrefe; Jan 1997.

54.   Neyer F, Felber J, Gebhardt C. Kurzskala zur Erfassung von Technikbereitschaft (technology commitment). Zusammenstellung sozialwissenschaftlicher Items und Skalen. In: Zusammenstellung sozialwissenschaftlicher Items und Skalen; 2016. URL: https://doi.org/10.6102/zis244 [accessed 2022-11-22]

55.   Laux L, Glanzmann P, Schaffner P, Spielberger CD. Das State-Trait-Angstinventar : STAI; theoretische Grundlagen und Handanweisung. Weinheim: Beltz Test; 1981.

56.   Neyer FJ, Felber J, Gebhardt C. Entwicklung und Validierung einer Kurzskala zur Erfassung von Technikbereitschaft. Diagnostica 2012 Apr;58(2):87-99. [doi: 10.1026/0012-1924/a000067]

57.   Ercan I, Hafizoglu S, Ozkaya G, Kirli S, Yalcintas E, Akaya C. Examining cut-off values for the State-Trait Anxiety Inventory/Examinando los puntajes de corte para el inventario de ansiedad estado-rasgo. Revista Argentina de Clinica Psicologica 2015 Aug;24(2):143-148.

58.   Schubert T, Friedmann F, Regenbrecht H. The experience of presence: factor analytic insights. Presence Teleoperators Virtual Environ 2001 Jun;10(3):266-281. [doi: 10.1162/105474601300343603]

59.   Regenbrecht H, Schubert T. Real and illusory interactions enhance presence in virtual environmentsillusory interactions enhance presence in virtual environments. Presence Teleoperators Virtual Environ 2002 Aug;11(4):425-434. [doi: 10.1162/105474602760204318]

60. Van SM, Barnard YF, Sandberg JJLA. The Think Aloud Method - A Practical Guide to Modelling CognitiveProcesses. Cambridge, MA: Academic Press; 1994.

61. Taber KS. The use of Cronbach's alpha when developing and reporting research instruments in science education. Res Sci Educ 2017 Jun 7;48(6):1273-1296. [doi: 10.1007/s11165-016-9602-2]

62. Ortuño-Sierra J, García-Velasco L, Inchausti F, Debbané M, Fonseca-Pedrero E. New approaches on the study of the psychometric properties of the STAI. Actas Esp Psiquiatr 2016 May;44(3):83-92 [FREE Full text] [Medline: 27254400]

63. Lee S, Lee DK. What is the proper way to apply the multiple comparison test? Korean J Anesthesiol 2020 Dec;73(6):572-572 [FREE Full text] [doi: 10.4097/kja.d.18.00242.e1] [Medline: 33271011]

64. MAXQDA, Software für qualitative Datenanalyse, 1989-2021. Berlin, Germany: VERBI Software. Consult. Sozialforschung GmbH; 2020. URL: https://www.maxqda.com/ [accessed 2022-11-21]

65. Mayring P. Einführung in die qualitative Sozialforschung : eine Anleitung zu qualitativem Denken, 6. Auflage edn. Weinheim; Basel: Beltz; 2016.

66. Vaismoradi M, Turunen H, Bondas T. Content analysis and thematic analysis: implications for conducting a qualitative descriptive study. Nurs Health Sci 2013 Sep 11;15(3):398-405. [doi: 10.1111/nhs.12048] [Medline: 23480423]

67. Botella C, Fernández-Álvarez J, Guillén V, García-Palacios A, Baños R. Recent progress in virtual reality exposure therapy for phobias: a systematic review. Curr Psychiatry Rep 2017 Jul 24;19(7):42. [doi: 10.1007/s11920-017-0788-4] [Medline: 28540594]

68. Bandelow B, Michaelis S. Epidemiology of anxiety disorders in the 21st century. Dialogues Clin Neurosci 2022 Apr 01;17(3):327-335. [doi: 10.31887/dcns.2015.17.3/bbandelow]

69. 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]

70. Beutel ME, Greenberg L, Lane RD, Subic-Wrana C. Treating anxiety disorders by emotion-focused psychodynamic psychotherapy (EFPP): an integrative, transdiagnostic approach. Clin Psychol Psychother 2019 Jan 25;26(1):1-13. [doi: 10.1002/cpp.2325] [Medline: 30255535]

71. Park MJ, 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 Jul 19;10:505 [FREE Full text] [doi: 10.3389/fpsyt.2019.00505] [Medline: 31379623]

72. Shin B, Oh J, Kim B, Kim HE, Kim H, Kim S, et al. Effectiveness of self-guided virtual reality-based cognitive behavioral therapy for panic disorder: randomized controlled trial. JMIR Ment Health 2021 Nov 22;8(11):e30590 [FREE Full text] [doi: 10.2196/30590] [Medline: 34813486]

73. Korbel J, Meywirth S, Zarnekow R. A taxonomy of virtual reality applications for the treatment of anxiety disorders. 2021 Presented at: 29th European Conference on Information Systems (ECIS); 2021; Marrakech.

74. Angern C. German Startup Awards: Sympatient-Gründer Julian Angern als "Newcomer des Jahres" ausgezeichnet.: Invirto; 2021. URL: https://www.presseportal.de/pm/151123/4906904 [accessed 2022-11-17]

75. Meyerbröker K. Virtual reality in clinical practice. Clin Psychol Psychother 2021 May 14;28(3):463-465 [FREE Full text] [doi: 10.1002/cpp.2616] [Medline: 34048628]

76. Gilpin G, Gain J, Lipinska G. The physiological signature of sadness: a comparison between text, film and virtual reality. Brain Cogn 2021 Aug;152:105734. [doi: 10.1016/j.bandc.2021.105734] [Medline: 34023614]

77. Anderson PL, Molloy A. Maximizing the impact of virtual reality exposure therapy for anxiety disorders. Curr Opin Psychol 2020 Dec;36:153-157. [doi: 10.1016/j.copsyc.2020.10.001] [Medline: 33176268]

Abbreviations

CBT: cognitive behavioral therapy
ET: exposure therapy
HMD: head-mounted display
IPQ: Igroup Presence Questionnaire
SCID: Structured Clinical Interview for DSM-5
STAI-S: State-Anxiety Scale of the State-Trait Anxiety Inventory
TCS: technology commitment scale
VR: virtual reality
VRE: virtual reality exposure

XSL•FO

RenderX

Review

Meeting the Unmet Needs of Individuals With Mental Disorders: Scoping Review on Peer-to-Peer Web-Based Interactions

Dawid Storman[1], MD; Paweł Jemioło[2], MSc, MA; Mateusz Jan Swierz[1], MD; Zuzanna Sawiec[3]; Ewa Antonowicz[3]; Anna Prokop-Dorner[4], PhD; Marcelina Gotfryd-Burzyńska[5], MA; Malgorzata M Bala[1], MD, PhD, Prof Dr

[1]Chair of Epidemiology and Preventive Medicine, Department of Hygiene and Dietetics, Jagiellonian University Medical College, Krakow, Poland

[2]AGH University of Science and Technology, Krakow, Poland

[3]Students' Scientific Research Group of Systematic Reviews, Jagiellonian University Medical College, Krakow, Poland

[4]Chair of Epidemiology and Preventive Medicine, Jagiellonian University Medical College, Krakow, Poland

[5]Institute of Psychology, Jagiellonian University, Krakow, Poland

Corresponding Author:
Dawid Storman, MD
Chair of Epidemiology and Preventive Medicine
Department of Hygiene and Dietetics
Jagiellonian University Medical College
Kopernika 7a
Krakow, 31-034
Poland
Phone: 48 502846363
Email: dawid.storman@doctoral.uj.edu.pl

Abstract

Background: An increasing number of online support groups are providing advice and information on topics related to mental health.

Objective: This study aimed to investigate the needs that internet users meet through peer-to-peer interactions.

Methods: A search of 4 databases was performed until August 15, 2022. Qualitative or mixed methods (ie, qualitative and quantitative) studies investigating interactions among internet users with mental disorders were included. The ϕ coefficient was used and machine learning techniques were applied to investigate the associations between the type of mental disorders and web-based interactions linked to seeking help or support.

Results: Of the 13,098 identified records, 44 studies (analyzed in 54 study-disorder pairs) that assessed 82,091 users and 293,103 posts were included. The most frequent interactions were noted for people with eating disorders (14/54, 26%), depression (12/54, 22%), and psychoactive substance use disorders (9/54, 17%). We grouped interactions between users into 42 codes, with the *empathy or compassion* code being the most common (41/54, 76%). The most frequently coexisting codes were *request for information* and *network* (35 times; $\phi=0.5$; $P<.001$). The algorithms that provided the best accuracy in classifying disorders by interactions were decision trees (44/54, 81%) and logistic regression (40/54, 74%). The included studies were of moderate quality.

Conclusions: People with mental disorders mostly use the internet to seek support, find answers to their questions, and chat. The results of this analysis should be interpreted as a proof of concept. More data on web-based interactions among these people might help apply machine learning methods to develop a tool that might facilitate screening or even support mental health assessment.

XSL•FO
RenderX

Introduction

Background

It is estimated that 38.2% of Europeans and 26.2% of Americans experience mental disorders annually [1,2]. Unfortunately, social perception of these disorders is largely based on stereotypes [3]. Despite antistigma campaigns [4], stigmatization and discriminatory practices are reinforced by media discourse that reproduces false and simplified mental representations of people with mental disorders [5-8]. Therefore, many individuals seek information or support on the web. The internet is an essential platform for creating web-based communities that provide a venue to ask questions, share experiences, and offer mutual emotional support [9]. Most studies reporting evidence that websites can provide meaningful help focused on people with physical disorders such as cancer [10], diabetes mellitus [11], and Alzheimer disease [12]. However, research suggests that more attention should be paid to people with psychological conditions, including those who self-harm [13], those who experience eating disorders [14], and those with various other mental disorders [15] because these conditions affect different aspects of daily functioning. According to the literature, people with mental disorders are willing to connect with others using social media [9] even though they have greater difficulties in establishing relationships offline than people without such disorders [16]. This trend was further reinforced during the COVID-19 pandemic when access to face-to-face professional help became limited and was replaced by remote support services [17]. At the same time, the number of internet users grew from 4.1 billion in 2019 to 4.9 billion in 2021 [18], which means that a higher number of people could benefit from our research.

According to several studies, both people who generate content and those who interact with creators may benefit from such an interaction [19-21]. Unlike spontaneous offline meetings, web-based interactions do not require the same level of engagement or instant reactions. Thus, this type of interactions may help people with mental disorders overcome increased levels of social anxiety or face information-processing challenges [22]. This, in turn, may provide a sense of empowerment and lead to shorter recovery times. In addition, the internet can offer anonymity, making web-based interactions with strangers less threatening than in-person contact [23].

Self-esteem is built on several key factors, one of which is a sense of belonging to a group [24]. Therefore, connecting with similar individuals (peers) may result in better recovery and social integration among people with mental disorders [25]. However, stigmatization and rejection can happen even within the communities themselves [26] but also in web-based interactions. Thus, it is critical for internet users with mental disorders to join the right web-based groups to avoid rejection from their peers. Unfortunately, there are still an insufficient number of mental health professionals who can provide necessary assistance within a web-based community. Therefore, internet users become organized into self-help groups. Available evidence demonstrates that web-based interactions between peers have enormous potential to help bridge the gap between the identified need for services and the limited resources available for conventional treatment [27].

A peer is defined as a person who has the same social position or abilities as other members of a group [28]. There are several types of peer relationships, such as (1) between a peer and another individual (dyad), (2) between a peer and a group, and (3) a hybrid of both types [29,30]. Furthermore, the types of peer-to-peer interactions are heterogeneous and may include mutual support or participation in consumer assistance or peer-run programs [25,31]. Some of these can occur web-based via different platforms available, such as support groups, forums, discussion groups, bulletin boards, social media, and chats [32].

Peer-to-peer interactions allow people to share experiences, exchange information, and provide advice and emotional support in a natural and spontaneous manner. Therefore, they constitute an exciting subject of research. There is evidence showing that relationships between peers promote behavioral changes [33], improve coping strategies [34], and alleviate social isolation and loneliness among people with mental disorders [35-37]. For many people, social networking on the internet is the major form of communication that facilitates social interactions [38]. This is especially true for individuals who experience difficulties in direct contact with others because of stigmatization [39]. Barak et al [40] reported lower levels of emotional distress among adolescents when they were involved in a web-based forum. However, peer-to-peer interactions on the internet may also negatively affect mental health. Generally, internet use raises concerns, such as user behavior control, accurate risk assessment, privacy, and confidentiality [41].

Currently, new technologies are being developed for people with mental disorders, including artificial intelligence (AI) that already plays a major role in general medicine and research [42-44]. Techniques based on AI are widely applied in medical imaging diagnostics [45-47], but they can also be used for personalization purposes [48-50]. By identifying patterns in the types of interactions linked to specific types of disorders, these techniques could help individualize interventions provided by moderators of web-based forums. AI might also serve as a supporting tool in situations where there are no forum administrators (eg, owing to high costs). It can tailor the content to individual needs and concerns of the users.

Objectives

So far, studies assessing peer-to-peer interactions, including systematic reviews [51-53], have focused on the efficacy of such interactions. However, studies that summarize qualitative research are lacking. To fill this gap, we conducted a scoping review that addressed the following research questions:

1. What are the needs that individuals with mental disorders fulfill through web-based peer-to-peer interactions?
2. What are the categories of peer-to-peer interactions and how can they be used in further research?
3. Is it possible to use machine learning (ML) techniques to assess and classify mental disorders based on the types of peer-to-peer interactions?

In our opinion, heterogeneous and multidimensional data can be best handled using ML techniques (or even deep learning if

XSL•FO

RenderX

sufficient data are available). Therefore, the aim of this proof-of-concept study was to explore the potential of ML in such an analysis.

Methods

The study was conducted in accordance with the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews; Multimedia Appendix 1 [54,55]). The study protocol was registered in the Open Science Framework (added on August 24, 2020, and registered on November 11, 2020 [56]).

Eligibility Criteria

For this analysis, we considered studies performed according to qualitative or mixed (ie, qualitative and quantitative) methodology that evaluated the following: (1) web-based interactions between participants with any mental disorder that is defined according to any standard diagnostic criteria and (2) interactions between a peer and another individual (dyad) [57]. Studies that assessed only family members or caregivers of people with mental disorders were excluded. No language or date restrictions were applied. In addition, the eligibility criteria were not limited to a specific location, publication status, or any other characteristic.

Search Strategy

We searched 4 electronic databases (Ovid MEDLINE, Embase, Cochrane Library, and Web of Science) until August 15, 2022. The search was performed without any restrictions on the language or publication date of the studies. All search strategies are available in Multimedia Appendix 2. For additional papers, we manually searched the references of reviews that were obtained through the search.

Study Selection

To identify eligible studies, titles, abstracts, and full texts were individually assessed by any 2 of the 5 reviewers (DS, PJ, MJS, MG, and APD). Conflicts were resolved by discussion or involvement of a third reviewer (DS or MMB).

Data Charting

Data charting was performed independently by 2 of the 5 authors (DS, PJ, MJS, ZS, and EA). Disagreements were resolved by consensus or arbitration by a third reviewer (DS). All relevant data on research characteristics (eg, study design, country of origin, and funder), methodology (eg, type of coding and coding scheme), participants (eg, age, gender, and type of mental disorder), and results (interactions) were extracted since November 8, 2020.

Credibility Assessment

The study quality was assessed by 2 of the 4 independent reviewers (MJS, EA, ZS, and PJ) using the Critical Appraisal Skills Programme (CASP) checklist for qualitative research [58]. The tool included 10 questions about study validity, study results, and whether the results helped locally. They could be answered by selecting *yes*, *no*, or *can't tell*. We divided the final question (*How valuable is the research?*) into 3 subquestions according to the hints provided in the manual: (1) input into existing knowledge (10a); (2) identification of unexplored areas (10b); and (3) external validity of the findings (10c). These 3 criteria were scored as 0 (not fully met) or 1 (fully met). The general assessment of the study quality was based on the sum of the scores from the 3 subquestions. A score of 3 indicated a valuable study; 2, a moderately valuable study; 1, a study of some quality; and 0, a study of no quality. Any disagreements were resolved by the involvement of a third independent reviewer (DS).

Synthesis of Results

The essential data on the population and methodology of the included studies were summarized in tabular and descriptive forms. All types of interactions observed in the studies were grouped into several categories (codes), which were defined based on the previous literature. To describe the categories and the links between them, several models were used (both originally developed and derived from the literature). The models were created during the discussion between the coauthors (DS, PJ, and APD), and they were presented as partition trees (Multimedia Appendix 3). The models were evaluated based on the lowest SD value of the number of codes in the category, which was the most common measure of the dispersion of results [59]. The frequency of codes as well as the co-occurrence of codes and diseases were presented using heatmaps (means and sums) and a circular chart (co-occurrence frequencies) to investigate possible associations between interactions and specific disorders.

All graphs were prepared using Python 3.7.10 (Python Software Foundation) libraries: Matplotlib 3.2.2 (John Hunter), Seaborn 0.11.1 (Michael Waskom), NetworkX 2.5.1 (Aric Hagberg, Dan Schult and Pieter Swart), Graphviz 2.47.1 (John Ellson), VOSviewer 1.6.6 (Nees Jan van Eck and Ludo Waltman), or Microsoft Office 2004 (Microsoft Corp). The source code is available on GitHub.

Statistical Analysis

We used the ϕ coefficient [60] to examine the associations between the types of interactions and mental disorders. Using Pandas 1.1.5 (Wes McKinney) and NumPy 1.19.5 (Travis Oliphant), we represented the data as a data frame and then used SciPy 1.4.1 (Travis Oliphant, Pearu Peterson, and Eric Jones) to calculate associations and their statistical significance.

Associations within the following subgroups were evaluated: (1) type of disorder, (2) studies assessed as valuable versus other studies, and (3) types of disorder using only valuable (high-quality) studies. ML techniques were applied to classify mental disorders based on interactions between users. For this purpose, several basic algorithms were used. These algorithms were selected based on their strong mathematical background and resultant explainability properties, as we were interested in identifying the variables that contributed to performance [61]. More specifically, we incorporated decision trees (with minimum samples per leaf ranging from 1 to 3), logistic regression (with L2 regularization), support vector machines (with the radial basis function kernel), k-nearest neighbors algorithm (with k ranging from 2 to 5), and Gaussian naïve

Bayes classifier (default settings). For this analysis, Scikit-learn (version 0.22.2) was used.

Mapping the Terms

To examine the relations between the terms as well as construct and visualize bibliometric networks, we used the mapping software VOSviewer (version 1.6.16) [62]. We aimed to investigate the co-occurrence networks of important terms extracted from the full text of the included studies. Thus, we provided a visualization. The distance between any pair of objects reflects their similarity as accurately as possible. Objects with high similarity are located close to each other, whereas objects with low similarity are located far from each other. We created a co-occurrence map by applying the default counting method and choosing number 5 as the minimum number of occurrences considering the most advantageous setting in terms of resources, time, and data received [62]. A total of 2 independent reviewers (DS and PJ) screened the list of terms extracted from VOSviewer and selected those that described

the interactions. Any discrepancies were resolved through discussion. The final terms were used to create visualizations. In addition, we compared the terms selected from VOSviewer with the codes from our codebook and calculated the percentage of overlap.

Results

Overview

The search identified 13,098 original references, and the screening of titles and abstracts yielded 86 full-text papers. A total of 44 studies were included in the final analysis and 8 were labeled as ongoing (Multimedia Appendix 4 [63-105] and Multimedia Appendix 5 [104-111]). The study flow is presented as a PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram in Figure 1 [112]. A list of excluded studies with the reason for exclusion is provided in Multimedia Appendix 6 [113-146].

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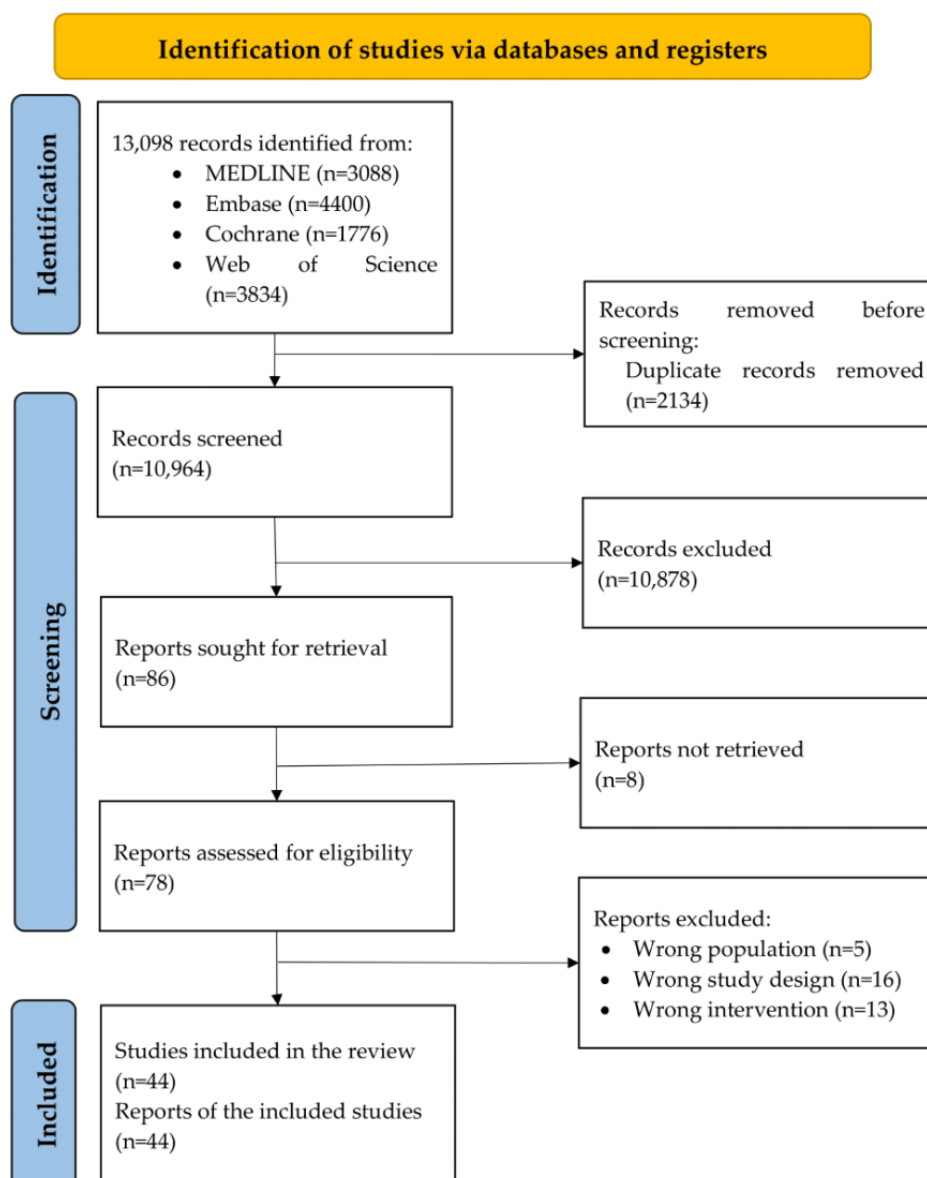
Included Studies

The detailed characteristics of the 44 included studies (analyzed in 54 study-disorder pairs) are presented in Table 1 and Multimedia Appendix 7 [63-105]. These studies were conducted between the years 2000 and 2022. Among the corresponding authors, 82% (36/44) originated from English-speaking countries, including the United States (16/44, 36%), the United Kingdom (9/44, 20%), Canada (6/44, 14%), Australia (2/44, 5%), Ireland (1/44, 2%), Singapore (1/44, 2%), and Hong Kong (1/44, 2%); whereas 18% (8/44) originated from non–English-speaking countries, including Sweden (3/44, 7%), Israel (2/44, 5%), Switzerland (1/44, 2%), Italy (1/44, 2%), and Hungary (1/44, 2%). None of the included studies provided information on whether the study protocol was registered in an appropriate registry.

Table 1. Characteristics of the included studies (N=44).

Variable	Value, n (%)[a]
Place of interaction	
Forum	26 (59)
Media (Facebook, Instagram, etc)	12 (27)
Support group	7 (16)
Blog	1 (2)
Chat	1 (2)
Access to the place of interaction	
Public access	25 (57)
Registration	4 (9)
Partial access (need to register to add comments)	6 (14)
Not reported	10 (23)
Type of analysis used	
Content analysis	23 (52)
Thematic analysis	13 (30)
Discourse analysis	5 (11)
Constant comparison	2 (5)
Conversational analysis	2 (5)
Other	13 (30)
Not reported	1 (2)
Coding schemes for social support and interaction	
Cutrona and Suhr	5 (11)
The self-reported coding scheme	3 (7)
Other	7 (16)
Not reported	33 (75)

[a]Some studies could be included in >1 subgroup.

Participants

The studies included 82,091 users (mean 3284, SD 9526; range 11-41,967) who posted 293,106 comments (mean 8374, SD 25,105; range 4-132,599) on 19,940 topics (mean 1173, SD 2916; range 5-10,169). The age of the participants ranged from 16 to 78 years, although most studies (37/44, 84%) did not report data on age. The proportion of women ranged from 0% to 100% (mean 65%, SD 34%); however, most studies (31/44, 70%) did not provide this information. The percentage of inactive to passive forum users was 47.62% (range 17%-85.47%). In 18% (8/44) of studies, participants were recruited to register on a forum created by the authors themselves.

There were 22 different mental disorders classified into 13 categories. Most commonly, studies have assessed peer-to-peer interactions among individuals with eating disorders (14/54, 26%), depression (12/54, 22%), and psychoactive substance use disorders (9/54, 17%). Next, there were the following disorders: postpartum depression (4/54, 7%), anxiety disorders (3/54, 6%), posttraumatic stress disorder (2/54, 4%). The remain disorders occurred only once (1/54, 2%): attention-deficit/hyperactivity disorder, bipolar affective disorder, mild cognitive impairment, obsessive compulsive disorder, schizoaffective disorder, and schizophrenia.

Type of Platform for Interaction

Of the 44 platforms, 11 (25%) platforms were developed specifically for the mental health setting, 3 (7%) platforms were intended for more general health, and for 30 (68%) platforms, this information was not reported. Most studies assessed web-based forum interactions (26/44, 59%). In most cases, access to the place of interaction was free and registration was not required to add comments (25/44, 57%). However, in some cases, registration was required to add and read comments (4/44, 9%). However, in some other cases, forums offered partially free access, with registration required to add, but not read, comments (6/44, 14%).

Most studies reported the presence of moderators (22/44, 50%). Their roles included the following: (1) provision of advice or therapy, (2) monitoring and control of content (sensitive, legal, sexual, abuse, and eliminate spam), or (3) moderation of discussions.

Anonymity was ensured in most of the studies (35/44, 80%). The authors paraphrased participants' statements and comments, did not record any personal data, excluded nicknames from the analysis, replaced nicknames with initials, or did not include quotations in the text. Membership terms of use were specified in 36% (16/44) of studies. For example, by accepting the terms and conditions, users agreed to treat other members with respect, provide support, avoid profanity and unhelpful language, avoid detailed and vivid descriptions of self-harming techniques, not offer drugs to other members, and provide links to sites selling drugs.

Methods in the Included Studies

Of the 44 studies, 17 (39%) studies used only qualitative methodology, whereas 19 (43%) studies also used a frequency analysis. Mixed methods (ie, qualitative and quantitative) were applied in 18% (8/44) of studies.

Content and thematic analyses were the most common (23/44, 52% and 13/44, 30%, respectively). Other analyses included membership categorization, ethnomethodological, netnographic, rhetorical, framework, interpretative phenomenological, image, sequence, paths, and social networks, each applied in a single study.

A low level of precision regarding reporting on methodological approaches impeded comparisons of the analytic strategies used by the authors. Coding performed by 2 people was reported in 52% (23/44) of studies. In 27% (12/44) of studies, complete coding was performed independently, in 16% (7/44) of studies, it was performed only during the calibration process, and in 9% (4/44) of studies, one author coded part of the material and the other author checked that coding. Coding was applied inductively in 43% (19/44) of studies and deductively in 14% (6/44) of studies. In 11% (5/44) of studies, both approaches were used, and the remaining 32% (14/44) of studies did not report the coding strategy. A codebook was developed openly (without blinding) in 18% (8/44) of studies. Blinding was established in 2% (1/44) of papers. The coding scheme for social support designed by Cutrona and Suhr [147] was used in 11% (5/44) of studies, whereas 16% (7/44) of studies adopted

different approaches proposed by Cohen and Wills [148], Oakley [149], Tong et al [150], Bauer et al [151], Morse and Field [152], Gaysynsky et al [153], and Bales [57]. In 7% (3/44) of studies, the authors used their own system [63,154,155].

To determine interrater agreement, Cohen κ was used in 14% (6/44) of studies, Krippendorff α was used in 2% (1/44), and data were not reported in the remaining 84% (37/44) of papers. In 4 studies, the diagnosis of participants was confirmed using the Center for Epidemiologic Studies Depression Scale or by a specialist.

Types of Interactions

We distinguished 42 codes that described peer-to-peer interactions. The codes were organized into 15 categories from A to O (Table 2), and we proposed 14 different models of coding interventions among peers (Multimedia Appendix 3). Six models were based on existing models: 1 [147], 3 [156], 4 [153], 5 (adapted from Liu et al [157]), 6 [158], and 7 [57]. A total of 2 models were modified based on the models by Cutrona and Suhr [147] (model 2) and Greiner et al [64] (model 8). The remaining 6 models were developed by us (models 9, 10, 11, 12, 13, and 14). After calculating the means and SDs, the models were ranked based on the lowest SD (Multimedia Appendix 8). Model 14 was characterized by the lowest SD (9.32), and the corresponding tree is presented in Figure 2.

The most frequent interactions were *empathy or compassion* (41/54, 76%), *network* (40/54, 74%), and *sharing self-disclosure* (39/54, 72%). Heatmaps of the selected codes and disorders and their co-occurrence are presented in Figure 3 and Multimedia Appendices 9 and 10. We visualized the normalized means of code co-occurrence across disorders and all included studies. However, heatmaps should be interpreted with caution because of the unequal number of papers regarding individual disorders, which resulted in certain codes being used significantly more often.

The co-occurrence of all codes is shown in Figure 4. Co-occurrence was observed most often for *request for information* and *network* (35 times). There was a positive association between these 2 codes ($\phi=0.5$; $P<.001$). The strongest positive association was noted between *requesting engagement* and *disagreement*, *relationship* and *confidentiality*, and *referring to the rules* and *rejection* ($\phi=0.65$; $P<.001$). As for correlations between codes and disorders, the strongest correlation was observed between attention-deficit/hyperactivity disorder and *illegal advice* ($\phi=0.70$; $P<.001$). The remaining associations for the overall and subgroup analyses are presented in Multimedia Appendix 10.

We achieved the highest accuracy in classifying disorders by interactions using 2 methods: decision trees (44/54, 81%) and logistic regression (40/54, 74%). The confusion matrices (with absolute values and relative percentages) of the ML techniques with detailed results are presented in Multimedia Appendix 11. In addition, using decision trees, we visualized the possible pathways to identify mental disorders (Multimedia Appendix 12).

Table 2. Codebook.

Node	Codes	Meaning
A1	Referral	• Referring the recipient to other sources of information or help, other places in general, and nonprofessional • Providing the recipient with access to new people or other communication channels
A2	Request for opinions or suggestions	• Asking about any act that offers direction or action for how to engage in the task or advances a belief or the value that is relevant to the task
A3	Situation appraisal	• Helping reassess or redefine the situation faced by the recipient
B1	Positive	• Showing positive emotions
B2	Negative	• Showing negative emotions
C1	Sharing self-disclosure	• Speaking about oneself, one's experience, and one's disease (recovery reports, treatment, diagnosis, etc)
D1	Sarcastic comments	• Being disrespectful, insolent toward other members or statements that express being hurt
D2	Aggression	• Presenting hostile or violent attitudes toward another with or without readiness to attack or confront
D3	Disagreement	• Expressing a different opinion
D4	Rejection	• Expressing little desire to include a person in their groups and relationships or excluding a person
D5	Reluctance or aversion	• Expressing a strong dislike or disinclination
E1	Encouragement or motivation	• Providing the recipient with a motive for doing something and confidence
E2	Compliment	• Improving one's self-worth by saying positive things about the recipient
F1	Practical tricks	• Sharing advice (not necessarily based on facts and can be based on self-experience) • Providing ideas or suggestions for action
F2	Instrumental	• Offering help or a talk
F3	Tangible	• Sharing goods or services
G1	Appreciation or gratitude	• Expressing appreciation to another individual from the group or the group all in all
H1	Requesting engagement	• Asking for opportunity to participate or be involved in group's life
H2	Request for other kinds of support	• Asking about anything other than facts, opinions, or suggestions
H3	Small talks or socializing	• Greetings • Taking politely about unimportant or uncontroversial matters
H4	Encouraging disclosure	• Motivating to expose oneself, revealing information about oneself
I1	Informational	• Sharing information or theoretical knowledge (should be based on facts)
I2	Referring to the rules	• Mentioning and enforcing the applicable group norms and rules
I3	Illegal advice	• Mostly related to drugs—providing information about where one can buy drugs and how to deal with getting a prescription from a physician
I4	Warnings	• Indicating a possible danger, problem, or other unpleasant situation
J1	Request for information	• Asking questions to obtain an answer about facts

Node	Codes	Meaning
J2	Clarifications	• Asking to make a statement or situation less confusing and more comprehensible (eg, asking for explanation or asking additional questions)
J3	Verifying the authenticity	• Asking about proofs (eg, a code of diagnosis)
K1	Related to professional help	• Providing information about places where one can obtain help from specialists
K2	Related to medication	• Providing information about drugs, doses, and route of administration
K3	Related to side effects	• Providing information about an unpleasant effect of a drug that occurs in addition to the main effect
L1	Presence or companions	• Offering to be there
L2	Offering hope	• Providing the recipient with hope
L3	Spiritual	• Offering prayer for the recipient
L4	Tension Release or jokes	• Posting messages that include humor • Reducing the anxiety that a person or a group may be experiencing
M1	Empathy or compassion	• Showing that their feelings are seen
N1	Apologizing	• Expressing regret for doing something wrong
N2	Confidentiality	• Keeping the recipient's problem in confidence
N3	Behavior promotion	• Supportive of harmful behavior • Supportive of minimizing harmful behavior • Unsupportive of harmful behavior
N4	Acceptance	• Being received and admitted into a group
N5	Relationship	• Conveying the importance of closeness
O1	Network	• Providing agreement with the views of the recipient • Providing validation, normalizing the situation • Showing the problem or situation as affecting more people, helping in identification, solidarity, and group cohesion • Alleviating any feelings of guilt that the recipient may have about the situation

XSL•FO
RenderX

Figure 2. Categorization of peer-to-peer interactions.
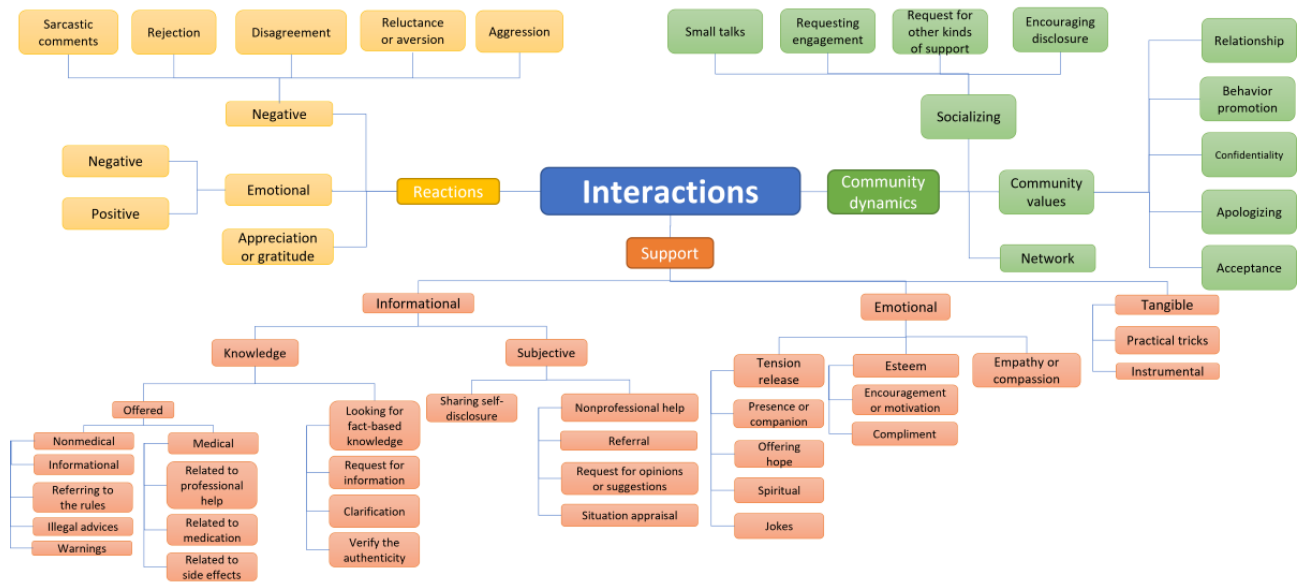
Figure 3. Heatmap with normalized means of code co-occurrence among mental disorders. ADHD: attention-deficit/hyperactivity disorder; AXD: anxiety disorders; BD: bipolar affective disorder; DEP: depression; ED: eating disorders; MCI: mild cognitive impairment; OCD: obsessive compulsive disorder; PPD: postpartum depression; PSU: psychoactive substance use; PTSD: posttraumatic stress disorder; SCZ: schizophrenia; SZA: schizoaffective disorder.
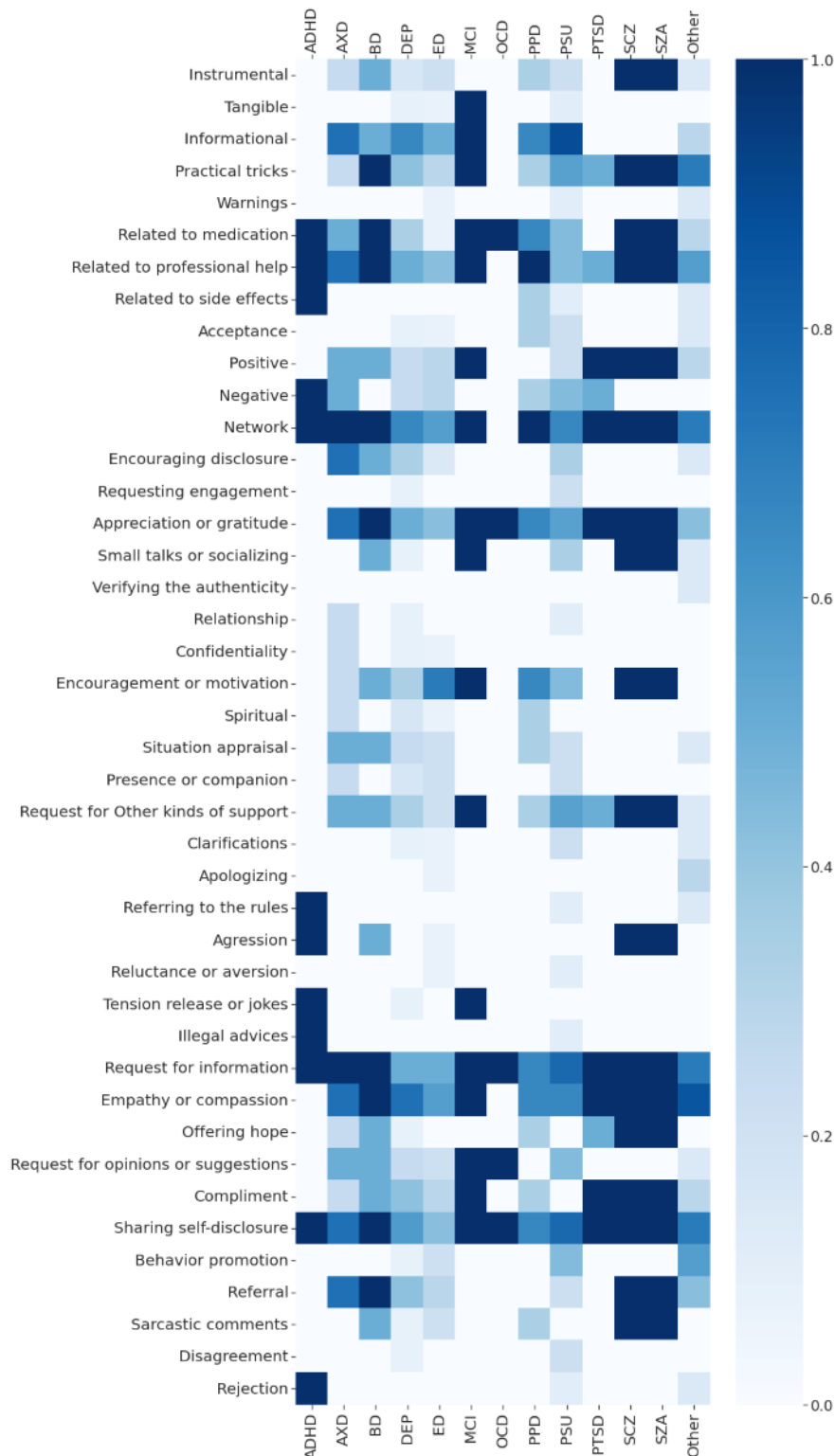
Figure 4. Circular chart of the co-occurrence of codes. The nodes represent specific types of interaction (Table 2). Node size corresponds to the number of primary studies that mention this type of interaction. The edges of the graphs indicate co-occurrence with other codes, and their size is proportional to the number of co-occurrences. A1: referral; A2: request for opinions or suggestions; A3: situation appraisal; B1: positive; B2: negative; C1: sharing self-disclosure; D1: sarcastic comments; D2: aggression; D3: disagreement; D4: rejection; D5: reluctance or aversion; E1: encouragement or motivation; E2: compliment; F1: practical tricks; F2: instrumental; F3: tangible; G1: appreciation or gratitude; H1: requesting engagement; H2: request for other kinds of support; H3: small talks or socializing; H4: encouraging disclosure; I1: informational; I2: referring to the rules; I3: illegal advice; I4: warnings; J1: request for information; J2: clarifications; J3: verifying the authenticity; K1: related to professional help; K2: related to medication; K3: related to side effects; L1: presence or companions; L2: offering hope; L3: spiritual; L4: tension release or jokes; M1: empathy or compassion; N1: apologizing; N2: confidentiality; N3: behavior promotion; N4: acceptance; N5: relationship; O1: network.
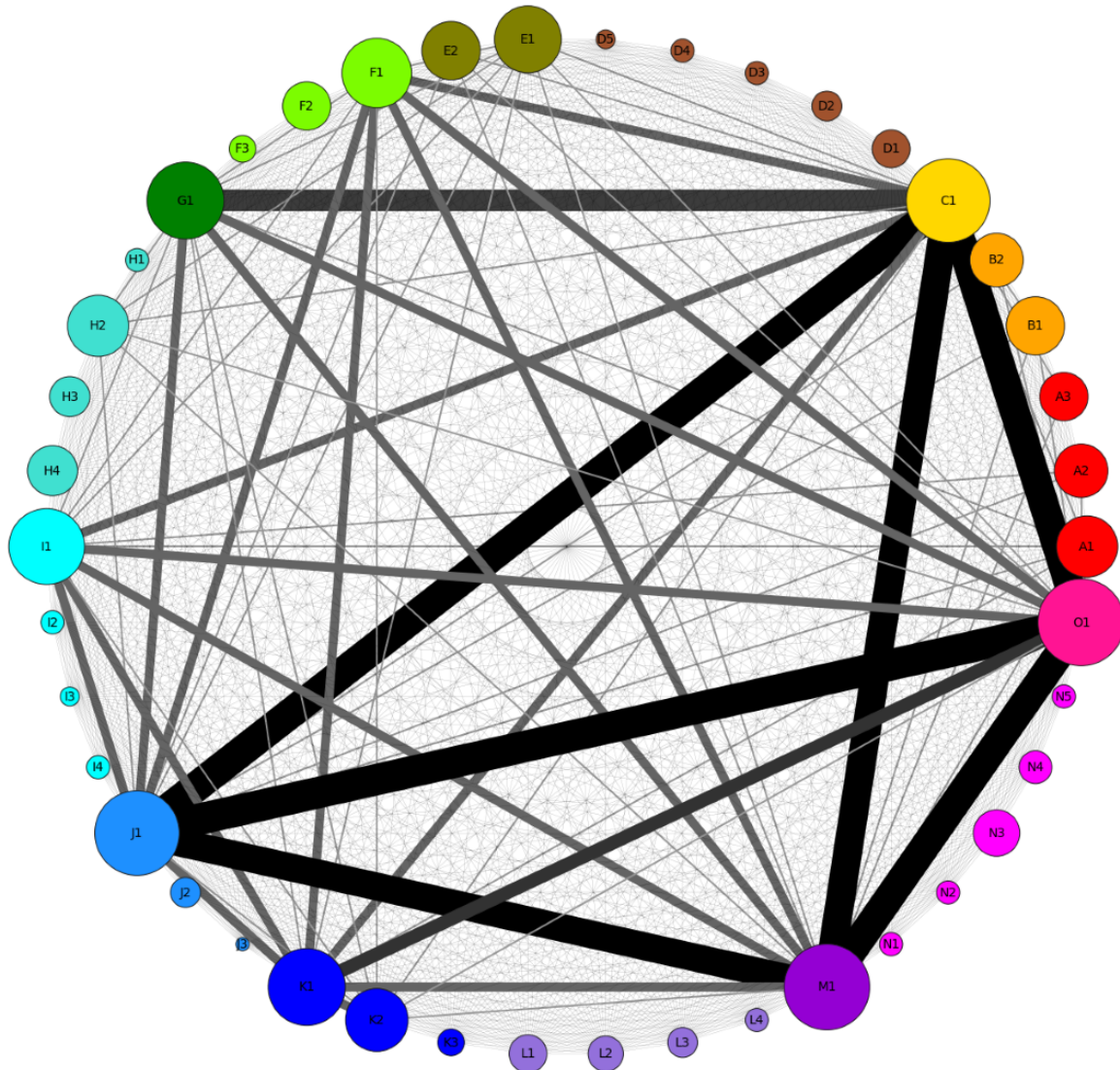
Quality of the Included Studies

A detailed credibility assessment of the individual studies is presented in the Multimedia Appendix 13 [63-105]. An overview of reviewer judgments for each CASP item across all the studies is presented in Figure 5. Of the 44 studies, 13 (30%) studies were assessed as *valuable*; 24 (55%), as *moderately valuable*; and 7 (16%), as being *of some value*. All (44/44, 100%) the studies assessed in this review used an appropriate qualitative methodology and discussed the contribution of the included studies to existing knowledge. The weakest domain included discussing the applicability of the results to other populations or considering other uses for research (16/44, 36%). The mean quality score was 2.14 (SD 0.67), which corresponded to *moderately valuable* or *valuable* studies.

Figure 5. Overview of reviewer judgments on each Critical Appraisal Skills Programme (CASP) item across all studies. N/A: not applicable.

Mapping the Terms

Of the 2068 terms extracted from full-text articles (collected in 18 clusters; Figure 6A), we selected 345 (occurring 11,039 times in included papers) to create a term co-occurrence map (Figure 6B). It produced 15 clusters, which are listed in Multimedia Appendix 14, along with a list of terms. The most frequent terms related to peer-to-peer interactions were support (629/11,039, 5.69%), information (576/11,039, 5.22%), and experience (372/11,039, 3.37%). The overlap of terms describing interactions identified using VOSviewer with codes from our codebook was 96.3% (180/187). Some (7/187, 3.7%) categories did not have an equivalent in the code (*avoidance*, *blame*, *competition*, *discrimination*, *shame*, *tolerance*, and *trust*).

Figure 6. VOSviewer heatmaps of terms; (A) a heatmap of 2068 terms and (B) a heatmap of selected 345 terms.

Discussion

Principal Findings

This systematic scoping review summarizes 44 studies that assessed peer-to-peer interactions among people with 22 different mental disorders. The interactions were categorized into 13 groups. The most common interactions such as *empathy or compassion*, *networking*, and *sharing self-disclosure* were observed on forums for people with eating disorders (14/54, 26%), depression (12/54, 22%), and psychoactive substance use (9/54, 17%). In this study, we focused on developing a codebook for future research. We believed that the reinterpretation of the data reported by the authors of the primary studies included in our review may have introduced bias. Therefore, we did not deliberately dwell on the coexistence of codes and specific disorders. For example, it seemed that sarcastic comments were present in most studies on schizophrenia and schizoaffective disorders. We could hypothesize that because of their condition, these people may be survivors of verbal aggression from their peers. However, it is possible that people with schizophrenia and schizoaffective disorders post sarcastic comments. Thus, this conclusion cannot be fully justified without looking into the primary data, but such an analysis was not the objective of this study. Nevertheless, the normalized means of code co-occurrence across disorders presented in Figure 3 are a good starting point to formulate hypotheses for our future primary studies.

Our study revealed the needs that prompt users to express themselves on the internet. For example, these may be information, emotional, or instrumental support needs. However, without primary research, it is difficult to determine whether these needs are fully (if at all) satisfied by internet use. We can

assume that the unmet need is fulfilled by another internet user; however, these interactions will be the subject of future research. We are aware that because a systematic scoping review aims to map and identify gaps in current knowledge, we will generate more questions than answers. Nevertheless, we hope that this study will inspire future qualitative research in the field.

Overall Completeness and Quality of Evidence

Our study provides evidence on the involvement of people with different mental disorders in online support groups. However, these disorders did not include personality disorders, organic mental disorders, or phobias, and there were no data on the experience of these individuals in seeking support on the internet. In addition, from the analysis of various interaction models, we noticed that some of the codes were not represented in our results because they might not have been assessed or the authors might have failed to report them (eg, owing to insufficient sample size). Therefore, we could not generalize the results to other populations because they may not fully reflect reality. As for the applied qualitative methodology, the authors used different approaches and analytical models. Owing to the diversity of theoretical perspectives, epistemological assumptions, and principles of conducting research, it can be challenging to apply a qualitative approach, including the comparison and synthesis of methods used in qualitative research [159]. Some authors did not comply with the available reporting guidelines [160-162]. Therefore, many studies lacked information on methodology, population characteristics, and outcomes.

Considering the level of adherence to the methodology and the applied methods themselves, we assessed the overall quality of the studies included in our analysis as moderate.

Codes and Correlations

We coded 42 types of interactions between the forum participants. We hierarchized the codes into a model consisting of 15 categories. Our model differs in structure compared with other models in the literature. Although merging some codes into one category may seem unintuitive at first, it resulted from a modified combination of different theoretical approaches and multidisciplinary backgrounds of authors (psychiatry, psychology, sociology, epidemiology, and public health). For example, even though proanorexia forums supported harmful behaviors and contained reinforcers for further weight loss or praise for achieving lower weight, we decided to include the *supportive of harmful behaviors* category in the *behavior promotion* category. A membership in a social group or a web-based community affects the beliefs, preferences, and behaviors of the members via various mechanisms of social influence [163]. Therefore, the category of *behavior promotion* embraces all acts that reinforce behavior patterns regardless of their health consequences.

Assessing the co-occurrence of interactions, we found that the *request for information* most commonly co-occurred with *network*, which stems from the reciprocal nature of conversation that involves the exchange of questions and replies. The strongest associations were found for *requesting engagement* and *disagreement*, *relationship* and *confidentiality*, and *referring*

to the rules and *rejection*. This may be explained by group processes that occur when new members join the group and are mobilized to share their story; the members are assured of confidentiality and presented with the rules that, for example, if violated, will result in the member being removed from the group [164].

A comparison of our codes with the terms identified in VOSviewer showed that our codebook may lack some interactions. However, these extra terms might have occurred in the background or discussion sections and do not apply to our study. Thus, VOSviewer clusters should always be interpreted together with content analysis.

Overall, our study showed that it is possible to use ML techniques to classify mental disorders based on secondary data. Although the results may seem satisfactory, as the accuracy for decision trees was >80%, we cannot consider them to be more than just a proof of concept because of several limitations.

Our Results in the Context of Previous Research

To our knowledge, this is the first scoping review that comprehensively summarizes evidence on all types of web-based peer-to-peer interactions among people with mental disorders. Previous reviews addressed only some types of peer interactions in the context of various nonpsychiatric health-related conditions, such as spinal cord injury [165,166], breastfeeding of hospitalized infants [167], or cancer [168]. In secondary research, peer-to-peer interactions are mostly assessed quantitatively (eg, efficacy assessment [51-53]).

Our study was not limited to specific mental health problems. This is in contrast to previous reviews on mental conditions, as they mainly addressed suicide prevention and dementia. Bowersox et al [169] conducted a scoping review on the function of peers in the prevention of suicidal behaviors. The authors concluded that peer-based interventions could play an important role in suicide prevention. Schlichthorst et al [170] studied peer support programs in suicide prevention and emphasized the usefulness of internet forums as support for people with a history of suicide attempt. Moreover, they alerted to the risks of unmoderated websites.

Carter et al [171] and Newman et al [172] also reviewed web-based peer support interventions in the context of a specific mental health problem (ie, dementia). However, unlike our study, they did not focus on people who directly experienced these problems but assessed individuals who cared for people with dementia. In addition, they did not assess the quality of the included studies. Similar to our approach, they searched several databases and applied similar guidelines for reporting scoping studies [173]. However, they attempted to answer different questions about the effectiveness of interventions and their cost-effectiveness, in addition to identifying the gaps in knowledge.

Limitations

Our study has several limitations. First, we believe that the use of ML techniques requires more data than those collected in this study. Nevertheless, we consider this analysis to be a proof of concept only. Second, the CASP tool was adapted to our

needs by dividing the last question into 3 subquestions. This makes it more challenging to compare the quality of the included studies with that of similar studies. Moreover, with a small data set available, we used the same data for training and calculating the accuracy of ML algorithms (without external validation), which limits the reliability of the results.

Strengths

The strengths of our study include the use of a broad question followed by comprehensive and rigorous search of eligible studies. We searched 4 databases and followed the reporting process provided by Tricco et al [54,55]. In addition, we also assessed the quality of all included studies. We proposed a codebook and partition tree based on the dispersion of the results and compared it with other models. This innovation helps standardize the evidence and allows for data comparison across studies. Finally, we applied ML techniques to identify mental disorders using interactions among peers. The results are quite satisfactory, and even though they are a proof of concept, they can be further explored in future studies.

Future Research

We believe that our codebook describing the categories of peer-to-peer interactions defined in this study can be used in future in-depth investigations of individual mental disorders. In addition, by using AI techniques and applying the rigorous validation of accuracy, this type of analysis could be used to facilitate the diagnosis or screening of mental disorders within web-based self-help groups. Moreover, the assessment of the co-occurrence of interactions and types of disorders could help identify adequate skills and communication styles to define the moderator's characteristics to meet the requirements of a particular forum. However, as this is a proof-of-concept investigation, more specific data are needed to achieve these goals. As only a few studies have investigated web-based peer-to-peer interactions in the setting of mental disorders, more primary research is needed to obtain more evidence. It would be helpful to develop an ML model to establish which interactions are associated with specific diseases and to use AI techniques to investigate more interactions. This could translate into creating a personalized health care experience for individuals with mental disorders.

Conclusions

Internet forums offering peer-to-peer support in mental health attract a heterogeneous group of people. Interactions between the members are predominately positive. Although the use of the internet to seek support for health problems has become commonplace, scientific evidence on this phenomenon is scarce. In the future, AI-based analysis of interactions between the members of mental health forums and a better understanding of their needs could help moderators provide personalized support to internet users.

Data Availability

The code is available on GitHub [174]. The data supporting the findings of this study are available in Multimedia Appendices 1-14.

Conflicts of Interest

None declared.

Multimedia Appendix 1
PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) checklist.
[DOCX File , 33 KB - mental_v9i12e36056_app1.docx]

Multimedia Appendix 2
Search strategies.
[DOCX File , 31 KB - mental_v9i12e36056_app2.docx]

Multimedia Appendix 3
Other partition trees.
[DOCX File , 73 KB - mental_v9i12e36056_app3.docx]

Multimedia Appendix 4
List of included studies.
[DOCX File , 33 KB - mental_v9i12e36056_app4.docx]

Multimedia Appendix 5
List of ongoing studies.
[DOCX File , 28 KB - mental_v9i12e36056_app5.docx]

Multimedia Appendix 6
List of excluded studies according to the reason for exclusion.
[[DOCX File , 33 KB](#) - [mental_v9i12e36056_app6.docx](#)]

Multimedia Appendix 7
Characteristics of included studies.
[[DOCX File , 46 KB](#) - [mental_v9i12e36056_app7.docx](#)]

Multimedia Appendix 8
Means of codes per category and SDs for analyzed models.
[[DOCX File , 29 KB](#) - [mental_v9i12e36056_app8.docx](#)]

Multimedia Appendix 9
Heatmaps with codes.
[[DOCX File , 146 KB](#) - [mental_v9i12e36056_app9.docx](#)]

Multimedia Appendix 10
Supplementary material.
[[XLSX File (Microsoft Excel File), 95 KB](#) - [mental_v9i12e36056_app10.xlsx](#)]

Multimedia Appendix 11
Machine-learning results with confusion matrices.
[[DOCX File , 612 KB](#) - [mental_v9i12e36056_app11.docx](#)]

Multimedia Appendix 12
Decision trees of prediction of diseases based on the presence of interaction.
[[DOCX File , 391 KB](#) - [mental_v9i12e36056_app12.docx](#)]

Multimedia Appendix 13
Quality of included studies.
[[DOCX File , 103 KB](#) - [mental_v9i12e36056_app13.docx](#)]

Multimedia Appendix 14
List of clusters and related terms from VOSviewer.
[[DOCX File , 41 KB](#) - [mental_v9i12e36056_app14.docx](#)]

References

1. Wittchen HU, Jacobi F, Rehm J, Gustavsson A, Svensson M, Jönsson B, et al. The size and burden of mental disorders and other disorders of the brain in Europe 2010. Eur Neuropsychopharmacol 2011 Sep;21(9):655-679. [doi: 10.1016/j.euroneuro.2011.07.018] [Medline: 21896369]
2. 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]
3. Angermeyer MC, Matschinger H. Public beliefs about schizophrenia and depression: similarities and differences. Soc Psychiatry Psychiatr Epidemiol 2003 Sep 1;38(9):526-534. [doi: 10.1007/s00127-003-0676-6] [Medline: 14504738]
4. Smith M. Anti-stigma campaigns: time to change. Br J Psychiatry Suppl 2013 Apr 02;55(s55):s49-s50. [doi: 10.1192/bjp.bp.113.126813] [Medline: 23553694]
5. Angermeyer MC, Schulze B. Reinforcing stereotypes: how the focus on forensic cases in news reporting may influence public attitudes towards the mentally ill. Int J Law Psychiatry 2001 Jul;24(4-5):469-486. [doi: 10.1016/s0160-2527(01)00079-6]
6. Becker AE, Kleinman A. Mental health and the global agenda. N Engl J Med 2013 Jul 04;369(1):66-73. [doi: 10.1056/nejmra1110827]
7. Loch AA, Wang Y, Guarniero FB, Lawson FL, Hengartner MP, Rössler W, et al. Patterns of stigma toward schizophrenia among the general population: a latent profile analysis. Int J Soc Psychiatry 2014 Sep 21;60(6):595-605. [doi: 10.1177/0020764013507248] [Medline: 24146365]
8. Kessler RC, Foster CL, Saunders WB, Stang PE. Social consequences of psychiatric disorders, I: educational attainment. Am J Psychiatry 1995 Jul;152(7):1026-1032. [doi: 10.1176/ajp.152.7.1026] [Medline: 7793438]

9. Gowen K, Deschaine M, Gruttadara D, Markey D. Young adults with mental health conditions and social networking websites: seeking tools to build community. Psychiatr Rehabil J 2012;35(3):245-250. [doi: 10.2975/35.3.2012.245.250] [Medline: 22246123]

10. Dickerson SS, Boehmke M, Ogle C, Brown JK. Seeking and managing hope: patients' experiences using the Internet for cancer care. Oncol Nurs Forum 2006 Jan 01;33(1):E8-17. [doi: 10.1188/06.ONF.E8-E17] [Medline: 16470231]

11. Barrera M, Glasgow R, McKay H, Boles S, Feil E. Do Internet-based support interventions change perceptions of social support?: an experimental trial of approaches for supporting diabetes self-management. Am J Community Psychol 2002 Oct;30(5):637-654. [doi: 10.1023/A:1016369114780] [Medline: 12188054]

12. White M, Dorman S. Receiving social support online: implications for health education. Health Educ Res 2001 Dec;16(6):693-707. [doi: 10.1093/her/16.6.693] [Medline: 11780708]

13. Murray CD, Fox J. Do Internet self-harm discussion groups alleviate or exacerbate self-harming behaviour? Aust e-J Advance Mental Health 2014 Dec 17;5(3):225-233. [doi: 10.5172/jamh.5.3.225]

14. Carrard I, Rouget P, Fernández-Aranda F, Volkart A, Damoiseau M, Lam T. Evaluation and deployment of evidence based patient self-management support program for Bulimia Nervosa. Int J Med Inform 2006 Jan;75(1):101-109. [doi: 10.1016/j.ijmedinf.2005.07.031] [Medline: 16115793]

15. Johnsen JK, Rosenvinge JH, Gammon D. Online group interaction and mental health: an analysis of three online discussion forums. Scand J Psychol 2002 Dec 28;43(5):445-449. [doi: 10.1111/1467-9450.00313] [Medline: 12500784]

16. Spinzy Y, Nitzan U, Becker G, Bloch Y, Fennig S. Does the Internet offer social opportunities for individuals with schizophrenia? A cross-sectional pilot study. Psychiatry Res 2012 Jul 30;198(2):319-320. [doi: 10.1016/j.psychres.2012.02.022] [Medline: 22440545]

17. Expanding telemental health in response to the COVID-19 pandemic. Psychiatric Times. 2020 Apr 08. URL: https://www.psychiatrictimes.com/view/expanding-telemental-health-response-covid-19-pandemic [accessed 2022-05-13]

18. Measuring Digital Development Facts and Figures 2021. Geneva: International Telecommunication Union; 2021.

19. van Uden-Kraan CF, Drossaert CH, Taal E, Seydel ER, van de Laar MA. Self-reported differences in empowerment between lurkers and posters in online patient support groups. J Med Internet Res 2008 Jun 30;10(2):e18 [FREE Full text] [doi: 10.2196/jmir.992] [Medline: 18653442]

20. Mo PK, Coulson NS. Empowering processes in online support groups among people living with HIV/AIDS: a comparative analysis of 'lurkers' and 'posters'. Comput Human Behav 2010 Sep;26(5):1183-1193. [doi: 10.1016/j.chb.2010.03.028]

21. Chung JE. Social networking in online support groups for health: how online social networking benefits patients. J Health Commun 2014;19(6):639-659. [doi: 10.1080/10810730.2012.757396] [Medline: 23557148]

22. Schrank B, Sibitz I, Unger A, Amering M. How patients with schizophrenia use the internet: qualitative study. J Med Internet Res 2010 Dec 19;12(5):e70 [FREE Full text] [doi: 10.2196/jmir.1550] [Medline: 21169176]

23. Kang R, Brown S, Kiesler S. Why do people seek anonymity on the internet?: informing policy and design. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2013 Presented at: CHI '13: CHI Conference on Human Factors in Computing Systems; Apr 27-May 2, 2013; Paris France. [doi: 10.1145/2470654.2481368]

24. McKenna KY, Bargh JA. Coming out in the age of the internet: identity "demarginalization" through virtual group participation. J Personal Social Psychol 1998 Sep;75(3):681-694. [doi: 10.1037/0022-3514.75.3.681]

25. Davidson L, Chinman M, Kloos B, Weingarten R, Stayner D, Tebes JK. Peer support among individuals with severe mental illness: a review of the evidence. Clin Psychol Sci Pract 1999;6(2):165-187. [doi: 10.1093/clipsy.6.2.165]

26. Link BG, Struening EL, Rahav M, Phelan JC, Nuttbrock L. On stigma and its consequences: evidence from a longitudinal study of men with dual diagnoses of mental illness and substance abuse. J Health Soc Behav 1997 Jun;38(2):177-190. [Medline: 9212538]

27. Lal S, Adair CE. E-mental health: a rapid review of the literature. Psychiatr Serv 2014 Jan 01;65(1):24-32. [doi: 10.1176/appi.ps.201300009] [Medline: 24081188]

28. Explore the Cambridge Dictionary. Cambridge Dictionary. URL: https://dictionary.cambridge.org/us/ [accessed 2022-05-13]

29. Moore SE, McMullan M, McEvoy CT, McKinley MC, Woodside JV. The effectiveness of peer-supported interventions for encouraging dietary behaviour change in adults: a systematic review. Public Health Nutr 2018 Dec 03;22(4):624-644. [doi: 10.1017/s1368980018003294]

30. Webel AR, Okonsky J, Trompeta J, Holzemer WL. A systematic review of the effectiveness of peer-based interventions on health-related behaviors in adults. Am J Public Health 2010 Feb;100(2):247-253. [doi: 10.2105/ajph.2008.149419]

31. Davidson L, Chinman M, Sells D, Rowe M. Peer support among adults with serious mental illness: a report from the field. Schizophr Bull 2006 Jul 28;32(3):443-450 [FREE Full text] [doi: 10.1093/schbul/sbj043] [Medline: 16461576]

32. Welch V, Petkovic J, Pardo Pardo J, Rader T, Tugwell P. Interactive social media interventions to promote health equity: an overview of reviews. Health Promot Chronic Dis Prev Can 2016 Apr;36(4):63-75 [FREE Full text] [doi: 10.24095/hpcdp.36.4.01] [Medline: 27077792]

33. Burns JM, Durkin LA, Nicholas J. Mental health of young people in the United States: what role can the internet play in reducing stigma and promoting help seeking? J Adolesc Health 2009 Jul;45(1):95-97. [doi: 10.1016/j.jadohealth.2008.12.006] [Medline: 19541256]

34. Tanis M. Health-related on-line forums: what's the big attraction? J Health Commun 2008 Oct 28;13(7):698-714. [doi: 10.1080/10810730802415316] [Medline: 18958781]

35. Galinsky MJ, Schopler JH, Abell MD. Connecting group members through telephone and computer groups. Health Soc Work 1997 Aug 01;22(3):181-188. [doi: 10.1093/hsw/22.3.181] [Medline: 9260082]

36. Nicholas D. Participant perceptions of online group work with fathers of children with Spina Bifida. In: Social Work with Groups. Milton Park, Abingdon-on-Thames, Oxfordshire, England, UK: Routledge; 2003.

37. Weinberg N, Schmale J, Uken J, Wessel K. Online help: cancer patients participate in a computer-mediated support group. Health Soc Work 1996 Feb;21(1):24-29. [doi: 10.1093/hsw/21.1.24] [Medline: 8626154]

38. Internet seen as positive influence on education but negative on morality in emerging and developing nations. Pew Research Center. 2015 Mar 19. URL: https://www.pewresearch.org/global/2015/03/19/internet-seen-as-positive-influence-on-education-but-negative-influence-on-morality-in-emerging-and-developing-nations/ [accessed 2022-02-23]

39. Dickerson FB, Sommerville J, Origoni AE, Ringel NB, Parente F. Outpatients with schizophrenia and bipolar I disorder: do they differ in their cognitive and social functioning? Psychiatry Res 2001 May;102(1):21-27. [doi: 10.1016/s0165-1781(01)00247-5]

40. Barak A, Dolev-Cohen M. Does activity level in online support groups for distressed adolescents determine emotional relief. Counsell Psychother Res 2006 Sep;6(3):186-190. [doi: 10.1080/14733140600848203]

41. Robinson J, Rodrigues M, Fisher S, Bailey E, Herrman H. Social media and suicide prevention: findings from a stakeholder survey. Shanghai Arch Psychiatry 2015 Feb 25;27(1):27-35 [FREE Full text] [doi: 10.11919/j.issn.1002-0829.214133] [Medline: 25852253]

42. Monnier D. Woebot: a continuation of and an end to psychotherapy? Psychotherapies 2020 Apr;40(2):71-78.

43. Myszczynska MA, Ojamies PN, Lacoste AM, Neil D, Saffari A, Mead R, et al. Applications of machine learning to diagnosis and treatment of neurodegenerative diseases. Nat Rev Neurol 2020 Aug 15;16(8):440-456. [doi: 10.1038/s41582-020-0377-8] [Medline: 32669685]

44. Rajkomar A, Dean J, Kohane I. Machine Learning in Medicine. N Engl J Med 2019 Apr 04;380(14):1347-1358. [doi: 10.1056/nejmra1814259]

45. Velichko E, Nepomnyashchaya E, Baranov M, Galeeva M, Pavlov V, Zavjalov S. A concept of smart medical autonomous distributed system for diagnostics based on machine learning technology. In: Internet of Things, Smart Spaces, and Next Generation Networks and Systems. Cham: Springer; 2019.

46. Halama N. Machine learning for tissue diagnostics in oncology: brave new world. Br J Cancer 2019 Sep 09;121(6):431-433 [FREE Full text] [doi: 10.1038/s41416-019-0535-1] [Medline: 31395951]

47. Lundervold AS, Lundervold A. An overview of deep learning in medical imaging focusing on MRI. Z Med Phys 2019 May;29(2):102-127 [FREE Full text] [doi: 10.1016/j.zemedi.2018.11.002] [Medline: 30553609]

48. Brisson A, Pereira G, Prada R, Paiva A, Louchart S, Suttie N, et al. Artificial intelligence and personalization opportunities for serious games. Proc AAAI Conf Artificial Intell Interact Digital Entertainment 2021 Jun 30;8(5):51-57. [doi: 10.1609/aiide.v8i5.12576]

49. Mobasher B. Intelligent Techniques for Web Personalization IJCAI 2003 Workshop, ITWP 2003, Acapulco, Mexico, August 11, 2003, Revised Selected Papers. Berlin, Heidelberg: Springer; 2005.

50. Yoganarasimhan H. Search personalization using machine learning. Manag Sci 2020 Mar;66(3):1045-1070. [doi: 10.1287/mnsc.2018.3255]

51. Ali K, Farrer L, Gulliver A, Griffiths KM. Online peer-to-peer support for young people with mental health problems: a systematic review. JMIR Ment Health 2015 May 19;2(2):e19 [FREE Full text] [doi: 10.2196/mental.4418] [Medline: 26543923]

52. Biagianti B, Quraishi SH, Schlosser DA. Potential benefits of incorporating peer-to-peer interactions into digital interventions for psychotic disorders: a systematic review. Psychiatr Serv 2018 Apr 01;69(4):377-388 [FREE Full text] [doi: 10.1176/appi.ps.201700283] [Medline: 29241435]

53. Eysenbach G, Powell J, Englesakis M, Rizo C, Stern A. Health related virtual communities and electronic support groups: systematic review of the effects of online peer to peer interactions. BMJ 2004 May 15;328(7449):1166 [FREE Full text] [doi: 10.1136/bmj.328.7449.1166] [Medline: 15142921]

54. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 02;169(7):467-473 [FREE Full text] [Medline: 30178033]

55. Tricco AC, Lillie E, Zarin W, O'Brien K, Colquhoun H, Kastner M, et al. A scoping review on the conduct and reporting of scoping reviews. BMC Med Res Methodol 2016 Feb 09;16:15 [FREE Full text] [doi: 10.1186/s12874-016-0116-4] [Medline: 26857112]

56. Storman D, Jemioło P, Swierz MJ, Bala MM, Prokop-Dorner A, Gotfryd M. Peer-to-peer online interventions for people with mental illnesses. Open Science Framework. 2020 Aug 18. URL: https://osf.io/j3azv/ [accessed 2022-11-30]

57. Greenwood E, Bales RF. Interaction process analysis: a method for the study of small groups. Am Sociol Rev 1950 Oct;15(5):693. [doi: 10.2307/2086941]

58. CASP Checklist. Critical Appraisal Skills Programme. URL: https://casp-uk.net/casp-tools-checklists/ [accessed 2022-05-13]

59. Manikandan S. Measures of dispersion. J Pharmacol Pharmacotherapeutics 2022 Apr 11;2(4):315-316. [doi: 10.4103/0976-500x.85931]

60. Krysinska K, Lester D. Post-traumatic stress disorder and suicide risk: a systematic review. Arch Suicide Res 2010 Jan 29;14(1):1-23. [doi: 10.1080/13811110903478997] [Medline: 20112140]

61. Vilone G, Longo L. Explainable artificial intelligence: a systematic review. arXiv 2020. [doi: 10.48550/arXiv.2006.00093]

62. van Eck NJ, Waltman L. VOSviewer manual. Univeristeit Leiden. 2013 Jan 1. URL: https://www.vosviewer.com/documentation/Manual_VOSviewer_1.5.4.pdf [accessed 2022-05-13]

63. Eghdam A, Hamidi U, Bartfai A, Koch S. Facebook as communication support for persons with potential mild acquired cognitive impairment: a content and social network analysis study. PLoS One 2018;13(1):e0191878 [FREE Full text] [doi: 10.1371/journal.pone.0191878] [Medline: 29377930]

64. Greiner C, Chatton A, Khazaal Y. Online self-help forums on cannabis: a content assessment. Patient Educ Couns 2017 Oct;100(10):1943-1950. [doi: 10.1016/j.pec.2017.06.001] [Medline: 28602568]

65. Andalibi N, Ozturk P, Forte A. Sensitive self-disclosures, responses, and social support on Instagram: the case of #depression. In: Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing. 2017 Presented at: CSCW '17: Computer Supported Cooperative Work and Social Computing; Feb 25-Mar 1, 2017; Portland Oregon USA. [doi: 10.1145/2998181.2998243]

66. Barney LJ, Griffiths KM, Banfield MA. Explicit and implicit information needs of people with depression: a qualitative investigation of problems reported on an online depression support forum. BMC Psychiatry 2011 May 18;11(1):88 [FREE Full text] [doi: 10.1186/1471-244X-11-88] [Medline: 21592335]

67. Beck SJ, Paskewitz EA, Anderson WA, Bourdeaux R, Currie-Mueller J. The task and relational dimensions of online social support. Health Commun 2017 Mar;32(3):347-355. [doi: 10.1080/10410236.2016.1138383] [Medline: 27268509]

68. Bronstein J. Is this OCD?: exploring conditions of information poverty in online support groups dealing with obsessive compulsive disorder. Inform Res 2014;19(4).

69. Brown S, Altice FL. Self-management of buprenorphine/naloxone among online discussion board users. Subst Use Misuse 2014 Jun 29;49(8):1017-1024 [FREE Full text] [doi: 10.3109/10826084.2014.888449] [Medline: 24779501]

70. Chen AT, Slattery K, Tomasino KN, Rubanovich CK, Bardsley LR, Mohr DC. Challenges and benefits of an internet-based intervention with a peer support component for older adults with depression: qualitative analysis of textual data. J Med Internet Res 2020 Jun 16;22(6):e17586 [FREE Full text] [doi: 10.2196/17586] [Medline: 32543448]

71. Cunningham JA, van Mierlo T, Fournier R. An online support group for problem drinkers: AlcoholHelpCenter.net. Patient Educ Couns 2008 Feb;70(2):193-198. [doi: 10.1016/j.pec.2007.10.003] [Medline: 18022340]

72. Doran J. Posting incognito ... males with eating problems: online emotional expression and support. J Cyber Ther Rehabil 2011;4(3):341.

73. Edward K, Robins A. Dual diagnosis, as described by those who experience the disorder: using the Internet as a source of data. Int J Ment Health Nurs 2012 Dec;21(6):550-559. [doi: 10.1111/j.1447-0349.2012.00833.x] [Medline: 22830579]

74. Evans M, Donelle L, Hume-Loveland L. Social support and online postpartum depression discussion groups: a content analysis. Patient Educ Couns 2012 Jun;87(3):405-410. [doi: 10.1016/j.pec.2011.09.011] [Medline: 22019021]

75. Gajaria A, Yeung E, Goodale T, Charach A. Beliefs about attention-deficit/hyperactivity disorder and response to stereotypes: youth postings in Facebook groups. J Adolesc Health 2011 Jul;49(1):15-20. [doi: 10.1016/j.jadohealth.2010.09.004] [Medline: 21700151]

76. Gavin J, Rodham K, Poyer H. The presentation of "pro-anorexia" in online group interactions. Qual Health Res 2008 Mar 01;18(3):325-333. [doi: 10.1177/1049732307311640] [Medline: 18235156]

77. Giles DC, Newbold J. Self- and other-diagnosis in user-led mental health online communities. Qual Health Res 2011 Mar 25;21(3):419-428. [doi: 10.1177/1049732310381388] [Medline: 20739589]

78. Horgan A, McCarthy G, Sweeney J. An evaluation of an online peer support forum for university students with depressive symptoms. Arch Psychiatr Nurs 2013 Apr;27(2):84-89. [doi: 10.1016/j.apnu.2012.12.005] [Medline: 23540518]

79. Juarascio AS, Shoaib A, Timko CA. Pro-eating disorder communities on social networking sites: a content analysis. Eat Disord 2010 Sep 30;18(5):393-407. [doi: 10.1080/10640266.2010.511918] [Medline: 20865593]

80. Kantrowitz-Gordon I. Internet confessions of postpartum depression. Issues Ment Health Nurs 2013 Dec 25;34(12):874-882. [doi: 10.3109/01612840.2013.806618] [Medline: 24274243]

81. Kendal S, Kirk S, Elvey R, Catchpole R, Pryjmachuk S. How a moderated online discussion forum facilitates support for young people with eating disorders. Health Expect 2017 Feb 03;20(1):98-111 [FREE Full text] [doi: 10.1111/hex.12439] [Medline: 26725547]

82. Keski-Rahkonen A, Tozzi F. The process of recovery in eating disorder sufferers' own words: an Internet-based study. Int J Eat Disord 2005;37 Suppl(S1):S80-6; discussion S87. [doi: 10.1002/eat.20123] [Medline: 15852327]

83. Lavis A, Winter R. #Online harms or benefits? An ethnographic analysis of the positives and negatives of peer-support around self-harm on social media. J Child Psychol Psychiatry 2020 Aug 27;61(8):842-854. [doi: 10.1111/jcpp.13245] [Medline: 32459004]

84. Lerman BI, Lewis SP, Lumley M, Grogan GJ, Hudson CC, Johnson E. Teen depression groups on Facebook: a content analysis. J Adolesc Res 2016 Oct 22;32(6):719-741. [doi: 10.1177/0743558416673717]

85. Lord VM, Reiboldt W, Gonitzke D, Parker E, Peterson C. Experiences of recovery in binge-eating disorder: a qualitative approach using online message boards. Eat Weight Disord 2018 Feb 31;23(1):95-105. [doi: 10.1007/s40519-016-0335-z] [Medline: 27796846]

86. Moore D, Ayers S, Drey N. A thematic analysis of stigma and disclosure for perinatal depression on an online forum. JMIR Ment Health 2016 May 19;3(2):e18 [FREE Full text] [doi: 10.2196/mental.5611] [Medline: 27197516]

87. Mudry TE, Strong T. Doing recovery online. Qual Health Res 2013 Mar;23(3):313-325. [doi: 10.1177/1049732312468296] [Medline: 23208202]

88. Naslund JA, Grande SW, Aschbrenner KA, Elwyn G. Naturally occurring peer support through social media: the experiences of individuals with severe mental illness using YouTube. PLoS One 2014 Oct 15;9(10):e110171 [FREE Full text] [doi: 10.1371/journal.pone.0110171] [Medline: 25333470]

89. Park A, Conway M, Chen AT. Examining thematic similarity, difference, and membership in three online mental health communities from reddit: a text mining and visualization approach. Comput Human Behav 2018 Jan;78:98-112 [FREE Full text] [doi: 10.1016/j.chb.2017.09.001] [Medline: 29456286]

90. Prescott J, Rathbone AL, Brown G. Online peer to peer support: qualitative analysis of UK and US open mental health Facebook groups. Digit Health 2020 Dec 10;6:2055207620979209 [FREE Full text] [doi: 10.1177/2055207620979209] [Medline: 33354335]

91. Salzmann-Erikson M, Hiçdurmaz D. Use of social media among individuals who suffer from post-traumatic stress: a qualitative analysis of narratives. Qual Health Res 2017 Jan 09;27(2):285-294. [doi: 10.1177/1049732315627364] [Medline: 26825480]

92. Sharkey S, Smithson J, Hewis E, Jones R, Emmens T, Ford T, et al. Supportive interchanges and face-work as 'protective talk' in an online self-harm support forum. Commun Med 2012 Feb 19;9(1):71-82. [doi: 10.1558/cam.v9i1.71] [Medline: 23763238]

93. Smithson J, Sharkey S, Hewis E, Jones RB, Emmens T, Ford T, et al. Membership and boundary maintenance on an online self-harm forum. Qual Health Res 2011 Nov 29;21(11):1567-1575. [doi: 10.1177/1049732311413784] [Medline: 21715606]

94. Soussan C, Kjellgren A. Harm reduction and knowledge exchange-a qualitative analysis of drug-related Internet discussion forums. Harm Reduct J 2014 Sep 08;11(1):25 [FREE Full text] [doi: 10.1186/1477-7517-11-25] [Medline: 25200686]

95. Sowles SJ, McLeary M, Optican A, Cahn E, Krauss MJ, Fitzsimmons-Craft EE, et al. A content analysis of an online pro-eating disorder community on Reddit. Body Image 2018 Mar;24:137-144 [FREE Full text] [doi: 10.1016/j.bodyim.2018.01.001] [Medline: 29414146]

96. Walstrom MK. "You know, who's the thinnest?": combating surveillance and creating safety in coping with eating disorders online. CyberPsychol Behav 2000 Oct;3(5):761-783. [doi: 10.1089/10949310050191755]

97. Worley J, Krishnan D. Fighting the battle of recovery together: a content analysis of anonymous posts in an online substance use forum. Issues Ment Health Nurs 2020 Feb 06;41(2):102-112. [doi: 10.1080/01612840.2019.1646364] [Medline: 31386598]

98. Elran-Barak R. Analyses of posts written in online eating disorder and depression/anxiety moderated communities: emotional and informational communication before and during the COVID-19 outbreak. Internet Interv 2021 Dec;26:100438 [FREE Full text] [doi: 10.1016/j.invent.2021.100438] [Medline: 34401396]

99. Flickinger TE, Waselewski M, Tabackman A, Huynh J, Hodges J, Otero K, et al. Communication between patients, peers, and care providers through a mobile health intervention supporting medication-assisted treatment for opioid use disorder. Patient Educ Couns 2022 Jul;105(7):2110-2115 [FREE Full text] [doi: 10.1016/j.pec.2022.02.014] [Medline: 35260260]

100. Goh AQ, Lo NY, Davis C, Chew EC. #EatingDisorderRecovery: a qualitative content analysis of eating disorder recovery-related posts on Instagram. Eat Weight Disord 2022 May 18;27(4):1535-1545. [doi: 10.1007/s40519-021-01279-1] [Medline: 34537927]

101. Kim H, Chung M, Rhee ES, Kim Y. Is it reciprocating or self-serving?: understanding coping strategies for postpartum depression in an online community for Korean mothers. Health Care Women Int 2022 Feb 23:1-18. [doi: 10.1080/07399332.2022.2037604] [Medline: 35195498]

102. Albano G, Bonfanti RC, Gullo S, Salerno L, Lo Coco G. The psychological impact of COVID-19 on people suffering from dysfunctional eating behaviours: a linguistic analysis of the contents shared in an online community during the lockdown. Res Psychother 2021 Dec 20;24(3):557 [FREE Full text] [doi: 10.4081/ripppo.2021.557] [Medline: 35047429]

103. Sik D, Németh R, Katona E. Topic modelling online depression forums: beyond narratives of self-objectification and self-blaming. J Ment Health 2021 Sep 28:1-10. [doi: 10.1080/09638237.2021.1979493] [Medline: 34582309]

104. Bendall S, Truss K, Liao Siling J, Phillips L. What stops young people from seeking help for the effects of trauma? A qualitative analysis of internet forums. In: Proceedings of the 11th International Conference on Early Intervention in Mental Health –"Prevention and Early Intervention: Broadening the Scope". 2018 Presented at: IEPA 11th International Conference on Early Intervention in Mental Health –"Prevention and Early Intervention: Broadening the Scope"; 7–10 October 2018; Boston, Massachusetts, USA. [doi: 10.1111/eip.12723]

105. McGregor KA, Clancy O. 100. Starving for support: natural language processing and machine learning analysis of anorexia nervosa in pro-eating disorder communities. J Adolesc Health 2019 Feb;64(2):S53. [doi: 10.1016/j.jadohealth.2018.10.116]

106. Parsons C, McGrath D. Medication information needs of people with Parkinson's disease and their carers: a qualitative study using an online discussion forum. Int J Pharm Pract 2019;27(S2):9.

107. Žaja N, Rukavina TV, Brborović O, Uzun S, Mimica N. Reasons for using depression Internet forums in Croatia. In: Proceedings of the 26th European Congress of Psychiatry. 2018 Presented at: 26th European Congress of Psychiatry; Mar, 2018; Nice, France. [doi: 10.26226/morressier.5a6ef3edd462b80290b58142]

108. Zhu L, Mehta N, Wu W, Savage R, Lam K, Ghuman I, et al. P4-401: dementia 2.0: an exploratory analysis of twitter and health forum data. Alzheimer's Dementia 2019 Jul 01;15:P1456. [doi: 10.1016/j.jalz.2019.06.4073]

109. Dixon-Ward KC, Chan SW. 'Faking it': exploring adolescent perceptions of depression (in)authenticity and 'attention seeking'. Br J Clin Psychol 2022 Jun 29;61(2):177-196. [doi: 10.1111/bjc.12339] [Medline: 34716598]

110. Kepner W, Meacham MC, Nobles AL. Types and sources of stigma on opioid use treatment and recovery communities on reddit. Subst Use Misuse 2022 Jul 09;57(10):1511-1522. [doi: 10.1080/10826084.2022.2091786] [Medline: 35815614]

111. Bickerstaff JM, Karim S, Whitman RK, Cypher A, Wiener S, Radovic A. "You have people here to help you, people like me." a qualitative analysis of a blogging intervention for adolescents and young adults with depression or anxiety. J Technol Behav Sci 2021 Dec 15;6(4):578-588. [doi: 10.1007/s41347-021-00210-w] [Medline: 34926793]

112. Page MJ, Moher D, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. BMJ 2021 Mar 29;372:n160 [FREE Full text] [doi: 10.1136/bmj.n160] [Medline: 33781993]

113. Abel S, Machin T, Brownlow C. Support, socialise and advocate: an exploration of the stated purposes of Facebook autism groups. Res Autism Spectrum Disorder 2019 May;61:10-21. [doi: 10.1016/j.rasd.2019.01.009]

114. Athanasopoulou C, Sakellari E. Facebook and health information: content analysis of groups related to schizophrenia. Stud Health Technol Inform 2015;213:255-258. [Medline: 26153008]

115. Fraga B, da Silva A, Murai F. Online social networks in health care: a study of mental disorders on Reddit. In: Proceedings of the 2018 IEEE/WIC/ACM International Conference on Web Intelligence (WI). 2018 Presented at: 2018 IEEE/WIC/ACM International Conference on Web Intelligence (WI); Dec 03-06, 2018; Santiago, Chile. [doi: 10.1109/wi.2018.00-36]

116. Lamy FR, Daniulaityte R, Nahhas RW, Barratt MJ, Smith AG, Sheth A, et al. Increases in synthetic cannabinoids-related harms: results from a longitudinal web-based content analysis. Int J Drug Policy 2017 Jun;44:121-129 [FREE Full text] [doi: 10.1016/j.drugpo.2017.05.007] [Medline: 28578250]

117. Linden M, Pirsich C. Negative and helpful statements in a patient only or therapist guided internet forum in the aftercare for psychosomatic inpatients. Psychiatr Danub 2017 Dec 30;29(4):446-450 [FREE Full text] [doi: 10.24869/psyd.2017.446] [Medline: 29197201]

118. Lyons M, Aksayli ND, Brewer G. Mental distress and language use: linguistic analysis of discussion forum posts. Comput Human Behav 2018 Oct;87:207-211. [doi: 10.1016/j.chb.2018.05.035]

119. Martínez-Pérez B, de la Torre-Díez I, Bargiela-Flórez B, López-Coronado M, Rodrigues JJ. Content analysis of neurodegenerative and mental diseases social groups. Health Informatics J 2015 Dec 03;21(4):267-283. [doi: 10.1177/1460458214525615] [Medline: 24698768]

120. Mazel S, Zisman-Ilani Y, Hennig S, Garnick D, Nicholson J. Virtual engagement in a social media community of mothers with substance use disorders: content analysis. JMIR Form Res 2021 Jun 24;5(6):e24353 [FREE Full text] [doi: 10.2196/24353] [Medline: 34184993]

121. McCaig D, Elliott MT, Siew CS, Walasek L, Meyer C. Profiling commenters on mental health-related online forums: a methodological example focusing on eating disorder-related commenters. JMIR Ment Health 2019 Apr 22;6(4):e12555 [FREE Full text] [doi: 10.2196/12555] [Medline: 31008715]

122. Moessner M, Feldhege J, Wolf M, Bauer S. Analyzing big data in social media: text and network analyses of an eating disorder forum. Int J Eat Disord 2018 Jul 10;51(7):656-667. [doi: 10.1002/eat.22878] [Medline: 29746710]

123. Mokkenstorm JK, Mérelle SY, Smit JH, Beekman AT, Kerkhof AJ, Huisman A, et al. Exploration of benefits and potential harmful effects of an online forum for visitors to the suicide prevention platform in the netherlands. Crisis 2020 May;41(3):205-213. [doi: 10.1027/0227-5910/a000627] [Medline: 31657643]

124. Nguyen T, O'Dea B, Larsen M, Phung D, Venkatesh S, Christensen H. Using linguistic and topic analysis to classify sub-groups of online depression communities. Multimed Tools Appl 2015 Dec 21;76(8):10653-10676. [doi: 10.1007/s11042-015-3128-x]

125. Selby P, van Mierlo T, Voci SC, Parent D, Cunningham JA. Online social and professional support for smokers trying to quit: an exploration of first time posts from 2562 members. J Med Internet Res 2010 Aug 18;12(3):e34 [FREE Full text] [doi: 10.2196/jmir.1340] [Medline: 20719739]

126. Tan YT, Rehm IC, Stevenson JL, De Foe A. Social media peer support groups for obsessive-compulsive and related disorders: understanding the predictors of negative experiences. J Affect Disord 2021 Feb 15;281:661-672. [doi: 10.1016/j.jad.2020.11.094] [Medline: 33234279]

127. Winkel S, Groen G, Petermann F. [Social support in suicide forums]. Prax Kinderpsychol Kinderpsychiatr 2005 Nov;54(9):714-727. [Medline: 16305019]

128. Xu R, Zhang Q. Understanding online health groups for depression: social network and linguistic perspectives. J Med Internet Res 2016 Mar 10;18(3):e63 [FREE Full text] [doi: 10.2196/jmir.5042] [Medline: 26966078]

129. Ahlström BH, Wentz E. Difficulties in everyday life: young persons with attention-deficit/hyperactivity disorder and autism spectrum disorders perspectives. A chat-log analysis. Int J Qual Stud Health Well being 2014 May 28;9(1):23376 [FREE Full text] [doi: 10.3402/qhw.v9.23376] [Medline: 24875238]

130. Carron-Arthur B, Reynolds J, Bennett K, Bennett A, Griffiths KM. What's all the talk about? Topic modelling in a mental health internet support group. BMC Psychiatry 2016 Oct 28;16(1):367 [FREE Full text] [doi: 10.1186/s12888-016-1073-5] [Medline: 27793131]

131. Cavazos-Rehg PA, Krauss MJ, Sowles S, Connolly S, Rosas C, Bharadwaj M, et al. A content analysis of depression-related Tweets. Comput Human Behav 2016 Jan 01;54:351-357 [FREE Full text] [doi: 10.1016/j.chb.2015.08.023] [Medline: 26392678]

132. Cohn AM, Amato MS, Zhao K, Wang X, Cha S, Pearson JL, et al. Discussions of alcohol use in an online social network for smoking cessation: analysis of topics, sentiment, and social network centrality. Alcohol Clin Exp Res 2019 Jan;43(1):108-114 [FREE Full text] [doi: 10.1111/acer.13906] [Medline: 30326140]

133. Davis S, Lewis CA. Addiction to self-harm? The case of online postings on self-harm message boards. Int J Ment Health Addiction 2018 Sep 27;17(4):1020-1035. [doi: 10.1007/s11469-018-9975-8]

134. Dempsey M, Foley S, Frost N, Murphy R, Willis N, Robinson S, et al. Am I lazy, a drama queen or depressed? A journey through a pluralistic approach to analysing accounts of depression. Qual Res Psychol 2019 Oct 17;19(2):473-493. [doi: 10.1080/14780887.2019.1677833]

135. Feuston J. Algorithms, oppression, and mental illness on social media. In: Proceedings of the Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems. 2019 Presented at: CHI '19: CHI Conference on Human Factors in Computing Systems; May 4 - 9, 2019; Glasgow Scotland UK. [doi: 10.1145/3290607.3299072]

136. Mancera KC, De Santacruz C, Salamanca MA. ["Accepting demented minds". Opinion group, information and support on stigma of mental illness on Facebook]. Rev Colomb Psiquiatr 2014 Jul;43(3):139-145. [doi: 10.1016/j.rcp.2014.02.007] [Medline: 26575127]

137. Mitchell JT, Sweitzer MM, Tunno AM, Kollins SH, McClernon FJ. "I use weed for my ADHD": a qualitative analysis of online forum discussions on cannabis use and ADHD. PLoS One 2016 May 26;11(5):e0156614 [FREE Full text] [doi: 10.1371/journal.pone.0156614] [Medline: 27227537]

138. Schmidt M, Dillon PJ, Jackson BM, Pirkey P, Kedia SK. "Gave me a line of ice and I got hooked": exploring narratives of initiating methamphetamine use. Public Health Nurs 2019 Jan 18;36(1):18-27. [doi: 10.1111/phn.12568] [Medline: 30565331]

139. Summers RF. Adaptive and maladaptive means of using facebook: a qualitative pilot study to inform suggestions for development of a future intervention for depression. J Psychiatr Pract 2015 Nov;21(6):73-74 [FREE Full text] [doi: 10.1097/PRA.0000000000000109] [Medline: 26554330]

140. Valeriani G, Corazza O, Bersani FS, Melcore C, Metastasio A, Bersani G, et al. Olanzapine as the ideal "trip terminator"? Analysis of online reports relating to antipsychotics' use and misuse following occurrence of novel psychoactive substance-related psychotic symptoms. Hum Psychopharmacol 2015 Jul 28;30(4):249-254. [doi: 10.1002/hup.2431] [Medline: 26216558]

141. Sedgewick F, Leppanen J, Austin A, Tchanturia K. Different pathways, same goals: a large-scale qualitative study of autistic and non-autistic patient-generated definitions of recovery from an eating disorder. Eur Eat Disord Rev 2022 Sep 24;30(5):580-591. [doi: 10.1002/erv.2873] [Medline: 34820930]

142. Coulson NS, Buchanan H, Aubeeluck A. Social support in cyberspace: a content analysis of communication within a Huntington's disease online support group. Patient Educ Couns 2007 Oct;68(2):173-178. [doi: 10.1016/j.pec.2007.06.002] [Medline: 17629440]

143. Hucker A, McCabe MP. A qualitative evaluation of online chat groups for women completing a psychological intervention for female sexual dysfunction. J Sex Marital Ther 2013 Jun 14;40(1):58-68. [doi: 10.1080/0092623x.2012.675020]

144. Li A, Jiao D, Liu X, Zhu T. A comparison of the psycholinguistic styles of schizophrenia-related stigma and depression-related stigma on social media: content analysis. J Med Internet Res 2020 Apr 21;22(4):e16470 [FREE Full text] [doi: 10.2196/16470] [Medline: 32314969]

145. Mbao M, Collins-Pisano C, Fortuna K. Older adult peer support specialists' age-related contributions to an integrated medical and psychiatric self-management intervention: qualitative study of text message exchanges. JMIR Form Res 2021 Mar 02;5(3):e22950 [FREE Full text] [doi: 10.2196/22950] [Medline: 33650979]

146. Evans K, Rennick-Egglestone S, Cox S, Kuipers Y, Spiby H. Remotely delivered interventions to support women with symptoms of anxiety in pregnancy: mixed methods systematic review and meta-analysis. J Med Internet Res 2022 Feb 15;24(2):e28093 [FREE Full text] [doi: 10.2196/28093] [Medline: 35166688]

147. Cutrona CE, Suhr JA. Controllability of stressful events and satisfaction with spouse support behaviors. Commun Res 2016 Jun 30;19(2):154-174. [doi: 10.1177/009365092019002002]

148. Cohen S, Wills TA. Stress, social support, and the buffering hypothesis. Psychol Bull 1985 Sep;98(2):310-357. [Medline: 3901065]

149. Oakley A. Social Support and Motherhood (Reissue): The Natural History of a Research Project. Old Park Hill Bristol: Bristol University Press, Policy Press; 2018.

150. Tong ST, Heinemann-Lafave D, Jeon J, Kolodziej-Smith R, Warshay N. The use of pro-ana blogs for online social support. Eat Disord 2013 Oct;21(5):408-422. [doi: 10.1080/10640266.2013.827538] [Medline: 24044597]

151. Bauer R, Bauer M, Spiessl H, Kagerbauer T. Cyber-support: an analysis of online self-help forums (online self-help forums in bipolar disorder). Nord J Psychiatry 2013 Jun;67(3):185-190. [doi: 10.3109/08039488.2012.700734] [Medline: 22817138]

152. Morse J, Field P. Nursing Research The Application of Qualitative Approaches. Cham: Springer; 1996.

153. Gaysynsky A, Romansky-Poulin K, Arpadi S. "My YAP family": analysis of a Facebook group for young adults living with HIV. AIDS Behav 2015 Jun 4;19(6):947-962. [doi: 10.1007/s10461-014-0887-8] [Medline: 25186783]

154. Yip JW. Evaluating the communication of online social support: a mixed-methods analysis of structure and content. Health Commun 2020 Sep 03;35(10):1210-1218. [doi: 10.1080/10410236.2019.1623643] [Medline: 31154856]

155. Struik LL, Baskerville NB. The role of Facebook in Crush the Crave, a mobile- and social media-based smoking cessation intervention: qualitative framework analysis of posts. J Med Internet Res 2014 Jul 11;16(7):e170 [FREE Full text] [doi: 10.2196/jmir.3189] [Medline: 25016998]

156. Rimé B. Interpersonal emotion regulation. In: Handbook of Emotion Regulation. New York: The Guilford Press; 2007.

157. Liu Y, Kornfield R, Shaw BR, Shah DV, McTavish F, Gustafson DH. When support is needed: social support solicitation and provision in an online alcohol use disorder forum. Digit Health 2017 May 22;3:2055207617704274 [FREE Full text] [doi: 10.1177/2055207617704274] [Medline: 29942595]

158. Wang Y, Kraut RE, Levine JM. Eliciting and receiving online support: using computer-aided content analysis to examine the dynamics of online social support. J Med Internet Res 2015 Apr 20;17(4):e99 [FREE Full text] [doi: 10.2196/jmir.3558] [Medline: 25896033]

159. Willis J, Jost M, Nilakanta R. Foundations of Qualitative Research Interpretive and Critical Approaches. Thousand Oaks, California, United States: SAGE Publications; 2007.

160. O'Brien BC, Harris IB, Beckman TJ, Reed DA, Cook DA. Standards for reporting qualitative research: a synthesis of recommendations. Acad Med 2014 Sep;89(9):1245-1251 [FREE Full text] [doi: 10.1097/ACM.0000000000000388] [Medline: 24979285]

161. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]

162. Levitt HM, Bamberg M, Creswell JW, Frost DM, Josselson R, Suárez-Orozco C. Journal article reporting standards for qualitative primary, qualitative meta-analytic, and mixed methods research in psychology: the APA Publications and Communications Board task force report. Am Psychol 2018 Jan;73(1):26-46. [doi: 10.1037/amp0000151] [Medline: 29345485]

163. Cialdini R. Influence: Science and practice. London, United Kingdom: Pearson Education; 1984.

164. Ethical Guidelines for group psychotherapists. International Association for Group Psychotherapy and Group Processes. 2009. URL: https://www.iagp.com/about-new/about-new/ethical-guidelines-for-group-psychotherapists/ [accessed 2022-05-13]

165. Barclay L, Hilton GM. A scoping review of peer-led interventions following spinal cord injury. Spinal Cord 2019 Aug 23;57(8):626-635. [doi: 10.1038/s41393-019-0297-x] [Medline: 31123333]

166. Chaffey L, Bigby C. Health education by peers with spinal cord injury: a scoping review. J Dev Phys Disabil 2017 Sep 21;30(1):141-154. [doi: 10.1007/s10882-017-9569-6]

167. Chepkirui D, Nzinga J, Jemutai J, Tsofa B, Jones C, Mwangome M. A scoping review of breastfeeding peer support models applied in hospital settings. Int Breastfeed J 2020 Nov 14;15(1):95 [FREE Full text] [doi: 10.1186/s13006-020-00331-7] [Medline: 33189155]

168. Kowitt SD, Ellis KR, Carlisle V, Bhushan NL, Black KZ, Brodar K, et al. Peer support opportunities across the cancer care continuum: a systematic scoping review of recent peer-reviewed literature. Support Care Cancer 2019 Jan 6;27(1):97-108 [FREE Full text] [doi: 10.1007/s00520-018-4479-4] [Medline: 30293093]

169. Bowersox NW, Jagusch J, Garlick J, Chen JI, Pfeiffer PN. Peer-based interventions targeting suicide prevention: a scoping review. Am J Community Psychol 2021 Sep 15;68(1-2):232-248 [FREE Full text] [doi: 10.1002/ajcp.12510] [Medline: 33720444]

170. Schlichthorst M, Ozols I, Reifels L, Morgan A. Lived experience peer support programs for suicide prevention: a systematic scoping review. Int J Ment Health Syst 2020 Aug 12;14(1):65 [FREE Full text] [doi: 10.1186/s13033-020-00396-1] [Medline: 32817757]

171. Carter G, Monaghan C, Santin O. What is known from the existing literature about peer support interventions for carers of individuals living with dementia: a scoping review. Health Soc Care Community 2020 Jul 09;28(4):1134-1151. [doi: 10.1111/hsc.12944] [Medline: 31919974]

172. Newman K, Wang AH, Wang AZ, Hanna D. The role of internet-based digital tools in reducing social isolation and addressing support needs among informal caregivers: a scoping review. BMC Public Health 2019 Nov 09;19(1):1495 [FREE Full text] [doi: 10.1186/s12889-019-7837-3] [Medline: 31706294]

173. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Social Res Methodol 2005 Feb;8(1):19-32. [doi: 10.1080/1364557032000119616]

174. p2p-online-interactions. GitHub. URL: https://github.com/pawljmlo/p2p-online-interactions [accessed 2022-11-28]

Abbreviations

AI: artificial intelligence
CASP: Critical Appraisal Skills Programme
ML: machine learning
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews

XSL•FO
RenderX

XSL•FO
RenderX