JMIR Mental Health

Journal Impact Factor (JIF) (2023): 4.8
Volume 8 (2021), Issue 8 ISSN 2368-7959 Editor in Chief: John Torous, MD, MBI

Contents

Original Papers

Viewpoint

Reviews

Letter to the Editor

Original Paper

Deep Learning With Anaphora Resolution for the Detection of Tweeters With Depression: Algorithm Development and Validation Study

Akkapon Wongkoblap[1,2,3*], MSc; Miguel A Vadillo[4,5*], PhD; Vasa Curcin[1,4*], PhD

[1]Department of Informatics, King's College London, London, United Kingdom
[2]DIGITECH, Suranaree University of Technology, Nakhon Ratchasima, Thailand
[3]School of Information Technology, Suranaree University of Technology, Nakhon Ratchasima, Thailand
[4]School of Population Health and Environmental Sciences, King’s College London, London, United Kingdom
[5]Departamento de Psicología Básica, Universidad Autónoma de Madrid, Madrid, Spain
[*]all authors contributed equally

Corresponding Author:
Akkapon Wongkoblap, MSc
DIGITECH
Suranaree University of Technology
111 University Avenue, Muang
Nakhon Ratchasima, 30000
Thailand
Phone: 66 44224336
Email: wongkoblap@sut.ac.th

Abstract

Background: Mental health problems are widely recognized as a major public health challenge worldwide. This concern highlights the need to develop effective tools for detecting mental health disorders in the population. Social networks are a promising source of data wherein patients publish rich personal information that can be mined to extract valuable psychological cues; however, these data come with their own set of challenges, such as the need to disambiguate between statements about oneself and third parties. Traditionally, natural language processing techniques for social media have looked at text classifiers and user classification models separately, hence presenting a challenge for researchers who want to combine text sentiment and user sentiment analysis.

Objective: The objective of this study is to develop a predictive model that can detect users with depression from Twitter posts and instantly identify textual content associated with mental health topics. The model can also address the problem of anaphoric resolution and highlight anaphoric interpretations.

Methods: We retrieved the data set from Twitter by using a regular expression or stream of real-time tweets comprising 3682 users, of which 1983 self-declared their depression and 1699 declared no depression. Two multiple instance learning models were developed—one with and one without an anaphoric resolution encoder—to identify users with depression and highlight posts related to the mental health of the author. Several previously published models were applied to our data set, and their performance was compared with that of our models.

Results: The maximum accuracy, F1 score, and area under the curve of our anaphoric resolution model were 92%, 92%, and 90%, respectively. The model outperformed alternative predictive models, which ranged from classical machine learning models to deep learning models.

Conclusions: Our model with anaphoric resolution shows promising results when compared with other predictive models and provides valuable insights into textual content that is relevant to the mental health of the tweeter.

(JMIR Ment Health 2021;8(8):e19824) doi:10.2196/19824

KEYWORDS

depression; mental health; Twitter; social media; deep learning; anaphora resolution; multiple-instance learning; depression markers

Introduction

Background

Mental health problems are widely recognized as major public health challenges worldwide. According to the World Health Organization, 264 million people were affected by depression globally in 2020 [1]. Mental illness, in general, is one of the leading causes of the global burden of this disease. It was estimated that in England, 105 billion British pounds (US $145 billion) were spent on mental health services and treatments or lost in productivity at work in 2018 [2], with the global costs expected to rise to US $6 trillion by 2030 [3]. A significant contributor to this cost is that people living with mental health problems sometimes receive inaccurate assessments [1]. This highlights the need for effective mental health services and a novel approach for diagnosing mental health disorders.

User-generated content on social media, reviews, blogs, and message board platforms offers an opportunity for researchers to explore and classify the huge amount of content in different domains, such as marketing [4], politics [5], and health [6-8], thereby providing a rapid method to understand user-created text and expressed emotion using text classification algorithms. Social networking (eg, Facebook and LinkedIn) and microblogging platforms (eg, Twitter and Tumblr) provide internet users with a safe space to post their feelings, thoughts, and activities. With some users publicly expressing their mental health statuses on their profiles, it becomes possible to train classification engines to detect internet users with mental health problems [9,10]. Using Twitter data, in particular, studies have examined users with depression [11-14], postpartum depression [15], anxiety, obsessive compulsive disorder, and posttraumatic stress disorder [11,16]. In addition, Facebook data were also used to detect users with depression [17,18] and postpartum depression [19].

Generally, text classifiers and user classification models tend to be developed separately. This presents a challenge for researchers who want to simultaneously understand both text sentiment analysis and user sentiment analysis. In this paper, we present a predictive model that can detect users with depression and identify their tweets as those related to health. An ideal technique for developing this type of model is multiple instance learning (MIL) [20], where the model can learn from a set of labeled bags or users instead of a set of individual instances or user-generated messages.

Anaphora resolution is an established natural language processing (NLP) problem and an emerging field in the analysis of social media content that helps with determining which previously mentioned person is the subject of a subsequent statement and understanding references to someone in the content on social media. This is particularly relevant to social media, as posts may frequently refer to individuals other than the tweeter [21].

Objectives

To the best of our knowledge, no study has focused on detecting users with depression on social networks with an anaphoric interpretation of the content. In this study, we aim to address the problem of anaphora resolution in user-generated content and present a predictive model that can reliably identify statements, thoughts, and attitudes relating to the tweeter, rather than a third party.

The objective of this study is to investigate whether user-generated content from Twitter can be used to detect users with depression. This raises three research questions:

1. Can MIL be used to develop a predictive model for detecting users with depression from their tweets?
2. Can sentiments of unlabeled tweets be predicted from the labels of users with depression?
3. Can anaphora resolution be combined with MIL to eliminate false positives?

This paper introduces MIL models with and without anaphora resolution to detect users with depression from their generated textual content on Twitter and predictive models that can highlight posts relevant to mental health. The results show that our algorithm outperforms the major recently published algorithms in the field. We further illustrate the differences in the tweets related to mental health from users with self-declared depression and users with no depression.

This Study

This study focuses on text analysis, predictive models for detecting social network users with mental disorders, and MIL. The most relevant studies published to date are reviewed below.

Text analysis is an NLP approach for identifying information within text. This technique has been developed to understand the textual content automatically and computationally. During the early stages of sentiment and emotion analysis, researchers manually annotated the text [22]. With the possibility of identifying emotions in text, the content has been computationally analyzed using a keyword or corpus-based approach and a learning-based approach [23,24].

The learning-based approach uses a predictive model to determine the relationship between an input and output word. Word embedding is a common learning-based technique that transforms the words of a document into dimensional vectors for word representation and determines word similarity. Global Vectors for Word Representation (GloVe) is a word-embedding approach that computes and aggregates word co-occurrence for representing the closest linguistic or semantic similarity between co-occurrent words as vectors [25]. GloVe was trained on several textual data sets, such as Wikipedia and common crawl (a copy of web content), and supported 50D, 100D, 200D, and 300D vectors.

Anaphora resolution is another text analysis problem related to determining which person is mentioned within textual content. There are three reference resolution algorithms [26]. The rule-based entity resolution extracts syntactic rules and semantic knowledge from the text. The statistical and machine learning–based entity resolution is a method to understand the coreference of a reference to an early entity. Deep learning for entity resolution reduces handcrafted feature requirements and represents words as vectors conveying semantic units. Aktaş et al [21] investigated anaphora resolution for conversations on

Twitter using a corpus and manual annotation. Twitter conversations revealed the cues of anaphora resolution to identify a mentioned person and provide context.

De Choudhury and Gamon [13] pioneered NLP and machine learning approaches for developing predictive models to detect users with mental disorders from social network data using a mental health screening questionnaire and linguistic analysis tools to extract emotional words and web-based behaviors from users' posts. However, the screening and data collection process was time consuming, and Coppersmith et al [11] introduced an automatic data gathering method using keywords to find the target users and programmatically retrieve the posts.

Following these initial studies, a number of novel methods have emerged for predicting mental disorders in social network users. The early work focused on classical supervised machine learning techniques and traditional text analysis approaches.

The psychometric analysis of textual content was used to compute the percentage of emotional, functional, and social concern words [13,15]. Linguistic inquiry and word count (LIWC) was used to compute the percentage of words relevant to categories from each tweet. The extracted percentages were then used to train a predictive model based on a support vector machine with a radial basis function [13].

Language models have been applied to analyze social media texts to address spelling errors, shortenings, and emoticons [11]. The language model was developed from an n-gram, which learns from the sequences of text and computes the probability of unseen text relevant to a category of the trained model. This model scored the probabilities of users with depression based on a higher probability of the positive class language model trained from the tweets of users with depression or the negative class language model developed from the tweets of control users [11].

A predictive model based on topic models was developed from the social network profiles of clinically diagnosed patients [17]. The topic model used latent Dirichlet allocation to extract topics from the text. All tweets from each user were used to compute 200 topics, which were then used to develop a logistic regression model for classifying the users with depression [17].

Building on the popularity of neural networks, novel models have been developed using word embedding [27,28] and deep neural network models [28]. The Usr2Vec model transformed text into an embedding matrix, where words commonly used together were represented in closely dimensional spaces for classifying users. The embeddings were learned from users' tweets and then summarized as user representations. The embedding matrices were used to train a predictive model using a multinomial logistic regression technique [27].

The deep learning model uses word embeddings to represent the sequential words of users' tweets. A predictive model was trained using a 1D convolutional neural network (CNN) and a global max pooling layer [28].

In addition to the textual content of the posts, a number of writing features can be analyzed: post or blog lengths, time gap between consecutive posts, and day of the week and time of the day of postings. Further network features of interest include likes, numbers of followers or following, characteristics of comments on other users' posts compared with original posts, and numbers of shares or retweets. Image analysis was used to characterize user posts [29,30].

To develop a predictive model, this study focused on MIL. It is a supervised learning technique first proposed by Keeler et al [31,32]. Although classical supervised learning requires an instance and a single label to learn during the training process, MIL can learn from a bag of instances $X=x_1, x_2,\dots, x_N$. Each instance x_n can be independent and has its own individual label, y_n, where $y_n \in \{0, 1\}$ for $n=1,\dots, N$, and it is assumed that each y_n is unknown during the training process. On the basis of these assumptions, an MIL classifier can predict a label Y for a given bag X as follows:

On the basis of these assumptions, MIL can provide an extreme result $Y=1$ in the case of having a predicted positive-instance label $y_n=1$ in a given input X. The relaxation of the MIL assumption can be computed using aggregated probabilistic distributions of instances, where $Y=P(x_n)$ for $n=1,\dots, N$.

The purpose of MIL is to facilitate the development of a predictive model for detecting social media users with depression and instantly label each of the posts associated with either mental health or other topics. Normally, data sets from social networking are labeled at the user level but not at the post level. This makes it difficult to find a change in patterns in the message topics posted on social networks.

MIL models have been widely applied to image classification [32], object detection [33], image annotation [34], medical image and video analysis [35,36], sentence selection [37], and document classification [38]. In document sentiment analysis, Angelidis and Lapata [20] proposed the MIL network (MILNET) to classify web-based review documents and instantly identify the sentiment polarity of each segment of given documents. MILNET comprises segment encoding, segment classification, and document classification via an attention mechanism. Segment encoding transformed sentences in a document into segments via word-embedding matrices and a CNN. Each segment representation was classified using a softmax classifier. An attention mechanism based on a bidirectional gated recurrent unit (GRU) was used to weight the important segments to make a final document prediction as the weighted sum of the segment distributions. MILNET performed well in predicting the sentiment of a document and identifying the sentiment of the text segments but was not as successful in identifying a person mentioned in the document.

In this study, we adopt the MIL approach to develop two models, namely multiple instance learning for social network (MIL-SocNet) and multiple instance learning with an anaphoric resolution for social network (MILA-SocNet), to classify users with depression and highlight published posts associated with the mental health topic of a tweeter. Both models use novel document segment encoding, a tweet encoder, and user representation rather than a document vector. The latter model also includes the anaphora resolution, which further improves the performance.

Methods

Data Set

The data set was retrieved from Twitter, which provides an application programming interface (API) to search public tweets using regular expressions or stream real-time tweets. This study collected only tweets and users set as public. All collected tweets and users were anonymized. This study was approved by the King's College Research Ethics Committee (reference number LRS-16/17-4705).

We selected a group of users with depression using the method proposed by Coppersmith et al [11]. Specifically, a regular expression was used to search tweets that contained the statement "I was diagnosed with depression" between January and May 2019. This resulted in 4892 tweets from 4545 unique users, who were then manually screened to ensure that the tweets did not refer to jokes, quotes, or someone else's depression symptoms. After removing these messages, all tweets in the profiles of the users who posted the tweets were downloaded. After verification, 2132 unique users were included in this data set.

A control group was randomly selected from a list of 2036 users who posted tweets between June 1 and June 7, 2019. Users from the group with depression were removed from the list of the control group.

The limits imposed by the Twitter API allowed us to only download the 3200 most recent tweets of all verified users from the depressed and control groups. In total, 5 million tweets were collected from the 2132 users with depression and 4.2 million tweets from the 2036 users with no declared depression.

Preprocessing

Before developing our MIL model, several transformations were performed on the data set. First, the user ID in each tweet was replaced by a generic *user*. Similarly, any numbers mentioned in tweets were replaced by the *number* and any specific URLs by *url*. The # character in each hashtag was replaced by the string *hashtag* (eg, *#depression* became *hashtag depression*). Finally, users with fewer than 100 tweets or less than 80% of tweets in English were removed from the data set, resulting in 3682 users, 1983 with declared depression and 1699 with no declared depression, as depicted in on the left-hand side of Figure 1. In addition, other dimensions of the data set were explored, as shown in Figure 1. Figure 2 illustrates the distribution of the number of tweets between the depressed and control groups. Slight differences were present between the control and depressed groups.

All tweets in our final data set were embedded from pretrained GloVe word vectors. GloVe is an unsupervised machine learning approach and an NLP technique that represents a word as a set of word vectors. GloVe computes and aggregates word co-occurrences to create a vector representation of the closest linguistic or semantic similarity between co-occurrent words [22]. As explained earlier, GloVe was trained on several textual data sets, for example, Wikipedia and common crawl (a copy of web content), and supported 50D, 100D, 200D, and 300D vectors. However, our study used pretrained word vectors trained on 2 billion tweets and 100D vectors to transform our tweets into word embedding.

Figure 1. Analysis of data set statistics. The left side shows the percentages of users and tweets between control users and users with depression, where the inner circle presents the number of users and the outer circle presents the number of posts. The middle shows the number of words per post between 2 groups. The right side shows the ratio of retweets to tweets per user between the classes. Blue denotes the control group, and orange represents the depressed group.

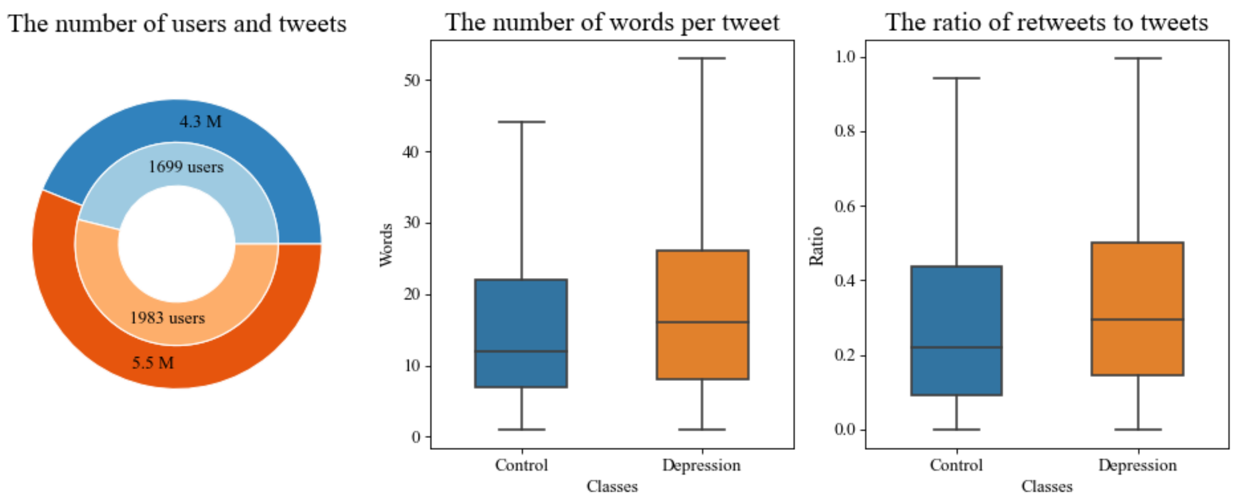
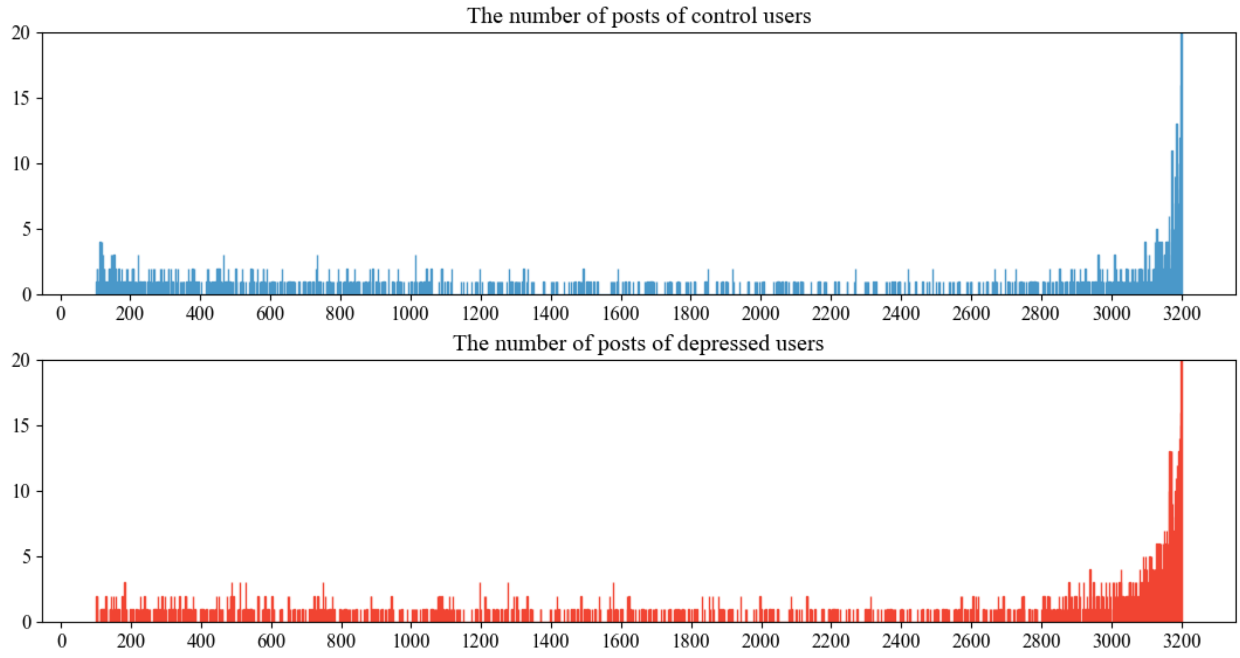


Figure 2. The distribution of the number of tweets between the depressed and control groups. This only shows a maximum of 20 tweets for clarity. The depression group with 3200 tweets had 436 users, and the control group with 3200 tweets had 485 users.



Predictive Model

Overview

This section describes the structure of our predictive model to classify a Twitter user with depression. This section will explain how an MIL model with supervised neural networks classifies users and provides a changing pattern of generated text associated with mental health or other topics.

Our proposed MIL-SocNet architecture comprises a tweet encoder, word attention on a tweet, tweet classification, a user encoder, tweet attention, and user classification (Figure 3). The differences between MIL-SocNet and the basic MILNET architecture are the tweet encoder and word attention, respectively. Our model uses a GRU, whereas MILNET uses a CNN and does not have an attentional mechanism.

Furthermore, the MIL-SocNet model was extended with an anaphoric resolution to create the MILA-SocNet model. We present this model to improve performance by adding an anaphora resolution encoder to ensure that the algorithm focuses on posts related to the author (Figure 4).

Figure 3. The structure of our proposed multiple instance learning-SocNet.

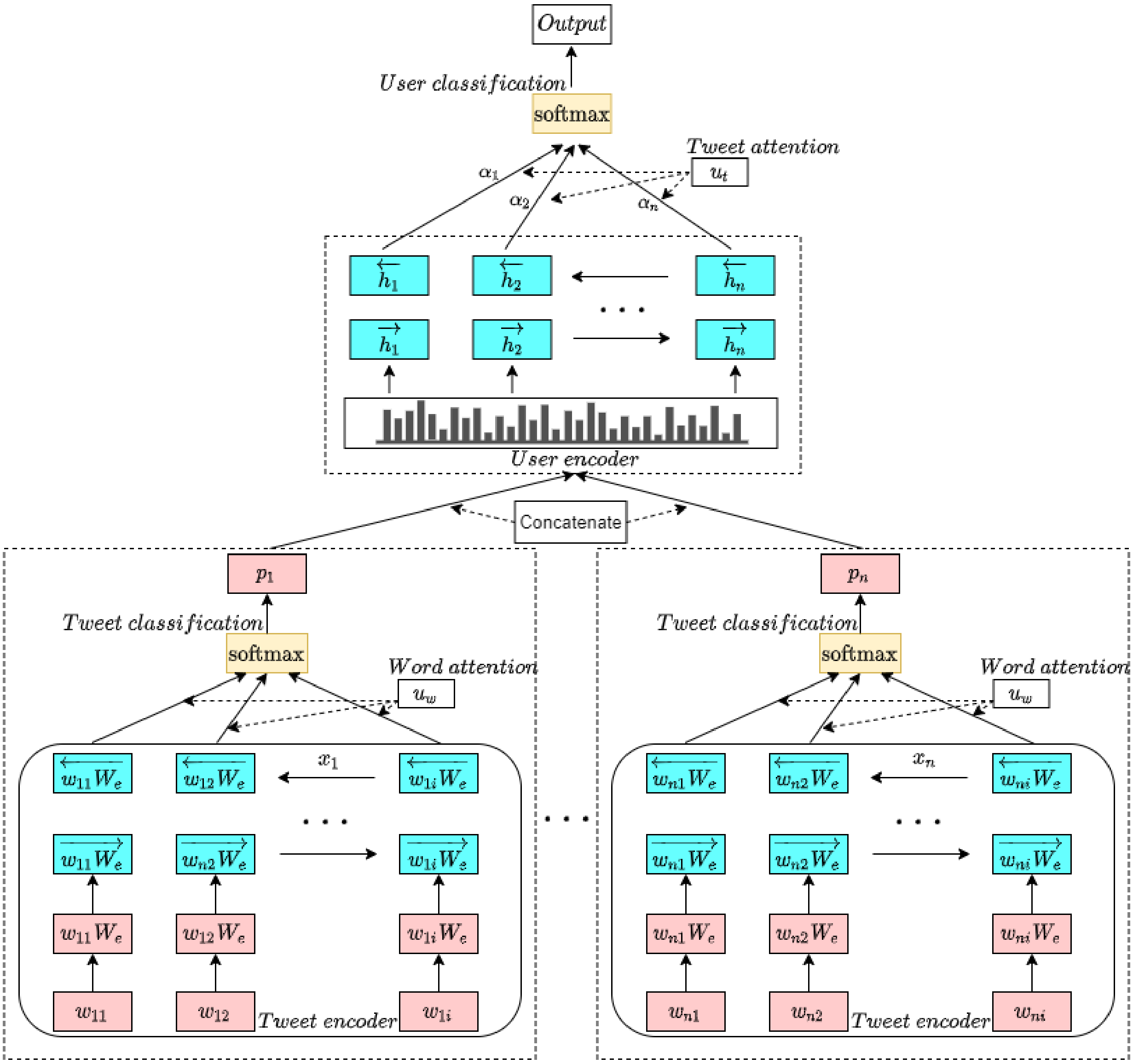
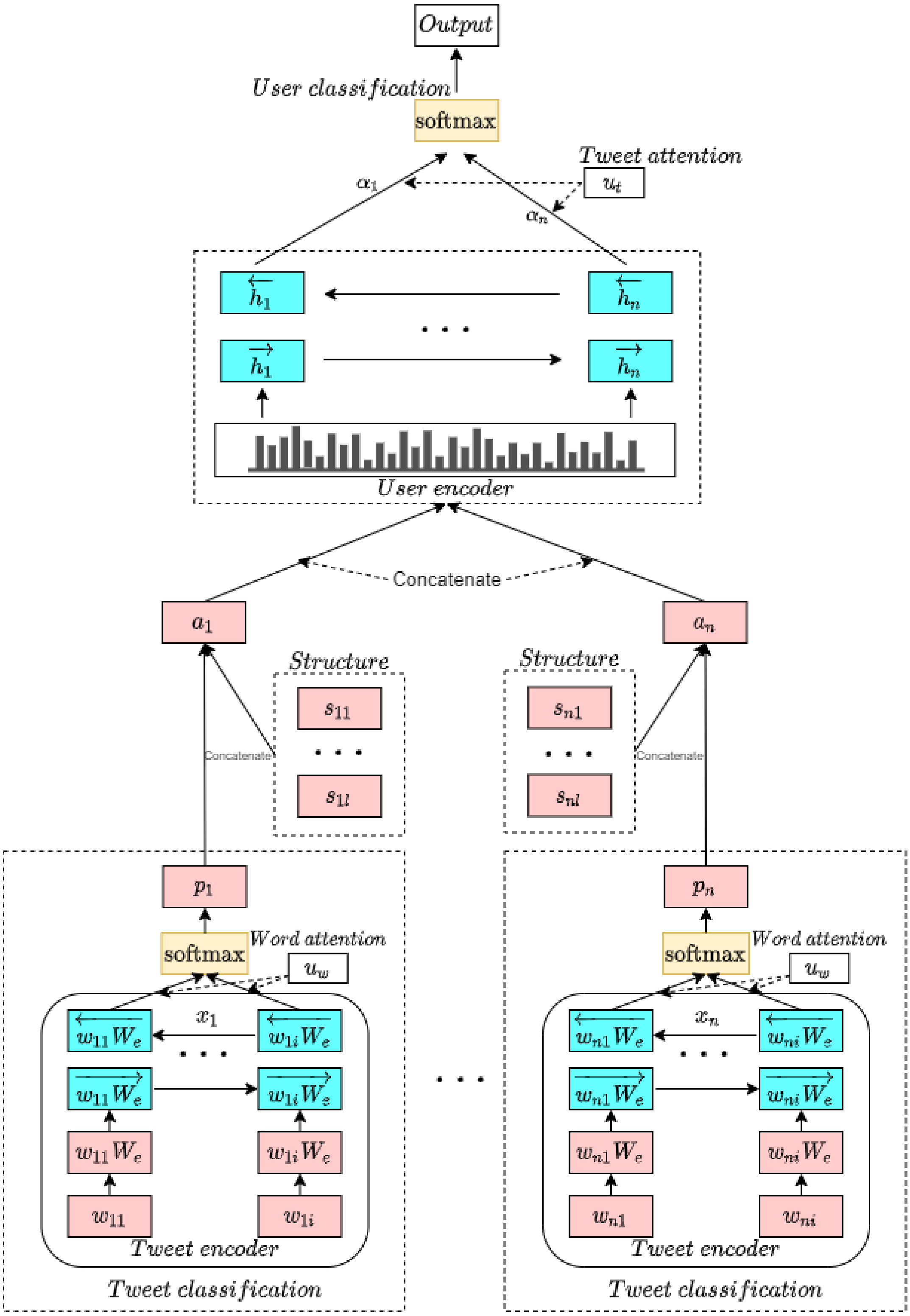


Figure 4. The structure of our proposed MILA-SocNet.



Tweet Encoder

The first layer of our proposed model transforms each tweet into a machine-readable form. First, tweets were transformed into word-embedding matrices. Each user publishes *j*=1, 2,…, *n* tweets, where *n* is the number of tweets used to train the model. Each tweet contains *k*=1, 2,…, *i* words, where *i* is the number of words in each tweet and varies from post to post. W_{jk} represents the *k*th word in the *j*th tweet. Every w_{jk} is then embedded through an embedding matrix W_e to be received a word vector x_{jk}. This layer embeds all words w_{jk} of jth post to the word vector:

$$x_{jk} = w_{jk}W_e, j \in [1, n] \text{ and } k \in [1, i]$$

The abovementioned equation operates times. After embedding all words, a bidirectional GRU is used to encode the vector:

The bidirectional GRU presents a hidden representation of h_{jk}, which is concatenated from and . The word hidden vector h_{jk} is then sent to an attention mechanism to select the important words.

Word Attention on a Tweet

Not every word equally represents tweet meanings. An attention mechanism is used to select words that best capture the relevant meaning of a tweet. The attention layer comprises a tanh function to produce an attention vector u_{jk} of the *k*th word in the *j*th tweet, where W_w and b_w are weights and bias, respectively.

$$u_{jk} = tanh\ (W_w h_{jk} + b_w)$$

The importance of words or attention weights a_{jk} is calculated via the normalized similarity of u_{jk} with the context vector of the word level u_w, which is learned and updated during the training step.

Finally, the tweet vector t_j is computed using the weighted sum of word importance with the hidden representation of h_{jk} generated from the bidirectional GRU.

Tweet Classification

To make a prediction about a tweet related to either a mental health or another topic, each tweet vector t_1, t_2,..., t_n from the attention layer is classified through a softmax function [39].

The function generates the probabilities of tweet labels , where with 1 denoting a mental health–related post and 0 denoting a non–mental health–related post. The labels used to train this layer are derived and computed from the labels of the user level only. The parameters W_c and b_c are learned and updated during the training step. Every predicted tweet label is used to teach a predictive model and detect a user with depression.

User Encoder

Detecting users with depression requires a pattern to differentiate between user groups. To predict these users, this study used a temporal pattern of posting generated from the tweet classification layer. This layer concatenates the probabilities of every classified tweet label into a single list of label probabilities called *user representation*. The user representations between the 2 groups are expected to differ, which will be explored and illustrated in the Discussion section. Then, user representation is passed through a bidirectional GRU to learn the changing patterns of text categories over the observation time. This generates the forward hidden state and the backward hidden state of the user representation. Finally, they were concatenated to h_j.

Anaphora Resolution Encoder

For the MILA-SocNet model with anaphora resolution, pronoun features from LIWC [40] are used to add informative interpretations to each tweet. Every tweet was analyzed for emotions, thinking styles, social states, parts of speech, and psychological dimensions.

Each tweet is combined between the extracted pronoun features s_j from the LIWC and a tweet classified label p_j from the tweet classification layer, where with represents the extracted features in the *j*th tweet. This yields the following anaphora resolution vector:

The vector is then passed through a bidirectional GRU to learn the text category and anaphoric features. This generates h_j combined from the forward hidden state and backward hidden state .

Tweet Attention

Not all user tweets were equally associated with depression. Some tweets may contain cues relevant to depression, whereas others may not. For this purpose, an attention mechanism is applied to reward tweets that correctly represent the characteristics and are important for correctly detecting a user with depression. This layer performs similarly in both MIL-SocNet and MILA-SocNet. A multilayer perceptron (MLP) produces the attention vector u_j of the *j*th tweet. The parameter W_t denotes the weights of the tweet and parameter b_t represents the bias of the tweet.

The attention weights of tweets or important tweets α_j are computed through the similarity of u_j with the context vector of tweet level u_t, which is learned and updated during the training step.

The user vector *v* is achieved by summarizing all the information of the tweet label possibilities of a user.

User Classification

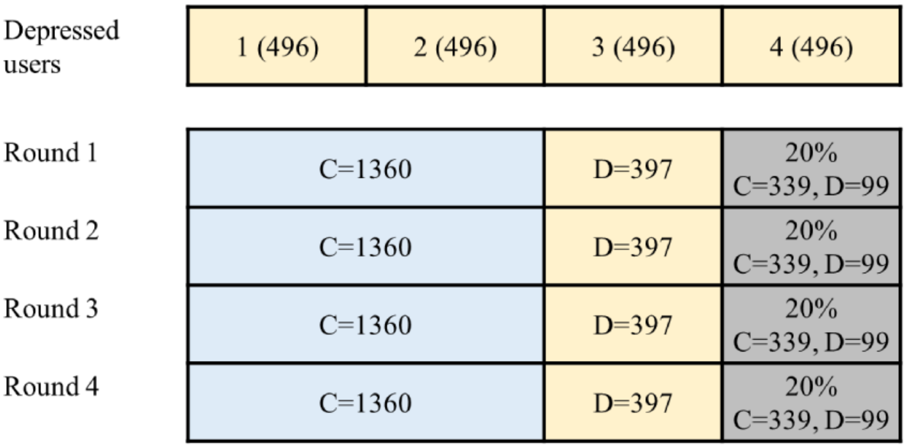
Finally, a predictive model for detecting a user with depression can be achieved through the user vector *v* derived from encoding the concatenation of the probabilities and the attention weights of the classified tweet labels from the user. A softmax function was again used to perform the classification.

Training the MIL Model

To train MILA-SocNet and MIL-SocNet, we used the Keras library with TensorFlow backend, a Python library for neural network APIs. We used an adaptive and momental bound method (AdaMod) [41], and the binary cross-entropy loss function to minimize loss. Every tweet from each user was tokenized and limited to 55 tokens or words. The model was trained using 2000 recent tweets from each user, with users with fewer than 2000 tweets having empty tweets padded with 0 values to achieve the matching length. To eliminate overfitting, dropout and early stopping were applied to the model during the training step.

Both our models and replicated models were trained and tested with holdout cross-validation. We split the users experiencing depression into four equal chunks and trained the models against all control users. Thus, each round used 496 users experiencing depression (22.60%) and 1699 control users (77.40%), mirroring the real-world incidence of depression. From the total users included in each round, 20% were used as test sets to evaluate the performance of the models. To reserve the same proportions of classes between the training and test tests, stratified cross-validation was used. Figure 5 shows the cross-validation process.

Figure 5. Holdout cross-validation on our experiment. C denotes control users and D represents users with depression. Blue, yellow, and gray represent control data, chunks of users with depression, and test sets, respectively.



Model Evaluation

To predict whether each Twitter user was likely to be depressed, we also trained a set of published predictive models ranging from classical machine learning to deep learning techniques by using user-generated textual content. Accuracy, precision, recall, and F1 scores were averaged across the test sets. Each model was trained and tested with the same samples in each round; however, data transformations differed in some cases, as explained in the Background section.

To compute the predictive performance of models for detecting social network users with depression, we used the following metrics:

To further compare the performance of MILA-SocNet and MIL-SocNet with the other published models, Akaike information criterion (AIC) was applied across all the models. AIC is a commonly used tool for model comparison and selection [42,43] that measures the information loss in each model, considering the model's complexity as well. AIC is defined as follows:

where *n* is the number of samples and *K* is the number of parameters or features of a model. denotes the natural logarithm of likelihood [44]. The equation also uses bias adjustment because of the small sample size [45,46]. A lower AIC value indicates better performance.

Results

This section shows the performance of MILA-SocNet and MIL-SocNet and compares their results in terms of accuracy, precision, recall, and F1 score against several published models including LIWC [13], language [11], topic [17], Usr2Vec [27], and deep learning [28] models, as explained in the Background section.

Table 1 shows the performance of our proposed MILA-SocNet and MIL-SocNet models against the alternative models. As observed, the MILA-SocNet achieves a maximum accuracy (92%), precision (92%), recall (92%), and F1 score (92%), immediately followed by the MIL-SocNet. The MIL-SocNet yielded an accuracy, precision, recall, and F1 score of 90%, 91%, 90%, and 90%, respectively. Each model was evaluated using the area under the curve of the receiver operating characteristic curve. As can be seen in Figure 6, the MILA-SocNet and MIL-SocNet models achieved the highest areas under the curve—93% in both cases. It should be noted that in those studies, the replicated models were reported with different proportions of classes. These results might be higher or lower than our reported results. In our study, the baseline result was 77% in the case of predicting the majority class in all cases. As can be observed, all the models achieved results that were above this baseline.

Table 2 lists the AIC values for each model. The likelihood was computed from the model-based probabilities of the observed labels. The number of parameters of the MILA-SocNet, MIL-SocNet, and deep learning models were recovered from the number of trainable parameters reported by the Keras library. The number of parameters of the language model was taken from the number of vocabularies in the positive and negative language models. The number of parameters of LIWC, Usr2Vec, and topic models were features in the models. The likelihoods and AICs were averaged from cross-validation, as explained earlier. As can be observed, MILA-SocNet achieves the lowest AIC, reflecting the best performance.

Table 1. Performance of our proposed MILA-SocNet (multiple instance learning with an anaphoric resolution for social network) and MIL-SocNet (multiple instance learning for social network) models and all replicated models.

Model	Accuracy, %	Precision	Recall	F1 score
MILA-SocNet	92.14	0.92	0.92	0.92
MIL-SocNet	90.49	0.91	0.90	0.90
Deep learning	89.07	0.89	0.89	0.89
Usr2Vec	84.38	0.84	0.84	0.83
LIWC[a]	83.31	0.83	0.83	0.81
Language	81.61	0.80	0.82	0.79
Topic	80.13	0.78	0.80	0.78

[a]LIWC: linguistic inquiry and word count.

Figure 6. Receiver operating characteristic curves of each model. Area under the curve with SDs of each model are denoted by different colors. The x-axis shows the false-positive rate, and the y-axis presents the true-positive rate. The dashed line indicates a random guess. AUC: area under receiver operating curve; DL: deep learning model; LIWC: linguistic inquiry and word count; LM: language model.

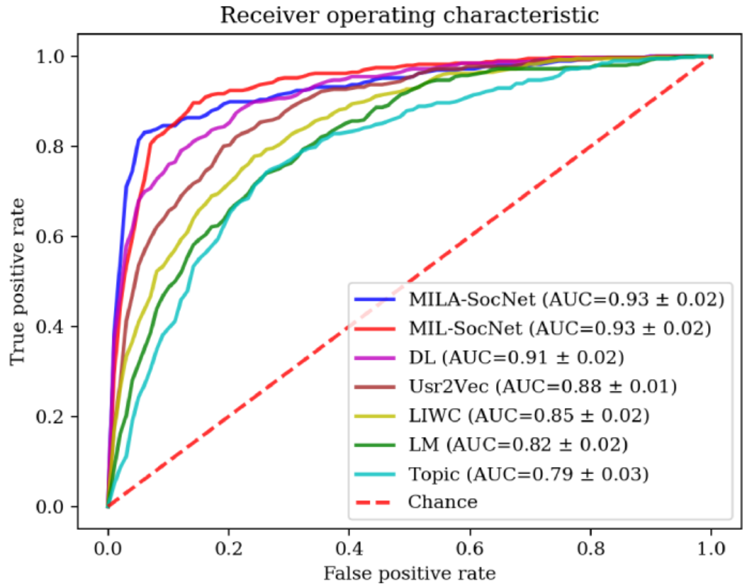


Table 2. The Akaike information criterion (AIC) results against all models. Each row is reported with the number of parameters (K), the residual sum of squares, and the AIC. A lower AIC is better.

Model	Number of parameters, K	Likelihood	AIC
MILA-SocNet[a]	59,668	−143.72	−597.05
MIL-SocNet[b]	56,296	−210.22	−464.45
Deep learning	138,502	−309.97	−260.84
Language	16695.5	−420.31	−61.03
LIWC[c]	93	−169.62	575.92
Usr2Vec	100	−190.28	640.32
Topic	200	−276.42	1290.66

[a]MILA-SocNet: multiple instance learning with an anaphoric resolution for social network.

[b]MIL-SocNet: multiple instance learning for social network.

[c]LIWC: linguistic inquiry and word count.

Discussion

Principal Findings

In this study, we presented two novel MIL models for detecting social network users with depression based on their self-identifying tweets. The original MIL-SocNet model was extended with anaphoric resolution to produce the second MILA-SocNet model. We also compared the performance of both models with that of several previously published models. As can be seen from Tables 1 and 2, MILA-SocNet and MIL-SocNet outperformed all other models in all metrics. We now look at several potential reasons for this result.

Although deep learning models can be trained on raw textual data, traditional machine learning models (eg, the LIWC, language, topic, and Usr2Vec models) require feature extraction to be performed using external tools, which may introduce the additional risk of losing useful information from short textual data [47,48]. For instance, misspelled and abbreviated words in tweets may not be present in the dictionary of an extraction tool, resulting in the mislabeling of words. This may be one of the reasons why traditional machine learning techniques performed worse than our proposed models.

Another reason for the performance gap may be that the sequential ordering of words in a tweet and tweets posted on a timeline may influence model performance. Training a predictive model with traditional machine learning methods requires aggregated data, which may cause the loss of contextual information compared with deep neural networks that can learn from the sequential information in the data [49-52].

Unlike the deep learning model that we have compared against [28], MILA-SocNet and MIL-SocNet used an attention mechanism that highlights words and tweets relevant to mental health. This attention mechanism may have contributed to our proposed models outperforming the deep learning model, even though our approach is also based on deep learning techniques.

Another important point to consider is that the addition of anaphoric resolution improves the performance of the base MIL model. The difference between MILA-SocNet and MIL-SocNet is only in anaphora resolution encoding, which highlights posts related to the tweeters rather than someone else. This is an important feature that has not been widely investigated in the field and should be considered while designing future studies.

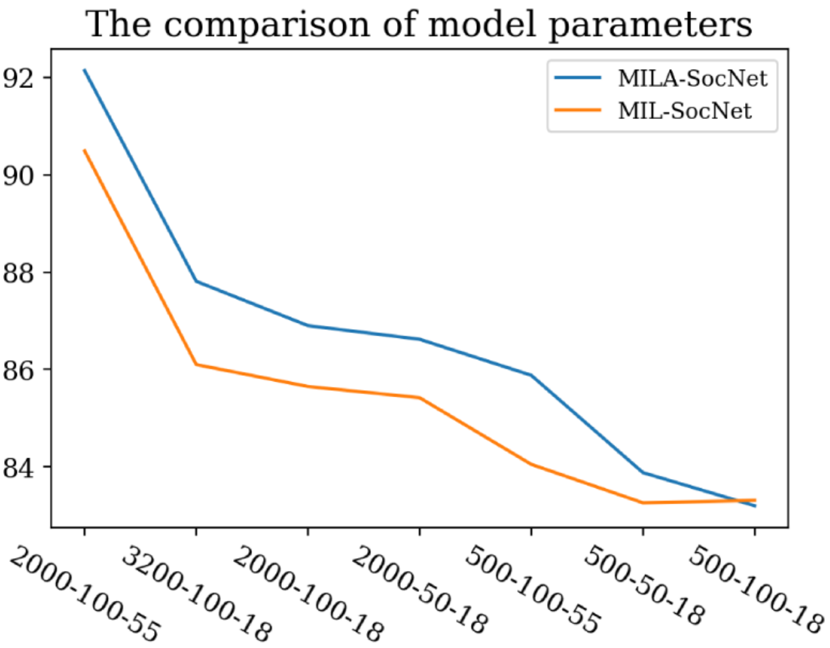
We further explored our proposed models by comparing the model performance under different conditions. A set of different parameters was used to train the models. The number of each user's posts used to train a model ranged from 500 to 3200 posts. The numbers of embedded dimensions were 50 and 100. The lengths of word tokens used to train the models were 18 and 55 tokens, respectively. Table 3 and Figure 7 show the predictive results of MILA-SocNet and MIL-SocNet with different parameters. Longer post length and longer word token provide better results, which is expected as these provide more textual content. Furthermore, models with fewer embedded dimensions perform worse than models with more dimensions.

After training the models, we investigated their interpretability by observing the attention weights to find out which tweets the model paid most attention to. Two users from each group were randomly selected from those correctly labeled by our model, and attention weights were extracted from the tweet attention layer. Textbox 1 highlights the tweets that achieved the highest and lowest weights for these 4 users, offering some insight into model's decision-making. Our predictive model with anaphoric resolution can identify tweets related to the tweeters' own experiences.

Table 3. Performance of MILA-SocNet (multiple instance learning with an anaphoric resolution for social network) and MIL-SocNet (multiple instance learning for social network) with different parameters. The first number in the model name (first column) represents the number of posts, the second is the number of embedded dimensions, and the last is the number of word tokens.

Model name	MILA-SocNet models				MIL-SocNet models			
	Accuracy, %	Precision	Recall	F1 score	Accuracy, %	Precision	Recall	F1 score
2000-100-55	92.14	0.92	0.92	0.92	90.49	0.91	0.90	0.90
500-100-55	85.88	0.86	0.86	0.84	84.05	0.83	0.84	0.83
3200-100-18	87.81	0.87	0.88	0.88	86.10	0.85	0.86	0.86
2000-100-18	86.90	0.86	0.87	0.86	85.65	0.85	0.86	0.85
500-100-18	83.20	0.82	0.83	0.82	83.31	0.83	0.83	0.81
2000-50-18	86.62	0.86	0.87	0.86	85.42	0.85	0.85	0.85
500-50-18	83.88	0.83	0.84	0.83	83.26	0.83	0.83	0.82

Figure 7. Results from different model parameters. Y-axis is the accuracy of the models. X-axis represents the number of posts, embedded dimensions, and post tokens in each model.



Textbox 1. Attention weights of posts. The “text” was paraphrased to anonymize users’ identities.

Users with depression

- User 1
  - Highest weight: I was also dealing with depression and anxiety badly. School was hell.
  - Lowest weight: @user Exam without someone’s supervision is bad.
- User 2
  - Highest weight: I get some rest, take medication, and engage with what I like. These help me and I do not force myself to do things.
  - Lowest weight: Talk about offensive things to physical harm: url.

Users with no depression

- User 1
  - Highest weight: The lady christmas jumper: url.
  - Lowest weight: All the best for your match and hope to see you play.
- User 2
  - Highest weight: He reminds me someone in a football team. He can play many positions and he is our best player.
  - Lowest weight: People believe you when you have evidence.

A recent survey on using social media data to identify users with depression showed that users from the United Kingdom expressed serious concerns about privacy risks and did not see the potential societal benefits outweighing these risks [53]. Thus, if these technologies are to have a meaningful impact on people’s lives, increased importance must be placed on the transparency and trust of the analytics performed.

Achieving this trust is, to an extent, helped by the compliance of any research with ethical codes and with the General Data Protection Regulation (GDPR), which helps in raising confidence in data safety and transparent analysis. However, *GDPR Article 9: Processing of special categories of personal data* specifically mentions that consent is not required if permission relates to personal data that are manifestly made public by the data subject. A core problem is the perception that any data in the public domain are automatically available for research. This is highly controversial from an ethical point of view, as the disruption presented by the wide availability of social network data impacts the norms that guide our perception of the usage of our data for research. Ultimately, GDPR is focused on process, not on the *objective* of the research, which is fundamental to shaping any research consent and the social consensus around it.

This study had some limitations. Collecting control group data is challenging because the samples may contain users with depression who do not publicly express their mental health state on their profiles. Although keyword-based self-declaration is a popular way of asserting depression [11,12], social media users with depression may use more complex ways of communicating their mental health state [54]. There is evidence that social media users post less frequently when they feel low, suggesting that there may be less data available for modeling depression [53].

With regard to technical limitations, this study used additional features from a language analysis tool, which counts words in psychological and word function categories. This may prevent our models from learning word functions directly from sentences. Our future work will use sentence structures extracted from text and train a predictive model with those features [55], which may produce further performance improvements.

The availability of data for model validation is another major concern. Owing to potential ethical issues, there are currently no open data sets to evaluate the performance of predictive models on social network data, making it difficult to compare the model performance. The alternative benchmarking approach used in this study is to replicate well-known study models in the field and apply them to the same data set as the new model being investigated.

Another source of potential bias is the pages that publish tweets about mental health information (eg, mental health charities) and users who report depression experiences of other people (eg, users’ friends, family, or a celebrity). Although we filtered those instances in our study, a significant concern still exists for similar work in the field.

Conclusions

This paper proposes two novel MIL models with and without anaphoric resolution to detect Twitter users with depression. Anaphoric resolution is introduced to address the problem of identifying the subject of a statement made in the post. The classifiers developed comprise a tweet encoder, word attention, tweet classification, user encoder, anaphoric resolution encoder, tweet attention, and user classification layers. Bidirectional long short-term memory layers were used to learn the sequence of words and order of tweets posted on a timeline. Word embedding was applied to transform the textual content into vectors. Additional pronoun features were used to add informative dimensions to our proposed model and highlight posts relevant to the posters themselves. The approach was evaluated against previously published traditional machine learning and deep learning techniques, and the experimental results show that our proposed model produces notably better results. Anaphoric resolution, in particular, improved the performance of our model further and should be considered for inclusion in future studies.

The potential impact of this research lies in its ability to offer social media users exhibiting signs of depression that are suitable for their formal diagnosis. As in other mental health disorders, treatments for depression produce better outcomes and at a lower cost of treatment, the earlier patients get into services. Targeted advertising by mental health charities may be seen as intrusive but is no different than companies advertising any other products to potential consumers based on their web activity.

Early research into public perception of this type of data usage shows that there is public skepticism about this approach. To overcome this animosity toward using social media data for mental health prediction modeling, we believe that future research in this area should focus on explainability and interpretability. We have shown that deep learning MIL models perform well, but they offer no explanation of their decision-making processes [56,57]. Extraction of patterns from the models can provide interpretability, as we demonstrated with tweet weight examples, and systematic sampling should be used to achieve the levels of trust acceptable to users. To gauge how acceptable these techniques are to the public, we intend to work with citizen juries to explore the change in opinion that such explainability can deliver [58].

Acknowledgments

AW is fully funded by a scholarship from the Royal Thai Government to study for a PhD. MAV was supported by Comunidad de Madrid (grants 2016-T1/SOC-1395 and 2020-5A/SOC-19723) and AEI /UE FEDER (grant PSI2017-85159-P). This work was partially supported by the UK Engineering and Physical Sciences Research Council (EPSRC) under grant EP/P010105/1 (CONSULT: Collaborative Mobile Decision Support for Managing Multiple Morbidities). VC is also supported by the National Institute for Health Research (NIHR) Biomedical Research Centre based at Guy's and St Thomas' National Health Service NHS) Foundation Trust and King's College London, and the Public Health and Multimorbidity Theme of the National Institute for Health Research’s Applied Research Collaboration (ARC) South London. The opinions in this paper are those of the authors and do not necessarily reflect the opinions of the funders.

Conflicts of Interest

None declared.

References

1. Depression. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/depression [accessed 2021-05-14]
2. Health Matters: Reducing Health Inequalities in Mental Illness. Public Health England. URL: https://www.gov.uk/government/publications/health-matters-reducing-health-inequalities-in-mental-illness/health-matters-reducing-health-inequalities-in-mental-illness [accessed 2021-07-14]
3. The Global Economic Burden of Noncommunicable Diseases. World Economic Forum. 2011. URL: http://www3.weforum.org/docs/WEF_Harvard_HE_GlobalEconomicBurdenNonCommunicableDiseases_2011.pdf [accessed 2021-07-14] [WebCite Cache ID 6CSThUnbF]
4. Lee JH, Jung SH, Park J. The role of entropy of review text sentiments on online WOM and movie box office sales. Electr Comm Res Appl 2017 Mar;22:42-52. [doi: 10.1016/j.elerap.2017.03.001]
5. Wilkerson J, Casas A. Large-scale computerized text analysis in political science: opportunities and challenges. Annu Rev Polit Sci 2017 May 11;20(1):529-544. [doi: 10.1146/annurev-polisci-052615-025542]
6. Korkontzelos I, Nikfarjam A, Shardlow M, Sarker A, Ananiadou S, Gonzalez GH. Analysis of the effect of sentiment analysis on extracting adverse drug reactions from tweets and forum posts. J Biomed Inform 2016 Aug;62:148-158 [FREE Full text] [doi: 10.1016/j.jbi.2016.06.007] [Medline: 27363901]
7. Ive J, Gkotsis G, Dutta R, Stewart R, Velupillai S. Hierarchical Neural Model With Attention Mechanisms for the Classification of Social Media Text Related to Mental Health. In: Proceedings of the Fifth Workshop on Computational Linguistics and Clinical Psychology: From Keyboard to Clinic. 2018 Presented at: CLPsych'18; June 7-9, 2018; New Orleans, LA. [doi: 10.18653/v1/W18-0607]
8. Gkotsis G, Oellrich A, Velupillai S, Liakata M, Hubbard TJ, Dobson RJ, et al. Characterisation of mental health conditions in social media using Informed Deep Learning. Sci Rep 2017 Mar 22;7(1):-. [doi: 10.1038/srep45141]
9. Wongkoblap A, Vadillo MA, Curcin V. Researching mental health disorders in the era of social media: systematic review. J Med Internet Res 2017 Jun 29;19(6):e228 [FREE Full text] [doi: 10.2196/jmir.7215] [Medline: 28663166]
10. Reece AG, Danforth CM. Instagram photos reveal predictive markers of depression. EPJ Data Sci 2017 Aug 8;6(1):-. [doi: 10.1140/epjds/s13688-017-0110-z]

11. Coppersmith G, Dredze M, Harman C. Quantifying Mental Health Signals in Twitter. In: the Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality. 2014 Presented at: CLPsych; 2014; Baltimore, Maryland, USA. [doi: 10.3115/v1/W14-3207]
12. Yazdavar AH, Al-Olimat HS, Ebrahimi M, Bajaj G, Banerjee T, Thirunarayan K, et al. Semi-Supervised Approach to Monitoring Clinical Depressive Symptoms in Social Media. 2017 Presented at: the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2017; 2017; Sydney, Australia p. 1191-1198 URL: http://europepmc.org/abstract/MED/29707701 [doi: 10.1145/3110025.3123028]
13. De Choudhury M, Gamon M, Counts S, Horvitz E. Predicting Depression via Social Media. In: the International AAAI Conference on Web and Social Media. 2013 Presented at: ICWSM; July 8-11, 2013; Cambridge, Massachusetts USA.
14. Leis A, Ronzano F, Mayer MA, Furlong LI, Sanz F. Detecting Signs of Depression in Tweets in Spanish: Behavioral and Linguistic Analysis. J Med Internet Res 2019 Jun 27;21(6):e14199. [doi: 10.2196/14199]
15. De Choudhury M, Counts S, Horvitz E. Predicting postpartum changes in emotion and behavior via social media. 2013 Presented at: CHI '13: CHI Conference on Human Factors in Computing Systems; 27 April 2013- 2 May 2013; Paris France p. 3267-3276. [doi: 10.1145/2470654.2466447]
16. Coppersmith G, Dredze M, Harman C, Hollingshead K. From ADHD to SAD: Analyzing the Language of Mental Health on Twitter through Self-Reported Diagnoses. 2015 Presented at: the 2nd Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality; 5 June 2015; Denver, Colorado, USA. [doi: 10.3115/v1/w15-1201]
17. Eichstaedt JC, Smith RJ, Merchant RM, Ungar LH, Crutchley P, Preoţiuc-Pietro D, et al. Facebook language predicts depression in medical records. Proc Natl Acad Sci U S A 2018 Oct 30;115(44):11203-11208 [FREE Full text] [doi: 10.1073/pnas.1802331115] [Medline: 30322910]
18. Schwartz HA, Eichstaedt J, Kern ML, Park G, Sap M, Stillwell D, et al. Towards Assessing Changes in Degree of Depression through Facebook. 2014 Presented at: the Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality; June 2014; Baltimore, Maryland, USA p. 118-125.
19. De Choudhury M, Counts S, Horvitz E, Hoff A. Characterizing Predicting Postpartum Depression from Shared Facebook Data. 2014 Presented at: CSCW '14: Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing; February 15 - 19, 2014; Baltimore Maryland USA.
20. Angelidis S, Lapata M. Multiple Instance Learning Networks for Fine-Grained Sentiment Analysis. TACL 2018 Dec;6:17-31. [doi: 10.1162/tacl_a_00002]
21. Aktaş B, Scheffler T, Stede M. Anaphora Resolution for Twitter Conversations: An Exploratory Study. 2018 Presented at: the First Workshop on Computational Models of Reference, Anaphora and Coreference; June 2018; New Orleans, Louisiana, USA. [doi: 10.18653/v1/w18-0701]
22. Mohammad SM. 9 - Sentiment Analysis: Detecting Valence, Emotions, and Other Affectual States from Text. In: Meiselman HL, editor. Emotion Measurement. Cambridge, UK: Woodhead Publishing; 2016:201-237.
23. Kim E, Klinger R. A Survey on Sentiment and Emotion Analysis for Computational Literary Studies. 2018. URL: https://arxiv.org/abs/1808.03137 [accessed 2021-07-14]
24. Al-Saqqa S, Abdel-Nabi H, Awajan A. A Survey of Textual Emotion Detection. 2018 Presented at: 2018 8th International Conference on Computer Science and Information Technology (CSIT); 11-12 July 2018; Amman, Jordan.
25. Pennington J, Socher R, Manning C. GloVe: Global Vectors for Word Representation. 2014 Presented at: the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP); October 2014; Doha, Qatar. [doi: 10.3115/v1/d14-1162]
26. Sukthanker R, Poria S, Cambria E, Thirunavukarasu R. Anaphora and coreference resolution: a review. Inform Fusion 2020 Jul;59:139-162. [doi: 10.1016/j.inffus.2020.01.010]
27. Amir S, Coppersmith G, Carvalho P, Silva MJ, Wallace BC. Quantifying Mental Health from Social Media with Neural User Embeddings. 2017 Presented at: the 2nd Machine Learning for Healthcare Conference; 2017; -.
28. Orabi AH, Buddhitha P, Orabi MH, Inkpen D. Deep Learning for Depression Detection of Twitter Users. 2018 Presented at: the Fifth Workshop on Computational Linguistics and Clinical Psychology: From Keyboard to Clinic; 2018; New Orleans, LA. [doi: 10.18653/v1/w18-0609]
29. Kang K, Yoon C, Kim EY. Identifying depressive users in Twitter using multimodal analysis. 2016 Presented at: BIGCOMP '16: Proceedings of the 2016 International Conference on Big Data and Smart Computing (BigComp); January 2016; NW Washington, DC United States. [doi: 10.1109/bigcomp.2016.7425918]
30. Chancellor S, Lin Z, Goodman EL, Zerwas S, De Choudhury M. Quantifying and Predicting Mental Illness Severity in Online Pro-Eating Disorder Communities. 2016 Presented at: CSCW '16: Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing; 27 February 2016- 2 March 2016; New York NY United States. [doi: 10.1145/2818048.2819973]
31. Keeler J, Rumelhart D, Leow W. Integrated Segmentation and Recognition of Hand-Printed Numerals. In: Advances in Neural Information Processing Systems. -: Morgan-Kaufmann; 1991.
32. Ilse M, Tomczak JM, Welling M. Attention-based Deep Multiple Instance Learning. 2018. URL: https://arxiv.org/abs/1802.04712 [accessed 2021-07-14]

33. Cinbis RG, Verbeek J, Schmid C. Weakly Supervised Object Localization with Multi-Fold Multiple Instance Learning. IEEE Trans. Pattern Anal. Mach. Intell 2017 Jan 1;39(1):189-203. [doi: 10.1109/tpami.2016.2535231]
34. Wu J, Yu Y, Huang C, Yu K. Deep multiple instance learning for image classification and auto-annotation. 2015 Presented at: 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 7-12 June 2015; Boston, MA, USA. [doi: 10.1109/cvpr.2015.7298968]
35. Xu Y, Mo T, Feng Q, Zhong P, Lai M, Chang E. Deep learning of feature representation with multiple instance learning for medical image analysis. In: Yan Xu; Tao Mo; Qiwei Feng; Peilin Zhong; Maode Lai; Eric I-Chao Chang. 2014 Presented at: 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP); 4-9 May 2014; Florence, Italy. [doi: 10.1109/icassp.2014.6853873]
36. Quellec G, Cazuguel G, Cochener B, Lamard M. Multiple-Instance Learning for Medical Image and Video Analysis. IEEE Rev. Biomed. Eng 2017;10:213-234. [doi: 10.1109/rbme.2017.2651164]
37. Wang W, Ning Y, Rangwala H, Ramakrishnan N. A Multiple Instance Learning Framework for Identifying Key Sentences and Detecting Events. 2016 Presented at: CIKM'16: ACM Conference on Information and Knowledge Management; October 24 - 28, 2016; Indianapolis Indiana USA. [doi: 10.1145/2983323.2983821]
38. Yan S, Zhu X, Liu G, Wu J. Sparse multiple instance learning as document classification. Multimed Tools Appl 2016 May 16;76(3):4553-4570. [doi: 10.1007/s11042-016-3567-z]
39. Bishop C. Pattern Recognition and Machine Learning. New York: Springer-Verlag; 2006.
40. Tausczik YR, Pennebaker JW. The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods. Journal of Language and Social Psychology 2009 Dec 08;29(1):24-54. [doi: 10.1177/0261927X09351676]
41. Kingma DP, Ba J. Adam: A Method for Stochastic Optimization. 2014 Presented at: the 3rd International Conference for Learning Representations; 2015; San Diego, USA.
42. Hauenstein S, Wood SN, Dormann CF. Computing AIC for black-box models using generalized degrees of freedom: A comparison with cross-validation. Communications in Statistics - Simulation and Computation 2017 Jul 03;47(5):1382-1396. [doi: 10.1080/03610918.2017.1315728]
43. Vrieze SI. Model selection and psychological theory: A discussion of the differences between the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). Psychological Methods 2012;17(2):228-243. [doi: 10.1037/a0027127]
44. Akaike H. A new look at the statistical model identification. IEEE Trans. Automat. Contr 1974 Dec;19(6):716-723. [doi: 10.1109/TAC.1974.1100705]
45. Panchal G, Ganatra A, Kosta Y, Panchal D. Searching Most Efficient Neural Network Architecture Using Akaike's Information Criterion (AIC). IJCA 2010 Feb 25;1(5):54-57. [doi: 10.5120/126-242]
46. HURVICH CM, TSAI C. Regression and time series model selection in small samples. Biometrika 1989;76(2):297-307. [doi: 10.1093/biomet/76.2.297]
47. Wang J, Wang Z, Zhang D, Yan J. Combining Knowledge with Deep Convolutional Neural Networks for Short Text Classification. 2017 Presented at: the Twenty-Sixth International Joint Conference on Artificial Intelligence (IJCAI-17); 2017; -. [doi: 10.24963/ijcai.2017/406]
48. Uysal A, Murphey Y. Sentiment Classification: Feature Selection Based Approaches Versus Deep Learning. 2017 Presented at: 2017 IEEE International Conference on Computer and Information Technology (CIT); 21-23 Aug. 2017; Helsinki, Finland. [doi: 10.1109/cit.2017.53]
49. Lee J, Dernoncourt F. Sequential Short-Text Classification with Recurrent and Convolutional Neural Networks. 2016 Presented at: the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies; 2016; San Diego, California. [doi: 10.18653/v1/n16-1062]
50. Lai S, Xu L, Liu K, Zhao J. Recurrent convolutional neural networks for text classification. 2015 Presented at: AAAI'15: Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence; 25 January 2015; Austin Texas.
51. Zhang X, Zhao J, LeCun Y. Character-level convolutional networks for text classification. 2015 Presented at: the 28th International Conference on Neural Information Processing Systems; December 7 - 12, 2015; Montreal Canada p. 657.
52. Wang P, Xu B, Xu J, Tian G, Liu C, Hao H. Semantic expansion using word embedding clustering and convolutional neural network for improving short text classification. Neurocomputing 2016 Jan;174:806-814. [doi: 10.1016/j.neucom.2015.09.096]
53. Ford E, Curlewis K, Wongkoblap A, Curcin V. Public Opinions on Using Social Media Content to Identify Users With Depression and Target Mental Health Care Advertising: Mixed Methods Survey. JMIR Ment Health 2019 Nov 13;6(11):e12942. [doi: 10.2196/12942]
54. Berry N, Lobban F, Belousov M, Emsley R, Nenadic G, Bucci S. #WhyWeTweetMH: Understanding Why People Use Twitter to Discuss Mental Health Problems. J Med Internet Res 2017 Apr 05;19(4):e107 [FREE Full text] [doi: 10.2196/jmir.6173] [Medline: 28381392]
55. Li J, Luong T, Jurafsky D, Hovy E. When Are Tree Structures Necessary for Deep Learning of Representations? 2015 Presented at: the 2015 Conference on Empirical Methods in Natural Language Processing; September 2015; Lisbon, Portugal. [doi: 10.18653/v1/d15-1278]
56. Holzinger A, Biemann C, Pattichis CS, Kell DB. What do we need to build explainable AI systems for the medical domain? arXiv. 2017. URL: https://arxiv.org/abs/1712.09923 [accessed 2021-07-14]

57. Gilpin LH, Bau D, Yuan Z, Bajwa A, Specter M, Kagal L. Explaining Explanations: An Overview of Interpretability of Machine Learning. 2018 Presented at: 2018 IEEE 5th International Conference on Data Science and Advanced Analytics (DSAA); 1-3 Oct. 2018; Turin, Italy. [doi: 10.1109/DSAA.2018.00018]
58. Tully MP, Hassan L, Oswald M, Ainsworth J. Commercial use of health data-A public "trial" by citizens' jury. Learn Health Sys 2019 Aug 18;3(4):e10200. [doi: 10.1002/lrh2.10200]

Abbreviations

AIC: Akaike information criterion
API: application programming interface
CNN: convolutional neural network
GDPR: General Data Protection Regulation
GloVe: Global Vectors for Word Representation
GRU: gated recurrent unit
LIWC: linguistic inquiry and word count
MIL: multiple instance learning
MILA-SocNet: multiple instance learning with an anaphoric resolution for social network
MILNET: multiple instance learning network
MIL-SocNet: multiple instance learning for social network
NLP: natural language processing

Edited by J Torous; submitted 03.05.20; peer-reviewed by V Gupta, T Loncar-Turukalo; comments to author 12.07.20; revised version received 02.09.20; accepted 31.03.21; published 06.08.21.

Please cite as:
Wongkoblap A, Vadillo MA, Curcin V
Deep Learning With Anaphora Resolution for the Detection of Tweeters With Depression: Algorithm Development and Validation Study
JMIR Ment Health 2021;8(8):e19824
URL: https://mental.jmir.org/2021/8/e19824
doi:10.2196/19824
PMID:34383688

Original Paper

Understanding People's Use of and Perspectives on Mood-Tracking Apps: Interview Study

Stephen M Schueller[1,2], PhD; Martha Neary[1], MSc; Jocelyn Lai[1], MA; Daniel A Epstein[2], PhD

[1]Department of Psychological Science, University of California, Irvine, Irvine, CA, United States

[2]Department of Informatics, University of California, Irvine, Irvine, CA, United States

Corresponding Author:
Stephen M Schueller, PhD
Department of Psychological Science
University of California, Irvine
4201 Social and Behavioral Sciences Gateway
Irvine, CA, 92697
United States
Phone: 1 9498243850
Email: s.schueller@uci.edu

Abstract

Background: Supporting mental health and wellness is of increasing interest due to a growing recognition of the prevalence and burden of mental health issues. Mood is a central aspect of mental health, and several technologies, especially mobile apps, have helped people track and understand it. However, despite formative work on and dissemination of mood-tracking apps, it is not well understood how mood-tracking apps used in real-world contexts might benefit people and what people hope to gain from them.

Objective: To address this gap, the purpose of this study was to understand motivations for and experiences in using mood-tracking apps from people who used them in real-world contexts.

Methods: We interviewed 22 participants who had used mood-tracking apps using a semistructured interview and card sorting task. The interview focused on their experiences using a mood-tracking app. We then conducted a card sorting task using screenshots of various data entry and data review features from mood-tracking apps. We used thematic analysis to identify themes around why people use mood-tracking apps, what they found useful about them, and where people felt these apps fell short.

Results: Users of mood-tracking apps were primarily motivated by negative life events or shifts in their own mental health that prompted them to engage in tracking and improve their situation. In general, participants felt that using a mood-tracking app facilitated self-awareness and helped them to look back on a previous emotion or mood experience to understand what was happening. Interestingly, some users reported less inclination to document their negative mood states and preferred to document their positive moods. There was a range of preferences for personalization and simplicity of tracking. Overall, users also liked features in which their previous tracked emotions and moods were visualized in figures or calendar form to understand trends. One gap in available mood-tracking apps was the lack of app-facilitated recommendations or suggestions for how to interpret their own data or improve their mood.

Conclusions: Although people find various features of mood-tracking apps helpful, the way people use mood-tracking apps, such as avoiding entering negative moods, tracking infrequently, or wanting support to understand or change their moods, demonstrate opportunities for improvement. Understanding why and how people are using current technologies can provide insights to guide future designs and implementations.

(JMIR Ment Health 2021;8(8):e29368) doi:10.2196/29368

KEYWORDS

mental health; mobile apps; mHealth; emotions; affect; self-tracking

Introduction

Background

Mental health and wellness are common and widespread concerns. Nearly 1 in 4 people will experience a diagnosable mental health condition in their life [1], and many more will experience levels of mental distress that do not rise to clinical levels [2]. The chronic underfunding of mental health services means that not enough trained providers exist to meet the need for evidence-based mental health care. As people look for tools to help them understand and improve their mental health and wellness, they are turning to a multitude of nontraditional options, including health technologies. About two-thirds of teens and young adults [3] and nearly half of adults [4] report having used a mobile health app, with many using apps specifically for mental health and wellness. The landscape of health technology is growing rapidly; estimates suggest that over 300,000 health app exist, with over 10,000 of those for mental health and wellness [5]. One of the most common reasons people use a health app is to track some aspect of their health [6]. In the mental health and wellness space, tracking usually involves either specific tracking, which focuses on disorder-related symptoms, or general tracking, which relates to transdiagnostic factors—or those that transcend disorders or that everyone experiences to various degrees [7]. The most common transdiagnostic factor in tracking of mental health and wellness is mood. Reviews conducted in 2020 [8] and 2021 [9] have found that tracking of moods, thoughts, or behaviors is present in over half of mental health and depression apps, and is one of most common features of these apps. Thus, mood tracking appears to be a common, relevant health feature that people track through apps.

Despite the prevalence of mood-tracking features in mental health apps, most research on mood-tracking apps has focused on (1) novel system design or (2) advancing methods for mood tracking in the context of treatment for specific clinical disorders (eg, [10-12]). Less is known about people's experience with publicly available mood-tracking apps, that is, apps that people can download directly from the app stores and start using on their own. The few studies that have explored people's experiences with mood-tracking apps have built understanding from app store reviews [13,14], but these are limited in the type of information they contain, and also tend to come from a biased sample of users with particularly positive or negative experiences to share [5].

This study aims to fill this gap by understanding more deeply the experience and views of those who have used publicly available mood-tracking apps in their daily lives. To gain more understanding of how and why people use apps to track their moods, we conducted an interview study with people who had used or were currently using a mood-tracking app. Data were collected in a semistructured interview that enabled us to learn about people's experiences using mood-tracking apps "in the wild," as well as through a card-sorting task to learn about people's reactions to screens from various publicly available mood-tracking apps. In both the interview and card-sorting task, we explored what people find useful and challenging about data entry and review approaches in existing mood-tracking apps. Our procedures, including the questions in the semistructured interviews and the types of screens selected, were guided by models of personal informatics; thus, we first discuss personal informatics including considerations of entering and tracking data. We then discuss the existing literature on mood-tracking apps to better frame our contribution to current knowledge.

Self-Tracking and Personal Informatics

Technology designed to support people in tracking aspects of themselves and their behaviors, known as personal informatics systems [15], has become increasingly widespread. Personal informatics systems to support health and wellness are particularly common, supporting people in monitoring their physical activity [16], diet [17,18], sleep [19], and mood [14]. Research in personal informatics has primarily focused on understanding people's needs and practices around tracking behaviors they have some direct control over, such as what and how much a person is eating or walking [20], often with the aim of making these behaviors more salient to satisfy curiosity or promote awareness [15], or to facilitate behavior change or habit formation [16].

By contrast, mood is not a behavior, rather it is a bodily and cognitive experience encompassing both physiological reactions and thoughts [21]. Accordingly, people may not have direct control over their mood and might be interested in understanding *why* certain mood states occur. This might hold some similarity to other bodily or cognitive experiences people track including chronic health conditions such as migraine [22], multiple sclerosis [23], irritable bowel syndrome [24], or recurrent bodily experiences such as menstrual cycles [25]. In these domains, people often aim to predict when an event might come (eg, when a person is due to have their next period), monitor a symptom's frequency (eg, whether migraines are occurring more often), or learn what triggers a symptom (eg, whether a food triggers an intolerance).

Self-tracking domains vary in *what* kinds of data they collect and *how* they present that data back to the user. Aligning with the quantified self-mantra of "self-knowledge through numbers" [26], many personal informatics systems strive to collect *quantitative* and *objective* data which can be sensed and are particularly well-suited to quantification, such as steps walked or food eaten. Other apps aim to support quantification of journaled data, such as validated questionnaire scales for stress [27] and calorie lookups for foods [18]. Measurable goals can be useful for goal setting and monitoring [28], and are often visualized to promote self-reflection [15]. Other personal informatics systems instead leverage qualitative or multimedia data, such as ratings, explanations, and photos. These systems can help promote awareness such as of food choices [29]. Systems involving open-ended data can also support people in curating a record for later reflection, reminiscence, or sensemaking [30]. Researchers who contribute to the personal informatics literature often look to understand and overcome barriers to *collection* of self-tracked data and support *reflection* on such data [15]. Reflection is often supported through visual summaries or feedback to promote self-understanding or encourage desirable behaviors [31]. Systems have proposed that self-understanding and persuasion can be promoted by summarizing tracked data in natural language sentences [32], presenting data passively on the home or lock screen of a phone [33], forecasting future data based on current data [34], encoding data into abstract shapes or images [35], or encouraging reflective thought on specific events from the past [36]. However, it is unclear what data form(s) people seek and prefer for mood.

Mood-Tracking Apps

Most research on mood-tracking apps has attempted to advance novel system design or to understand their benefits. Some examples of novel system designs include using various visual representations of mood [37], tangible modalities [11], and reducing traditional assessments of mood [38]. Although mood is tracked manually in these instances, other work is exploring whether tracking could be done automatically through passive collection of data from smartphones or wearable devices [39-41]. One way to determine the benefit of mood-tracking apps is to compare them with traditional forms of tracking, such as paper-and-pencil tracking. Comparisons between mood-tracking apps and traditional paper versions of mood assessments for individuals with mood disorders have shown similar outcomes [12,42,43]. For example, MobiMood, a mood-tracking app developed for tablet devices, was compared with traditional measures of depression, anxiety, and rumination and MobiMood was found to have good validity, while improving the ability to capture daily fluctuations in mood over those measures [44]. Other mood-tracking apps, such as Aurora and Monarca, have been reported by users to support greater awareness of their emotion and moods and provide useful information to reflect on [45,46]. However, translational barriers mean that apps designed and evaluated in research settings are rarely available on app stores. Moreover, commercially available mental health technologies vary considerably in the actual features they offer [47]. Thus, there is a need to study publicly available mood-tracking apps, and how people perceive their benefits, as people's experiences with publicly available mood-tracking apps differ significantly from experiences with research prototypes.

Some work has begun to understand publicly available mood-tracking apps by characterizing their features. Caldeira et al [14] reviewed mood-tracking apps in light of Li et al's model of personal informatics [15] and found that the majority of features mapped on to the collection (entry) and reflection (review) phases. For mood entry, some apps allow people to select from faces or emojis, typically labeled with text descriptors (eg, "happy," "angry," "content"). Others use 1 or more scales, sometimes with textual anchors (eg, "terrible," "ok," "great") or without, or allow people to select from a list of words or describe their own mood in free text. For mood review, many apps use a form or list to summarize the moods a person has logged, for example, a scrolling feed of all entries or a daily list summarizing the moods a person has journaled in a day. Others provide graphs which summarize the relative frequency of different logged moods, showing how moods have varied over time; or calendars to describe the typical mood each day, week, or month; or simply surface counts of the number of times various moods have been logged. Although research has demonstrated that people like apps that facilitate mood tracking [6], little research has understood why and what features of mood-tracking apps drive this interest.

Some work has explored people's perceptions of app features and general impressions of mood-tracking apps, and has demonstrated that tracking one's mood might be motivated by desires to produce awareness and become more mindful of one's own mental health [48]. Although Drake and colleagues [49] found only moderate enthusiasm for the *usefulness* of an online mood-tracking platform, participants found the ability to see one's mood displayed over time beneficial. However, some of the features participants rated as less useful, such as automatic feedback provided about one's mood, might have been rated lower not because that feature could not be helpful, but instead due to the feature's implementation. Two studies, Caldeira and colleagues [14] and Widnall and colleagues [13], explored users' experiences of a wide variety of publicly available mood-tracking apps; however, they did so by analyzing app reviews from the app stores.

In sum, past work has proposed and evaluated diverse features facilitating data entry and review for mood-tracking apps. However, this work has focused primarily on the design and feasibility of novel mood-tracking technologies, rather than on understanding why and how people are using publicly available and popular mood-tracking technologies in the real world. These novel mood-tracking technologies seldomly become available to the public after their feasibility assessments, leading to a research-to-product gap. Feedback on specific apps and app reviews are both somewhat limited in providing understanding of people's experiences with apps. App reviews tend to offer open-ended and general reflections rather than specific aspects of app features, while reflections on single apps can relate to specific implementation challenges. Therefore, we need deeper understanding of people's experience with various mood-tracking apps and reactions to the different ways mood-tracking features including data entry and review are implemented. We therefore sought to better understand why people who choose to use mood-tracking apps in real-world settings do so and how useful or problematic people find aspects of these apps, especially features of data entry and review. We did so by conducting semistructured interviews and a card sorting task with people who were or had previously used a mood-tracking app. We were specifically interested in answering the following research questions:

- Why and how do people use mood-tracking apps?
- What do people find useful about the approaches that mood apps take to entry and review?
- What do people find challenging about the approaches that mood apps take to entry and review?

Answering these questions will help better understand the use of mood-tracking apps in real-world contexts and might inform the design and implementation of mental health technologies to account for people's everyday perspectives and experiences.

Methods

Procedure

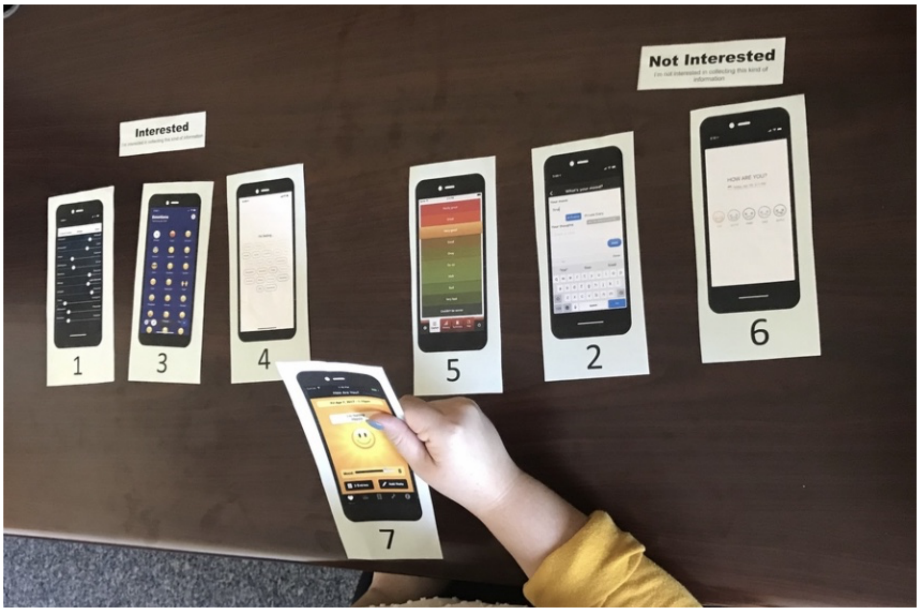
We conducted in-person sessions with people who were currently using or had previously used a mood-tracking app. The study was approved by the Institutional Review Board at the University of California, Irvine. Consent was obtained via oral confirmation including receipt and review of a study information sheet during the in-person sessions. These sessions were performed by research team members including student trainees, masters-level research staff, and faculty with expertise in psychology and computer science/informatics. Each session consisted of a semistructured interview and a card sorting activity, split roughly evenly between the interview and the activity. Participants granted verbal consent to participate and to have the interview audio recorded. We compensated participants US $30 for their time. Combining the semistructured interview and the card sorting activity, the average total length of these sessions was 54 minutes (SD 10.1 minutes; range 24.3-68.7 minutes).

The semistructured interviews aimed to understand what participants who used mood-tracking apps hoped to gain from using them, what data they typically entered, and whether or how they used the information. We asked probing questions to further understand what motivated them to enter and review their mood data and their thoughts about their app's approach to entry and review. Multimedia Appendix 1 provides the semistructured interview guide used for the interview portion. Prior to starting the interview study, we conducted a systematic review of publicly available mood-tracking apps to understand the types of entry and review features they contained. Although not the focus of this study, this feature review informed the development of the interview guide and screens selected for our card sorting activity. Additionally, we incorporated concepts of data entry and review from personal informatics and from previous reviews of features of mood-tracking apps [14]. We discussed these questions as a study team to arrive at an initial semistructured interview guide and iterated on the guide during training activities and practice sessions described below.

The card sorting activity aimed to understand how people felt about the data entry and review methods typical of mood-tracking apps. We chose 7 data entry and 5 data review screens representative of common strategies used in commercial mood-tracking apps. We selected apps that had representative combinations of mode of entry and mood granularity (ie, number of different options for entering one's mood) for entry and types of displays, visual representations, and features to facilitate interpretation. For example, data entry techniques included sliders with words at each endpoint, selection of relevant emojis, and open-ended text description fields. Data review techniques included feeds (ie, an updated and ongoing list of all the data entered by the user) and line charts of logged data over time, and pie and bar charts aggregating data from the past week or month. Multimedia Appendix 2 shows the data entry screens and Multimedia Appendix 3 shows the data review screens. Participants first categorized entry screens and then review screens into those they were "Interested" or "Not interested" in; participants were permitted to also create their own category (eg, "kind of interested" or "really not interested"). While sorting these cards, participants were asked to explain aloud why they found a particular entry or review screen interesting or not. We then asked participants to further categorize the screens they were interested in as "Too Specific," "Too General," or "Just Right." Again, we asked participant to explain the rationale for their choices aloud. Once participants had finished sorting, we asked probing questions to further understand why participants had categorized screens in particular ways. Figure 1 displays an example of the card sorting task for entry screens.

In order to prepare members of the research team for the interviews and card-sorting task, we had several training sessions including group didactics, one-on-one interview practice, a mock session with a supervisor, supervisor participation in initial session(s), and ongoing group supervision. These sessions were facilitated by 3 members (SMS, DAE, and MN) of the lead research team having extensive and diverse experience with interviewing, including 1 PhD-level clinical psychologist (SMS), 1 PhD-level human–computer interaction researcher/informaticist (DAE), and 1 master's-level project manager (MN). Group didactics included 2 sessions; the first session on interviewing tips included topics such as taking effective notes, facilitating conversation, using identified probes to ask follow-up questions, and addressing difficulties that might arise in interviews, and the second session was on observing an experiencing interviewer conduct the session protocol with a mock participant. An interviewing guide was provided to all research team members following the session on interviewing tips to help support conducting the interviews, in addition to a 1-page tip sheet with examples of open and probing questions which they could refer to during interviews. All interviewers then conducted one-on-one practice sessions with another research team member that was audio recorded and later reviewed by supervisors. Portions of these practice sessions were played in group meetings to help support training on interview methods. Prior to conducting a session with a participant, each research team member conducted a mock session with a supervisor to practice completing the full session protocol and to receive feedback from a supervisor. Lastly, a supervisor attended at least one session for each team member, to observe interviewing skills with participants and to ask follow-up questions as needed. Supervisors were required to sign off on team members training in order for them to start interviewing participants without a supervisor present, and in some cases interviewers had a supervisor present for more than 1 interview. Given that all participant sessions were recorded, portions of sessions were reviewed in group supervision meetings on an ongoing basis to support consistency with the session protocol.

Figure 1. Example card sorting task demonstrating how participants sorted data entry screens into those they were “Interested” or “Not Interested” in.



Participants

We recruited 22 participants from May to August 2019 through flyers on and around our university campus, local community centers, and coffee shops, and by email to our university-affiliated research registry. Participants were required to be at least 18 years old and have used an app that allowed them to track their mood, feelings, or mental well-being for at least two weeks. Flyers and emails listed these eligibility criteria, the length of study participation, and the study compensation. Eligibility criteria were checked prior to scheduling their session, including the name of the app the participant had used. Table 1 presents participant demographics. Although our sample skewed young (mean 24.7 [SD 8.9]), participants had a range of ages; 82% (18/22) of participants were women, consistent with data suggesting that women are more likely to use health [50] and mental health apps [51]. Our sample was highly educated, with all but 1 participant indicating completion of some college and many participants having advanced degrees. Although our participants reported using a diverse set of mood-tracking apps, Daylio was commonly used with over 40% (9/22, 41%) of the participants using it. We quote participants in the “Results” section with PXX.

Table 1. Demographic data of participants (N=22).

Demographic	Participants, n (%)
Gender	
Female	18 (82)
Male	3 (14)
Nonbinary	1 (5)
Age (years)	
Range	18-58
Mean (SD)	24.7 (8.9)
Race/Ethnicity	
Asian	11 (50)
White	6 (27)
Hispanic-Latinx	2 (9)
Mixed	3 (14)
Education level	
High-school graduate	2 (9)
Some college	10 (45)
Bachelor’s degree	6 (27)
Master’s degree	4 (18)
App used	
Daylio	9 (41)
Pacifica (Sanvello)	2 (9)
Flo	2 (9)
Moodpath	2 (9)
Color Calendar	1 (5)
Clue	1 (5)
MoodPanda	1 (5)
T2 Mood Tracker	1 (5)
Perspective	1 (5)
Pixels	1 (5)
Thought Journal	1 (5)
Length of app use	
>2 years	4 (18)
>6 months	2 (9)
>3 months	5 (23)
>1 month	8 (36)
>2 weeks	3 (14)
Timing of app use	
Current users	16 (73)
Past users	6 (27)

Analytic Strategy

We analyzed 22 interviews using a bottom–up thematic analysis [52]. Interviewers composed memos following each interview, which summarized each participant’s perspective. An external company transcribed our recordings. Five members (SMS, MN, JL, DAE, and Lisa Vasquez) of the research team collectively affinity diagrammed [53] a subset of the interview transcripts to develop a codebook. Our codebook contained 13 codes (eg, desired benefit from using the app, preferred or disliked data entry methods) with 37 subcodes (eg, for self-awareness, for advice, a preference for gradients scales, a disliking for numeric scales). Two researchers (JL and Lisa Vasquez) coded the interview transcripts, discussing and refining codes with input from the rest of the research team. The sample size was determined based on reaching thematic saturation for the themes [54], which occurred after about 15 interviews. We interviewed another 7 participants to ensure no new trends emerged, and then stopped recruiting.

We also counted which categories participants sorted each entry and review screen into. Participants often created their own categories as "in between" the categories we provided or would describe screens in relation to one another (eg, "I'm more interested in card five than card four but less than the other options" [P3]). One researcher (DAE) grouped participant-made categories based on their description and discussed the results with the rest of the research team.

Results

We examined why people seek to use mood-tracking apps and what they find beneficial and challenging about their strategies to supporting data entry and review.

Why Do People Use Mood-Tracking Apps?

Much like other personal informatics domains, participant motivations for tracking were both intrinsic and extrinsic. Some participants began tracking due to life events, while others started due to recommendations, prompting, or suggestions from other people. Participants generally felt that mood tracking helped improve their self-awareness, promoted self-reflection, allowed them to relate their mood to other factors, and aided them in intervening or changing their mood.

Thirteen participants talked about starting to use a mood-tracking app based on life events. These events included particularly stressful times in a person's life, as a New Year's resolution, or due to an awareness that one's mental health seemed to be suffering. P1 noted, "I first started using it when I was in high school when I was experiencing a lot of anxiety, and like some depression because I was really stressed out through the coursework and I thought that by tracking my mood and doing some meditation that are available in the app, it would have helped me to kind of see my emotions and see the pattern of it."

Interestingly, no participant mentioned using a mood-tracking app in response to a positive time in their life with the goal of capturing or documenting positive emotions or feelings. Most of the events seemed to result in a desire to improve their mood or awareness of their negative moods.

Participants frequently found apps through recommendations from digital tools and social connections. Nine participants conducted their own searches of the app store, 7 participants received recommendations from other people, and 3 chose the mood app from an advertisement. These recommendations included personal and professional sources such as friends, family members, and therapists. As P21 noted, "So, someone I talked to online, she said she used it. She said it was pretty good. I wasn't feeling very down when I started using it, I just figured it was like a cutesy little way to track how I've been doing over the course of the month or a year. So, I started using it."

P20 similarly expressed using an app based on a recommendation; however, their recommendation came from a professional rather than someone online, "I'd been using actually mood apps for a while, because they were recommended by my psychiatrist. She pretty much told me was some people use journals and stuff like that, but that it's easier to use an app just because everybody has a phone nowadays."

Seventeen participants particularly appreciated how apps helped facilitate self-awareness and self-reflection. For example, P9 stated, "I think it's good to reflect upon it. If it's a lot of downs over a period of time, I want to sit and think, 'Why was...What was the problem there?' I think it's a good start to just trying to be happier overall."

When reflecting on their mood later, participant's perspectives often changed: For example, P10 stated, "I write it as it is, exactly how I feel at that moment and then later I can look at it from a different perspective when I'm not angry anymore and see like was that worth it? I can see a different perspective on my emotions from a different time."

Although apps tended not to provide support for identifying correlations, participants were still able to use their apps to identify relationships between their mood and contextual factors. Participants noted finding potential triggers for their mood including people, events, and aspects of one's physical health. P12 reflected that the mood-tracking app she used helped: "basically, tracking my actions, seeing what I'm surrounding my environment with, and if I can eliminate any negative...just basically anything that's not good for me, I see. It's basically math. You see what you're doing to have a better, healthy lifestyle. You have to subtract what's not adding to that."

Eight participants talked about insights they uncovered through these relationships and saw correlations which suggested specific triggers for their moods. For example, P11 stated, "I realized through tracking my emotions and through looking at the correlation and everything of when it was happening. I realized that, first of all, I was getting stressed and unhappy and just not doing well too much overall, and especially during the PMS phase."

Although participants often aimed to better understand what led to negative mood states with the hope of potentially being able to alleviate or avoid those negative moods, 6 participants reported that they did not like to record or see indications of negative moods. Although participants appreciated relating negative moods to other information, the dislike of recording or seeing one's negative mood presented challenges to realizing these benefits. As P3 stated, "Yeah, the thing is whenever I had a bad experience, I didn't want to put it in the app. I just didn't feel an inclination to open the app and put it in there, I just wanted to lay there and cry. You don't use it for the bad times. Put all the good times on it."

Participants also appreciated how mood-tracking apps facilitated conversations in their daily life, including with friends, family members, and professionals. Eight participants noted the use of mood-tracking apps with formal mental health support. P1 discussed her use of a mood-tracking app with her therapist: “I was seeing a therapist at the time...And she was the one who actually told me to look at the monthly progress with the app, so I can get a better picture of how I felt because I didn’t have bad weeks all the time. It’s just one here and there.”

Conversations with other people about mood-tracking apps helped provide ideas on how and why to use the app and added meaning or perspective when interpreting the data. P11 talked about the ability to use this information to have more meaningful conversations with her mom: “I feel she’s someone who has a hard time understanding other people’s emotions and mental health kind of stuff unless she has direct evidence of things. Which is unfortunate, but also I found it beneficial to be able to say, ‘Hey, look at this. This is happening during these times. There is evidence here now.’ I think that [the app] was helpful to kind of help her to understand as well.”

Social features within apps also helped a few participants normalize their experience. For example, P9 used MoodPanda, and appreciated that she could see the mood descriptions that others wrote in the app: “I mean, it’s just also not a feeling of aloneness, just seeing other people’s problems. I mean, I kind of get that, that kind of thing. I know that if I feel a certain way, it’s just a reminder of, oh, there’s other people. I’m not the first one to feel this way.”

What Do People Like About Entry and Review in Mood-Tracking Apps?

Although we focused our questions and analysis to identify the functional benefits of entry and review, participants did comment on appreciating aspects such as aesthetics. Participants also frequently mentioned their preferences depending on other features or properties of mood-tracking apps, such as the inclusion of social features or the ability to preserve privacy.

With respect to data entry, participants seemed to appreciate simplicity in conjunction with flexibility. Participants disliked when they felt overwhelmed by the number of options of moods and emotions to track, but also wanted to be able to capture the nuance in the moods. Eight participants noted that sliders were a particularly effective way to balance nuance and options; for example, P13 stated, “same thing I said before. I like the fact that there’s a dial, a spectrum that lets you pick somewhere along the range.”

P10 similarly reflected on the benefits of sliders, noting that they are able to capture that moods vary in intensity in addition to valence, and that the discrete emotions could be mixed or felt to varying degrees: “it’s not just clicking Tired or Energetic, but it’s like on a gradient. I think that’s interesting...Because you don’t always feel completely stressed or completely relaxed. Sometimes you feel like a little bit of both or in between, so I think the gradient is important to differentiate that feeling.”

Participants thought customization could balance the desire for simplicity and flexibility. For example, P7 stated: “I prefer the having the sort of structure of having a few moods to choose from, and then if the user doesn’t really relate to those moods, then they can add their own because that adds a layer of customization now also, like sort of a personal connection, they specifically feel this way so they can put them down.”

These preferences toward simple, but flexible entry methods were also reflected in the card sorting task. Four of the 7 data entry techniques were categorized by most of the participants (59% [13/22]-73% [16/22]) as methods they would be interested in using to enter data. All of the well-regarded screens supported multiple ways of identifying or describing moods, such as by using sliders for multiple mood states, having multiple emojis or words that could be used to represent moods, or combining emojis and words in a simple emotion scale. The 2 data entry strategies that participants found “just right” presented multiple words or emojis and allowed participants to endorse which they were experiencing (Figure 2). P12 appreciated the simplicity of entry in these screens, saying: “It’s very brief like Color Calendar. I just pick an emoji, and that’s it? Like I said, that sake of condensation. That’s very appealing.” Such screens often reduced the cognitive complexity of tracking as P16 noted: “But, I feel like it’s just right in that it has four choices so you don’t have to think about it too much. The language is accessible, again, to a casual user and if you do have diagnosis, it has a sense of humor about it.”

By contrast, screens that were less well-regarded often felt inflexible or inappropriate. For example, P15 noted, “I’m not a robot, so I just can’t go off of this scale.”

Whereas P19 indicated a specific dislike for numbers: “Oh rating by numbers, I don’t think I would use that.” This further reinforces the preferences we saw for flexibility and multiple ways to enter data. In the card sorting task, the 2 data review screens that participants were most interested in used a line graph (14/22, 64% of participants) and a feed (12/22, 54%) to summarize the mood entries a person had journaled. These approaches helped capture not only what moods occurred, but when, and also each mood in relation to other entries (Figure 3). For example, of a screen with a line graph, P20 said, “I like it. It shows you in a graph how your mood goes up and down…it shows you the days in a row how you’ve been feeling like that, so you can look back and see what you can do or what you were doing that made you feel that certain type of way for those couple of days.”

P12 appreciated how a screen with a feed could help her recall what triggered the mood on a specific day: “This is more specific on the day, like off the bat I can see, and then recall, ‘Oh maybe this day, I’m doing something.’ This I would go into and see, ‘Oh what did I write down that day?’”

Figure 2. Participants preferred entry screens which allowed them to select words (Youper, left) or emojis that represent words (Mood - Journal & Anxiety Chat, right).

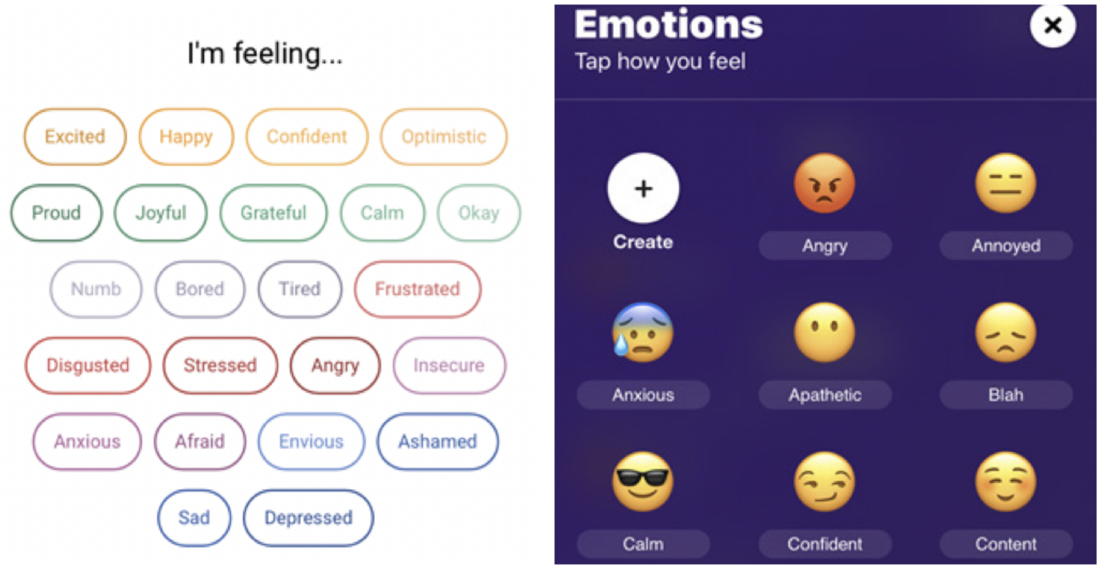
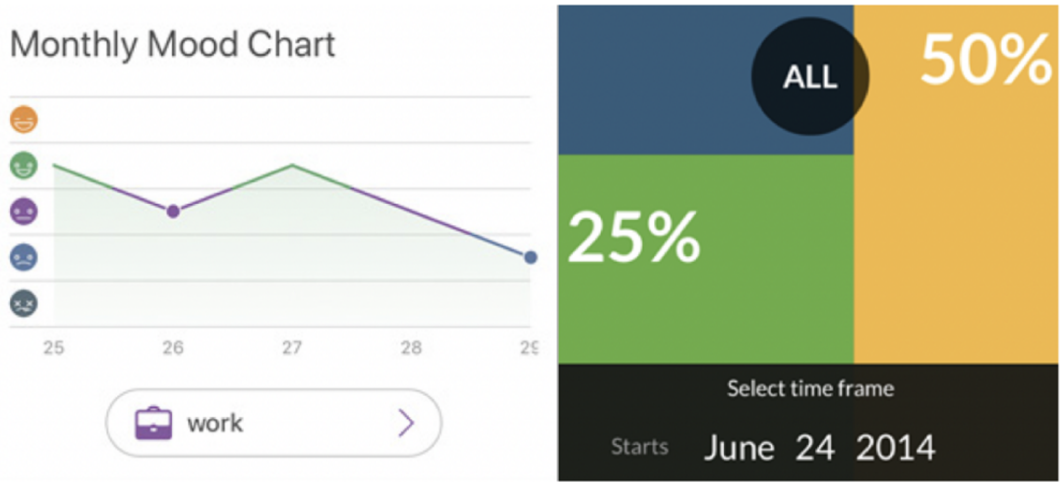


Figure 3. Participants preferred review screens which allowed them to see trends in their mood over time (Daylio, left) to ones which aggregated time ranges (Mood Meter, right).



What Do People Find Challenging About Mood Entry and Review?

Many participants had concerns about how meaningful or representative the mood data they logged felt to them. Some of these concerns related to how the app presented information back to them, although other participants had deeper concerns around whether the way they entered mood data might diminish its value.

Although 1 participant (P11) indicated that she was comfortable drawing inferences from the tracked data, it was much more common for participants to express discomfort or confusion. Nine participants commented on different features of the app that made interpretations from the data challenging, including an overreliance on displays that required effort to scroll or search through, or displays that merely presented data without summary or interpretation. P2's statement was representative of other participant perspectives: "I don't really know how to analyze it. Maybe if they gave me a method they're like, 'Oh, count the days you did this, divide by that.' I could have done it - numbers, or a chart or something more mathematical, because I don't really know how to approach this, I don't how to do it."

Overall, participants noted that although the apps provided the opportunity for them to collect a lot of information about their emotional life, there was little guidance to help them translate that information to something valuable. As P16 said: "I have confirmation that I am feeling shitty. But, it doesn't really help you do anything about it."

Not surprisingly, as 13 participants had started tracking in response to negative events or in attempts to improve their mood, participants wanted to see more interventional components in mood-tracking apps. For example, P7 stated: "I think it would be cool if it sort of had a bit more than just your own input like it took your input. Not only does it compile it into data but sort of give you solutions or recommendations based on the data that you provide."

From this vantage point, it seemed that people wanted to see how mood-tracking apps could better help them manage and improve their own mood. Relatedly, participants also voiced concern as to whether the data collected related to mood could actually be useful, given that mood and emotions are complex phenomena that are difficult to capture and quantify. As P9 summarized: "I don't think any app can realistically capture the human feelings."

Six participants noted that tracking one's mood just once a day may not effectively represent the full depth of their emotional experience or may miss important moments that would be useful to track and understand. As one example, P2 said: "one of the problems I had with it was my mood changes a lot throughout the day. I wake up and I'm in a good mood, and then 20 minutes later I'm in a bad mood, and then I'm in a good mood, and it doesn't account for that. So, it didn't really...I don't know. It was kind of not helpful in that aspect."

Many mood-tracking apps only allow the users to enter their mood once per day. However, even those that allow tracking multiple times per day may not strongly encourage that because multiple reminders and repeated entries over a day increase user burden. They might then fail to capture actual variance in some people's mood. As P17 stated, "Mine doesn't move too often. There's a better way to present that information. A line graph doesn't...Maybe for someone who's really cycling, sure. But for me when I'm more like this, it doesn't make much sense."

Participants commonly reported that they would track their mood at a regular time every day, as P5 stated, "At the end of the day usually....Usually before I sleep too, because it gives me the whole day to just think about how my day went and just reflect on my day." P7, however, noted a challenge with this in terms of what could be learned "what I've learned recently is that usually at 8:00 pm I'm feeling really good, but sometimes later at night, I can feel a little worse about myself, my mood can go down a little. And that doesn't...the mood having only one daily mood can sort of ignore or doesn't focus on the fact that my mood changed later on."

Many apps summarize moods logged over days or weeks. Given that participants tended to only enter their mood a few times per day or less, they wondered whether the summaries presented from their data are actually meaningful. P10 stated: "You have statistics. You can tell basically how often you're feeling happy on days, but it wasn't right."

In the card sorting activity, half or more of participants (50% [11/22]-59% [13/22]) indicated that they would not be interested in using the 3 data review screens which used calendars, stacking bar graphs, and pie charts to summarize the days or the amount of time they expressed a particular mood (Figure 3) because these screens tended to aggregate how frequently a mood level or mood term was indicated at varying internals (eg, day, week, or month). For example, P3 disliked how a screen summarized the percentage of entries which fit a mood category ("green" for good moods, "red" for bad): "It just shows these are all of the moods you felt during all this time…I guess not showing a trend, or specific days. You don't know what days you felt yellow, or what days you felt green." P19 felt similarly about a stacked bar chart: "again, it's labeling my whole month as something. I'm not sure how I feel about that. I don't think there's been a month in my life where I've followed one trend."

Discussion

Principal Results

Our findings provide an overview of what people gain from the use of mood-tracking apps that they adopted as well as the shortfalls or challenges in such tools. In general, our participants turned to mood-tracking apps to increase self-awareness or self-reflection, especially during challenging times in their lives. Participants liked simple screens for data entry and appreciated opportunities for review that linked their information, such as moods and triggers, to produce insights, but desired opportunities for customization around how such information was linked. Furthermore, our participants disliked the high burden of mood tracking and the lack of opportunities to gain self-awareness when data were sparse. Many of our participants therefore questioned the usefulness or accuracy of insights. Table 2 highlights key findings from our thematic analysis.

We highlight some potentials for mood tracking through technology, leading to design challenges and implications for future digital mood-tracking tools including (1) understanding the dynamics of mood; (2) customizability and personalization; and (3) undersupported tracking methods. As mood is an important component of mental health and wellness, these implications also have relevance to creating digital tools that can help people understand and manage their mental health and wellness.

Table 2. Key findings from our thematic analysis.

Research question	Key findings
Why do people use these mood-tracking apps?	Life events which triggered stressful or challenging moments motivated some participants, while others sought out mood tracking for greater self-awareness and self-reflection.
What do people find useful about the approaches that mood apps take to entry and review?	Most participants preferred data entry screens with fewer options, but which supported the ability to customize what they logged. For data review, participants wanted to be able to compare a logged mood with previous entries, preferring graphs and feeds which enabled this.
What do people find challenging about the approaches that mood apps take to entry and review?	The burden of mood tracking led participants to track once per day (or less often), and they questioned how well that mood reflected the moments they did not track and whether visualizations of that data were representative. They also found it challenging to identify if or how their mood improved based on the trends they observed.

Understanding the Dynamics of Mood

People's moods fluctuate and asking people how they feel at a moment may not hold much value as an evaluation [55], especially if only done once a day. Rather than journaling and reflecting on their mood in a particular moment, participants were more interested in patterns, connections, or insights. However, generating patterns, connections, or insights about one's mood requires sufficient data collected across time that can capture the dynamics of how a person's mood persists or changes over time, and can reveal actual, rather than illusory, correlations. Furthermore, understanding the dynamics of a person's mood could help them understand whether moments reflect highs or lows. Mood tracking, therefore, could help unpack the extreme values that a person experiences and better support ways for people to track and understand this information.

Psychological research has demonstrated that mood variability, or the variance in one's mood over regular intervals, is a stable trait that varies across individuals and as such might be able to be captured and represented when one starts tracking [56]. In the practice of mental health treatments, mood tracking usually begins with a short period of intense tracking to better identify one's variability and patterns [57]. Similar approaches of intense tracking over short periods have been used to help people identify food intolerances [24] and make sleep recommendations [19]. Mood-tracking apps, however, typically do not vary their entry or review processes by incentivizing more intense tracking when beginning a program. Many health trackers do enjoy a novelty period, in which people engage in more frequent and regular use immediately after starting to use a health tracking technology [58]; however, this novelty period is based on people's motivations rather than technological features. Intentionally structuring intense and maintenance periods of mood self-tracking could better support developing an overall understanding of a person's mood tendencies.

In addition to variance, an important aspect of mood dynamics is people's highs and lows. Participants in our study reported a hesitation to record extreme values when using scales or intense negative moods when using words or emojis, in part to avoid viewing these moods when reviewing their data. Similar challenges have been reported in other tracking domains, such as food, where tracking instances of making unhealthy choices might produce guilt or shame [17]. In experiential tracking domains such as fertility, people often stop tracking when the data only present struggles or lack of improvement [59]. It might be possible for data entry features to normalize or contextualize extremes, especially intense negative mood states. For example, mood-tracking apps could trigger a follow-up notification to better understand the duration of such moods, in the same way a friend or family member might check-in on a person who disclosed to them feeling particularly bad. Gathering deeper nuance during negative mood states to unpack valence as well as intensity (ie, how bad and how strong is the emotion), or providing normative data about how common such moods are either compared with one's own or others' data, may help encourage users to track these intense mood states. Other experiential domains may similarly benefit from normalizing lack of success or improvement.

Customizability and Personalization

Mood is ultimately a subjective phenomenon, with potential variations in what it means across people. This is different from other health-related aspects that might be tracked objectively such as weight, calories, or steps. Although a step may vary slightly from person to person (eg, stride length, time to take a step), there is general agreement as to whether a person took a step or not. People strive for relatively accurate accounts of their experience through tracking activities [16,60], but accuracy in mood is challenging to achieve given its subjective nature. Furthermore, people vary in the language they use to describe their moods; customizability could address these issues.

Our findings parallel Ayobi and colleagues' [61] work noting considerable variability in the way people tracked their habits and mood using paper bullet journals. Similarly, our participants had varied preferences for granularity in mood-tracking apps and the way to enter information such as emojis, words, numbers, or scales. Although participants often shared general interest in screens in our card sorting task, they varied in whether they characterized data entry screens as "too general," "too specific," or "just right" or data review screens as "too simple," "too complicated," or "just right" for review. Many apps include ways to customize and our participants reflect positively on customizability as a way to balance simplicity and flexibility. However, customizing features also requires more effort on the part of the user. This amount of effort is worrisome given that many people tend to not use features that allow customization when it fails to fit into their workflow or it requires extra time to explore the app [62].

Although customizability will help a variety of people meet their needs with an app, it is worth noting that customization is often limited to specific aspects of entry or review. For example, many mood-tracking apps allow users to customize the specific word or image used for a mood or how many different moods are displayed. Less customizability exists in terms of the form of entry. In the same way that some people find value in food journaling through calorie-driven apps, while others instead prefer using food photos to remain mindful of food choices, such variability might be useful for moods. It is worth considering whether novel apps could be developed to support diverse approaches to data entry and review. Under supported approaches to entry and review are prime opportunities for exploring and contributing new designs.

Undersupported Tracking Methods

Given that few participants tracked their mood more than once a day, if even that, it is worth understanding how mood-tracking apps could be designed in ways to facilitate self-reflection and insight from such sparse data. Mood-tracking apps often attempt to interpret data on the assumption that they represent a complete record of a person's mood. It is worth considering how reflection could be promoted in a mood-tracking app which explicitly promotes once-a-day mood tracking, or an app which aims to support logging during just the "difficult" moments that people appear interested in understanding yet are hesitant to track. This might be accomplished by having people reflect on their day overall rather than their mood in the moment, or have people rate their highest and lowest moods that occurred on a given day or since the last assessment, perhaps even identifying when they occurred.

Alternatively, further research could leverage how to encourage people to journal their moods more often, or to add more details when they do log. Mood-tracking apps could leverage classic techniques examined in other personal informatics domains, such as journaling from the phone lock screen [63,64], leveraging notifications [65], experience sampling and day reconstruction methods [66], or passive sensing [39,40]. When people do decide to track their current mood, conversational agents or on-screen prompts could also encourage people to reflect and log on their moods for earlier in the day, week, or month to facilitate creation of more representative mood histories. However, care needs to be taken to ensure such strategies do not interfere with people's general desire for simple and flexible approaches to mood tracking.

Past work has demonstrated that people find photographs to be an especially useful way to track one's mood [67]. However, commercial apps tend not to leverage this, instead using scales, emojis, or text for mood entry. People frequently take pictures of a variety of mood-relevant events using their phones, such as photos of joyful moments with family members or of stressful whiteboard meetings at work. Leveraging these photos in mood-tracking apps might be another way to maximize information while reducing burden. Sharing pictures and other visual content is an important part of various technologies, especially social technologies such as social media [68]. As it has been noted that one affordance of technology for mental health purposes is the ability to be deeply integrated into people's daily lives [69], mood-tracking technologies might leverage this affordance to design better tracking experiences.

Limitations

Our sample was predominantly women, however, given that we recruited people who had previously used mood-tracking apps and that women tend to use health and mental health apps [50,51], our demographics track with the general rates of adoption of such apps. We also oversampled for college students, with 45% (10/22) of our sample currently in college. College students are more likely to be tech-savvy, early adopters of technology, and may have different experiences or motivations to use these apps than the general population. However, a majority of our sample consisted of noncollege students, as we also recruited from the broader community. All participants were local due to sessions being conducted in-person. We also note that overall, our participants had relatively high levels of education and therefore might have higher levels of data literacy. Therefore, their preferences—including what they found helpful or problematic—might differ from people with less education or lower data literacy. Indeed, past research has demonstrated that people from some backgrounds find current mood-tracking tools less usable, in the case of this study those from low-income and traditionally minoritized racial and ethnic groups [70]. Further challenges around stigma, literacy, and access may prevent other demographic groups and backgrounds from getting the desired benefit from mood tracking.

Our participants' experience with mood-tracking apps may not be representative of wider experiences. For example, many of our participants had used Daylio, which may not hold in other populations. We attempted to mitigate the limitation of the specific app a participant had used by showing participants multiple apps within our card sorting task and getting their feedback. A few screens did come from Daylio, which a few participants indicated preference toward simply due to familiarity. More often, participants weighed whether Daylio, or whichever app they had used, would be more or less helpful than the app they were considering in the activity. Most of our participants (16/22, 73%) were currently using mood-tracking apps, whereas others were past users; as such, the information shared might vary based on their ability to recall mood-tracking apps.

The benefits our participants reported were based on their perceived benefits and we did not test whether or not changes were actually accrued in these constructs over time. Nevertheless, future work could formally evaluate whether mood tracking leads to better understanding of one's mood or the causes and consequences of one's mood. Our results and findings should be considered in light of these limitations.

Future Directions

As the number of technologies for mental health and wellness continue to expand, we are likely to see an increased interest in technologies that address broad aspects of mental health such as mood. We examined the perceptions that people held as to how such mood-tracking apps benefited them, but future work could test such benefits experimentally by either assigning people to use mood-tracking apps or manipulating the conditions under which they used them (eg, looking at entering and reviewing data or the number of times per day a person might be prompted to record their moods). Other work could explore the design of mood-tracking technologies from the perspective of how people currently use technologies generally and mood-tracking apps specifically. For example, mood-tracking capabilities might be integrated into widely used platforms such as messaging or social media to facilitate support and connectedness around mood. Another option would be to build mood-tracking apps optimized for once a day tracking as discussed.

Our findings also highlighted that people's interest in mood-tracking apps was often driven by life events or circumstances and as such we should better understand how the social and cultural factors in people's everyday lives impact perspectives and desired features of mood-tracking apps, specifically and mental health technologies. Technology use for the purpose of understanding and promoting one's mental health needs to be placed into the context of people's lives and the broader way they use technology. Some work has explored the way individuals with depression use technology and social media for emergent practices of self-regulation [71], and we could better understand emergent practices of self-insight and self-knowledge with regard to various affective processes including mood.

Conclusion

In this paper, we explored mood-tracking apps from the vantage point of people's experiences with apps they adopted in real-world settings. We found people often turned to these apps during shifts in their lives—such as negative events or changes in their mental health—with the desire to gain insight and potentially support self-awareness and behavior change. We found various instances in which these apps provided useful support for these goals, such as visualizations (figures or calendars) to help illustrate trends. However, we also found other instances in which these apps fell short, including a lack of recommendations or suggestions to support data interpretation and limited intervention features. Therefore, although mood-tracking apps have helped translate concepts from mood and emotion theories into widely available digital tools, various aspects of their design could be improved to help create tools that people would find more useful.

Acknowledgments

SMS receives funding from One Mind for the operation and management of One Mind PsyberGuide. DAE is partially supported by NSF IIS-1850389. We thank Lisa Vasquez for her help in conducting this study including participating in the affinity diagramming and conducting the coding of transcripts.

Conflicts of Interest

SMS serves on the Scientific Advisory Board for Headspace, for which he receives compensation. Other authors have nothing to declare.

Multimedia Appendix 1
Semi-structured interview guide.
[PDF File (Adobe PDF File), 89 KB - mental_v8i8e29368_app1.pdf]

Multimedia Appendix 2
Data entry screens for card sorting task.
[PDF File (Adobe PDF File), 422 KB - mental_v8i8e29368_app2.pdf]

Multimedia Appendix 3
Data review screens for card sorting task.
[PDF File (Adobe PDF File), 259 KB - mental_v8i8e29368_app3.pdf]

References

1. Substance Abuse and Mental Health Services Administration. Key substance use and mental health indicators in the United States: Results from the 2018 National Survey on Drug Use and Health (HHS Publication No. PEP19-5068, NSDUH Series H-54). Rockville, MD: Center for Behavioral Health Statistics and Quality, Substance Abuse and Mental Health Services Administration; 2019.
2. Horwath E, Johnson J, Klerman G, Weissman M. What are the public health implications of subclinical depressive symptoms? Psych Quart 1994 Dec;65(4):323-337 [FREE Full text] [doi: 10.1007/bf02354307]
3. Rideout V, Fox S, Peebles A, Robb MB. Coping with COVID-19: How young people use digital media to manage their mental health. Coping with COVID-19; 2021. URL: https://hopelab.org/research/coping-with-covid/ [accessed 2021-07-18]
4. Krebs P, Duncan DT. Health App Use Among US Mobile Phone Owners: A National Survey. JMIR Mhealth Uhealth 2015 Nov;3(4):e101 [FREE Full text] [doi: 10.2196/mhealth.4924] [Medline: 26537656]
5. Neary M, Schueller SM. State of the Field of Mental Health Apps. Cogn Behav Pract 2018 Nov;25(4):531-537 [FREE Full text] [doi: 10.1016/j.cbpra.2018.01.002] [Medline: 33100810]
6. Rubanovich CK, Mohr DC, Schueller SM. Health App Use Among Individuals With Symptoms of Depression and Anxiety: A Survey Study With Thematic Coding. JMIR Ment Health 2017 Jun 23;4(2):e22 [FREE Full text] [doi: 10.2196/mental.7603] [Medline: 28645891]
7. Kring AM, Sloan DM. Emotion regulation and psychopathology: A transdiagnostic approach to etiology and treatment. New York, NY: Guilford Press; 2009.
8. Qu C, Sas C, Daudén Roquet C, Doherty G. Functionality of Top-Rated Mobile Apps for Depression: Systematic Search and Evaluation. JMIR Ment Health 2020 Jan 24;7(1):e15321 [FREE Full text] [doi: 10.2196/15321] [Medline: 32012079]
9. Lagan S, D'Mello R, Vaidyam A, Bilden R, Torous J. Assessing mental health apps marketplaces with objective metrics from 29,190 data points from 278 apps. Acta Psychiatr Scand 2021 Apr 09:e1 (forthcoming). [doi: 10.1111/acps.13306] [Medline: 33835483]
10. Khue L, Ouh E, Jarzabek S. Mood self-assessment on smartphones. 2015 Oct Presented at: Wireless Health; 2015 Oct 14; Bethesda, MD p. 1-8 URL: https://doi.org/10.1145/2811780.2811921 [doi: 10.1145/2811780.2811921]

11. Lee K, Hong H. Designing for self-tracking of emotion and experience with tangible modality. 2017 Jun Presented at: Designing Interactive Systems; 2017 Jun 10; Edinburgh, United Kingdon p. 465-475 URL: https://doi.org/10.1145/3064663.3064697 [doi: 10.1145/3064663.3064697]
12. Nahum M, Van Vleet TM, Sohal VS, Mirzabekov JJ, Rao VR, Wallace DL, et al. Immediate Mood Scaler: Tracking Symptoms of Depression and Anxiety Using a Novel Mobile Mood Scale. JMIR Mhealth Uhealth 2017 Apr 12;5(4):e44. [doi: 10.2196/mhealth.6544] [Medline: 28404542]
13. Widnall E, Grant CE, Wang T, Cross L, Velupillai S, Roberts A, et al. User Perspectives of Mood-Monitoring Apps Available to Young People: Qualitative Content Analysis. JMIR Mhealth Uhealth 2020 Oct 10;8(10):e18140 [FREE Full text] [doi: 10.2196/18140] [Medline: 33037875]
14. Caldeira C, Chen Y, Chan L, Pham V, Zheng K. Mobile apps for mood tracking: an analysis of features and user reviews. 2017 Apr Presented at: AMIA Annual Symposium Proceedings; 2017; Washington, DC p. 495-504.
15. Li I, Dey A, Forlizzi J. A stage-based model of personal informatics systems. 2010 Apr Presented at: the SIGCHI conference on human factors in computing systems; 2010; Atlanta, Georgia p. 557-566. [doi: 10.1145/1753326.1753409]
16. Consolvo S, Everitt K, Smith I, Landay J. Design requirements for technologies that encourage physical activity. 2006 Apr Presented at: the SIGCHI conference on Human Factors in computing systems; 2006; Montreal, Quebec, Canada p. 457-466. [doi: 10.1145/1124772.1124840]
17. Cordeiro F, Epstein D, Thomaz E, Bales E, Jagannathan A, Abowd G. Barriers and negative nudgese xploring challenges in food journaling. New York: Association for Computing Machinery; 2015 Apr Presented at: the SIGCHI conference on Human Factors in computing systems; April 18 - 23 2015; Seoul, Republic of Korea p. 1159-1162 URL: http://doi.org/10.1145/2702123.2702155 [doi: 10.1145/2702123.2702155]
18. Tsai CC, Lee G, Raab F, Norman GJ, Sohn T, Griswold WG, et al. Usability and Feasibility of PmEB: A Mobile Phone Application for Monitoring Real Time Caloric Balance. Mobile Netw Appl 2007 Jul 15;12(2-3):173-184. [doi: 10.1007/s11036-007-0014-4]
19. Daskalova N, Metaxa-Kakavouli D, Tran A, Nugent N, Boergers J, McGeary J. SleepCoacher: A personalized automated self-experimentation system for sleep recommendations. 2016 Oct Presented at: User Interface Softwarte and Technology; 2016; Tokyo, Japan p. 347-358 URL: https://dl.acm.org/doi/10.1145/2984511.2984534 [doi: 10.1145/2984511.2984534]
20. Choe E, Lee N, Lee B, Pratt W, Kientz J. Understanding quantified-selfers' practices in collecting and exploring personal data. 2014 Apr Presented at: the SIGCHI Conference on Human Factors in Computing Systems; 2014; Toronto, ON, Canada p. 1143-1152 URL: https://doi.org/10.1145/2556288.2557372 [doi: 10.1145/2556288.2557372]
21. Russell J. Core affect and the psychological construction of emotion. Psychol Rev 2003 Jan;110(1):145-172 [FREE Full text] [doi: 10.1037/0033-295x.110.1.145] [Medline: 12529060]
22. Schroeder J, Chung C, Epstein D, Karkar R, Parsons A, Murinova N, et al. Examining self-tracking by people with migraine: goals, needs, and opportunities in a chronic health condition. In: Proceedings of the 2018 Interactive Systems Conference. 2018 Jul Presented at: Designing Interactive Systems Conference; June 9-13, 2018; Hong Kong p. 135-148. [doi: 10.1145/3196709.3196738]
23. Ayobi A, Marshall P, Cox A, Chen Y. Quantifying the body and caring for the mind: Self-tracking in multiple sclerosis. In: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. Quantifying the body and caring for the mind; 2017 May Presented at: Conference on Human Factors in Computing Systems; May 6-11, 2017; Denver, CO p. 6889-6901. [doi: 10.1145/3025453.3025869]
24. Karkar R, Schroeder J, Epstein D, Pina L, Scofield J, Fogarty J, et al. Tummytrials: a feasibility study of using self-experimentation to detect individualized food triggers. In: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. 2017 May Presented at: CHI Conference on Human Factors in Computing Systems; May 6-11, 2017; Denver, CO p. 6850-6863. [doi: 10.1145/3025453.3025480]
25. Epstein D, Lee N, Kang J, Agapie E, Schroeder J, Pina L, et al. Examining menstrual tracking to inform the design of personal informatics tools. In: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. 2017 May Presented at: the CHI Conference on Human Factors in Computing Systems; 2017; Denver, CO p. 6876-6888. [doi: 10.1145/3025453.3025635]
26. Wolf G. Know thyself: Tracking every facet of life, from sleep to mood to pain, 24/7/365. Wired Magazine. 2009 Jun 22. URL: https://www.wired.com/2009/06/lbnp-knowthyself/ [accessed 2021-06-01]
27. Wang R, Chen F, Chen Z, Li T, Harari G, Tignor S, et al. StudentLife: assessing mental health, academic performance and behavioral trends of college students using smartphones. In: Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing. 2014 Sep Presented at: ACM international joint conference on pervasive and ubiquitous computing; September 13 - 17, 2014; Seattle, WA p. 3-14. [doi: 10.1145/2632048.2632054]
28. Locke EA, Latham GP. Building a practically useful theory of goal setting and task motivation. A 35-year odyssey. Am Psychol 2002 Sep;57(9):705-717. [doi: 10.1037//0003-066x.57.9.705] [Medline: 12237980]
29. Epstein D, Cordeiro F, Fogarty J, Hsieh G, Munson S. Crumbs: Lightweight daily food challenges to promote engagement and mindfulness. In: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems. 2016 May Presented at: CHI Conference on Human Factors in Computing Systems; May 7 - 12, 2016; San Jose, CA p. 5632-5644. [doi: 10.1145/2858036.2858044]

30. Elsden C, Durrant AC, Kirk DS. It's just my history isn't it?: Understanding smart journaling practices. In: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems. 2016 May Presented at: CHI Conference on Human Factors in Computing Systems; May 7 - 12, 2016; San Jose, CA p. 2819-2831. [doi: 10.1145/2858036.2858103]
31. Baumer E, Khovanskaya V, Matthews M, Reynolds L, Sosik V, Gay G. Reviewing reflection: on the use of reflection in interactive system design. In: Proceedings of the 2014 Conference on Designing Interactive Systems. 2014 Jun Presented at: Designing interactive systems; June 7 - 11, 2014; Vancouver, BC, Canada p. 93-102. [doi: 10.1145/2598510.2598598]
32. Epstein D, Cordeiro F, Bales E, Fogarty J, Munson S. Taming data complexity in lifelogs: Exploring visual cuts of personal informatics data. In: Proceedings of the 2014 Conference on Designing Interactive Systems. 2014 Jun Presented at: Designing Interactive Systems; June 21 - 25, 2014; Vancouver BC, Canada p. 667-676. [doi: 10.1145/2598510.2598558]
33. Lane N, Mohammod M, Lin M, Yang X, Lu H, Ali S, et al. Bewell: A smartphone application to monitor, model and promote wellbeing. 2011 May Presented at: 5th international ICST conference on pervasive computing technologies for healthcare; 2011 May 23; Dublin, Ireland p. 23-26. [doi: 10.4108/icst.pervasivehealth.2011.246161]
34. Hollis V, Konrad A, Springer A, Antoun M, Antoun C, Martin R, et al. What Does All This Data Mean for My Future Mood? Actionable Analytics and Targeted Reflection for Emotional Well-Being. Human–Computer Interaction 2017 Jan 04;32(5-6):208-267. [doi: 10.1080/07370024.2016.1277724]
35. Snyder J, Murnane E, Lustig C, Voida S. Visually encoding the lived experience of bipolar disorder. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. 2019 May 02 Presented at: CHI Conference on Human Factors in Computing Systems; 2019; Glasgow, Scotland, UK p. 1-14. [doi: 10.1145/3290605.3300363]
36. Isaacs E, Konrad A, Walendowski A, Lennig T, Hollis V, Whittaker S. Echoes from the past: how technology mediated reflection improves well-being. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2013 Apr 27 Presented at: SIGCHI conference on human factors in computing systems; April 27 - May 2, 2013; Paris, France p. 1071-1080. [doi: 10.1145/2470654.2466137]
37. Schreiber M, Jenny GJ. Development and validation of the 'Lebender emoticon PANAVA' scale (LE-PANAVA) for digitally measuring positive and negative activation, and valence via emoticons. Personality and Individual Differences 2020 Jul;160:109923. [doi: 10.1016/j.paid.2020.109923]
38. Torkamaan H, Ziegler J. Mobile mood tracking: An investigation of concise and adaptive measurement instruments. Proc. ACM Interact. Mob. Wearable Ubiquitous Technol 2020 Dec 17;4(4):1-30. [doi: 10.1145/3432207]
39. Sükei E, Norbury A, Perez-Rodriguez MM, Olmos PM, Artés A. Predicting Emotional States Using Behavioral Markers Derived From Passively Sensed Data: Data-Driven Machine Learning Approach. JMIR Mhealth Uhealth 2021 Mar 22;9(3):e24465 [FREE Full text] [doi: 10.2196/24465] [Medline: 33749612]
40. Yan S, Hosseinmardi H, Kao H, Narayanan S, Lerman K, Ferrara E. Affect Estimation with Wearable Sensors. J Healthc Inform Res 2020 Mar 11;4(3):261-294. [doi: 10.1007/s41666-019-00066-z]
41. 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 Dec 20;20(7):e241 [FREE Full text] [doi: 10.2196/jmir.9775] [Medline: 30030209]
42. Kumar A, Wang M, Riehm A, Yu E, Smith T, Kaplin A. An Automated Mobile Mood Tracking Technology (Mood 24/7): Validation Study. JMIR Ment Health 2020 May 20;7(5):e16237 [FREE Full text] [doi: 10.2196/16237] [Medline: 32432558]
43. Torous J, Staples P, Shanahan M, Lin C, Peck P, Keshavan M, et al. Utilizing a personal smartphone custom app to assess the Patient Health Questionnaire-9 (PHQ-9) depressive symptoms in patients with major depressive disorder. JMIR Ment Health 2015 Mar;2(1):e8 [FREE Full text] [doi: 10.2196/mental.3889] [Medline: 26543914]
44. Church K, Hoggan E, Oliver N. A study of mobile mood awareness and communication through mobimood. In: Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries. 2010 Oct 16 Presented at: 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries; October 16 - 20, 2010; Reykjavik, Iceland p. 128-137. [doi: 10.1145/1868914.1868933]
45. Gay G, Pollak JP, Adams P, Leonard JP. Pilot study of Aurora, a social, mobile-phone-based emotion sharing and recording system. J Diabetes Sci Technol 2011 Mar 01;5(2):325-332 [FREE Full text] [doi: 10.1177/193229681100500219] [Medline: 21527101]
46. Bardram J, Frost M, Szántó K, Faurholt-Jepsen M, Vinberg M, Kessing L. Designing mobile health technology for bipolar disorder: a field trial of the monarca system. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2013 Apr 27 Presented at: SIGCHI Conference on Human Factors in Computing Systems; April 27 - May 2, 2013; Paris France p. 2627-2636. [doi: 10.1145/2470654.2481364]
47. Huguet A, Rao S, McGrath PJ, Wozney L, Wheaton M, Conrod J, et al. A systematic review of cognitive behavioral therapy and behavioral activation apps for depression. PLoS One 2016;11(5):e0154248. [doi: 10.1371/journal.pone.0154248] [Medline: 27135410]
48. Kelley C, Lee B, Wilcox L. Self-tracking for mental wellness: Understanding expert perspectives and student experiences. In: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. 2017 May 02 Presented at: CHI Conference on Human Factors in Computing Systems; May 6 - 11, 2017; Denver, CO p. 629-641. [doi: 10.1145/3025453.3025750]

49. Drake G, Csipke E, Wykes T. Assessing your mood online: acceptability and use of Moodscope. Psychol Med 2013 Jul;43(7):1455-1464 [FREE Full text] [doi: 10.1017/S0033291712002280] [Medline: 23149120]
50. Carroll JK, Moorhead A, Bond R, LeBlanc WG, Petrella RJ, Fiscella K. Who uses mobile phone health apps and does use matter? A secondary data analytics approach. J Med Internet Res 2017 Apr 19;19(4):e125 [FREE Full text] [doi: 10.2196/jmir.5604] [Medline: 28428170]
51. Anguera JA, Jordan JT, Castaneda D, Gazzaley A, Areán PA. Conducting a fully mobile and randomised clinical trial for depression: access, engagement and expense. BMJ Innov 2016 Jan;2(1):14-21 [FREE Full text] [doi: 10.1136/bmjinnov-2015-000098] [Medline: 27019745]
52. Corbin J, Strauss A. Basics of qualitative research: Techniques and procedures for developing grounded theory, 4th Edition. Thousand Oaks, CA: SAGE Publications, Inc; 2014.
53. Simonsen J, Friberg K. Collective analysis of qualitative data. In: Situated design methods. Cambridge, MA: MIT Press; 2014:99-117.
54. Fusch P, Ness L. Are we there yet? Data saturation in qualitative research. The Qualitative Report 2015 Sep;20(9):1408-1416 [FREE Full text] [doi: 10.46743/2160-3715/2015.2281]
55. Davies W. How are we now? Real-time mood-monitoring as valuation. Journal of Cultural Economy 2016 Dec 05;10(1):34-48. [doi: 10.1080/17530350.2016.1258000]
56. McConville C, Cooper C. Personality correlates of variable moods. Personality and Individual Differences 1998 Jan;26(1):65-78. [doi: 10.1016/s0191-8869(98)00139-1]
57. Beck JS. Cognitive behavior therapy: Basics and beyond. New York, NY: The Guilford Press; 2021.
58. Shin G, Feng Y, Jarrahi M, Gafinowitz N. Beyond novelty effect: a mixed-methods exploration into the motivation for long-term activity tracker use. JAMIA Open 2019 Apr;2(1):62-72 [FREE Full text] [doi: 10.1093/jamiaopen/ooy048] [Medline: 31984346]
59. Costa Figueiredo M, Caldeira C, Eikey EV, Mazmanian M, Chen Y. Engaging with health data: The interplay between self-tracking activities and emotions in fertility struggles. Proceedings of the ACM on Human-Computer Interaction 2018 Nov 01;2(CSCW):1-20. [doi: 10.1145/3274309]
60. Rooksby J, Rost M, Morrison A, Chalmers M. Personal tracking as lived informatics. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2014 Apr 26 Presented at: SIGCHI Conference on Human Factors in Computing Systems; April 26 - May 1, 2014; Toronto, ON, Canada p. 1163-1172. [doi: 10.1145/2556288.2557039]
61. Ayobi A, Sonne T, Marshall P, Cox A. Flexible and mindful self-tracking: Design implications from paper bullet journals. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. 2018 Apr 19 Presented at: CHI Conference on Human Factors in Computing Systems; April 21-26, 2018; Montreal, QC, Canada p. 1-14. [doi: 10.1145/3173574.3173602]
62. Mackay W. Triggers and barriers to customizing software. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 1991 Apr Presented at: SIGCHI conference on Human factors in computing systems; April 27 - May 2, 1991; New Orleans, LA p. 153-160. [doi: 10.1145/108844.108867]
63. Choe E, Lee B, Kay M, Pratt W, Kientz J. SleepTight: Low-burden, self-monitoring technology for capturing and reflecting on sleep behaviors. In: Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing. 2015 Sep 07 Presented at: ACM International Joint Conference on Pervasive and Ubiquitous Computing; 2015; Osaka, Japan p. 121-132. [doi: 10.1145/2750858.2804266]
64. Zhang X, Pina L, Fogarty J. Examining unlock journaling with diaries and reminders for in situ self-report in health and wellness. In: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems. 2016 May 07 Presented at: CHI Conference on Human Factors in Computing Systems; 2016; San Jose, CA p. 5658-5664. [doi: 10.1145/2858036.2858360]
65. Bentley F, Tollmar K. The power of mobile notifications to increase wellbeing logging behavior. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2013 Apr 27 Presented at: SIGCHI Conference on Human Factors in Computing Systems; April 27 - May 2, 2013; Paris, France p. 1095-1098. [doi: 10.1145/2470654.2466140]
66. Dockray S, Grant N, Stone AA, Kahneman D, Wardle J, Steptoe A. A comparison of affect ratings obtained with ecological momentary assessment and the day reconstruction method. Soc Indic Res 2010 Jan 19;99(2):269-283. [doi: 10.1007/s11205-010-9578-7]
67. Wallbaum T, Heuten W, Boll S. Comparison of in-situ mood input methods on mobile devices. In: Proceedings of the 15th International Conference on Mobile and Ubiquitous Multimedia. 2016 Dec 12 Presented at: International Conference on Mobile and Ubiquitous Multimedia; December 12 - 15, 2016; Rovaniemi, Finland p. 123-127. [doi: 10.1145/3012709.3012724]
68. Highfield T, Leaver T. Instagrammatics and digital methods: studying visual social media, from selfies and GIFs to memes and emoji. Communication Research and Practice 2016 Apr 25;2(1):47-62. [doi: 10.1080/22041451.2016.1155332]
69. Schueller SM, Muñoz R, Mohr DC. Realizing the Potential of Behavioral Intervention Technologies. Curr Dir Psychol Sci 2013 Dec 03;22(6):478-483. [doi: 10.1177/0963721413495872]

70. Sarkar U, Gourley GI, Lyles CR, Tieu L, Clarity C, Newmark L, et al. Usability of Commercially Available Mobile Applications for Diverse Patients. J Gen Intern Med 2016 Dec;31(12):1417-1426 [FREE Full text] [doi: 10.1007/s11606-016-3771-6] [Medline: 27418347]
71. Eschler J, Burgess E, Reddy M, Mohr D. Emergent self-regulation practices in technology and social media use of individuals living with depression. In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. 2020 Apr 21 Presented at: CHI Conference on Human Factors in Computing Systems; April 25 - 30, 2020; Honolulu, HI p. 1-13. [doi: 10.1145/3313831.3376773]

Edited by J Torous, G Eysenbach; submitted 03.04.21; peer-reviewed by C Escobar-Viera, R Mallick; comments to author 19.04.21; revised version received 09.06.21; accepted 24.06.21; published 11.08.21.

Please cite as:
Schueller SM, Neary M, Lai J, Epstein DA
Understanding People's Use of and Perspectives on Mood-Tracking Apps: Interview Study
JMIR Ment Health 2021;8(8):e29368
URL: https://mental.jmir.org/2021/8/e29368
doi:10.2196/29368
PMID:34383678

Original Paper

The Persistence of the Impact of COVID-19–Related Distress, Mood Inertia, and Loneliness on Mental Health During a Postlockdown Period in Germany: An Ecological Momentary Assessment Study

Matthias Haucke[1,2], BSc, MSc; Shuyan Liu[1*], BA, MSc, PhD; Stephan Heinzel[2*], Dipl-Psych, PhD

[1]Department of Psychiatry and Psychotherapy, Charité – Universitätsmedizin Berlin, Campus Charité Mitte, Berlin, Germany

[2]Department of Education and Psychology, Clinical Psychology and Psychotherapy, Freie Universität Berlin, Berlin, Germany

[*]these authors contributed equally

Corresponding Author:
Shuyan Liu, BA, MSc, PhD
Department of Psychiatry and Psychotherapy
Charité – Universitätsmedizin Berlin
Campus Charité Mitte
Charitéplatz 1
Berlin
Germany
Phone: 49 30450517002
Email: siyan908@hotmail.com

Abstract

Background: The first wave of the COVID-19 pandemic in early 2020 increased mental health problems globally. However, little is known about mental health problems during a low-incidence period of the pandemic without strict public health measures.

Objective: We aim to investigate whether COVID-19–related risk factors for mental health problems persist beyond lockdown measures. We targeted a vulnerable population that is at risk of developing low mental health and assessed their daily dynamics of mood and emotion regulation after a strict lockdown.

Methods: During a postlockdown period in Germany (between August 8, 2020, and November 1, 2020), we conducted an ecological momentary assessment with 131 participants who experienced at least mild COVID-19–related distress and loneliness. To estimate negative mood inertia, we built a lag-1 three-level autoregressive model.

Results: We found that information exposure and active daily COVID-19 cases did not have an impact on negative mood amid a postlockdown period. However, there was a day-to-day carryover effect of negative mood. In addition, worrying about COVID-19, feeling restricted by COVID-19, and feeling lonely increased negative mood.

Conclusions: The mental health of a vulnerable population is still challenged by COVID-19–related stressors after the lifting of a strict lockdown. This study highlights the need to protect mental health during postpandemic periods.

(JMIR Ment Health 2021;8(8):e29419) doi:10.2196/29419

KEYWORDS

COVID-19; outbreaks; epidemics; pandemics; psychological responses and emotional well-being; ecological momentary assessment; risk and protective factors; low incidence and restrictions

Introduction

The COVID-19 pandemic and its associated socioeconomic consequences increased global mental health problems [1,2]. Negative mental health outcomes of the COVID-19 pandemic are associated with fear of becoming infected [3,4] and various mitigation strategies to curb the spread of COVID-19 (eg, curfews and restrictions to public life). These measures can disrupt regular routines, impair mood homeostasis [5-7], and impose economic hardship (eg, income loss and unemployment) [8], which can fuel anxiety, depression, and loneliness [9-12]. However, it is unclear whether these effects continue after lockdown measures have been eased. As variants emerge and cause sudden spikes in COVID-19 case numbers (eg, the B.1.1.7 variant in the United Kingdom in late 2020), fear of getting infected and/or another lockdown could persist. Moreover, after the pandemic and lockdown measures end, socioeconomic uncertainty remains [13]. Chronic psychological distress and social isolation are risk factors for developing mental disorders such as psychosis, substance abuse disorder, and affective disorder [14-17]. To investigate whether COVID-19–related stressors remain beyond lockdown measures, we set up an ecological momentary assessment (EMA) study in Germany during a postlockdown period. We focus on a group at high risk of poor mental health: those who experienced at least mild psychological distress and loneliness amid the COVID-19 pandemic. We expect a carryover effect of negative mood from one measurement to the next (mood inertia) and assume that COVID-19–related stressors (ie, momentary COVID-19–related worry, COVID-19 information seeking and perceived restriction, loneliness, and daily reported COVID-19 cases) result in an increase in momentary negative mood (Multimedia Appendix 1).

Methods

Study Design and Sampling

We conducted an EMA that involves repeated sampling of individuals' current behaviors and experiences in real time and in their natural environments [18] during a postlockdown period (from August 8, 2020, to November 1, 2020) in Germany, when restrictions were lenient (eg, no private or public meeting restrictions, reopening of most leisure facilities, bars, and catering facilities; see Multimedia Appendix 1). EMA aims to minimize recall bias, maximize ecological validity, and approximate temporal causality (ie, Granger causality) and allows researchers to study microprocesses that influence behavior in real-world contexts [19]. Participants were recruited via online advertisements on universities' websites, Twitter, and eBay classifieds. Participants had to fill in an online prequestionnaire on the Siuvo Intelligent Psychological Assessment Platform. After an initial contact via phone or email, we sent participants our study information, informed consent, and a QR code (to install a smartphone app) by mail.

We targeted vulnerable individuals who reported at least mild psychological distress and sometimes felt lonely amid the COVID-19 pandemic. We used the COVID-19 Peritraumatic Distress Index (CPDI [20]; cutoff score=28, indicating mild distress) questionnaire and the short-form version of the UCLA Loneliness Scale (ULS-8 [21]; cutoff score=16, indicating mild loneliness), respectively. Other inclusion criteria were being at least 18 years of age, not working night shifts, not currently infected with COVID-19, using an Android smartphone, and speaking fluent German. The CPDI was designed to evaluate changes in mental health status, cognitive skills, avoidance and compulsive behavior, physical symptoms, and loss of social functioning due to the COVID-19 pandemic. The questionnaire has been previously validated in a sample in Germany [20].

Data Collection

We used a smartphone app called "movisensXS" (movisens GmbH), which was developed for research purposes. The app is compliant with the General Data Protection Regulation (European Union) and Berlin Data Protection Act (Berliner Datenschutzgesetz – BlnDSG). Participants completed a 20-minute baseline assessment, followed by 7 consecutive days in which they received 8 randomized prompts between 8 AM and 10 PM. The study procedure was approved by the Ethics Committees of Charité – Universitätsmedizin Berlin (ref: EA2/143/20) and Freie Universität Berlin (ref: 030/2020).

Measurements

To quantify COVID-19–related distress, we measured worries about the COVID-19 pandemic, perceived restrictions due to the COVID-19 pandemic, COVID-19 information exposure, and feelings of loneliness. Finally, we measured respondents' momentary negative mood (anxiety, depression, fatigue, stress, and unhappiness). All questions were measured on a visual analogue scale ranging from 0 (not at all) to 100 (very much). To account for the steady increase in active COVID-19 cases in Germany during the time of measurement [22], we included daily COVID-19 cases as a predictor in our analysis. Our smartphone study consisted of a sociodemographic assessment (ie, age, gender, years of education) and the EMA. The exact EMA items can be found in Multimedia Appendix 1 and online at [23].

Statistical Analysis

All statistical analyses were conducted in R (version 3.5.3; R Foundation for Statistical Computing [24]). To consider the hierarchical data structure and autoregressive parameters, we performed model selection using autoregressive (AR) multilevel models with the dependent variable negative mood. We followed the approach by Haan-Rietdijk et al [25]; details about the model selection procedure can be found in Multimedia Appendix 1 and online at [23].

Results

We assessed 775 people for eligibility in an online questionnaire. The final sample size was 131 (18%; recruitment flow is shown in Figure 1 and sample characteristics are shown in Table 1; for power estimation, see Multimedia Appendix 1). No participant filled in less than 28 (50%) of the daily questionnaires, while 40 (<0.01%) of the total sent daily questionnaires were not answered by the participants.

Figure 1. Recruitment flow.

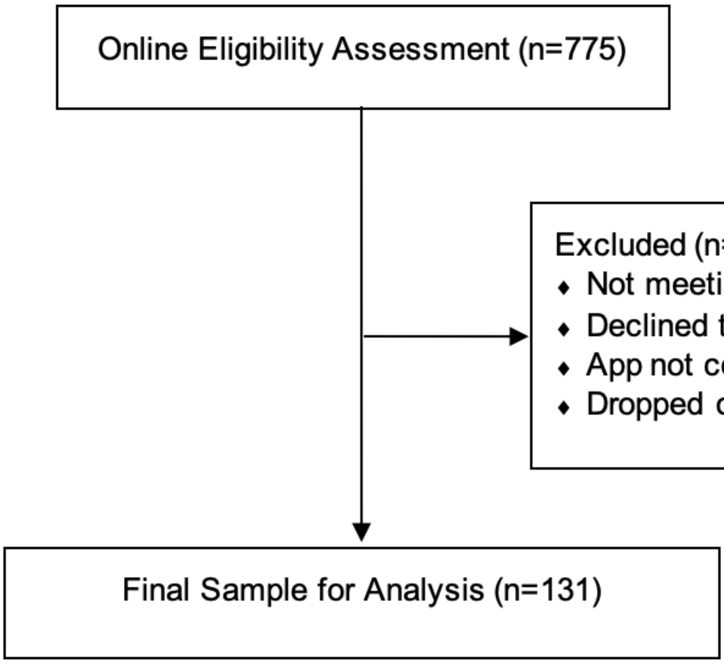


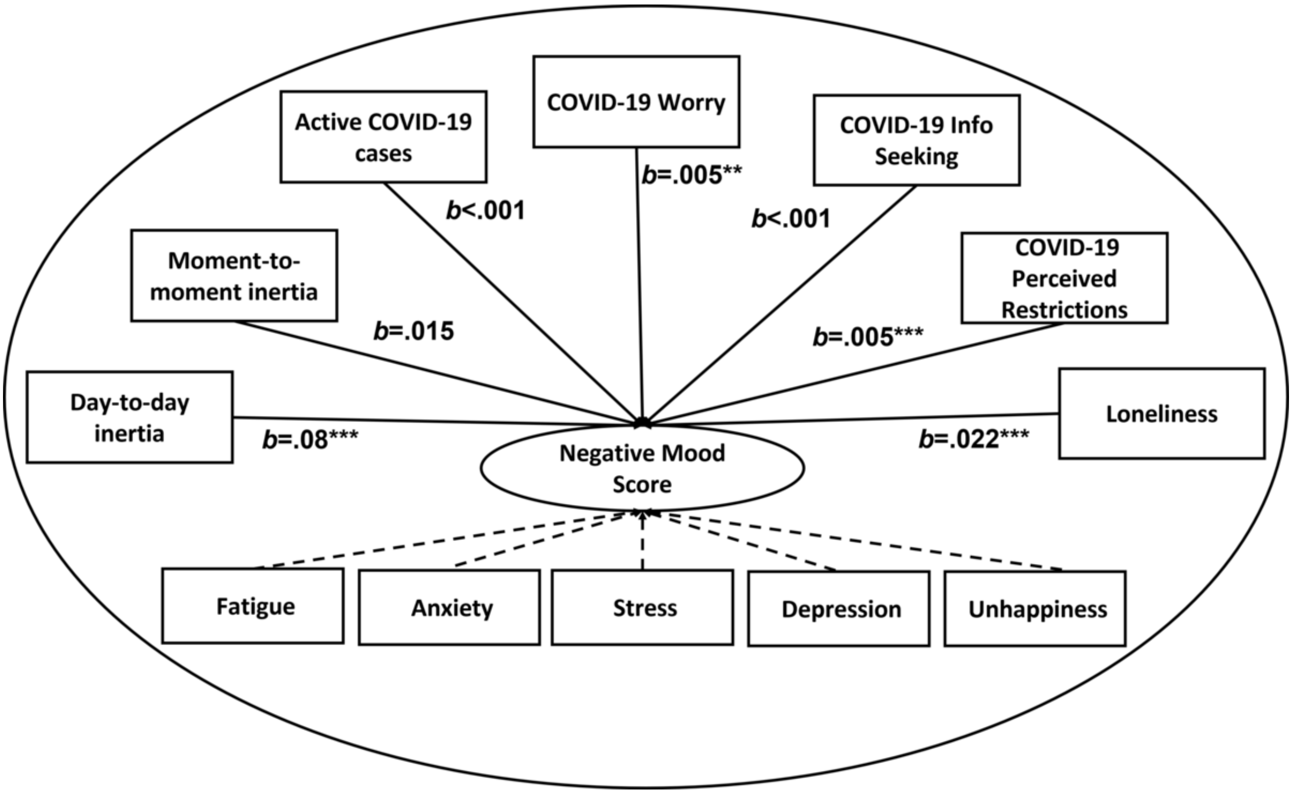
Table 1. Demographics and sample characteristics.

Parameter	Values
COVID-19 Peritraumatic Distress Index score, mean (SD)	48.42 (16.31)
UCLA Loneliness Scale score, mean (SD)	22 (4.03)
Education (in years), mean (SD)	15.08 (3.66)
Age (in years), mean (SD)	31.62 (10.76)
Gender, n (%)	Male: 49 (37); female: 82 (63)

We used a lag-1 three-level AR model, which allows us to separate the variance of negative mood scores into variance at the person level (level 3), variance at the day level (level 2), and variance at the questionnaire level (level 1). We created two lagged variables, a within-day centered predictor at questionnaire level and a within-person centered lagged predictor at the day level. The very first beep of each day (ie, the time period between the previous day's beep and next day's beep) was excluded from the analysis to remove possible unexplained carryover effects resulting from the night (eg, lack of sleep). This model includes mood inertias, COVID-19 worries, COVID-19 information seeking, perceived restrictions, and loneliness during the last hour, as well as daily active COVID-19 cases as random effects. The momentary negative mood score was built by averaging momentary feelings of fatigue, anxiety, depression, unhappiness, and stress. A graphical check indicated a positive skew of negative mood; therefore, we performed a square root transformation on this variable. The analysis script can be found online at [23].

We found that loneliness (b=.022, $t_{3713.83}$=18.68, P<.001), COVID-19 perceived restriction (b=.005, $t_{129.84}$=3.65, P<.001), COVID-19–related worry (b=.005, $t_{132.74}$=2.87, P=.001), and day-to-day mood inertia (b=.078, $t_{134.58}$=3.96, P=.001) increased negative mood scores. Active daily COVID-19 case numbers (b<.001, $t_{92.17}$=–0.27, P=.87), COVID-19–related information seeking (b<.001, $t_{88.41}$=0.73, P=.47), and moment-to-moment inertia (b=.015, $t_{42.19}$=0.17, P=.87) did not increase negative mood scores (see Figure 2).

Figure 2. Loneliness, COVID-19 worries, feelings of restriction, and day-to-day mood inertia increased negative mood. Moment-to-moment mood inertia, active COVID-19 cases, and COVID-19 information seeking did not increase negative mood. *P<.05, **P<.01, ***P<.0001 (two-tailed). N=131.



Discussion

Principal Findings

We found that negative effects of the COVID-19 pandemic on mental health outlast lockdown measures. In line with findings from the first COVID-19 wave [8,26-30], we found that loneliness, worrying about COVID-19, and perceived restriction increased negative mood during a postlockdown period. Similar to the Ebola pandemic [31], possible reasons for the lasting effect of the COVID-19 pandemic might be worries about the negative economic consequences, concern about resurgence of the virus, struggles to rebuild social networks, and/or deliberately withdrawing from social contacts to avoid infection.

Furthermore, we found a negative carryover effect of mood between days (mood inertia), indicating dysfunctional mood regulation. Restrictive policies during the COVID-19 pandemic can impact mental health, possibly due to impaired mood homeostasis (ie, failure to positively regulate mood via mood-modifying activities) [7]. Importantly, our results show that even when the acute threat and restrictive measures are less pronounced, negative daily mood inertia remains.

Neither COVID-19 information seeking nor active COVID-19 cases increased negative mood. This contrasts with previous findings from lockdown periods [6,32]. For example, an EMA study during the first lockdown in Germany and Austria reported increased perceived COVID-19–related restrictions that were positively associated with increased daily news consumption, especially in individuals living alone [32]. In addition, an EMA study conducted in New Jersey in the United States between April 24 and May 26, 2020, showed that undergraduates felt more anxious about COVID-19 on days when the number of new cases and deaths due to COVID-19 were higher [6]. Our opposing finding might be caused by the relatively low domestic case numbers and associated news during the postlockdown period. Moreover, negative COVID-19 news might have less impact on mood over the course of the pandemic as people get accustomed to it.

Limitations

We did not make explicit comparisons to participant status before the COVID-19 outbreak or to a control group, which limits generalizability to other populations. Furthermore, we did not measure adaptability, which has been associated with positive mood (eg, optimism and satisfaction) [33]. Finally, we did not assess the nature of COVID-19 worries and reasons for feeling restricted.

Conclusions

Even if cases are low and lockdown policies are lenient, mental health is still challenged by COVID-19–related stressors. Although information exposure to COVID-19 and daily COVID-19 cases had no impact on mood, we found a day-to-day carryover effect of negative mood. Moreover, COVID-19–related restriction, worry about COVID-19, and loneliness increased negative mood. Thus, the negative impact of the COVID-19 pandemic on mental health outlasts lockdown measures and mental health challenges will likely continue after the pandemic.

Acknowledgments

This research was supported by the Berlin University Alliance (grant to SL and SH).

Conflicts of Interest

None declared.

Multimedia Appendix 1

Supplementary materials.

[DOCX File , 29 KB - mental_v8i8e29419_app1.docx]

References

1. Jacobson NC, Lekkas D, Price G, Heinz MV, Song M, O'Malley AJ, et al. Flattening the Mental Health Curve: COVID-19 Stay-at-Home Orders Are Associated With Alterations in Mental Health Search Behavior in the United States. JMIR Ment Health 2020 Jun 01;7(6):e19347 [FREE Full text] [doi: 10.2196/19347] [Medline: 32459186]
2. Reger MA, Stanley IH, Joiner TE. Suicide Mortality and Coronavirus Disease 2019-A Perfect Storm? JAMA Psychiatry 2020 Nov 01;77(11):1093-1094. [doi: 10.1001/jamapsychiatry.2020.1060] [Medline: 32275300]
3. Nelson BW, Pettitt A, Flannery JE, Allen NB. Rapid assessment of psychological and epidemiological correlates of COVID-19 concern, financial strain, and health-related behavior change in a large online sample. PLoS One 2020 Nov 11;15(11):e0241990 [FREE Full text] [doi: 10.1371/journal.pone.0241990] [Medline: 33175882]
4. Cao W, Fang Z, Hou G, Han M, Xu X, Dong J, et al. The psychological impact of the COVID-19 epidemic on college students in China. Psychiatry Res 2020 May;287:112934 [FREE Full text] [doi: 10.1016/j.psychres.2020.112934] [Medline: 32229390]
5. Saha K, Torous J, Caine ED, De Choudhury M. Psychosocial Effects of the COVID-19 Pandemic: Large-scale Quasi-Experimental Study on Social Media. J Med Internet Res 2020 Nov 24;22(11):e22600 [FREE Full text] [doi: 10.2196/22600] [Medline: 33156805]
6. Kleiman EM, Yeager AL, Grove JL, Kellerman JK, Kim JS. Real-time Mental Health Impact of the COVID-19 Pandemic on College Students: Ecological Momentary Assessment Study. JMIR Ment Health 2020 Dec 15;7(12):e24815 [FREE Full text] [doi: 10.2196/24815] [Medline: 33207308]
7. Taquet M, Quoidbach J, Fried EI, Goodwin GM. Mood Homeostasis Before and During the Coronavirus Disease 2019 (COVID-19) Lockdown Among Students in the Netherlands. JAMA Psychiatry 2021 Jan 01;78(1):110-112 [FREE Full text] [doi: 10.1001/jamapsychiatry.2020.2389] [Medline: 32725176]
8. Liu S, Heinzel S, Haucke MN, Heinz A. Increased Psychological Distress, Loneliness, and Unemployment in the Spread of COVID-19 over 6 Months in Germany. Medicina (Kaunas) 2021 Jan 09;57(1):53 [FREE Full text] [doi: 10.3390/medicina57010053] [Medline: 33435368]
9. Huckins JF, daSilva AW, Wang W, Hedlund E, Rogers C, Nepal SK, et al. Mental Health and Behavior of College Students During the Early Phases of the COVID-19 Pandemic: Longitudinal Smartphone and Ecological Momentary Assessment Study. J Med Internet Res 2020 Jun 17;22(6):e20185 [FREE Full text] [doi: 10.2196/20185] [Medline: 32519963]
10. Rajkumar RP. COVID-19 and mental health: A review of the existing literature. Asian J Psychiatr 2020 Aug;52:102066 [FREE Full text] [doi: 10.1016/j.ajp.2020.102066] [Medline: 32302935]
11. Nekliudov NA, Blyuss O, Cheung KY, Petrou L, Genuneit J, Sushentsev N, et al. Excessive Media Consumption About COVID-19 is Associated With Increased State Anxiety: Outcomes of a Large Online Survey in Russia. J Med Internet Res 2020 Sep 11;22(9):e20955 [FREE Full text] [doi: 10.2196/20955] [Medline: 32788143]
12. Zhang W, Yang X, Zhao J, Yang F, Jia Y, Cui C, et al. Depression and Psychological-Behavioral Responses Among the General Public in China During the Early Stages of the COVID-19 Pandemic: Survey Study. J Med Internet Res 2020 Sep 04;22(9):e22227 [FREE Full text] [doi: 10.2196/22227] [Medline: 32886066]
13. McKibbin W, Fernando R. Economics in the time of COVID-19. London: CEPR Press URL: https://www.incae.edu/sites/default/files/covid-19.pdf [accessed 2021-06-22]
14. Deng W, Cheung S, Tsao S, Wang X, Tiwari A. Telomerase activity and its association with psychological stress, mental disorders, lifestyle factors and interventions: A systematic review. Psychoneuroendocrinology 2016 Feb;64:150-163. [doi: 10.1016/j.psyneuen.2015.11.017] [Medline: 26677763]
15. Heinz A, Deserno L, Reininghaus U. Urbanicity, social adversity and psychosis. World Psychiatry 2013 Oct 04;12(3):187-197 [FREE Full text] [doi: 10.1002/wps.20056] [Medline: 24096775]
16. Narita Z, Stickley A, DeVylder J. Loneliness and psychotic experiences in a general population sample. Schizophr Res 2020 Apr;218:146-150. [doi: 10.1016/j.schres.2020.01.018] [Medline: 32014362]
17. Beutel ME, Klein EM, Brähler E, Reiner I, Jünger C, Michal M, et al. Loneliness in the general population: prevalence, determinants and relations to mental health. BMC Psychiatry 2017 Mar 20;17(1):97 [FREE Full text] [doi: 10.1186/s12888-017-1262-x] [Medline: 28320380]

18. Shiffman S, Stone AA, Hufford MR. Ecological momentary assessment. Annu Rev Clin Psychol 2008 Apr;4(1):1-32. [doi: 10.1146/annurev.clinpsy.3.022806.091415] [Medline: 18509902]
19. Hamilton JD. Time series analysis. Princeton, NJ: Princeton University Press; 1994.
20. Liu S, Heinz A. Cross-Cultural Validity of Psychological Distress Measurement During the Coronavirus Pandemic. Pharmacopsychiatry 2020 Sep 24;53(5):237-238. [doi: 10.1055/a-1190-5029] [Medline: 32583390]
21. Hays R, DiMatteo MR. A short-form measure of loneliness. J Pers Assess 1987 Mar 1;51(1):69-81. [doi: 10.1207/s15327752jpa5101_6] [Medline: 3572711]
22. Germany COVID. Worldometer. URL: https://www.worldometers.info/coronavirus/country/germany [accessed 2021-08-11]
23. The Persistence of COVID-19 Related Distress, Mood Inertia and Loneliness effects on mental health during a post-lockdown phase: An ecological momentary assessment study. OSFHome. URL: https://osf.io/58xqd [accessed 2021-08-11]
24. The Comprehensive R Archive Network. URL: http://cran.r-project.org [accessed 2021-08-11]
25. de Haan-Rietdijk S, Kuppens P, Hamaker EL. What's in a Day? A Guide to Decomposing the Variance in Intensive Longitudinal Data. Front Psychol 2016 Jun 14;7:891 [FREE Full text] [doi: 10.3389/fpsyg.2016.00891] [Medline: 27378986]
26. Li LZ, Wang S. Prevalence and predictors of general psychiatric disorders and loneliness during COVID-19 in the United Kingdom. Psychiatry Res 2020 Sep;291:113267 [FREE Full text] [doi: 10.1016/j.psychres.2020.113267] [Medline: 32623266]
27. Hu Z, Lin X, Chiwanda Kaminga A, Xu H. Impact of the COVID-19 Epidemic on Lifestyle Behaviors and Their Association With Subjective Well-Being Among the General Population in Mainland China: Cross-Sectional Study. J Med Internet Res 2020 Aug 25;22(8):e21176 [FREE Full text] [doi: 10.2196/21176] [Medline: 32759103]
28. Galea S, Merchant RM, Lurie N. The Mental Health Consequences of COVID-19 and Physical Distancing: The Need for Prevention and Early Intervention. JAMA Intern Med 2020 Jun 01;180(6):817-818. [doi: 10.1001/jamainternmed.2020.1562] [Medline: 32275292]
29. Cosco TD, Fortuna K, Wister A, Riadi I, Wagner K, Sixsmith A. COVID-19, Social Isolation, and Mental Health Among Older Adults: A Digital Catch-22. J Med Internet Res 2021 May 06;23(5):e21864 [FREE Full text] [doi: 10.2196/21864] [Medline: 33891557]
30. Cohen-Mansfield J, Muff A, Meschiany G, Lev-Ari S. Adequacy of Web-Based Activities as a Substitute for In-Person Activities for Older Persons During the COVID-19 Pandemic: Survey Study. J Med Internet Res 2021 Jan 22;23(1):e25848. [doi: 10.2196/25848]
31. Van Bortel T, Basnayake A, Wurie F, Jambai M, Koroma AS, Muana AT, et al. Psychosocial effects of an Ebola outbreak at individual, community and international levels. Bull World Health Organ 2016 Mar 01;94(3):210-214 [FREE Full text] [doi: 10.2471/BLT.15.158543] [Medline: 26966332]
32. Arend A, Blechert J, Pannicke B, Reichenberger J. Increased Screen Use on Days With Increased Perceived COVID-19-Related Confinements-A Day Level Ecological Momentary Assessment Study. Front Public Health 2020 Feb 2;8:623205 [FREE Full text] [doi: 10.3389/fpubh.2020.623205] [Medline: 33634062]
33. Besser A, Flett GL, Nepon T, Zeigler-Hill V. Personality, Cognition, and Adaptability to the COVID-19 Pandemic: Associations with Loneliness, Distress, and Positive and Negative Mood States. Int J Ment Health Addict 2020 Nov 18:1-25 [FREE Full text] [doi: 10.1007/s11469-020-00421-x] [Medline: 33230393]

Abbreviations

AR: autoregressive
CPDI: COVID-19 Peritraumatic Distress Index
EMA: ecological momentary assessment
ULS-8: short-form UCLA Loneliness Scale

Edited by J Torous; submitted 06.04.21; peer-reviewed by M Lotto, A Teles; comments to author 04.05.21; revised version received 24.05.21; accepted 18.07.21; published 26.08.21.

Please cite as:
Haucke M, Liu S, Heinzel S
The Persistence of the Impact of COVID-19–Related Distress, Mood Inertia, and Loneliness on Mental Health During a Postlockdown Period in Germany: An Ecological Momentary Assessment Study
JMIR Ment Health 2021;8(8):e29419
URL: https://mental.jmir.org/2021/8/e29419
doi:10.2196/29419
PMID:34347622

Viewpoint

Advancing Health Equity in Digital Mental Health: Lessons From Medical Anthropology for Global Mental Health

Ellen Elizabeth Kozelka[1*], PhD; Janis H Jenkins[1,2], PhD; Elizabeth Carpenter-Song[3*], PhD

[1]Department of Anthropology, University of California San Diego, La Jolla, CA, United States

[2]Department of Psychiatry, University of California San Diego, La Jolla, CA, United States

[3]Department of Anthropology, Dartmouth College, Hanover, NH, United States

[*]these authors contributed equally

Corresponding Author:
Ellen Elizabeth Kozelka, PhD
Department of Anthropology
University of California San Diego
9500 Gilman Drive
La Jolla, CA, 92093
United States
Phone: 1 (858) 534 4145
Email: ekozelka@ucsd.edu

Abstract

Digital health engenders the opportunity to create new effective mental health care models—from substance use recovery to suicide prevention. Anthropological methodologies offer a unique opportunity for the field of global mental health to examine and incorporate contextual mental health needs through attention to the lived experience of illness; engagement with communities; and knowledge of context, structures, and systems. Attending to these diverse mental health needs and conditions as well as the limitations of digital health will allow global mental health researchers, practitioners, and patients to collaboratively create new models for care in the service of equitable, accessible recovery.

(JMIR Ment Health 2021;8(8):e28555) doi:10.2196/28555

KEYWORDS

qualitative methods; digital health; mental health; health equity

Introduction

The past year has seen a dramatic shift in the use of digital health with the rapid adoption of telehealth within mental health settings in response to COVID-19 [1]. It has accelerated processes of patient and provider acceptance of this form of care as well as the institutionalization of digital health within the US health care system [2,3]. There is a growing evidence base regarding the potential for digital mental health [3,4]. At the same time, some scholars have raised concerns regarding the growth of digital health, noting privacy concerns and potential commodification of marginalized populations [5] along with socioeconomic and structural limitations to access [6,7]. As such, researchers and practitioners are experiencing a critical moment of opportunity to gain a deeper understanding of the promise, perils, and limitations for digital mental health. In this viewpoint, we aim to illuminate how the skills and orientations from a medical anthropological approach provide crucial insights into the development and implementation of digital health tools in the service of supporting equitable and meaningful recovery in the field of global mental health. Specifically, we argue that medical anthropology's emphasis on illness experience; engagement with communities; and context, structures, and systems offer lessons for the field of digital mental health.

Attention to Lived Experience of Illness

Psychological and medical anthropology have a long history of investigating cultural conceptions of mental health and illness as well as their relationship to health-seeking behaviors (eg, [8-10]). Cultural conceptions of illness are the "common sense" knowledge used to interpret experience [11]. This framework has shown that biomedical and psychiatric understandings of illness and treatment exist alongside social, spiritual, and supernatural explanations of distress as well as diverse modalities to ameliorate suffering [eg, 12-14]. Cultural conceptions of mental health and illness must be understood because they play a significant role in shaping health-seeking behaviors, treatment options, therapeutic experience, and recovery [15-17].

Attending to lived experiences and understandings of mental illness offer insight into people's expectations and desires in relation to the various therapeutic models and goals presented in digital mental health tools. For example, an ethnographic study on community-based, residential substance use treatment in the United States-México border found spontaneous incorporation of existing digital tools to aid recovery [18]. However, the problems experienced in relation to substance use and the desired solutions could not always or easily be understood through the current biomedical framework. People and their families were seeking spiritual healing for sins or mutual aid on their recovery journey [18]. If a research team were to design novel digital mental health tools for this community, a purely biomedical framework would not fully reflect local conceptions of illness or the majority of available treatment options in Tijuana; only 8.5% of the centers self-identified their treatment as secular and clinical [19]. Ethnographic methods allow research teams to understand the ways people experience mental illness and how they may incorporate digital tools to foster meaningful mental health recovery in ways that match their desires and needs.

Qualitative and ethnographic methods facilitate engagement with first-person perspectives of mental illness and treatment. This approach goes beyond user-centered design [20,21]. Ethnography for digital mental health focuses on understanding more than just the design, acceptability, and functionality, although those are key elements. To bring in ethnographic methods refocuses attention on what it actually feels like to experience and move through the world with mental illness as well as what people's ideas and goals are for recovery, driving innovation toward health equity. This methodology provides deeply contextualized understanding through intensive engagement with individuals and communities rather than a rapid assessment of needs and experiences that may not capture the full context of illness experience. First-person perspectives garnered through ethnographic methods thus provide opportunities to observe if and how technologies are used and made meaningful by people with mental illness, families/caregivers, and care providers.

Ethnographic methods are thus critical for the design and adaptation of digital mental health interventions. Automatic reminders from smart watches to "breathe" do little to create mindfulness habits if the barrier to care is belief (meditation will not help me) rather than memory (I keep forgetting to meditate). Better ethnographic or on-the-ground knowledge of care needs are crucial to informing the best possible avenues for aid (eg, access to smartphones, language options, stigma reduction, incorporating cultural understandings of health and illness, facilitating familial/social support systems, etc). In other words, ethnographic, qualitative methods provide a richer perspective of the fundamental human processes related to mental health, illness, and recovery in context [22], which will positively shape the design and utility of digital tools.

Engagement With Communities

Although digital health offers a critical opportunity to decrease the global burden of mental illness, its content and form must equally reflect the conceptions of illness and priorities of people within the target population alongside evidence-based models for care. Qualitative methods are often applied to examine acceptability in the development and evaluation of digital health interventions [23]. A study examining people's perspectives on and experiences with technology in a community mental health context found high interest in using technology as well as experience using existing smartphone features and mainstream apps to support mental health recovery [23]. For example, patients described using the alarm function as a medication reminder and daily affirmations on Instagram as helpful supports. Through close attention to first-person perspectives, scholars can better understand how underserved and marginalized populations conceptualize their illness and recovery processes to cocreate meaningful digital health services that address their real needs. As such, researchers can harness existing technology to their advantage as well as diminish redundancies by not designing products or apps with limited novel benefits to what people already have [24].

Brewer et al [25] call for community engagement to build technology that responds to the particular needs and desires of minoritized communities. Cocreating the infrastructure for its inclusion could allow vital mental health care to reach the most marginalized populations. Investment in robust digital health systems now will result in broader, long-term access to high quality health care that responds to community needs [2]. Although access and utilization have been increasing steadily even among the most marginalized [26], anthropological methods offer critical insight into the lived experience of mental illness and its intersections with technology in the service of treatment and recovery (eg, [27]).

More research is needed to understand how underserved and marginalized populations conceptualize mental illnesses and digital modalities of treatment as well as frame expectations of and motivations for seeking care. If this work is overlooked, then any scaling up of treatments may fail to result in increased utilization of services since those services will not match desired care. Frank [28] highlights that current models of addiction treatment individualize and depoliticize addiction, separating it from the context in which it occurs as well as the drivers for entering care. We argue that incorporating ethnography and other participatory methods into the design, implementation, evaluation, and adaptation of digital mental health will help address the research to practice gap [29,30].

Further, community-based participatory research can help conceptualize new models of care and recovery that address current problems of equity in mental health care. Including minoritized people with mental illness through participatory methods would allow for digital mental health to mitigate both the power differentials in health care that seem to ignore or devalue patient experiences and perspectives [31,32] as well as racial and ethnic disparities in access to health care systems [33]. These methodologies should include care providers and interdisciplinary researchers along with patients. Digital health creates opportunities for all stakeholders to participate in events they otherwise could not, such as seminars or conferences out of state aimed at shaping best care practices. However, participatory methods must go beyond incorporating people in settings where their insights have historically been marginalized. Expanding digital options to better suit the context of care will allow digital health to “bend the curve,” [2] accelerating access to high quality mental health care that suits the beliefs, needs, and recovery goals of all mental health stakeholders.

Emphasis on Context, Structures, and Systems

Although studies suggest strong interest in using technology to support mental health among people with lived experience of mental illness, incorporating digital mental health has not been straightforward for the average person. Studies have outlined problems in uptake ranging from the overwhelming number of similar options and misalignment between application design and patient needs to the lack of an evidence base and personal data security [34,35]. Context is a key element to shaping the best, coconstructed care practices for marginalized communities (eg, care recipients, family members, Black, Latinx, indigenous, other peoples of color). This includes the cultural, structural, financial, geographic, and material differences that shape a person’s ability to encounter and engage with digital mental health. Thus, context must be the driving source of knowledge if we as global mental health researchers and practitioners hope to increase access to mental health care through digital tools.

For example, despite living on both sides of that border, from as far north as San Francisco and as far south as México City, many women in the community-based substance use treatment study previously mentioned wanted to remain in contact with their peers after they left treatment, following the 12-step orientation of the center. However, due to economic and political conditions, many of them could not travel to the center in Tijuana, México to attend weekly peer-based support meetings for former patients. As such, they began a WhatsApp group in which they hosted text message–based 12-step meetings. In this way, they were able to maintain their support networks, regardless of the type of smartphone they had or their data plans. They could participate whenever they had wireless connection available, be it at home, work, the library, or a coffee shop. Peer-to-peer care via digital platforms provides support and hope to those initially seeking out help to understand their conditions [36] as well as those whose access to their existing support community is limited by structural factors.

Throughout 2020, conversations with participants have illuminated some of the major drawbacks that center around equity in digital mental health care access. For example, many of the women in the community-based substance use treatment study did actively engage in recovery-based activities via their phones. Yet for some, their cell service was regularly turned off due to their inability to pay their bill or they were locked out of Facebook and WhatsApp accounts (primary means of communication) as abusive ex-partners hacked them. This led to women creating multiple accounts under different names, often losing touch with long-distance relatives and friends. As many of their relationships were strained due to previous or ongoing drug use, these changes or loss of contact often signaled relapse and a sharp decrease in both tangible and intangible social support. Context plays a key role in understanding the dynamic effect technologies aimed at providing aid or connection can have on people’s lives.

Similarly, another author found that strong interest in using digital health tools was tempered among people accessing community mental health programs by the limited data plans on the prepaid cell phones that were commonly owned in the population [34]. Although patients expressed interest in using new forms of technology, they were reluctant to download mental health apps that might exceed their data allotments and compromise their ability to use their phones for other needs. This underscores the critical obligation for researchers and practitioners to consider the broader context in which proposed technologies will be used and how such constraints should be foregrounded in the design and development process.

In another study led by one author focusing on adolescent mental health in southern California, the research team designed a smartphone app to help introduce adolescents to mindfulness, after discussing their needs and mental health–seeking behaviors. However, their phones would regularly be taken away by parents for a variety of reasons, from punishment to incentivizing focus on schoolwork. This meant the app, aimed at helping manage stress and anxiety, often had limited utility during periods typically associated with stress, anger, and anxiety for this particular population, mirroring findings by other teams on parental monitoring of mobile technologies by adolescents [37-39]. Cultural context thus relates to not only generalized ideas about digital health but also the dynamic social interactions that shape access to digital health tools.

If the aim of digital health is to expand care to underserved populations, then that care must take into account their differential access. Although inequitable access to mobile phones and other technology has been narrowed for those with mental illness in the United States and some other places in the world [40,41], unreliable access to the internet or continually disrupted phone service due to economic precarity still limits the feasibility of these models of care, typically for those who need it most [42]. Further, even as clinicians aim to help, differential patterns of mental health diagnoses across racial and ethnic groups change the tools people receive and the treatments they are offered, ultimately sharpening gender, class, and racial inequities in care [43,44]. As Hansen [44] aptly argues, “before we talk about how psychiatry can address the social determinants of health, we have to ask how psychiatry itself already is a social determinant of health.” Digital mental health can only address inequities in access if researchers and care providers confront how many identity groups (eg, race, ethnicity, gender, sexual orientation) are structurally marginalized within health care. Torous and Roberts [6] have called for ethical guidelines to understand shifting care needs. We similarly urge for cultural guidelines that do not ignore the very real differences in framings for mental health and necessary care. All stakeholders can then help shape novel treatments to meet their needs from the outset.

Digital Health Aimed at Equity

Digital health has the potential to mitigate power differentials as well as disparities in mental health care systems. Qualitative and ethnographic methods have the ability to inform research and practice for novel mental health care strategies by illuminating how, in what ways, and for whom interventions work [45,46]. Understanding why people choose (or not) to use digital mental health tools will aid attempts by scholars and practitioners to make sense of the ways people evaluate care options. Further, ethnographic methods can help understand how all health care models, including novel ones such as digital mental health, assume certain values and methods for care based on cultural context [47]. In this viewpoint, we seek to point out that these assumed values may not be shared by the diverse populations that digital mental health is intended to serve. Using certain methodologies such as user-centered design [20,21] is a necessary step in mitigating this. However, our call to incorporate ethnography extends beyond the creation of an app. An app alone without community involvement in the design along with research commitment to long-term community care will not change the conditions that increase the risk for mental illness among the most marginalized. Using ethnographic methods in digital mental health projects may help sidestep the individualization of structural inequities as well as increase partnerships to improve both individual health and broader social conditions, in line with both health equity and social justice.

We believe that the conceptual and methodological approach of medical anthropology provides a strong foundation for understanding people's dynamic conceptions of mental illness and experiences within therapeutic models; bridging gaps in understanding that can exist between care providers, patients, family members, and other stakeholders through community engagement to adapt the best possible collaborative care aimed at health equity; and broadening the focus of care and its access beyond the individual to rethink the ways social and structural barriers must be addressed when integrating novel digital health care into health systems. Medical anthropology provides a framework and methodology through which culturally humble [48,49] research and practice aimed at health equity can be conducted to incorporate digital health into the field of global mental health.

Acknowledgments

Funding for EEK's research was provided by The Center for Iberian and Latin American Studies Tinker Field Research Grant and El Colegio de la Frontera Notre (2014), the Society for Psychological Anthropology's Lemelson Fellowship (2015), the UCSD Frontiers of Innovation Scholars Program (2017), and the F.G. Bailey Dissertation Research Grant (2018). EEK's study would not have been possible without the support of Dr Olga Odgers Ortiz and the research team at Colegio de la Frontera Notre. ECS acknowledges the generous support of the research by the Natalia Mental Health Foundation and the West Family Foundation.

Authors' Contributions

EEK and ECS cowrote the manuscript and provided data from their respective research studies throughout the manuscript. JHJ provided data from her study on adolescent mental health. All authors reviewed the final manuscript.

Conflicts of Interest

None declared.

References

1. Wosik J, Fudim M, Cameron B, Gellad Z, Cho A, Phinney D, et al. Telehealth transformation: COVID-19 and the rise of virtual care. J Am Med Inform Assoc 2020 Jun 01;27(6):957-962 [FREE Full text] [doi: 10.1093/jamia/ocaa067] [Medline: 32311034]
2. Torous J, Jän Myrick K, Rauseo-Ricupero N, Firth J. Digital mental health and COVID-19: using technology today to accelerate the curve on access and quality tomorrow. JMIR Ment Health 2020 Mar 26;7(3):e18848 [FREE Full text] [doi: 10.2196/18848] [Medline: 32213476]
3. Ben-Zeev D, Brian RM, Jonathan G, Razzano L, Pashka N, Carpenter-Song E, et al. Mobile health (mHealth) versus clinic-based group intervention for people with serious mental illness: a randomized controlled trial. Psychiatr Serv 2018 Sep 01;69(9):978-985. [doi: 10.1176/appi.ps.201800063] [Medline: 29793397]
4. Naslund JA, Aschbrenner KA, Araya R, Marsch LA, Unützer J, Patel V, et al. Digital technology for treating and preventing mental disorders in low-income and middle-income countries: a narrative review of the literature. Lancet Psychiatry 2017 Jun;4(6):486-500 [FREE Full text] [doi: 10.1016/S2215-0366(17)30096-2] [Medline: 28433615]
5. Zuboff S. The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power. New York, NY: Public Affairs: Hachette Book Group; 2019.
6. Torous J, Roberts LW. Needed innovation in digital health and smartphone applications for mental health: transparency and trust. JAMA Psychiatry 2017 May 01;74(5):437-438. [doi: 10.1001/jamapsychiatry.2017.0262] [Medline: 28384700]

7. Bojorquez I, Aguilera-Guzmán RM, Aguilera A, Mejía-Arango S. Feasibility study of a cell phone intervention to promote mental health among deported mexican migrants. Salud Publica Mex 2018;60(4):379-380. [doi: 10.21149/9040] [Medline: 30137931]
8. Edgerton R. Conceptions of psychosis in Four East African Societies. Am Anthropologist 1966 Apr;68(2):408-425. [doi: 10.1525/aa.1966.68.2.02a00070]
9. Farmer P. Sending sickness: sorcery, politics, and changing concepts of AIDS in rural Haiti. Med Anthropology Q 1990 Mar;4(1):6-27. [doi: 10.1525/maq.1990.4.1.02a00020]
10. Good M. The biotechnical embrace. Cult Med Psychiatry 2001 Dec;25(4):395-410. [doi: 10.1023/a:1013097002487] [Medline: 11800072]
11. White GM, Marsella AJ. Introduction: cultural conceptions in mental health research and practice. In: White GM, Marsella AJ, editors. Cultural Conceptions of Mental Health and Therapy. Boston: Springer, Dordrecht; 1982:3-38.
12. Jenkins JH. Ethnopsychiatric interpretations of schizophrenic illness: the problem of nervios within Mexican-American families. Cult Med Psychiatry 1988 Sep;12(3):301-329. [doi: 10.1007/BF00051972] [Medline: 3234015]
13. Carpenter-Song E, Chu E, Drake RE, Ritsema M, Smith B, Alverson H. Ethno-cultural variations in the experience and meaning of mental illness and treatment: implications for access and utilization. Transcult Psychiatry 2010 Apr;47(2):224-251. [doi: 10.1177/1363461510368906] [Medline: 20603387]
14. Ross N, Timura C, Maupin J. The case of curers, noncurers, and biomedical experts in Pichátaro, Mexico. Resiliency in folk-medical beliefs. Med Anthropol Q 2012 Jun;26(2):159-181. [doi: 10.1111/j.1548-1387.2012.01199.x] [Medline: 22905435]
15. Mechanic D. The concept of illness behavior. J Chronic Dis 1962 Mar;15:189-194. [doi: 10.1016/0021-9681(62)90068-1] [Medline: 14471950]
16. Kleinman A. The Illness Narratives: Suffering, Healing, and the Human Condition. New York, NY: Basic Books Inc; 1988.
17. Jenkins JH, Carpenter-Song E. The new paradigm of recovery from schizophrenia: cultural conundrums of improvement without cure. Cult Med Psychiatry 2005 Dec;29(4):379-413. [doi: 10.1007/s11013-006-9000-8] [Medline: 16773457]
18. Kozelka E. Investigating the Role of Cultural Environment in Addiction Treatment and Recovery in the United States-México Border Zone. La Jolla: University of California, San Diego; 2020.
19. Odgers Ortiz O, Olivas-Hernández OL. Productive misunderstandings: the ambivalent relationship between religious-based treatments and the lay state in Mexico. Int J Latin Am Religions 2019 Nov 21;3(2):356-369. [doi: 10.1007/s41603-019-00091-1]
20. De Vito Dabbs A, Myers BA, Mc Curry KR, Dunbar-Jacob J, Hawkins RP, Begey A, et al. User-centered design and interactive health technologies for patients. Comput Inform Nurs 2009;27(3):175-183 [FREE Full text] [doi: 10.1097/NCN.0b013e31819f7c7c] [Medline: 19411947]
21. Birnbaum F, Lewis D, Rosen RK, Ranney ML. Patient engagement and the design of digital health. Acad Emerg Med 2015 Jun;22(6):754-756. [doi: 10.1111/acem.12692] [Medline: 25997375]
22. Jenkins JH. Extraordinary Conditions: Culture and Experience in Mental Illness. Oakland, CA: University of California Press; 2015.
23. Carpenter-Song E, Jonathan G, Brian R, Ben-Zeev D. Perspectives on mobile health versus clinic-based group interventions for people with serious mental illnesses: a qualitative study. Psychiatr Serv 2020 Jan 01;71(1):49-56. [doi: 10.1176/appi.ps.201900110] [Medline: 31615368]
24. Carpenter-Song E, Noel VA, Acquilano SC, Drake RE. Real-world technology use among people with mental illnesses: qualitative study. JMIR Ment Health 2018 Nov 23;5(4):e10652. [doi: 10.2196/10652] [Medline: 30470681]
25. Brewer LC, Fortuna KL, Jones C, Walker R, Hayes SN, Patten CA, et al. Back to the future: achieving health equity through health informatics and digital health. JMIR Mhealth Uhealth 2020 Jan 14;8(1):e14512 [FREE Full text] [doi: 10.2196/14512] [Medline: 31934874]
26. Ben-Zeev D, Atkins DC. Bringing digital mental health to where it is needed most. Nat Hum Behav 2017 Dec;1(12):849-851. [doi: 10.1038/s41562-017-0232-0] [Medline: 31024173]
27. Jenkins JH, Carpenter-Song EA. Awareness of stigma among persons with schizophrenia: marking the contexts of lived experience. J Nerv Ment Dis 2009 Jul;197(7):520-529. [doi: 10.1097/NMD.0b013e3181aad5e9] [Medline: 19597360]
28. Frank D. "I Was Not Sick and I Didn't Need to Recover": methadone maintenance treatment (MMT) as a refuge from criminalization. Subst Use Misuse 2018 Jan 28;53(2):311-322 [FREE Full text] [doi: 10.1080/10826084.2017.1310247] [Medline: 28704148]
29. Hetrick SE, Robinson J, Burge E, Blandon R, Mobilio B, Rice SM, et al. Youth codesign of a mobile phone app to facilitate self-monitoring and management of mood symptoms in young people with major depression, suicidal ideation, and self-harm. JMIR Ment Health 2018 Jan 23;5(1):e9 [FREE Full text] [doi: 10.2196/mental.9041] [Medline: 29362208]
30. Unertl KM, Schaefbauer CL, Campbell TR, Senteio C, Siek KA, Bakken S, et al. Integrating community-based participatory research and informatics approaches to improve the engagement and health of underserved populations. J Am Med Inform Assoc 2016 Jan;23(1):60-73 [FREE Full text] [doi: 10.1093/jamia/ocv094] [Medline: 26228766]
31. Smith-Oka V. Microaggressions and the reproduction of social inequalities in medical encounters in Mexico. Soc Sci Med 2015 Oct;143:9-16. [doi: 10.1016/j.socscimed.2015.08.039] [Medline: 26331863]

32. Nyttingnes O, Ruud T, Rugkåsa J. 'It's unbelievably humiliating'-Patients' expressions of negative effects of coercion in mental health care. Int J Law Psychiatry 2016;49(Pt A):147-153. [doi: 10.1016/j.ijlp.2016.08.009] [Medline: 27726890]
33. Fiscella K, Sanders MR. Racial and ethnic disparities in the quality of health care. Annu Rev Public Health 2016;37:375-394. [doi: 10.1146/annurev-publhealth-032315-021439] [Medline: 26789384]
34. Carpenter-Song E. Promoting meaningful recovery with digital mental health care. Epidemiol Psychiatr Sci 2020 Mar 26;29:e105 [FREE Full text] [doi: 10.1017/S2045796020000165] [Medline: 32100662]
35. Torous J, Nicholas J, Larsen ME, Firth J, Christensen H. Clinical review of user engagement with mental health smartphone apps: evidence, theory and improvements. Evid Based Ment Health 2018 Aug;21(3):116-119. [doi: 10.1136/eb-2018-102891] [Medline: 29871870]
36. Naslund JA, Aschbrenner KA, Marsch LA, Bartels SJ. The future of mental health care: peer-to-peer support and social media. Epidemiol Psychiatr Sci 2016 Apr;25(2):113-122 [FREE Full text] [doi: 10.1017/S2045796015001067] [Medline: 26744309]
37. Israni ST, Matheny ME, Matlow R, Whicher D. Equity, inclusivity, and innovative digital technologies to improve adolescent and young adult health. J Adolesc Health 2020 Aug;67(2S):S4-S6 [FREE Full text] [doi: 10.1016/j.jadohealth.2020.05.014] [Medline: 32718514]
38. Holtz BE, Murray KM, Hershey DD, Dunneback JK, Cotten SR, Holmstrom AJ, et al. Developing a patient-centered mHealth app: a tool for adolescents with type 1 diabetes and their parents. JMIR Mhealth Uhealth 2017 Apr 19;5(4):e53 [FREE Full text] [doi: 10.2196/mhealth.6654] [Medline: 28428167]
39. George MJ, Odgers CL. Seven fears and the science of how mobile technologies may be influencing adolescents in the digital age. Perspect Psychol Sci 2015 Nov;10(6):832-851 [FREE Full text] [doi: 10.1177/1745691615596788] [Medline: 26581738]
40. Mohan D, Bashingwa JJH, Tiffin N, Dhar D, Mulder N, George A, et al. Does having a mobile phone matter? Linking phone access among women to health in India: an exploratory analysis of the National Family Health Survey. PLoS One 2020;15(7):e0236078 [FREE Full text] [doi: 10.1371/journal.pone.0236078] [Medline: 32687527]
41. Marler W. Mobile phones and inequality: findings, trends, and future directions. New Media Soc 2018 Apr 07;20(9):3498-3520. [doi: 10.1177/1461444818765154]
42. Ben-Zeev D. Mobile health for all: public-private partnerships can create a new mental health landscape. JMIR Ment Health 2016 Jun 06;3(2):e26 [FREE Full text] [doi: 10.2196/mental.5843] [Medline: 27268283]
43. Mendoza S, Rivera-Cabrero AS, Hansen H. Shifting blame: buprenorphine prescribers, addiction treatment, and prescription monitoring in middle-class America. Transcult Psychiatry 2016 Aug;53(4):465-487 [FREE Full text] [doi: 10.1177/1363461516660884] [Medline: 27488225]
44. Hansen H. Substance - induced psychosis: clinical - racial subjectivities and capital in diagnostic apartheid. Ethos 2019 Mar 28;47(1):73-88. [doi: 10.1111/etho.12223]
45. Betancourt TS, Chambers DA. Optimizing an era of global mental health implementation science. JAMA Psychiatry 2016 Mar;73(2):99-100. [doi: 10.1001/jamapsychiatry.2015.2705] [Medline: 26720304]
46. De Silva MJ, Ryan G. Global mental health in 2015: 95% implementation. Lancet Psychiatry 2016 Jan;3(1):15-17. [doi: 10.1016/S2215-0366(15)00567-2] [Medline: 26772058]
47. Six S, Bilsen J, Deschepper R. Dealing with cultural diversity in palliative care. BMJ Support Palliat Care 2020 Aug 21:bmjspcare-2020-002511. [doi: 10.1136/bmjspcare-2020-002511] [Medline: 32826261]
48. Yeager KA, Bauer-Wu S. Cultural humility: essential foundation for clinical researchers. Appl Nurs Res 2013 Nov;26(4):251-256 [FREE Full text] [doi: 10.1016/j.apnr.2013.06.008] [Medline: 23938129]
49. Carpenter-Song EA, Nordquest Schwallie M, Longhofer J. Cultural competence reexamined: critique and directions for the future. Psychiatr Serv 2007 Oct;58(10):1362-1365. [doi: 10.1176/ps.2007.58.10.1362] [Medline: 17914018]

Edited by J Torous; submitted 05.03.21; peer-reviewed by D Ben-Zeev, S Six; comments to author 10.04.21; revised version received 20.05.21; accepted 25.05.21; published 16.08.21.

Please cite as:
Kozelka EE, Jenkins JH, Carpenter-Song E
Advancing Health Equity in Digital Mental Health: Lessons From Medical Anthropology for Global Mental Health
JMIR Ment Health 2021;8(8):e28555
URL: https://mental.jmir.org/2021/8/e28555
doi:10.2196/28555
PMID:34398788

Original Paper

Clinical Utility of Wearable Sensors and Patient-Reported Surveys in Patients With Schizophrenia: Noninterventional, Observational Study

Adrienne C Lahti[1], MD; Dai Wang[2], PhD; Huiling Pei[3], PhD; Susan Baker[3], PhD; Vaibhav A Narayan[3], PhD

[1]Department of Psychiatry and Behavioral Neurobiology, University of Alabama at Birmingham, Birmingham, AL, United States

[2]Janssen Research & Development, Raritan, NJ, United States

[3]Janssen Research & Development, Titusville, NJ, United States

Corresponding Author:
Adrienne C Lahti, MD
Department of Psychiatry and Behavioral Neurobiology
University of Alabama at Birmingham
SC 501, 1720 7th Ave S
Birmingham, AL 35233
United States
Phone: 1 205 996 7171
Email: alahti@uab.edu

Abstract

Background: Relapse in schizophrenia may be preceded by early warning signs of biological, sensory, and clinical status. Early detection of warning signs may facilitate intervention and prevent relapses.

Objective: This study aims to investigate the feasibility of using wearable devices and self-reported technologies to identify symptom exacerbation correlates and relapse in patients with schizophrenia.

Methods: In this observational study, patients with schizophrenia were provided with remote sensing devices to continuously monitor activity (Garmin vivofit) and sleep (Philips Actiwatch), and smartphones were used to record patient-reported outcomes. Clinical assessments of symptoms (Positive and Negative Syndrome Scale and Brief Psychiatric Rating Scale) were performed biweekly, and other clinical scales on symptoms (Clinical Global Impression-Schizophrenia, Calgary Depression Scale), psychosocial functioning, physical activity (Yale Physical Activity Survey), and sleep (Pittsburgh Sleep Quality Index) were assessed every 4 weeks. Patients were observed for 4 months, and correlations between clinical assessments and aggregated device metrics data were assessed using a mixed-effect model. An elastic net model was used to predict the clinical symptoms based on the device features.

Results: Of the 40 patients enrolled, 1 patient relapsed after being stable with evaluable postbaseline data. Weekly patient-reported outcomes were moderately correlated with psychiatric symptoms (Brief Psychiatric Rating Scale total score, *r*=0.29; Calgary Depression Scale total score, *r*=0.37; and Positive and Negative Syndrome Scale total score, *r*=0.3). In the elastic net model, sleep and activity features derived from Philips Actigraph and Garmin vivofit were predictive of the sitting index of the Yale Physical Activity Survey and sleep duration component of the Pittsburgh Sleep Quality Index. On the basis of the combined patient data, a high percentage of data coverage and compliance (>80%) was observed for each device.

Conclusions: This study demonstrated that wearable devices and smartphones could be effectively deployed and potentially used to monitor patients with schizophrenia. Furthermore, metrics-based prediction models can assist in detecting earlier signs of symptom changes. The operational learnings from this study may provide insights to conduct future studies.

Trial Registration: ClinicalTrials.gov NCT02224430; https://www.clinicaltrials.gov/ct2/show/NCT02224430

(JMIR Ment Health 2021;8(8):e26234) doi:10.2196/26234

KEYWORDS

activity; relapse; schizophrenia; sleep; wearable devices; mobile phone

Introduction

Background

Psychotic symptom exacerbation and relapse are frequently observed in patients with schizophrenia and can lead to a decline in social functioning, reduced treatment response, and worsening of clinical outcomes [1]. These patients and their caregivers experience an increased burden because of relapse and consequent hospitalization [2-4]. Relapse in schizophrenia may be preceded by early warning signs, including depressed mood; social withdrawal; and changes in physical activities, feelings, emotions, and sleep disturbances [5]. Therefore, identifying warning signs can enable early intervention to avoid subsequent relapse events [6]. Symptom onset can be rapid; however, continuous monitoring may provide an advantage for early intervention [7].

Web-based data capturing technologies such as Information Technology–Aided Relapse Prevention in Schizophrenia have been piloted to recognize warning signs based on patient reporting of prodromal symptoms of relapse [8]. However, the frequency at which it is practical to obtain this information and the subjective nature of patients' and caregivers' responses pose challenges. Small, unobtrusive remote sensing devices, along with existing mobile technologies, make it possible to capture real-time data on patients' activities, sleep patterns, behaviors, and symptoms. More recent studies have indicated the general availability and acceptability of devices for remote assessment and management [9-11]. Smartphones are commonly used and have multiple embedded sensors (eg, accelerometer, microphone, GPS, and camera). These can be leveraged to collect symptom reports through patient-reported outcome (PRO) surveys and to collect passive data to measure changes in behavior [9,12-19]. A consumer wrist-worn smartwatch or fitness band can additionally provide measurements of precise and objective activity patterns spanning sleep-rest and active-awake periods. These devices have an advantage of generating continuous streaming data that are more reproducible and less obtrusive than relying on patient and caregiver reports alone. In addition, changes in device compliance may itself be a signal and indicate a clinically relevant change in behavior [9,17].

Study Objectives

This clinical study was designed to explore the signatures of relapse. However, because patients were mainly recruited from outpatient clinics and followed up for 4 months—a short period to observe relapses in stabilized patients—there were insufficient relapses to perform the primary objective. We subsequently evaluated the feasibility of using wearable devices (singly and in combination) and self-reporting technologies to identify potentially predictive symptom correlates in patients with schizophrenia or schizoaffective disorder who are at increased risk of relapse. Continuous monitoring using wearable devices (eg, fitness bands and smartwatches) and self-reporting via smartphones were used in this study, and predictive modeling was applied to examine the correlations between clinical assessments and aggregated metrics data.

Methods

Overview

This study was conducted at the University of Alabama, Birmingham (UAB), from August 8, 2015, to March 28, 2016. The protocol was approved by the UAB Institutional Review Board, and all patients or legally authorized representatives provided written informed consent and Health Insurance Portability and Accountability Act authorization before the start of the study.

Patients

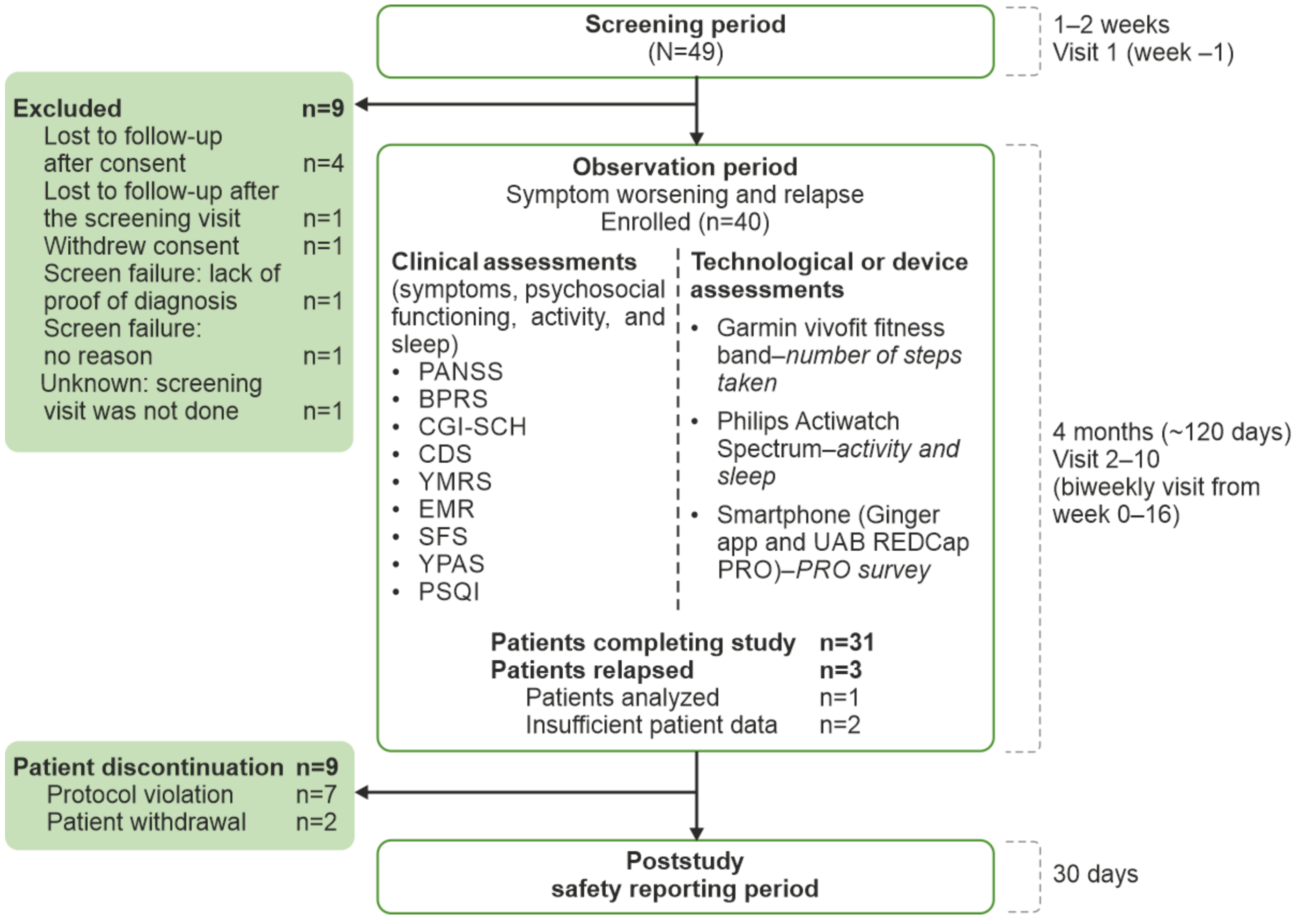
Men and women (aged ≥19 years) who met the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5), criteria for schizophrenia or schizoaffective disorder diagnosed based on the Structured Clinical Interview for DSM-5-Text Revision Axis I Disorders using the Diagnostic Interview for Genetic Studies-4.0 were included in this study. The target population included patients discharged from the inpatient psychiatry unit, emergency department, or outpatient clinics of UAB who were maintained on a stable dose of antipsychotic medication that remained unchanged for 2 weeks before the start of the study.

Patients were excluded if they had physical or clinical disabilities or both, such as hearing, vision, or motor impairment, leading to difficulties in operating a smartphone or responding to prompts (determined using a demonstration smartphone for screening); severe substance use disorder (≥6 symptoms) according to DSM-5 Level 2–Substance Use–Adult scale (adapted from the National Institute on Drug Abuse-Modified Alcohol, Smoking, and Substance Involvement Screening Test); or if they were enrolled or planning to enroll in an interventional study for the treatment or prevention of worsening of symptoms of schizophrenia.

Study Design

This was a noninterventional, observational, exploratory clinical study in which patients were allowed to continue with their usual standard of care and antipsychotic treatment as prescribed by their physician. Patients were screened for eligibility for up to 2 weeks. Enrolled patients were observed for relapse for 4 months (approximately 120 days; observation or study participation period), followed by a 30-day poststudy safety reporting period (Figure 1). Patients were considered to have experienced relapse if they had a rating of moderately severe, very severe, or extremely severe (item score ≥5)—in the previous 2 weeks—for ≥1 item on the Positive and Negative Syndrome Scale (PANSS) positive subscale (items P1-P7) or ≥2 items on the negative subscale (items N1-N7) [20] or if they had symptom exacerbation (increased PANSS total score) that required a change in antipsychotic medication or upward dosage adjustment.

Figure 1. Study design and patient disposition. Patient-reported outcome self-reported symptom questionnaire administered every other day (bidaily) or weekly. Patients were allowed to continue their usual standard of care and antipsychotic treatment as prescribed by their physician but maintained a stable dose, which had not changed for 2 weeks before enrollment. BPRS: modified Brief Psychiatric Rating Scale; CDS: Calgary Depression Scale; CGI-SCH: Clinical Global Impression-Schizophrenia; EMR: electronic medical record; PANSS: Positive and Negative Syndrome Scale; PRO: patient-reported outcome; PSQI: Pittsburgh Sleep Quality Index; REDCap: Research Electronic Data Capture; SFS: Social Functioning Scale; UAB: University of Alabama, Birmingham; YMRS: Young Mania Rating Scale; YPAS: Yale Physical Activity Survey.



Wearable remote sensing devices and a smartphone were provided to eligible patients for use during the observation period. All patients had to undergo a training tutorial for the devices and smartphone based on their individual learning needs. Repeated practice was performed until patients were comfortable using these items. The importance of refraining from tampering or attempting to deactivate the devices was conveyed to the patients. The confidence of patients with using the devices was assessed, and retraining was performed, if required. Potential predictors of symptom worsening or relapse (eg, sleep quality) were collected using remote sensing devices, and the results were subsequently compared with standard clinical assessments. Patients did not have access to the activity and sleep data generated by the wearable devices, so that behavior was independent of feedback.

The 30-day poststudy safety reporting period was per Janssen Adverse Event Reporting Requirements for Noninterventional Studies, wherein all adverse events were recorded from the first use of the Janssen products and for 30 days after the last use of these Janssen drug products within the study.

Clinical Assessments

During the observation period, clinical assessments were performed for symptom worsening and relapse identification. PANSS and modified Brief Psychiatric Rating Scale (BPRS) [21] were assessed every 2 weeks (biweekly), whereas Clinical Global Impression-Schizophrenia (CGI-SCH) [22], Calgary Depression Scale (CDS) [23], Young Mania Rating Scale [24], and electronic medical records were assessed every 4 weeks. Physical activity was assessed using the Yale Physical Activity Survey (YPAS) [25], and sleep was monitored using the Pittsburgh Sleep Quality Index (PSQI) [26] every 4 weeks. Physicians completed the CGI-SCH [22] severity scale for each patient, and a trained study coordinator administered the PANSS to patients. Clinical assessments performed at baseline for patient characterization included psychosocial functioning (assessed using the Social Functioning Scale [27]) or quality of life, cognitive functioning, impulsivity, and measure of addiction. Patient self-reports were also used to assess symptom status and relapse.

Device Data Collection and Processing

General Information

The study was monitored according to the sponsor's current standard operating procedures for the monitoring of clinical trials, and activities were implemented to ensure proper operational study oversight. These activities focused on identifying and resolving operational and quality issues to ensure data integrity, protocol compliance, and safety of the study participants. Written instructions were provided for collecting source documentation, which was reviewed for accuracy and completeness by the sponsor during on-site monitoring visits and underwent internal data reviews throughout the study and at the time of database lock. Discrepancies were resolved with the investigator or designees, as appropriate.

The UAB clinical site captured all clinical assessments using source documents, and these data were entered into the REDCap (Research Electronic Data Capture) system. The nature and location of all source documents were identified to ensure that all sources of original data required to complete data collection were known to the sponsor or investigator and study site personnel.

Smartwatches or fitness bands can monitor activity based on the time frame, duration, and intensity of movement. The unprocessed data collected from wearable devices (eg, accelerometry) may not directly represent variables or metrics that are amenable to patient-relevant interpretation or traditional prespecified statistical measures. The raw data are reduced by the device manufacturers to metrics representing the average state of an individual during specified periods (eg, steps taken, time spent resting or sleeping, and activity intensity levels). These behavioral and lifestyle measures provided a set of metrics that were tested independently and collectively as a pattern toward the specified aims. Analytics can be derived from the patterns that are normal (baseline) for a given patient to detect any relevant changes during clinical follow-up. For sleep, domains that could be assessed through remote sensing devices include onset time, duration, and quality; frequency and pattern of sleep disruptions can be monitored by the number and duration of movements. Similarly for activity, steps per day (mobility) and patterns of daily activity—distribution of high-, medium-, and low-intensity activities—can be assessed. Mobile phone reporting of clinically relevant metrics such as medication compliance; well-being; and the degree of symptom experience, such as seeing or hearing things, could also be assessed.

Device data and PRO responses collected between clinical visits were aggregated to test correlations with the clinical visits and outcomes. Compliance using the devices was also monitored (defined as using or wearing either all devices for 50% of the time in 24 hours or 2 of 4 devices during the 4-month observation period).

Garmin Vivofit Fitness Band

Garmin vivofit is a wristband with an easy-to-read display that was worn at all times by patients to track ambulatory activity (number of steps taken every 15 minutes). A single device was dispensed at visit 1 with instructions for use. At every visit to the clinic, data from the device were downloaded onto the site computer and stored in a secure location. The Garmin vivofit device is a consumer-grade fitness device. Consumer devices have the benefit of higher user acceptance as a social norm. The disadvantage is that the measurements are not clinically validated, and specific firmware versions need to be tracked, as they can impact the results. Despite these challenges, if the intended outcome is detecting relative changes in behavior for individuals rather than cross-sectional studies, they can provide some utility.

Philips Actiwatch Spectrum

The Philips Actiwatch is a wristwatch with an actigraphy system for tracking objective data on off-wrist status, sleep-wake, activity count, and light exposure. Data were collected in 30-second epochs. This actiwatch was worn 24×7, is designed for clinical trials and populations, and is well established for actigraphy-based sleep assessments [28]. A single wristwatch was dispensed to patients at visit 1. At every visit to the clinic, data from the device were downloaded from a proprietary docking station supplied by Philips.

As per the study protocol, patients were requested to wear both Philips Actiwatch and Garmin vivofit devices at all times without additional guidance on the arm preference.

Smartphone

Smartphone apps were used in the study to collect PROs consisting of self-assessment and symptom-tracking questionnaires. One set of questions was given every 2 days (bidaily), and the other set of questions was given weekly throughout the observation period. The Ginger app was used at the beginning of the study for collection, but patients were migrated to the UAB REDCap PRO survey system. The REDCap system collected the same information at the same frequency as the originally deployed Ginger app but had the advantage of direct capture in the clinical database rather than needing to capture the data periodically through a third-party upload. The reason for the switch was that the duration of the trial recruitment period exceeded that of the contracted services with Ginger. A smooth transition was easily developed and implemented, and no data were lost.

Patients who had personal smartphones were asked to download the Ginger app. The site provided a smartphone if the patient did not own one. Patients were restricted to the use of only one personal smartphone with the Ginger app during the course of the study. To prevent the erroneous collection of nonpatient data, the use of the smartphone was restricted to the patient only and was not to be shared with others.

In addition, patients were instructed on the completion of UAB REDCap PRO surveys via an email set up for each patient by the UAB. The email contained a link for access to the survey. Responses to the survey were collected within REDCap through a secure web-based app module to manage and build the web-based surveys.

Both Ginger and REDCap data were recorded on the Health Insurance Portability and Accountability Act–compliant UAB server. At the end of the study, all patients were asked to return their wristwatches. Patients completing the study were allowed to keep the site-issued smartphone, but the monthly plan was terminated. Patients did not have access to the data generated during the course of the study.

Analytical Methods

Statistical Analysis

The sample size for the study was estimated based on previous experience, assuming that 35% to 40% of patients may experience a full significant exacerbation of psychotic symptoms or relapse during the course of the study. The level of relapse anticipated within 120 days for patients treated for a current relapse was based on the experience of UAB in their clinics. For patient characteristics, all continuous variables were summarized using descriptive statistics, and the categorical variables were summarized using frequency measures. As the analyses were exploratory in nature, a two-sided significance level of 5% was used, unless specified otherwise, for all statistical tests. Multiplicity adjustments were not made for the analysis.

Without significant clinical changes to detect or model, we assessed the within-patient stability of the clinical scales using the intraclass correlation coefficient (ICC), a metric commonly used in psychometrics to assess the test-retest and interrater reliability. If the patient-level variance is small for stable patients, it is easier to detect potentially important deviations from individualized norms. The ICC was calculated using a mixed-effect model, in which patient was included as a random effect.

Table 1 summarizes all the metrics obtained from the devices. These metrics were aggregated biweekly to generate the feature sets for predictive modeling. Means and SDs of the device metrics during the 2 weeks immediately before the corresponding clinical assessment were calculated. The summary statistics of all device metrics were combined as the feature set for models predictive of clinical assessments.

Associations among the clinical symptoms and between clinical symptoms and device variables were assessed using mixed-effect models. For each pair, one variable was considered as the response variable and the other variable was considered as the independent variable. Patient was included in the model as a random effect to account for correlations among the repeated measures. Each data variable was scaled so that each had a variance of 1. Testing whether the dependent and independent variables are correlated is equivalent to testing whether the coefficient of the independent variable is 0.

Elastic net [29] was applied to build models using linear regression to predict the clinical assessments and the patient-reported activity and sleep scales using feature sets constructed from the device data. Ten-fold repeated cross-validation with 30 repeats was also applied to train the models and assess the performance of the model. The repeated cross-validation was conducted at the patient level, where data on 10% of patients were held out for validation. The performance of the predictive models was assessed using the root mean square error and the R^2. The R^2 was calculated as the difference between the total sum of squares and the sum of squares owing to the error sum of squares divided by the total sum of squares. When the R^2 is calculated based on cross-validation results and the sum of squares owing to error is calculated based on out-of-sample predictions, the estimated R^2 may be negative.

Table 1. Metrics obtained from Philips Actigraph, Garmin vivofit, and surveys.

Device and metric	Frequency
Philips Actigraph	
Time spent sedentary (minutes)	Daily
Time spent on low-intensity activities (minutes)	Daily
Time spent on moderate-intensity activities (minutes)	Daily
Time spent on vigorous-intensity activities (minutes)	Daily
Total activity count	Daily, nightly
Average activity count	Daily, nightly
Maximum activity count	Daily, nightly
Duration of active or sleeping (minutes)	Daily, nightly
Percentage of time with invalid sleep-wake status	Daily, nightly
Time spent awake (minutes)	Daily, nightly
Percentage of time spent awake	Daily, nightly
Number of wake bouts	Nightly
Average duration of wake bouts (minutes)	Nightly
Time spent sleeping (minutes)	Daily, nightly
Percentage of time spent sleeping	Daily, nightly
Sleep onset latency (minutes)	Nightly
Time spent resting after waking up (minutes)	Nightly
Time of valid rest (minutes)	Nightly
Sleep efficiency	Nightly
Sleep start time (hours:minutes)	Nightly
Rest start time (hours:minutes)	Nightly
Sleep fragmentation	Nightly
Garmin vivofit	
Time spent sedentary (seconds)	Daily
Time spent walking (seconds)	Daily
Total steps	Daily
Number of epochs	Daily
Number of steps during nighttime	Nightly
Survey	
Bidaily survey summary score	Bidaily
Feel down or depressed	Bidaily
Feel confused or have trouble with your thinking	Bidaily
Feel stressed or overwhelmed	Bidaily
See or hear things that other people could not see or hear	Bidaily
Feel suspicious or paranoid	Bidaily
Have trouble sleeping the night before	Bidaily
Weekly survey summary score	Weekly
Feel anxious or nervous	Weekly
Feel unmotivated	Weekly
Have trouble getting things done	Weekly
Missed any schizophrenia medications in the past 7 days	Weekly

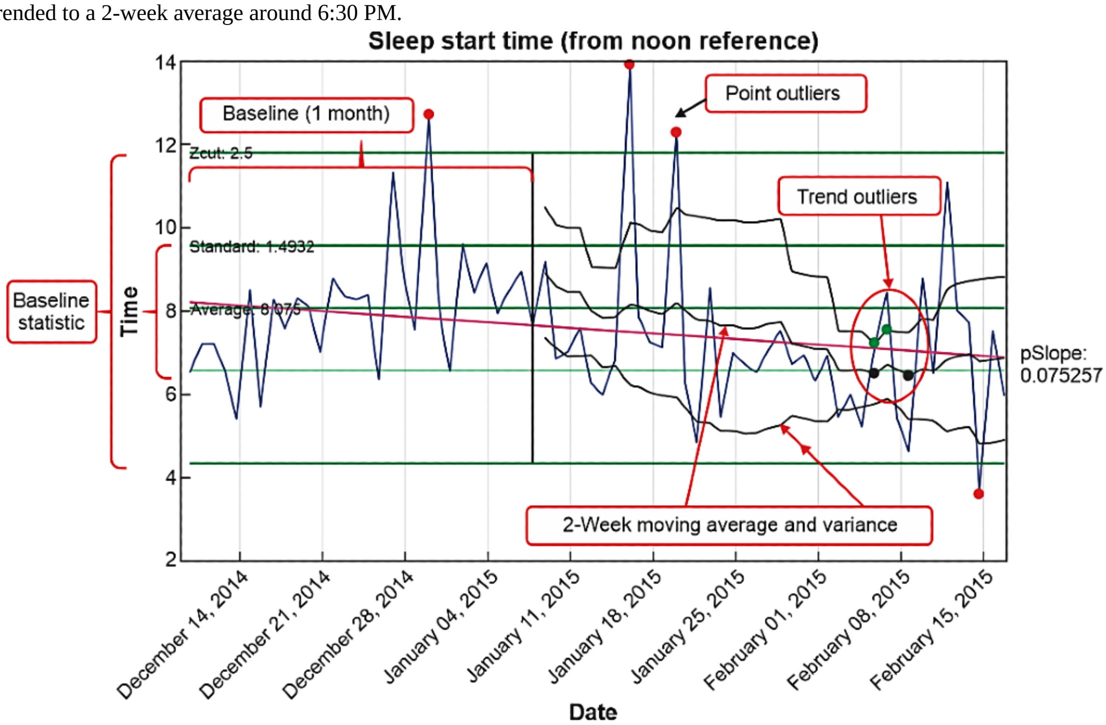
Detecting Statistically Relevant Change in Streaming Data

There is evidence that detectable signs of relapse may be specific to an individual [17]. As such, it is important to develop metrics that can detect statistically relevant changes based on a person's own data rather than what may be normal for individuals. Metrics derived from the devices or PRO can be monitored as a *real-time* streaming process similar to process control in manufacturing or even the stock market. A stable longitudinal period can be used to establish a baseline norm (mean and variance) for each metric of interest. At any subsequent point, statistical tests can be performed for point outlier status of the daily tracked measures or if the short-term mean and variance of a defined measurement period are statistically different from the baseline measures, indicating that a trend may be occurring.

Outliers and trends may have causal explanations other than relapse; however, detected changes create an opportunity to have conversations with the patient about behaviors, with objective data. For example, in Figure 2, examination of sleep onset data revealed a significant trend toward earlier sleep times as the patient approached relapse. The average sleep onset time at baseline was 8 PM. Near relapse, the data trended to a 2-week average around 6:30 PM.

It is relatively easy to detect statistically relevant changes in the streaming metrics. The challenge is assigning causality to the change without an understanding of the context [13,16,17]. Most detected changes are probably not relevant to relapse but could be used in a clinical setting for follow-up or conversations with caregivers based on these objective data.

Figure 2. Sleep onset time for a single relapse patient. Outliers and trends may have causal explanations other than relapse but allow an opportunity to have conversations about behaviors with objective data. The first month was used to establish baseline normal data. Outlier alerts can be set based on baseline mean and variance. In this case, outlier detection alerts were set at 2.5 SDs from the baseline population distribution. After the baseline, a moving average window of 2 weeks was calculated and overlayed on daily data and was used to look for significant changes or trends from the baseline mean or variance. Finally, statistically relevant changes in the slope over a defined interval can indicate relevant changes. The flagged points represent outliers or changes above a set threshold. Flags noted on the moving average mean indicate significant differences for the moving average mean, and flags noted on the moving average variance indicate significant differences in variance from the baseline. For example, examining the sleep onset data, there is a significant trend toward earlier sleep times as the patient approached relapse. The average sleep onset time at baseline was 8 PM. Near relapse, the data had trended to a 2-week average around 6:30 PM.



Results

Patient Disposition and Baseline Characteristics

A total of 49 patients were screened, and 40 patients were enrolled in the study. In total, 9 participants were screened out because of loss to follow-up after consent (n=4) and after the screening visit (n=1), consent withdrawal (n=1), screen failure—lack of proof of diagnosis (n=1) and no reason (n=1), and unknown—screening visit not being performed (n=1).

Of these 40 patients, 31 (78%) completed the study and 9 (23%) discontinued the study (protocol violations: 18%, 7/40 and withdrawal of consent: 5%, 2/40; Figure 1). Only 8% (3/40) of patients experienced a relapse during the study, of which only 1 patient had sufficient postscreening data to establish a reference baseline. The enrolled patients had a higher proportion of men (25/40; 63%), a median age of 40.3 years, and a mean baseline BMI of 34.9 kg/m^2. Most patients (38/40;95%) were on at least one antipsychotic medication during the course of the study, and the most common medication was Risperdal Consta (28/40; 70%). All patients continued their prescribed medication throughout the study (Table 2).

Table 2. Patients' demographics and baseline characteristics (N=40).

Characteristic	Values
Age (years), median (range)	40.3 (19-65)
Gender, n (%)	
Women	15 (38)
Men	25 (63)
Race, n (%)	
Black or African American	29 (73)
White	11 (28)
Baseline BMI (kg/m^2), mean (SD)	34.9 (8.65)
Baseline PANSS[a] total score, mean (SD)	35.7 (5.02)
Baseline CGI-S[b] score, n (%)	
Normal, not at all ill	1 (3)
Borderline mentally ill	4 (10)
Mildly ill	11 (28)
Moderately ill	17 (43)
Markedly ill	7 (18)

[a]PANSS: Positive and Negative Syndrome Scale.

[b]CGI-S: Clinical Global Impression-Severity scale.

Clinical Assessments

For the patient experiencing a relapse, the CGI disease severity scores worsened from 2 to 5 (borderline mentally to markedly ill), and the global improvement changed from 4 to 6 (*no change* to *much worse*). The PANSS domains that worsened at relapse included positive subscale items (delusions-P1 and conceptual disorganization-P2), negative subscale items (poor rapport-N3), and general psychopathology scale items (somatic concern-G1, anxiety-G2, tension-G4, depression-G6, uncooperative-G8, lack of judgment-G12, and disturbance of volition-G13). The BPRS scale worsened in the subdomains of somatic concern, anxiety, conceptual disorganization, tension, depressive mood, uncooperative, and disorientation. The Young Mania Rating Scale worsened in the subdomains of sleep, language, thought disorder, and insight, and the CDS worsened in the subdomains of depression and early wakening.

Correlations Among Clinical Symptoms

For the 40 enrolled patients, the within-patient stability of the clinical symptom scales as assessed by ICC showed either excellent or good agreement, suggesting that most of the assessed symptoms were stable throughout the study (Multimedia Appendix 1). When assessing the association among clinical symptoms, the PANSS total score and the BPRS total scores showed a very strong correlation (r=0.97). However, the Clinical Global Impression-Severity scale score did not strongly correlate with the PANSS and BPRS scores (Table 3).

Table 3. Associations among psychiatric symptoms (N=40).

Symptom scores	BPRS[a] total score	CGI-S[b] score	CDS[c] total score	YMRS[d] total score
PANSS[e] total score				
Coefficient	*0.97*[f]	0.12	*0.68*	*0.29*
P value	*<.001*	.09	*<.001*	*.03*
BPRS total score				
Coefficient	—[g]	*0.15*	*0.65*	*0.34*
P value	—	*.01*	*<.001*	*.002*
CGI-S score				
Coefficient	—	—	*0.24*	0.23
P value	—	—	*.003*	.18
CDS total score				
Coefficient	—	—	—	*0.44*
P value	—	—	—	*.003*

[a]BPRS: modified Brief Psychiatric Rating Scale.
[b]CGI-S: Clinical Global Impression-Severity scale.
[c]CDS: Calgary Depression Scale.
[d]YMRS: Young Mania Rating Scale.
[e]PANSS: Positive and Negative Syndrome Scale.
[f]Italic values indicate significant correlation (*P*<.05).
[g]Not applicable.

Device-Based Assessments

With 1 patient having a stable baseline before relapse, it is not possible to infer statistical indicators from the data that might have suggested an impending potential relapse. However, a few anecdotal observations have been made. Before relapse, the day-to-day variation in mobility was very high, ranging from 1 to 7 miles per day, as captured by the Garmin vivofit fitness band. Some disrupted sleeping was captured using the Philips Actiwatch. The patient reported a spike in *the feeling of suspicion* and missed his medication dose at this time.

Data Coverage and Patient Compliance

Across all patients, the range of days each device was in use provided the maximum number of observable days for each patient-device combination. On the basis of the combined patient data, a high percentage of data coverage and compliance was observed for each device. Using the Garmin vivofit, Philips Actiwatch Spectrum, PRO bidaily survey, and PRO weekly survey, the data coverage was 96%, 92%, 80%, and 89%, respectively, and the device compliance was 97%, 94%, 82%, and 88%, respectively.

Correlations Between Clinical Symptoms and Device Metrics

The low variability in clinical symptoms prevented assessments of individual-based changes in device data with clinical status changes. Subsequently, an analysis was performed to directly correlate the aggregated device metrics with related clinical measures. Correlations were observed between the Philips Actigraph sleep features and the PSQI sleep duration component. Features such as sleep duration, time spent sleeping, and time of valid rest significantly correlated with PSQI sleep duration (r=0.36, r=0.36, and r=0.34, respectively). Sleep start time and resting-state time showed more modest but significant correlations with PSQI sleep disturbances (r=−0.26 and r=−0.25, respectively; Table 4).

Correlations were observed between the Philips Actigraph activity features and the YPAS individual indices (Table 5). Garmin activity feature—total steps taken—showed a modest correlation with the BPRS total score (r=−0.23; P=.03). Garmin activity metrics also showed correlations with the YPAS global or individual indices (Tables 6 and 7).

Table 4. Associations between Philips Actigraph sleep features and the Pittsburgh Sleep Quality Index global and component scores.

Metric	PSQI[a] sleep duration		PSQI sleep disturbances	
	Coefficient	*P* value	Coefficient	*P* value
Total activity count	0.03	.80	0.24	.10
Average activity count	−0.14	.22	0.27	.08
Maximum activity count	0	.99	0.17	.23
Duration (minutes)	*0.36*[b]	*<.001*	−0.02	.88
Time spent awake (minutes)	0.11	.30	0.21	.10
Percentage of time spent awake	−0.1	.40	0.25	.09
Number of wake bouts	0.26	.01	0.09	.45
Average number of wake bouts	−0.18	.08	0.22	.11
Time spent sleeping	*0.36*	*.001*	−0.07	.58
Percentage of time spent sleeping	0.1	.40	−0.25	.09
Sleep onset latency (minutes)	0.05	.56	−0.02	.87
Time spent resting after waking up (minutes)	−0.09	.39	nc[c]	nc
Time of valid rest (minutes)	*0.34*	*<.001*	−0.01	.92
Sleep efficiency	0.21	.09	−0.23	.11
Sleep start time (hours:minutes)	−0.01	.93	*−0.26*	*.01*
Rest start time (hours:minutes)	−0.03	.78	*−0.25*	*.01*
Sleep fragmentation	0	.99	0.17	.18

[a]PSQI: Pittsburgh Sleep Quality Index.

[b]Italic values indicate significant correlation (*P*<.05).

[c]nc: the model did not converge.

Table 5. Associations between Philips Actigraph activity features and the Yale Physical Activity Survey global or individual indexes.

Activity features	YPAS[a] global index	YPAS vigorous activity index	YPAS leisure walking index	YPAS moving index	YPAS standing index	YPAS sitting index
Time spent sedentary (minutes)						
Coefficient	*−0.35*[b]	nc[c]	0.01	*−0.25*	*−0.36*	*0.27*
P value	*.01*	—[d]	.95	*.02*	*.01*	*.01*
Time spent on low-intensity activities (minutes)						
Coefficient	*0.40*	nc	0.07	*0.26*	*0.40*	*−0.3*
P value	*.003*	—	.61	*.03*	*.003*	*.004*
Time spent on moderate-intensity activities (minutes)						
Coefficient	0.82	0.93	nc	nc	1.03	−0.05
P value	.25	.17	—	—	.26	.97
Time spent on vigorous-intensity activities (minutes)						
Coefficient	0.22	0.29	0.05	nc	0.45	nc
P value	.61	.70	.89	—	.31	—
Total activity count						
Coefficient	*0.33*	nc	nc	0.22	*0.33*	−0.2
P value	*.02*	—	—	.14	*.02*	.06
Average activity count						
Coefficient	0.33	nc	0.06	0.25	*0.33*	−0.22
P value	.06	—	.63	.27	*.03*	.13
Maximum activity count						
Coefficient	0.25	nc	−0.02	0.12	*0.26*	−0.01
P value	.14	—	.89	.25	*.03*	.95
Duration (minutes)						
Coefficient	0.06	0.1	−0.03	0	0.07	0.02
P value	.46	.30	.70	.97	.48	.86
Time spent awake (minutes)						
Coefficient	nc	nc	−0.06	0.11	0.05	0.01
P value	—	—	.57	.26	.81	.91
Time spent awake (%)						
Coefficient	nc	nc	−0.01	0.19	0.06	0.01
P value	—	—	.93	.13	.77	.96
Time spent sleeping						
Coefficient	nc	−0.13	0.01	−0.18	−0.06	0.04
P value	—	.49	.95	.13	.78	.70
Time spent sleeping (%)						
Coefficient	nc	nc	0.01	−0.19	−0.06	−0.01
P value	—	—	.93	.13	.77	.96

[a]YPAS: Yale Physical Activity Survey.

[b]Italic values indicate significant correlation ($P<.05$).

[c]nc: the model did not converge.

[d]Not applicable.

Table 6. Associations between Garmin activity features and clinical scores.

Activity features	PANSS[a] total score	BPRS[b] total score	CGI-SCH[c] score	CDS[d] total score	YMRS[e] total score
Time spent sedentary (seconds)					
Coefficient	0.15	0.12	0.25	−0.03	nc[f]
P value	.66	.59	.27	.88	—[g]
Time spent walking (seconds)					
Coefficient	−0.15	−0.12	−0.25	0.03	nc
P value	.65	.58	.26	.88	—
Total steps					
Coefficient	−0.21	*−0.23*[h]	−0.07	nc	−0.13
P value	.19	*.03*	.44	—	.40

[a]PANSS: Positive and Negative Syndrome Scale.
[b]BPRS: modified Brief Psychiatric Rating Scale.
[c]CGI-SCH: Clinical Global Impression-Schizophrenia.
[d]CDS: Calgary Depression Scale.
[e]YMRS: Young Mania Rating Scale.
[f]nc: the model did not converge.
[g]Not applicable.
[h]Italic values indicate significant correlation (*P*<.05).

Table 7. Associations between Garmin activity features and the Yale Physical Activity Survey global or individual indexes.

Activity features	YPAS[a] global index	YPAS vigorous activity index	YPAS leisure walking index	YPAS moving index	YPAS standing index	YPAS sitting index
Time spent sedentary (seconds)						
Coefficient	−0.15	−0.14	0.09	*−0.3*[b]	*−0.29*	*0.25*
P value	.21	.26	.59	*.04*	*.03*	*.05*
Time spent walking (seconds)						
Coefficient	0.15	0.14	−0.1	*0.31*	*0.29*	−0.25
P value	.20	.26	.57	*.03*	*.03*	.05
Total steps						
Coefficient	*0.17*	0.1	*0.38*	0.26	0.13	*−0.21*
P value	*.05*	.26	*.03*	.11	.33	*.04*

[a]YPAS: Yale Physical Activity Survey.
[b]Italic values indicate significant correlation (*P*<.05).

A significant but moderate association between survey data features and psychiatric symptoms was observed. The bidaily survey summary score correlated with the BPRS total score (*r*=0.23; *P*=.05) and the CDS total score (*r*=0.37; *P*=.01). The weekly survey summary score was also associated with the PANSS total score (*r*=0.30; *P*=.03), the BPRS total score (*r*=0.29; *P*=.01), and the CDS total score (*r*=0.37; *P*=.01). Specific queries such as *feel down or depressed*, *feel confused or have trouble with your thinking*, *feel stressed or overwhelmed*, and *have trouble getting things done* correlated with the CDS total score (Table 8).

In the elastic net model, survey data (bidaily, weekly, and bidaily+weekly) were observed to be predictive of the PANSS total score, the BPRS total score, and the CDS total score (Table 9). The features derived from the survey data were found to be predictive of the total score, subjective sleep quality component, and sleep disturbance component of the PSQI (Table 4). The features from Philips Actigraph data were also found to be predictive of the sleep latency component of the PSQI. The sleep and activity features derived from the Philips Actigraph and Garmin activity data were observed to be predictors of the sleep duration component of the PSQI (Table 10) and the sitting index of the YPAS (Multimedia Appendix 2).

Table 8. Associations between survey data features and psychiatric symptoms.

Survey data features	PANSS[a] total score	BPRS[b] total score	CGI-SCH[c] score	CDSS[d] total score	YMRS[e] total score
Bidaily survey summary score					
Coefficient	0.27	*0.23* [f]	nc[g]	*0.37*	nc
P value	.09	*.05*	—[h]	*.01*	—
Feel down or depressed					
Coefficient	0.28	0.2	nc	*0.34*	0.01
P value	.11	.10	—	*.01*	.93
Feel confused or have trouble with your thinking					
Coefficient	0.25	0.19	nc	*0.35*	0.12
P value	.09	.08	—	*.01*	.34
Feel stressed or overwhelmed					
Coefficient	0.29	0.16	–0.04	*0.25*	0.02
P value	.06	.11	.49	*.04*	.86
See or hear things that other people could not see or hear					
Coefficient	0.26	0.2	0	0.11	nc
P value	.22	.13	.99	.42	—
Have trouble sleeping the night before					
Coefficient	0.10	0.10	–0.07	0.02	–0.04
P value	.39	.22	.48	.84	.69
Weekly survey summary score					
Coefficient	*0.30*	*0.29*	–0.2	*0.37*	0.09
P value	*.03*	*.01*	.26	*.01*	.47
Feel anxious or nervous					
Coefficient	0.09	0.18	–0.17	0.26	0.09
P value	.71	.12	.16	.06	.55
Feel unmotivated					
Coefficient	0.19	*0.23*	–0.06	0.26	0.1
P value	.16	*.04*	.70	.20	.47
Have trouble getting things done					
Coefficient	*0.35*	*0.25*	0.05	*0.28*	0.09
P value	*.04*	*.04*	.73	*.04*	.40
Missed any schizophrenia medications in past 7 days					
Coefficient	0.23	0.03	nc	0.11	0.37
P value	.46	.86	—	.19	.07

[a]PANSS: Positive and Negative Syndrome Scale.
[b]BPRS: modified Brief Psychiatric Rating Scale.
[c]CGI-SCH: Clinical Global Impression-Schizophrenia.
[d]CDS: Calgary Depression Scale.
[e]YMRS: Young Mania Rating Scale.
[f]Italic values indicate significant correlation (*P*<.05).
[g]nc: the model did not converge.
[h]Not applicable.

Table 9. Performance of the elastic net models for predicting clinical scores (N=40).

Feature set	RMSE[a]	SD	R^2	SD
PANSS[b] total score				
Survey bidaily	*4.27*[c]	*0.03*	*0.11*	*0.01*
Survey weekly	*4.23*	*0.05*	*0.12*	*0.02*
Survey (bidaily+weekly)	*4.20*	*0.04*	*0.13*	*0.02*
BPRS[d] total score				
Survey bidaily	*3.90*	*0.03*	*0.19*	*0.01*
Survey weekly	*3.94*	*0.06*	*0.17*	*0.02*
Survey (bidaily+weekly)	*3.86*	*0.05*	*0.21*	*0.02*
CGI[e]-severity				
Survey bidaily	0.84	0.01	0.03	0.03
Survey weekly	0.87	0.01	−0.03	0.02
Survey (bidaily+weekly)	0.84	0.01	0.03	0.02
CDS[f] total score				
Survey bidaily	*2.48*	*0.03*	*0.25*	*0.02*
Survey weekly	*2.52*	*0.03*	*0.23*	*0.02*
Survey (bidaily+weekly)	*2.45*	*0.03*	*0.27*	*0.02*

[a]RMSE: root mean square error.

[b]PANSS: Positive and Negative Syndrome Scale.

[c]Italic values indicate that the uncorrected *P* value for R^2 is <.05 for elastic net models.

[d]BPRS: modified Brief Psychiatric Rating Scale.

[e]CGI: Clinical Global Impression

[f]CDS: Calgary Depression Scale.

Table 10. Performance of the elastic net models for predicting the Pittsburgh Sleep Quality Index global and component scores.

Feature set	RMSE[a]	SD	R^2	SD
PSQI[b] global score				
Philips Actigraph sleep	2.68	0.01	–0.01	0.01
Philips Actigraph activity	2.67	0.01	–0.01	0.01
Philips Actigraph (sleep+activity)	2.67	0.01	–0.01	0.01
Garmin activity	2.62	0.02	0.03	0.01
Survey bidaily	*2.40*[c]	*0.03*	*0.19*	*0.02*
Survey weekly	*2.41*	*0.02*	*0.18*	*0.02*
Survey (bidaily+weekly)	*2.41*	*0.03*	*0.18*	*0.02*
PSQI subjective sleep quality				
Philips Actigraph sleep	0.76	0.01	–0.01	0.01
Philips Actigraph activity	0.75	0.00	–0.01	0.01
Philips Actigraph (sleep+activity)	0.76	0.00	–0.01	0.01
Garmin activity	0.76	0.00	–0.01	0.01
Survey bidaily	*0.73*	*0.01*	*0.05*	*0.02*
Survey weekly	*0.74*	*0.01*	*0.03*	*0.01*
Survey (bidaily+weekly)	*0.73*	*0.01*	*0.04*	*0.02*
PSQI sleep latency				
Philips Actigraph sleep	1.64	0.01	–0.01	0.01
Philips Actigraph activity	*1.58*	*0.01*	*0.05*	*0.01*
Philips Actigraph (sleep+activity)	*1.60*	*0.01*	*0.04*	*0.01*
Garmin activity	1.64	0.01	−0.01	0.01
Survey bidaily	1.62	0.01	0.02	0.01
Survey weekly	1.63	0.01	−0.00	0.01
Survey (bidaily+weekly)	1.61	0.01	0.02	0.01
PSQI sleep duration				
Philips Actigraph sleep	*1.43*	*0.02*	*0.13*	*0.02*
Philips Actigraph activity	*1.42*	*0.02*	*0.15*	*0.02*
Philips Actigraph (sleep+activity)	*1.44*	*0.01*	*0.12*	*0.02*
Garmin activity	*1.50*	*0.02*	*0.06*	*0.02*
Survey bidaily	*1.52*	*0.01*	*0.03*	*0.01*
Survey weekly	1.55	0.01	−0.01	0.01
Survey (bidaily+weekly)	1.52	0.01	0.03	0.01
PSQI habitual sleep efficiency				
Philips Actigraph sleep	7.38	0.02	−0.00	0.01
Philips Actigraph activity	7.38	0.03	−0.00	0.01
Philips Actigraph (sleep+activity)	7.38	0.02	−0.00	0.01
Garmin activity	7.38	0.02	−0.00	0.01
Survey bidaily	7.38	0.03	−0.00	0.01
Survey weekly	7.38	0.02	−0.00	0.01
Survey (bidaily+weekly)	7.38	0.02	−0.00	0.01
PSQI sleep disturbances				

Feature set	RMSE[a]	SD	R^2	SD
Philips Actigraph sleep	3.76	0.08	0.05	0.04
Philips Actigraph activity	3.86	0.01	−0.00	0.01
Philips Actigraph (sleep+activity)	3.83	0.04	0.01	0.02
Garmin activity	3.85	0.02	0.00	0.01
Survey bidaily	*3.34*	*0.04*	*0.25*	*0.02*
Survey weekly	*3.35*	*0.03*	*0.24*	*0.01*
Survey (bidaily+weekly)	*3.27*	*0.04*	*0.28*	*0.02*
PSQI use of sleep medication				
Philips Actigraph sleep	1.35	0.00	−0.00	0.01
Philips Actigraph activity	1.35	0.01	−0.01	0.01
Philips Actigraph (sleep+activity)	1.35	0.01	−0.01	0.01
Garmin activity	1.36	0.01	−0.01	0.01
Survey bidaily	1.33	0.02	0.02	0.03
Survey weekly	1.32	0.01	0.03	0.02
Survey (bidaily+weekly)	1.34	0.02	0.02	0.02
PSQI daytime dysfunction				
Philips Actigraph sleep	0.96	0.01	−0.04	0.01
Philips Actigraph activity	0.96	0.01	−0.04	0.02
Philips Actigraph (sleep+activity)	0.97	0.01	−0.04	0.01
Garmin activity	0.97	0.01	−0.04	0.01
Survey bidaily	0.96	0.01	−0.04	0.02
Survey weekly	0.95	0.03	−0.01	0.06
Survey (bidaily+weekly)	0.96	0.02	−0.03	0.04

[a]RMSE: root mean square error.

[b]PSQI: Pittsburgh Sleep Quality Index.

[c]Italic values indicate significant correlation (*P*<.05).

Discussion

Principal Findings

With the recent increase in the use of mobile technologies (smartphones) and remote sensing wearable devices (smartwatches) with inbuilt sensors, continuous tracking of activity, sleep, and self-reported parameters has been explored widely in various psychiatric disorders [9,12,13,17,19,30-33]. This study demonstrates that parameters that are potentially relevant to the detection of relapse, such as changes in activity and sleep [34], and self-reported symptoms can be collected using smart devices in patients with schizophrenia or schizoaffective disorder.

Several studies have shown that the vast majority of patients with schizophrenia report sleep abnormalities, which often tend to occur before the onset of illness and can predict an acute exacerbation of psychotic symptoms [35]. Individuals with schizophrenia also report significant changes in activity before a relapse episode [36,37]. Therefore, continuous sleep, activity level, and self-reported metrics provide an objective and time-stamped record of changes in behavior that may be clinically relevant and can become a part of the clinical dialog. Although there may be a potential in using remote sensing technologies to track physical activity and sleep behavior to detect relapses, patient adherence and acceptance can limit the effectiveness of these apps. Individuals with schizophrenia may often refuse to use such devices because of a lack of familiarity with the technologies, with using the device, and with the disease state (eg, paranoid behavior). In this study, high data coverage and compliance indicated the overall acceptance of all remote sensing mobile devices used. Our experience in this study suggests that the involvement of patients' physician, good training, and clear communication related to the research goals and potential benefits and risks of the study are important factors in ensuring good compliance and engagement.

In this study, correlations among the individual clinical symptoms as well as between clinical symptoms and device measures were evaluated. The assessment of within-patient stability of the clinical scales showed that most patients' symptoms were stable throughout the study. The correlation analysis among clinical symptoms indicated that the PANSS total score and the BPRS total score showed a strong correlation.

Patient-reported survey data features were correlated with clinical symptoms.

In this study, an elastic net model was used to predict correlations between clinical assessments and the device metrics. The Philips Actigraph and the Garmin activity data showed some utility in predicting the sitting index of the YPAS. The features derived from the patient-reported survey data were useful in predicting the total score, the subjective sleep quality component, and the sleep disturbance component of the PSQI. Features derived from the Philips Actigraph and Garmin vivofit were predictive of the sleep duration component of the PSQI. The implementation of technology-based measures along with metrics-based prediction methods could be effective in instituting early warning systems for symptom decline and relapse patterns in schizophrenia.

Limitations

With limited relapses or clinical changes, this study was not sufficiently powered to identify predictors of relapse or symptom exacerbation by comparing clinical and technological assessments. Although a higher relapse rate was expected based on historical rates in the clinic, we speculate that there may have been a recruitment bias for this study. Less stable subjects may have been less interested in participation. In addition, the frequent clinical assessments and questionnaires and the presence of the devices would have generated a heightened personal focus on relevant symptoms. Future trial designs should consider tracking recruitment bias in the design. In general, expanding the study to more patients and a longer duration would be required to develop robust models predictive of clinical change. The longitudinal stability of the patients would facilitate, however, the ability to detect relative changes and outliers when they occur. Although detecting change is facilitated by having a long-term stable baseline, patients that are unstable will inherently have more variability. Higher variability can make the detection of event outliers more difficult. Having continuous data, however, allows statistical averaging for longer periods than single-point estimates. Trends and changes in variability can be detected with streaming data, even in unstable subjects, more readily than with point estimates from clinical visits. This study demonstrates that remote sensing devices and mobile technology can be used to monitor metrics that are relevant for patients with schizophrenia. It will be important to capture the context of detected changes in future studies to determine the metrics most predictive of relapse or symptom exacerbation.

Conclusions

In summary, relapse prediction using remote sensing technologies may aid clinicians to be cautioned in advance to detect the approaching exacerbation of symptoms and patterns of relapse in patients with schizophrenia. The operational learnings from this study provide insights to conduct future studies with remote sensing devices in this patient population to devise earlier intervention strategies.

Acknowledgments

The authors thank the study participants, without whom this study would never have been accomplished, and the investigators for their participation in this study. The authors also thank Antonia Covin for project management and Michael Beyer for the information technology framework. Ramji Narayanan, MPharm, ISMPP CMPP (SIRO Clinpharm Pvt Ltd), provided writing assistance, and Ellen Baum, PhD (Janssen Global Services, LLC), provided additional editorial support for this manuscript.

Conflicts of Interest

This study is supported by Janssen Research & Development, LLC, United States. SB, DW, HP, and VAN are employees of Janssen Research & Development and hold company stocks. ACL has no conflicts of interest to declare.

Multimedia Appendix 1
Stability of the clinical scales during the study.
[DOCX File , 15 KB - mental_v8i8e26234_app1.docx]

Multimedia Appendix 2
Performance of the elastic net models for predicting the global and individual Yale Physical Activity Survey indices.
[DOCX File , 17 KB - mental_v8i8e26234_app2.docx]

References

1. Emsley R, Chiliza B, Asmal L, Harvey BH. The nature of relapse in schizophrenia. BMC Psychiatry 2013;13:50 [FREE Full text] [doi: 10.1186/1471-244X-13-50] [Medline: 23394123]
2. Caqueo-Urízar A, Miranda-Castillo C, Lemos GS, Lee MS, Ramírez PM, Mascayano TF. An updated review on burden on caregivers of schizophrenia patients. Psicothema 2014 May;26(2):235-243. [doi: 10.7334/psicothema2013.86] [Medline: 24755026]
3. Hor K, Taylor M. Suicide and schizophrenia: a systematic review of rates and risk factors. J Psychopharmacol 2010 Nov;24(4 Suppl):81-90 [FREE Full text] [doi: 10.1177/1359786810385490] [Medline: 20923923]
4. Chan SW. Global perspective of burden of family caregivers for persons with schizophrenia. Arch Psychiatr Nurs 2011 Oct;25(5):339-349. [doi: 10.1016/j.apnu.2011.03.008] [Medline: 21978802]

5. Jørgensen P. Early signs of psychotic relapse in schizophrenia. Br J Psychiatry 1998 Apr;172:327-330. [doi: 10.1192/bjp.172.4.327] [Medline: 9715335]
6. Birchwood M. Early intervention in psychotic relapse: cognitive approaches to detection and management. Behav Change 2014 Oct 06;12(1):2-19. [doi: 10.1017/s0813483900004344]
7. Wang D, Gopal S, Baker S, Narayan VA. Trajectories and changes in individual items of positive and negative syndrome scale among schizophrenia patients prior to impending relapse. NPJ Schizophr 2018 Jun 20;4(1):10 [FREE Full text] [doi: 10.1038/s41537-018-0056-6] [Medline: 29925851]
8. Spaniel F, Vohlídka P, Hrdlicka J, Kozený J, Novák T, Motlová L, et al. ITAREPS: Information technology aided relapse prevention programme in schizophrenia. Schizophr Res 2008 Jan;98(1-3):312-317. [doi: 10.1016/j.schres.2007.09.005] [Medline: 17920245]
9. 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]
10. Finucane MJ. Patients' attitudes to and experience of the use of technology in mental health in an Irish primary care setting. Int J Integr Care 2019 Aug 08;19(4):370. [doi: 10.5334/ijic.s3370]
11. Gay K, Torous J, Joseph A, Pandya A, Duckworth K. Digital technology use among individuals with schizophrenia: results of an online survey. JMIR Ment Health 2016;3(2):e15 [FREE Full text] [doi: 10.2196/mental.5379] [Medline: 27146094]
12. 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]
13. Kluge A, Kirschner M, Hager OM, Bischof M, Habermeyer B, Seifritz E, et al. Combining actigraphy, ecological momentary assessment and neuroimaging to study apathy in patients with schizophrenia. Schizophr Res 2018 May;195:176-182. [doi: 10.1016/j.schres.2017.09.034] [Medline: 29030262]
14. Buck B, Scherer E, Brian R, Wang R, Wang W, Campbell A, et al. Relationships between smartphone social behavior and relapse in schizophrenia: a preliminary report. Schizophr Res 2019 Jun;208:167-172. [doi: 10.1016/j.schres.2019.03.014] [Medline: 30940400]
15. Bradstreet S, Allan S, Gumley A. Adverse event monitoring in mHealth for psychosis interventions provides an important opportunity for learning. J Ment Health 2019 Oct 26;28(5):461-466. [doi: 10.1080/09638237.2019.1630727] [Medline: 31240970]
16. Meyer N, Kerz M, Folarin A, Joyce DW, Jackson R, Karr C, et al. Capturing rest-activity profiles in schizophrenia using wearable and mobile technologies: development, implementation, feasibility, and acceptability of a remote monitoring platform. JMIR Mhealth Uhealth 2018 Oct 30;6(10):e188. [doi: 10.2196/mhealth.8292] [Medline: 30377146]
17. Ben-Zeev D, Brian R, Wang R, Wang W, Campbell AT, Aung MS, et al. CrossCheck: Integrating self-report, behavioral sensing, and smartphone use to identify digital indicators of psychotic relapse. Psychiatr Rehabil J 2017 Sep;40(3):266-275. [doi: 10.1037/prj0000243] [Medline: 28368138]
18. Kerz M, Folarin A, Meyer N, Begale M, MacCabe J, Dobson RJ. SleepSight: A wearables-based relapse prevention system for schizophrenia. In: Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct. 2016 Presented at: UbiComp '16: The 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing; September 12 - 16, 2016; Heidelberg Germany p. 113-116. [doi: 10.1145/2968219.2971419]
19. Lunsford-Avery JR, LeBourgeois MK, Gupta T, Mittal VA. Actigraphic-measured sleep disturbance predicts increased positive symptoms in adolescents at ultra high-risk for psychosis: a longitudinal study. Schizophr Res 2015 May;164(1-3):15-20 [FREE Full text] [doi: 10.1016/j.schres.2015.03.013] [Medline: 25818627]
20. Kay SR, Fiszbein A, Opler LA. The positive and negative syndrome scale (PANSS) for schizophrenia. Schizophr Bull 1987;13(2):261-276. [doi: 10.1093/schbul/13.2.261] [Medline: 3616518]
21. Overall J, Gorham D. The Brief Psychiatric Rating Scale (BPRS): Recent developments in ascertainment and scaling. Psychopharmacol Bull 1988;24(1):97-99 [FREE Full text] [doi: 10.1037/t01554-000]
22. Haro JM, Kamath SA, Ochoa S, Novick D, Rele K, Fargas A, SOHO Study Group. The Clinical Global Impression-Schizophrenia scale: a simple instrument to measure the diversity of symptoms present in schizophrenia. Acta Psychiatr Scand Suppl 2003(416):16-23. [doi: 10.1034/j.1600-0447.107.s416.5.x] [Medline: 12755850]
23. Müller MJ, Brening H, Gensch C, Klinga J, Kienzle B, Müller K. The Calgary Depression Rating Scale for schizophrenia in a healthy control group: psychometric properties and reference values. J Affect Disord 2005 Sep;88(1):69-74. [doi: 10.1016/j.jad.2005.04.005] [Medline: 16040126]
24. Young RB, Ziegler V, Meyer D. Young Mania Rating Scale. In: Handbook of Psychiatric Measures. Washington, DC: American Psychiatric Association; 2000:1-820.
25. Dipietro L, Caspersen CJ, Ostfeld AM, Nadel ER. A survey for assessing physical activity among older adults. Med Sci Sports Exerc 1993 May;25(5):628-642. [Medline: 8492692]
26. Buysse DJ, Reynolds CF, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. Psychiatry Res 1989 May;28(2):193-213. [doi: 10.1016/0165-1781(89)90047-4] [Medline: 2748771]

27. Birchwood M, Smith J, Cochrane R, Wetton S, Copestake S. The Social Functioning Scale. The development and validation of a new scale of social adjustment for use in family intervention programmes with schizophrenic patients. Br J Psychiatry 1990 Dec 02;157(6):853-859. [doi: 10.1192/bjp.157.6.853] [Medline: 2289094]
28. Actiwatch spectrum. Philips. URL: https://www.usa.philips.com/healthcare/product/HC1046964/actiwatch-spectrum-activity-monitor [accessed 2020-10-06]
29. Zou H, Hastie T. Regularization and variable selection via the elastic net. J Royal Statistical Soc B 2005 Apr;67(2):301-320. [doi: 10.1111/j.1467-9868.2005.00503.x]
30. Nuij C, van Ballegooijen W, Ruwaard J, de Beurs D, Mokkenstorm J, van Duijn E, et al. Smartphone-based safety planning and self-monitoring for suicidal patients: rationale and study protocol of the CASPAR (Continuous Assessment for Suicide Prevention And Research) study. Internet Interv 2018 Sep;13:16-23 [FREE Full text] [doi: 10.1016/j.invent.2018.04.005] [Medline: 30206514]
31. Faurholt-Jepsen M, Bauer M, Kessing LV. Smartphone-based objective monitoring in bipolar disorder: status and considerations. Int J Bipolar Disord 2018 Jan 23;6(1):6 [FREE Full text] [doi: 10.1186/s40345-017-0110-8] [Medline: 29359252]
32. Biagianti B, Hidalgo-Mazzei D, Meyer N. Developing digital interventions for people living with serious mental illness: perspectives from three mHealth studies. Evid Based Ment Health 2017 Nov;20(4):98-101 [FREE Full text] [doi: 10.1136/eb-2017-102765] [Medline: 29025862]
33. Aledavood T, Torous J, Triana Hoyos AM, Naslund JA, Onnela J, Keshavan M. Smartphone-Based Tracking of Sleep in Depression, Anxiety, and Psychotic Disorders. Curr Psychiatry Rep 2019 Jun 04;21(7):49 [FREE Full text] [doi: 10.1007/s11920-019-1043-y] [Medline: 31161412]
34. Reeve S, Sheaves B, Freeman D. The role of sleep dysfunction in the occurrence of delusions and hallucinations: a systematic review. Clin Psychol Rev 2015 Dec;42:96-115 [FREE Full text] [doi: 10.1016/j.cpr.2015.09.001] [Medline: 26407540]
35. Kaskie R, Graziano B, Ferrarelli F. Schizophrenia and sleep disorders: links, risks, and management challenges. Nat Sci Sleep 2017 Sep;Volume 9:227-239. [doi: 10.2147/nss.s121076]
36. Birchwood M, Spencer E, McGovern D. Schizophrenia: early warning signs. Adv Psychiatr Treat 2018 Jan 02;6(2):93-101. [doi: 10.1192/apt.6.2.93]
37. Lindamer LA, McKibbin C, Norman GJ, Jordan L, Harrison K, Abeyesinhe S, et al. Assessment of physical activity in middle-aged and older adults with schizophrenia. Schizophr Res 2008 Sep;104(1-3):294-301 [FREE Full text] [doi: 10.1016/j.schres.2008.04.040] [Medline: 18550338]

Abbreviations

BPRS: Brief Psychiatric Rating Scale
CDS: Calgary Depression Scale
CGI-SCH: Clinical Global Impression-Schizophrenia
DSM-5: Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition
ICC: intraclass correlation coefficient
PANSS: Positive and Negative Syndrome Scale
PRO: patient-reported outcome
PSQI: Pittsburgh Sleep Quality Index
REDCap: Research Electronic Data Capture
UAB: University of Alabama, Birmingham
YPAS: Yale Physical Activity Survey

Edited by J Torous; submitted 10.12.20; peer-reviewed by B Buck, N Meyer; comments to author 21.01.21; revised version received 22.04.21; accepted 10.05.21; published 09.08.21.

Please cite as:
Lahti AC, Wang D, Pei H, Baker S, Narayan VA
Clinical Utility of Wearable Sensors and Patient-Reported Surveys in Patients With Schizophrenia: Noninterventional, Observational Study
JMIR Ment Health 2021;8(8):e26234
URL: https://mental.jmir.org/2021/8/e26234
doi:10.2196/26234
PMID:34383682

Original Paper

Validity of the Aktibipo Self-rating Questionnaire for the Digital Self-assessment of Mood and Relapse Detection in Patients With Bipolar Disorder: Instrument Validation Study

Jiří Anýž[1], PhD; Eduard Bakštein[1], PhD; Andrea Dally[2], PhDr; Marián Kolenič[1], MD, PhD; Jaroslav Hlinka[1], PhD; Tereza Hartmannová[1,2], Mgr; Kateřina Urbanová[1,2], Mgr; Christoph U Correll[3,4,5], MD, PhD; Daniel Novák[6], PhD; Filip Španiel[1], MD, PhD

[1]National Insitute of Mental Health, Klecany, Czech Republic

[2]Mindpax s.r.o, Prague, Czech Republic

[3]Department of Psychiatry, The Zucker Hillside Hospital, Glen Oaks, NY, United States

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

[5]Department of Child and Adolescent Psychiatry, Charité Universitätsmedizin Berlin, Berlin, Germany

[6]Department of Cybernetics, Faculty of Electrical Engineering, Czech Technical University in Prague, Prague, Czech Republic

Corresponding Author:
Jiří Anýž, PhD
National Insitute of Mental Health
Topolová 748
Klecany, 250 67
Czech Republic
Phone: 420 283 088 409
Email: jiri.anyz@nudz.cz

Abstract

Background: Self-reported mood is a valuable clinical data source regarding disease state and course in patients with mood disorders. However, validated, quick, and scalable digital self-report measures that can also detect relapse are still not available for clinical care.

Objective: In this study, we aim to validate the newly developed ASERT (Aktibipo Self-rating) questionnaire—a 10-item, mobile app–based, self-report mood questionnaire consisting of 4 depression, 4 mania, and 2 nonspecific symptom items, each with 5 possible answers. The validation data set is a subset of the ongoing observational longitudinal AKTIBIPO400 study for the long-term monitoring of mood and activity (via actigraphy) in patients with bipolar disorder (BD). Patients with confirmed BD are included and monitored with weekly ASERT questionnaires and monthly clinical scales (Montgomery-Åsberg Depression Rating Scale [MADRS] and Young Mania Rating Scale [YMRS]).

Methods: The content validity of the ASERT questionnaire was assessed using principal component analysis, and the Cronbach α was used to assess the internal consistency of each factor. The convergent validity of the depressive or manic items of the ASERT questionnaire with the MADRS and YMRS, respectively, was assessed using a linear mixed-effects model and linear correlation analyses. In addition, we investigated the capability of the ASERT questionnaire to distinguish relapse (YMRS≥15 and MADRS≥15) from a nonrelapse (interepisode) state (YMRS<15 and MADRS<15) using a logistic mixed-effects model.

Results: A total of 99 patients with BD were included in this study (follow-up: mean 754 days, SD 266) and completed an average of 78.1% (SD 18.3%) of the requested ASERT assessments (completion time for the 10 ASERT questions: median 24.0 seconds) across all patients in this study. The ASERT depression items were highly associated with MADRS total scores (*P*<.001; bootstrap). Similarly, ASERT mania items were highly associated with YMRS total scores (*P*<.001; bootstrap). Furthermore, the logistic mixed-effects regression model for scale-based relapse detection showed high detection accuracy in a repeated holdout validation for both depression (accuracy=85%; sensitivity=69.9%; specificity=88.4%; area under the receiver operating characteristic curve=0.880) and mania (accuracy=87.5%; sensitivity=64.9%; specificity=89.9%; area under the receiver operating characteristic curve=0.844).

Conclusions: The ASERT questionnaire is a quick and acceptable mood monitoring tool that is administered via a smartphone app. The questionnaire has a good capability to detect the worsening of clinical symptoms in a long-term monitoring scenario.

(JMIR Ment Health 2021;8(8):e26348) doi:10.2196/26348

KEYWORDS

bipolar disorder; symptom monitoring; ecological mood assessment; relapse detection; mobile application; mobile phone

Introduction

Background

Bipolar disorder (BD) is a severe mental illness characterized by recurrent depressive, manic, and mixed episodes [1]. The World Health Organization has identified BD as one of the most frequent causes of disability in youth [2]. The existing literature suggests that measurement-based care may be an effective clinical strategy in most psychiatric disorders [3-6], and monitoring of the clinical course of BD is therefore of high clinical and research interest. Several organizations, such as the International Society for Bipolar Disorders, have recommended the implementation of measurement-based care in the treatment of BD [7].

Although behavioral analysis, including actigraphy and smartphone monitoring, is currently a topic of ongoing research [8-11], mood assessment is still the primary source of clinical data for BD status. In clinical practice, both clinician-observed instruments and patient-reported instruments are used to assess the course of BD [9,10,12]. In this paper, we focus on collecting self-reported mood data using a mobile phone app, which provides an opportunity to monitor the mental health status of the patient frequently and at low cost.

Self-reported Measurement Tools

Self-report instruments have been used by clinicians for decades, and there has been increasing interest in mental health technologies for self-report instruments in the past few years. The World Health Organization has reported that mobile technologies have the potential to transform health care in various medical specialties worldwide [13]. Currently, approximately 76% of adults in advanced economies and about 45% of adults in emerging economies own a smartphone [14], and it is assumed that this proportion will continue to increase in the next few years.

The growing interest in digital monitoring platforms for disease monitoring in BD has resulted in an increased demand for quick and focused symptom self-monitoring tools [15]. Self-monitoring tools have been well received by patients [16], and the general validity of smartphone-based monitoring has been demonstrated [17]. Reports from psychiatric clinics also demonstrate that patients using self-report instruments have a better outcome than patients undergoing usual care without the availability of self-report instruments [12].

As mentioned above, patient-reported assessments are used in clinical practice and research. Patient-reported measures are tools used to assess the patient's condition based on their subjective perspective. The main benefit is self-administration, which does not require the presence of clinical staff and can take place in the patient's natural environment. This process increases ecological validity and reduces the cost of data collection. However, these features also increase the requirements for patient adherence, which can, to some extent, be mitigated by smartphone-aided data collection by using notifications, reminders, or interactions with the study staff when there are missing questionnaires.

In BD, as with observer-reported assessments, there are patient-reported measurement tools available that either assess depressive symptoms and manic symptoms together, or only assess one of the mood polarities. In a meta-analysis [12], the authors summarized existing patient-reported measurement tools with high clinical utility scores and reported three tools that assessed only depressive symptoms, four that assessed only manic symptoms, and five that were used to assess both manic and depressive symptoms.

Currently, the most used self-reported measurement instruments that assess both manic and depressive symptoms are the Internal State Scale [18], Multidimensional Assessment of Thymic States [19], Affective Self-Rating Scale [20], National Institute of Mental Health's (NIMH's) Prospective Live Chart Methodology-Self [21], ChronoRecord [22], MoodZoom [23], and openSIMPLE [24,25].

For the purposes of long-term monitoring of patients with BD via a mobile app, we found no existing questionnaires that would meet our requirements, that is, (1) a questionnaire with short completion time with no more than 10 questions and only a limited number of options; (2) a questionnaire for assessing both manic and depressive symptoms on comparable scales; (3) a questionnaire designed for weekly sampling; (4) a questionnaire with high sensitivity to subclinical mood changes during remission periods, allowing early detection of deterioration of the clinical state; and (5) a questionnaire that was suitable and validated for user-friendly smartphone delivery.

This Study and the Aktibipo Self-rating Questionnaire

In this study, we present a novel tool for unassisted, app-based, self-evaluation of mood, the ASERT (Aktibipo Self-rating) questionnaire [26], aimed at the depressive and manic symptoms of patients with BD, following the five principles mentioned earlier.

We aimed to answer the following three principal questions: (1) what is the validity of the questionnaire for measuring depression and mania with respect to the respective standard clinical scales; (2) what is the internal structure of the questionnaire, and what is its consistency; and (3) can ASERT be used for detecting a symptomatic episode (a relapse), as defined by the corresponding clinical scales?

Methods

The ASERT Questionnaire

ASERT is a novel questionnaire for the ecological momentary assessment of mood in BD [26]. The questionnaire contains 10 items that map depressive symptoms (4 items), manic symptoms (4 items), and nonspecific symptoms (two items), with 5 possible response levels for each symptom. The observation period is the past week. The questionnaire is available in Czech, English, and German. The English and German versions were validated using back translation. The questionnaire was administered on a weekly basis through a smartphone app developed by Mindpax [27]. The questions and reply options are presented in Table 1. The presented results used the Czech version in a clinical study.

Table 1. The three available language versions of the ASERT (Aktibipo Self-rating) questionnaire.

Group and number[a]	English version[b]	German version[c]	Czech version[d]
Depressive			
1	I feel sad, downhearted	Ich fühle mich traurig und niedergeschlagen	Cítím se smutně, sklesle
2	I do not enjoy anything, and nothing pleases me	Ich genieße nichts und ich habe an nichts Gefallen	Nic mě nebaví, netěší
3	I have no energy	Ich habe keine Energie	Nemám energii
4	I feel gloomy and pessimistic about the future	Ich bin bedrückt und pessimistisch über die Zukunft	Budoucnost vidím černě, pesimisticky
Manic			
5	I feel unusually great, optimistic	Ich fühle mich ungewohnt großartig und bin ungewöhnlich optimistisch	Cítím se neobvykle skvěle, optimisticky
6	I have excess energy	Ich habe einen Überschuss an Energie	Mám nadmíru energie
7	My thinking is very fast, others cannot keep up with me	Mein Denken ist sehr schnell, andere können mit mir nicht mithalten	Myslí mi to hodně rychle, ostatní mě nestíhají
8	I need to sleep less than usual	Ich brauche weniger Schlaf als sonst	Potřebuji spát méně, než obvykle
Nonspecific			
9	I feel restless, tense	Ich fühle mich ruhelos und angespannt	Cítím neklid, napětí
10	I cannot focus	Ich kann mich nicht konzentrieren	Nemohu se soustředit

[a]Usage of the ASERT (Aktibipo Self-rating) questionnaire is subject to a written agreement with Mindpax s.r.o.

[b]Reply options: 0=I do not agree; 1=more likely I do not agree; 2=I probably agree; 3=I agree; 4=I completely agree.

[c]Möglichkeiten: 0=Trifft nicht zu; 1=Ich stimme eher nicht zu; 2=Teils/Teils; 3=Trifft eher zu; 4=Trifft zu.

[d]Možné odpovědi: 0=nesouhlasím; 1=spíše nesouhlasím; 2=asi souhlasím; 3=souhlasím; 4=naprosto souhlasím.

The AKTIBIPO400 Study

The questionnaire, as well as its validation, is part of the ongoing noninterventional and nonpharmacological AKTIBIPO400 study. The aim of the clinical study is to monitor the physical activity, mood, and clinical state of patients with BD and those of healthy controls using an actigraphy wristband with telemetric data transmission. The participation timeframe was 18 months for patients with BD and 3 months for healthy controls. The study population included men and women between 18 and 60 years of age undergoing standard clinical treatment for BD (International Classification of Diseases-10 diagnosis F31) and remitted at enrollment, meeting the thresholds of the Montgomery-Åsberg Depression Rating Scale (MADRS) [28] sum score ≤9 and the Young Mania Rating Scale (YMRS) [29] sum score ≤12. All study participants in all groups wore a wrist-worn actigraphy device at all times and used the accompanying mobile app (a proprietary system developed by Mindpax) using which they submitted weekly ASERT mood self-reports. Further study details can be found in the study website [30].

The study was approved by the institutional ethical committee of the NIMH (Czech Republic; case number: 101/17) and was conducted in compliance with the Declaration of Helsinki.

The Mindpax Mobile Health System

The study used a mobile health system developed by the Mindpax company, which included a mobile app installed in the participant's smartphone (Android or iOS) and the Mindpax wristband (see Figure 1 for system overview). The wristband version used in the study included a nonremovable battery with a battery life of approximately 12 months. Epoch-level actigraphy data were aggregated and stored by the wristband and regularly (several times a day) transferred via Bluetooth to the mobile app and through secured connection via mobile data or Wi-Fi to the Mindpax servers for evaluation. To avoid bias or confusion, the participants were blinded to their activity and mood history, in addition to visualizations of sleep length and total daily activity. Figure 1 shows a mobile app version dedicated to end users with full feedback. To ensure the privacy of patient data, the system uses secure data connections (secure shell protocol), and the system accounts are anonymized, with citizen names and clinical details available to the clinical study management team and participating psychiatrists only. The Mindpax system is a class I medical device notified by the Czech State Institute for Drug Control under reference number 00904675, and the company undergoes a regular certification process under International Organization for Standardization (ISO) 13485.

Figure 1. Snapshots from the Mindpax mHealth system. (A) Patient view of the mobile app showing activity and sleep in a spiral graph as well as the time course of mood. (B) Clinician view of the system showing various parameters reported by the patient and measured by wristband actigraphy. mHealth: mobile health.

Clinical State Assessment

For the purposes of this study, only the participants in the core group of the AKTIBIPO400 study [30] were included: those with a diagnosis of BD and bipolarity index >50 [31,32], as evaluated by an institutional psychiatrist during the initial visit to the NIMH at enrollment. The patients were assessed monthly using the MADRS and YMRS clinical scales via telephone interviews by trained psychologists. Additional information about clinical episodes, hospitalizations, and work status was collected retrospectively every 6 months from the caregiving psychiatrists.

Clinical scales were used as the primary measure of the clinical state. For relapse detection, we used the definition of a depressive relapse as the sum of MADRS≥15, and for a manic relapse as the sum of YMRS≥15; the nonrelapse state (remission) was defined as the sum of MADRS<15 or the sum of YMRS<15, in accordance with the International Society for Bipolar Disorders task force definition of an episode more severe than subsyndromal depression or mania [7].

Data Preprocessing

The different intervals of the sampling period for the self-reported ASERT questionnaire (weekly) and the clinical scales (monthly) pose a challenge to evaluating the validity of the ASERT questionnaire. Therefore, we created a matched data set, including only ASERT responses no further than 3 days from the nearest MADRS and YMRS assessment event of a given participant, thus containing all ASERT responses in a week-long window around each phone assessment. MADRS and YMRS assessments without such defined matching ASERT responses were excluded from the data set. All matchings were unique, and neither of the two clinical scale assessments was matched to more than one self-report and vice versa. No missing values in individual question responses were possible, as the mobile app allowed only complete questionnaires to be submitted. All clinical scale assessments were completed. Therefore, no imputation of missing values was needed, and completely missing scales and questionnaires were not imputed.

The Internal Structure and Basic Properties of the ASERT Questionnaire

We used the principal component analysis [33] to validate the content, that is, the assumed three-factor structure of 4 depression-related questions, 4 mania-related questions, and 2 nonspecific questions of the ASERT questionnaire. Cronbach α [34] was calculated to provide a consistency estimate for each factor.

We estimated the ASERT response rate of each patient to the platform as the number of filled-in self-reports divided by the expected number of filled-in reports (ie, the number of weeks spent in the study).

A Comparison Between ASERT Responses and Clinical Scales

Linear Mixed-Effects Models

The convergent validity between the ASERT questionnaire and the corresponding clinical scales was assessed using linear mixed-effects (LME) models [35], which consisted of fixed effects that quantified the dependency shared across the group, and random effects that quantified individual effects. For these analyses, data from a single patient were considered dependent; therefore, random effects parameters (random intercepts and slopes) were estimated at the individual patient level.

Correlation

Apart from the LME, we also calculated Pearson correlation coefficients to measure the association between the ASERT depression and mania item responses and the MADRS and YMRS, respectively.

In addition, we also computed correlations between selected individual items from the clinical scales and matching items in the ASERT questionnaire. As interindividual variation may lead to biased results when estimating correlation coefficients, we used a simple 2-step procedure to estimate the correlation coefficient in cases with substantial individual variation. In the first step, we computed the correlation coefficient for each patient separately. In the second step, we averaged the individual correlation coefficients to estimate the group-level correlation coefficient. Because each patient is represented by a different number of observations, due to the ongoing state of the AKTIBIPO400 study, we decided to introduce weighting into the second step of the procedure by assigning a weight that was proportional to the number of observations available for each patient.

Statistical Assessment

To assess the statistical significance of the estimated LME fixed-effects coefficients and group correlation coefficients, we performed a case bootstrap with 10,000 resamples. In each iteration, we sampled the data on both levels of the structure, which means that for each simulation, we took a random sample with replacement on the level of patients, and subsequently we sampled with replacement from the observations in each patient subset.

Detection of Scale-Based Relapses From the Self-reports

After evaluating the validity of the questionnaire, we assessed the performance of the questionnaire in identifying a relapse based on a score of ≥15 on the MADRS or YMRS scale. The MADRS or YMRS scores of <15 were considered a nonrelapse (interepisode or remission) state. Only patients with records of both relapse and nonrelapse states were included in the subset for this task. Similar to the comparison of the ASERT questionnaire and the clinical scales described above, we used a generalized mixed-effects model. In this task, we used a logistic mixed-effects model due to the binary outcome (relapse or nonrelapse).

Two separate models were trained and evaluated for the detection of depressive and manic relapses. To increase the predictive robustness of the relapse detection model, we extended the model inputs by adding the values of the ASERT questionnaire responses one week before the clinical scale assessment (matched in the window 4-10 days before the phone interview), and the ASERT responses 2 weeks before the clinical scale assessment (matched in the window 11-17 days before the phone interview).

As *false alarms* (ie, lack of specificity) significantly affect the usability of a relapse detection model, we used the receiver operating characteristic on the training set to modify the detection threshold of the trained model to achieve the desired specificity of 90%.

Model Validation and Model Selection

To obtain an unbiased estimate of the model relapse detection performance, we used a standard split-validation design in a ratio of 70:30. In this scenario, the data set is divided randomly into two subsets: the training data set (70%) on which the model is trained, and the testing data set (30%) on which the model is evaluated. The data splitting procedure is performed with respect to individual patients, resulting in each patient being represented by the same ratio of his or her data in the training and testing sets.

The model performance criteria were computed for the testing set. We chose relapse detection accuracy, area under the receiver operating characteristic curve (AUROC), sensitivity, and specificity as a set of model performance measurements. We repeated the split-validation procedure 999 times with randomly chosen splits to obtain a model performance distribution rather than a point estimate.

The best-scoring model was selected using a stepwise comparison procedure. The sequence of model structures consisted of a simple model without random effects (a logistic regression model), followed by models with various random effect structures and various predictor variables (see Table S1 in Multimedia Appendix 1 for a more detailed description of the models). In each comparison, the pair consisting of the basic (simpler) model and the more complex model was repeatedly trained and tested on randomly selected training and testing sets. Performance measures were recorded and compared in a pairwise manner. The basic model was rejected as being inferior to the more complex model if the performance measures were better in a substantial number of data set splits (≥90% of the random simulations).

Results

Data Description

The original data set contained 8438 ASERT self-reports from 103 patients and 2186 records of clinical scales (MADRS, YMRS) from 107 patients. For ASERT validation, we used only a subset of the ASERT self-reports for which a concurrent record of the clinical scales was available. As described above, the clinical scale and the self-report were considered concurrent or temporally matching if they were obtained no more than 3 days apart.

After applying these constraints to the data, we obtained 2159 paired records (ie, clinical scales with corresponding ASERT responses) from 99 patients, which were used for the questionnaire validation and for evaluating the internal structure of the questionnaire. The participant characteristics for this data set are summarized in Table 2, which summarizes patients' general demographics, psychometrics, and medications (for a full list of medications refer to Table S2 in Multimedia Appendix 1).

The median completion time of the ASERT, measured by the application and stored in the database, was 24 seconds. The average response rate was 78.1% (SD 18.3%).

Table 2. Demographic and diagnostic characteristics of the validation data set.

Demographics category	All patients (n=99), n (%)	Relapse detection (subset: depression; n=51), n (%)	Relapse detection (subset: mania; n=27), n (%)
General demographics			
Patients[a], n (%)	99 (100)	51 (52)[b]	27 (27)[b]
Age[c] (years), mean (SD)	37.7 (11)	38.0 (10.6)	35.7 (10.5)
Sex[c], n (%)			
Female	60 (61)	33 (65)	21 (78)
Male	39 (40)	18 (35)	6 (22)
Psychometrics			
Days of follow-up[a], mean (SD; range)	754.0 (266.4; 64-1116)	800.5 (229.6; 390-1116)	800.4 (236.5; 390-1113)
Response rate[a], mean (SD; range)	0.78 (0.18; 0.16-1)	0.76 (0.18; 0.32-0.98)	0.79 (0.16; 0.47-0.98)
Bipolarity index[c], mean (SD; range)	68.7 (10.6; 50-90)	69.0 (10.8; 52-90)	69.5 (10.4; 52-90)
MADRS[d], mean (SD; range; N)	4.78 (6.6; 0-48; 2159)	7.07 (7.9; 0-48; 1121)	5.9 (8.0; 0-48; 635)
YMRS[e], mean (SD; range; N)	2.14 (4.1; 0-35; 2159)	2.76 (4.7; 0-35; 1121)	4.45 (6.1; 0-35; 635)
Medication, n (%)			
Antidepressant	43 (43)	30 (59)	13 (48)
Antipsychotic	76 (77)	38 (75)	20 (74)
Mood stabilizer and anticonvulsant	81 (89)	46 (90)	23 (85)

[a]The items referring to the end of the study (number of patients, days of follow-up, and response rate).
[b]n=99.
[c]The items obtained at the beginning of the study (age, sex, and bipolarity index). The items without any markup were collected throughout the study.
[d]MADRS: Montgomery-Åsberg Depression Rating Scale.
[e]YMRS: Young Mania Rating Scale.

Content Validity of the ASERT Questionnaire

The principal component analysis of the ASERT responses showed two dominant components—pc_1 with 50.3% explained variability and pc_2 with 21% explained variability (Tables 3 and 4). The first principal component relates to the depression-related questions (pc_1 high values of loadings for questions 1-4) together with nonspecific questions (pc_1 high loadings for questions 9 and 10). The second principal component summarizes the mania-related questions (pc_2 high values of the loadings for questions 5-8) and nonspecific questions (pc_2 high values of loadings for questions 9 and 10), but their loading coefficients were smaller than those of the main mania-related questions. Each of the remaining principal components explained <7.4% of the variability.

The Cronbach α values showed high consistency for the sum of the depression-oriented questions Q1-Q4 (Cronbach α=.92; if nonspecific, Q9 and Q10 were also added, the value remained at Cronbach α=.92), and for the mania-oriented questions Q5-Q8 (Cronbach α=.85; if Q9 and Q10 were added, the consistency dropped to Cronbach α=.72). The two nonspecific questions Q9 and Q10 together also showed high consistency (Cronbach α=.85).

The results indicated that (1) the ASERT questionnaire demonstrates the intended two-factor structure with highly consistent components and (2) the two nonspecific questions are more correlated with the depressive cluster. Therefore, we concluded that it was reasonable to focus only on the summary characteristics (the sum of the questions about depression, the sum of the questions about mania, and the sum of the nonspecific questions), rather than the individual items, in our further analyses.

Table 3. Loadings of the principal component analysis of the ASERT[a] questionnaire self-reports.

ASERT questionnaire question number	PC^b_1	PC_2	PC_3	PC_4	PC_5	PC_6	PC_7	PC_8	PC_9	PC_{10}
Q1	0.433	—[c]	0.282	–0.281	0.111	–0.423	0.229	–0.178	–0.580	–0.190
Q2	0.425	—	0.310	–0.177	—	–0.111	0.147	–0.266	0.752	0.109
Q3	0.427	–0.135	0.207	0.805	0.229	0.136	—	0.108	–0.101	0.113
Q4	0.389	—	0.213	–0.385	–0.221	0.420	–0.435	0.481	—	—
Q5	—	0.506	0.288	0.236	–0.388	–0.101	–0.343	–0.261	—	–0.506
Q6	—	0.493	0.239	—	–0.163	—	—	–0.120	–0.181	0.787
Q7	—	0.415	0.145	—	—	—	0.614	0.608	0.128	–0.201
Q8	—	0.428	—	–0.188	0.805	0.255	–0.179	–0.149	—	–0.107
Q9	0.371	0.261	–0.586	—	—	–0.533	–0.292	0.238	0.124	—
Q10	0.397	0.204	–0.479	—	–0.242	0.502	0.347	–0.352	–0.110	—

[a]ASERT: Aktibipo Self-rating.

[b]*PC*: principal component.

[c]Values with an absolute value lower than 0.1 are not shown.

Table 4. Summary statistics of the principal component analysis of the ASERT (Aktibipo Self-rating) questionnaire self-reports.

Summary statistic	PC^a_1	PC_2	PC_3	PC_4	PC_5	PC_6	PC_7	PC_8	PC_9	PC_{10}
SD	1.378	0.890	0.527	0.416	0.397	0.341	0.327	0.322	0.271	0.267
Proportion of variance	0.503	0.210	0.074	0.046	0.042	0.031	0.028	0.028	0.020	0.019
Cumulative proportion of variance	0.503	0.714	0.787	0.833	0.875	0.906	0.934	0.962	0.981	1

[a]*PC*: principal component.

Convergent Validity of the ASERT Depression and Mania Items With the MADRS and YMRS, Respectively

Following the results of the previous section, we modeled the relationships between the clinical scales and ASERT self-reports using two LME models with random slopes and random intercepts for individual patients.

The first model showed a significant linear relationship between the sum of the depression-related and nonspecific questions of the ASERT questionnaire (DEPNSP) and the sum of the MADRS scale, with the slope β_{DEPNSP}=.87 (*P*<.001) and the intercept β_0=.71 (*P*<.001). The fixed-effects coefficient value means that a unit change in the ASERT depressive and nonspecific subscore corresponds to a 0.87 increase in the MADRS sum score on the study population level. The second model linked mania-related questions to the sum of the YMRS scale and showed a significant association, with a slope of β_{MAN}=.73 (*P*<.001) and intercept β_0=1.05 (*P*<.001). The fixed-effects coefficient value means that a unit change in the ASERT manic subscore corresponds to a 0.73 increase in the YMRS sum score on the study population level.

Alternatively, using Pearson correlation coefficient, the depression-related subscores of ASERT showed a significant correlation with the sum of MADRS, and the mania-related subscores also showed a significant correlation with the sum of the YMRS scores. Specifically, the weighted group correlation coefficient was ρ_{wgr}=0.53 (*P*<.001) for the relationship between the MADRS scale and the depression-related and nonspecific components of the ASERT questionnaire and the weighted group correlation coefficient was ρ_{wgr}=0.32 (*P*<.001) for the YMRS and the mania-related subset of ASERT. For the relationship between the individual ASERT item responses and matching items from the MADRS and YMRS, see Table 5.

Table 5. The correlation between Aktibipo Self-rating questionnaire subscores and the total scores of the Young Mania Rating Scale and Montgomery-Åsberg Depression Rating Scale clinical scales and the correlation between individual items of the Aktibipo Self-rating questionnaire and corresponding items from the Montgomery-Åsberg Depression Rating Scale and Young Mania Rating Scale clinical scales.

Subset of variables	Group correlation coefficient-weighted[a]	*P* value
ASERT[b] questionnaire depression and mania subscores versus sum of MADRS[c] or YMRS[d]		
MADRS: Sum of questions~ASERT: sum of questions about depression	0.51	<.001
MADRS: Sum of questions~ASERT: sum of questions about depression and nonspecific questions	0.53	<.001
YMRS: Sum of questions~ASERT: sum of questions about mania	0.32	<.001
YMRS: Sum of questions~ASERT: sum of questions about mania and nonspecific questions	0.25	<.001
ASERT questionnaire individual items versus MADRS or YMRS selected items		
MADRS: 2. sadness (subject)~ASERT: 1. sadness	0.49	<.001
MADRS: 9. pessimism~ASERT: 4. future	0.41	<.001
YMRS: 2. energy (+)~ASERT: 6. energy (+)	0.40	<.001
YMRS: 7. speech and thinking disorders~ASERT: 7. acceleration	0.46	<.001
YMRS: 4. sleep~ASERT: 8. sleep	0.31	<.001
MADRS: 3. internal tension~ASERT: 9. unrest	0.31	<.001
MADRS: 6. disturbance of concentration~ASERT: 10. concentration	0.46	<.001

[a]Weighted group-level correlation coefficients (ie, taking into consideration the different sample sizes for different patients, the weights are proportional to the number of observations for each patient) for sums of the clinical scales and of the ASERT (Aktibipo Self-rating) questionnaires and for a more detailed view of the relationship between the individual items of the clinical scales and their counterparts in the ASERT questionnaire.

[b]ASERT: Aktibipo Self-rating.

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

[d]YMRS: Young Mania Rating Scale.

Relapse Detection Data

Applying the additional constraints on the matched data for relapse detection resulted in the addition of 2069 paired self-reports 1 week before the clinical scale assessments and of 2046 self-reports 2 weeks before the clinical scale assessments.

For depression, 51 patients had data on both relapse and nonrelapse with 1121 paired observations for the coinciding MADRS scale and self-report, 1054 observations from the previous week, and 1072 observations from 2 weeks before. From the 1121 observations, 195 fulfilled the criteria for relapse (MADRS≥15).

For mania, 27 patients had data on both relapse and nonrelapse with 635 paired observations for the coinciding YMRS scales and self-reports, 591 observations from the previous week, and 604 observations from 2 weeks before. Of the 635 clinical scales, 63 fulfilled the criteria for relapse (YMRS≥15).

ASERT-Based Relapse Detection

For depression relapse (MADRS≥15), the best-performing model from the model selection included the sum of the depression-related questions, together with the nonspecific questions from the current week and from the previous week in the fixed-effects part, and random intercepts for each patient in the random effect part. This was the last model showing any improvement in the succession of hierarchically nested models in the model validation and model selection section and is listed in Table S1 (Multimedia Appendix 1), achieving an accuracy of 88.3%, an AUROC of 0.932, sensitivity of 79.5%, and specificity of 90.2% on the training data set and an accuracy of 85%, AUROC of 0.880, sensitivity of 69.9%, and specificity of 88.4% on the testing data set (Table 6). Inclusion of the ASERT depression score 2 weeks before the scale did not significantly improve this model.

The best-performing model for mania relapse (YMRS≥15) used the sum of mania-related questions from the current week in the fixed-effects part and random intercepts in the random effect part, achieving an accuracy of 88.6%, AUROC of 0.901, sensitivity of 71.3%, and specificity of 90.6% on the training set and an accuracy of 87.5%, AUROC of 0.844, sensitivity of 64.9%, and specificity of 89.9% on the testing set (Table 7). Inclusion of the ASERT mania subscore from the previous week or 2 weeks before the scale did not significantly improve this model.

Table 6. Performance of the ASERT (Aktibipo Self-rating) questionnaire in detecting depressive relapse[a].

Set	Accuracy	AUROC[b]	Sensitivity	Specificity
Training	0.883	0.932	0.795	0.902
Testing	0.850	0.880	0.699	0.884

[a]For depressive relapse, the results of the best-scoring logistic mixed-effects model in the detection of depression relapse from the sum of the depression-related questions and the nonspecific questions from the current week and from the previous week, as follows: MADRS≥15 ~ $\text{ASERT}_{\text{DEPNSP,current week}}$ + $\text{ASERT}_{\text{DEPNSP,previous week}}$ + (1 | $\text{ID}_{\text{patient}}$). MADRS: Montgomery-Åsberg Depression Rating Scale; DEP: depression; NSP: nonspecific question; (1 | $\text{ID}_{\text{patient}}$): random intercept. Note that the detection thresholds for all models were adjusted to a specificity of 90% by using the receiver operating characteristic curve of the training set.

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

Table 7. Performance of the ASERT (Aktibipo Self-rating) in detecting manic relapse[a].

Set	Accuracy	AUROC[b]	Sensitivity	Specificity
Training	0.886	0.901	0.713	0.906
Testing	0.875	0.844	0.649	0.899

[a]For manic relapse, results of the best-scoring logistic mixed-effects model in the detection of mania relapse used the sum of the mania-related questions from the current week, as follows: YMRS≥15 ~ $\text{ASERT}_{\text{MAN,current week}}$ + (1 | $\text{ID}_{\text{patient}}$). YMRS: Young Mania Rating Scale; MAN: mania; (1 | $\text{ID}_{\text{patient}}$): random intercept. Note that the detection thresholds for all models were adjusted to a specificity of 90% by using the receiver operating characteristic curve of the training set.

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

Discussion

Principal Findings

The ASERT questionnaire for ecological momentary mood assessment was designed to provide a quick and scalable way to report the self-perceived mood of patients affected by BD via a smartphone app, including mania-related symptoms and depression-related symptoms. We evaluated the content and convergent validity, internal structure, and consistency of the questionnaire, using an extensive data set consisting of 99 patients with a mean follow-up duration of almost 2 years, with weekly ASERT assessments and monthly YMRS and MADRS phone interviews.

The analysis of the internal structure using principal component analysis confirmed the assumed structure of 4 depression-oriented questions (Q1-Q4), 4 mania-oriented questions (Q5-Q8), and 2 nonspecific questions (Q9 and Q10) with high internal consistency (Cronbach α≥.85). Although higher scores on the nonspecific questions (Q9 and Q10) were associated with higher depression scores and higher manic scores, nonspecific questions were more consistently associated with depressive scores.

Although the two nonspecific questions (Q9: "I feel restless, tense" and Q10: "I cannot focus") did not increase consistency when summed up with the depressive questions, they increased the accuracy of the mixed-effects model in comparison with the MADRS and increased accuracy in the detection of depressive relapses. This improved performance is likely due to the matching items being present in the MADRS scale—the items number 3 (inner tension) and 6 (concentration difficulties)—which have previously been shown to form a separate cognitive-anxiety cluster [36] or were used as a proxy for anxiety [37]. However, both variants, including or excluding nonspecific questions, are reasonable when estimating the depressive state of the patient and may be used depending on the intended range of symptoms studied.

This symptomatic dimension, mirrored by both nonspecific ASERT questions and MADRS items 3 and 6, might be of specific heuristic value in further parsing the neurobiological heterogeneity of mood disorders. In a large sample of subjects with affective disorders, the presence of restlessness and distractibility identified subgroups with an increasing likelihood of BD diagnosis [38]. These findings provide an important empirical confirmation that these symptoms may play a role in discriminating between unipolar disorder and BD. This finding is of particular relevance as patients with BD with comorbid anxiety and rapid mood switches might represent a genetically distinct subtype of BD [39]. Primarily, continuous monitoring of anxiety might be of clinical significance due to the association of comorbid anxiety with greater severity, recurrence, and overall impairment in BD [40,41]. On the basis of these arguments, we decided to keep the nonspecific questions as part of the ASERT.

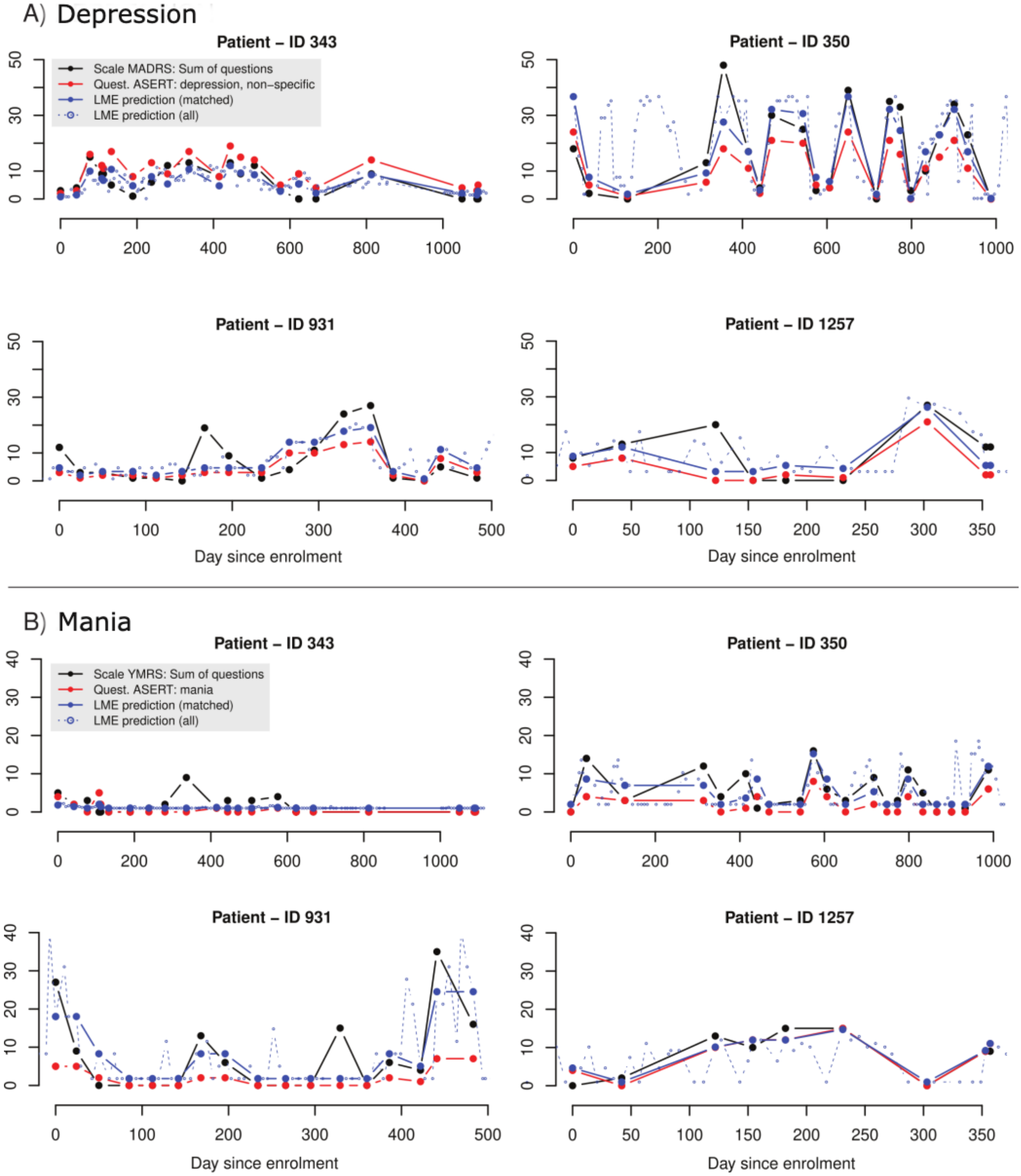
The main advantages of the ASERT questionnaire are its simple structure and coverage of the main symptoms of BD. The simple structure of ASERT allows it to be administered via a mobile app and to be evaluated quickly (the median filling time was 24.0 seconds), which should contribute positively to adherence to the self-report by the patients; in our case of a long-term follow-up, the mean response rate was 78.1%, which is comparable with previous long-term self-assessment studies [23]. This brevity and efficiency comes at the cost of a decreased level of detail in the description of the mood than can be provided by more extensive questionnaires, such as the Internal State Scale [18] with 17 items, ranging from 0 to 100; the Multidimensional Assessment of Thymic States [19] with 20 items, ranging from 0 to 10; or the Affective Self-Rating Scale

[20] with 18 items, ranging from 0 to 4. However, ASERT may arguably provide a higher level of detail than other less granular scales, aimed at both depression symptoms and manic symptoms, such as the NIMH Prospective Live Chart Methodology-Self [21], which has only two items, ranging from –4 to +4, and the ChronoRecord [22], which has six questions, ranging from 0 to 100. However, both of these scales are aimed at a shorter time frame of the past 24 hours, rather than the past week, as in the case of ASERT. This is also the case for the openSIMPLE platform for psychoeducation [24,25], where the patient rates their state in 5 visual analog scales. The daily assessments were compounded using more comprehensive weekly questionnaires.

A comparison between the ASERT subscores and the relevant clinical scales, using an LME model, showed a highly significant association both for the sum of the depression-oriented questions (Q1-Q4) versus the sum of the coincident MADRS scale ($P<.001$) and also for the sum of the mania-oriented questions (Q5-Q8) versus the sum of the coincident YMRS scale ($P<.001$). In addition, many of the ASERT items can be clearly and significantly mapped to the corresponding items from the MADRS or YMRS scales, which increases the interpretability of the collected data.

Self-reports measure a different perspective on a patient's current symptoms than that offered by the clinician interview-based scales. Self-reports are subjective and may be biased by impaired insight during symptomatic periods, especially during mania [42]. However, according to a recent review, several existing self-assessment tools for mania showed high clinical validity [43]. In addition, ASERT contains fewer items than MADRS and YMRS. These two reasons are likely responsible for the imperfect, weighted Pearson correlation coefficients, showing a ρ_{wgr}=0.53 ($P<.001$) for the ASERT depressive and nonspecific subscores compared with the MADRS total score and ρ_{wgr}=0.32 ($P<.001$) for the manic subscore compared with the YMRS total score. This result indicates that an approximately 30-minute clinical interview cannot be replaced by a 30-second self-report. However, there is a clear difference in the burden on the patient as well as in the collection cost. This difference makes it possible to monitor the course of a patient's symptoms frequently, without the need for clinician involvement. Figure 2 demonstrates the high overlap between the output of the ASERT-based mixed-effects model and the MADRS or YMRS, as well as its ability to capture additional mood changes between the MADRS and YMRS evaluations.

Figure 2. Time courses of clinical scales (MADRS and YMRS) and ASERT questionnaire subscores (depressive and nonspecific, manic) of 4 selected patients, demonstrating the ability of the patients to report their mood. The observations from the presented data set are shown in full lines. The dashed lines show additional ASERT questionnaire responses, unmatched to a near MADRS or YMRS for (A) depression and (B) mania. Full lines: black, MADRS or YMRS; red, ASERT depressive and nonspecific subscore or manic subscore; blue, model estimate of MADRS or YMRS; dashed line, model estimate on nonmatched assert questionnaires. ASERT: Aktibipo Self-rating; LME: linear mixed-effects model; MADRS: Montgomery-Åsberg Depression Rating Scale; YMRS: Young Mania Rating Scale.



ASERT-Based Relapse Detection

In addition to the linear relationship between the ASERT self-report and clinical rating scales, we evaluated the ability of ASERT to identify a relapse defined by MADRS≥15 for depression and YMRS≥15 for mania. The dichotomized MADRS and YMRS scores classified into remission and relapse were modeled using a logistic mixed-effects model. In contrast to the LME models used in the statistical comparison, we used more explanatory variables, including ASERT subscores from the previous weeks. The best-scoring model was selected using a repeated randomized split-validation procedure. The results for the separate models for depression and mania showed high accuracy in relapse detection. For depression, the test set accuracy for a model using the ASERT response from the same and the previous weeks was 85% (sensitivity 69.9%; specificity 88.3%; AUROC 0.880), and for mania, the accuracy of the best model using same-week data was 87.5% (sensitivity 64.9%; specificity 89.9%; AUROC 0.844). The high AUROC values show the good detection capabilities of the model even for highly imbalanced data, which is the case for the detection of relatively sparse relapses. In our data set, only 17.39% (195/1121) of events represented relapses for depression and 9.9% (63/635) represented relapses for mania. Owing to this class imbalance in the data set, it would be possible to obtain comparable accuracy values by classifying all the events as remission events (approximately 0.826 for depression and 0.900 for mania). However, this is only possible with a sensitivity of 0% and specificity of 100%. The relatively high values for sensitivity, together with the high AUROC values for both depression and mania relapse, indicate that the model correctly detects relapses.

For practical and demonstration reasons, we decided to alter the threshold probability for the model estimates to maintain 90% specificity, leading to one false alarm in 10 events. This threshold can be set to any arbitrary value based on the trade-off between false alarms and overlooked relapses.

Another consequence of the relatively rare occurrence of relapses in patients with BD is that very long observation periods are necessary to capture both the symptomatic and nonsymptomatic periods in outpatients with BD. Long-term studies investigating the relapse detection capability of self-assessment and other digital tools are therefore relatively scarce. In a cross-sectional setting (a single assessment per subject), Adler et al [20] achieved results comparable with this study: the depression or hypomania detection ability of the Affective Self-Rating Scale against a threshold reference of combined MADRS-based depression and Hypomania Interview Guide-Clinical–based hypomania showed a sensitivity of 0.90 and specificity of 0.71. It has also been shown that identification of a depressive subgroup among nonclinical samples of older adults was possible by a machine learning model using a combination of ecological momentary assessment records and actigraphy features at very high accuracy (AUROC 0.96) [11]. A similar model based on a combination of sleep and activity-related features was able to predict mood episodes in a short window of 3 days with high accuracy (AUROC between 0.79 and 0.93 in BD) [10].

Regarding comparable research in other severe psychiatric conditions, previous studies on patients with schizophrenia have identified questions targeted at basic psychosis symptoms as an efficient strategy to identify emerging relapses [44]. This strategy has also been successfully implemented by our group for schizophrenia in the Information Technology Aided Relapse Prevention Programme in Schizophrenia program [45,46]. Targeted questionnaires may thus complement passive sensing approaches and digital phenotyping in relapse detection in the future [47] and aid targeted care for patients in home environments.

In this study, the relapse prediction model was based only on the current and past week's mood assessments. To this end, we chose a conservative approach to track the relapse risk. Alternatively, even more sophisticated models could be used to enhance relapse prediction, for example, through data transformations and advanced statistical techniques, including estimating spectral density, periodicity [48], or using various nonlinear models [49].

The data could also be combined with general relapse risk predictors, including medication, bipolarity index score, and patient demographics. As these and similar factors have been shown to predict the risk of future relapse, it is likely that their inclusion in the model should improve its performance. These approaches hold promise for future improvement of the existing models and the design of new and more efficient relapse prevention algorithms in BD.

The results presented earlier demonstrate that the relapse detection capability of ASERT is promisingly high. However, to obtain an estimate of the detection performance in a clinical setting, an additional prospective study is necessary.

Limitations

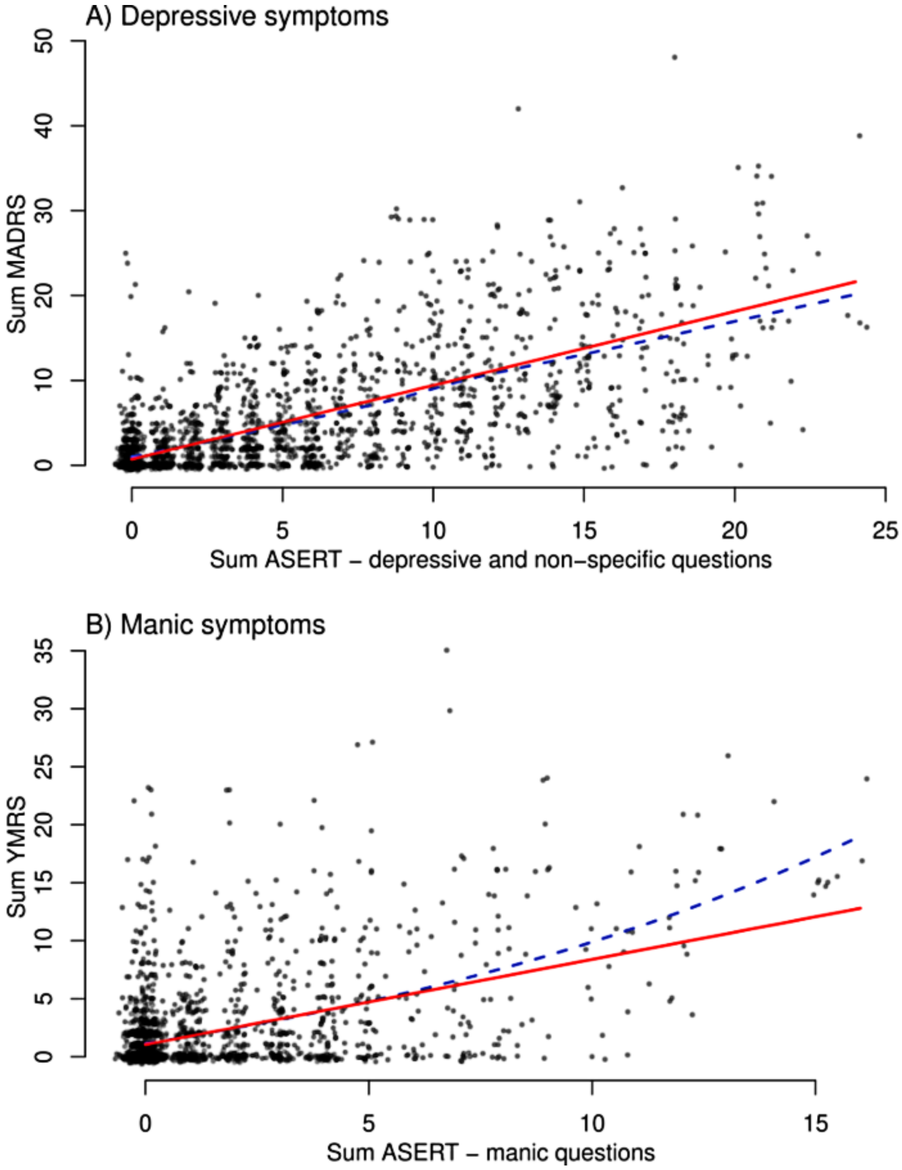
Our study has several limitations. First, ASERT is designed to monitor subtle mood changes during remission periods and to react quickly to any worsening. Therefore, the range may not be sufficient to distinguish between different levels of more severe symptomatic periods. However, the results do not suggest saturation for the combination of depressive and nonspecific questions and show moderate saturation for mania (dashed line in Figure 3). The saturation in the ASERT questionnaire for mania manifests as nonlinearity of the relationship to the clinical scales and may have introduced a negative bias into the validation models presented earlier. The saturation thus renders our modeling results conservatively, which promises further possible improvement by nonlinear models.

Second, the MADRS and YMRS interviews with the psychologist may have affected the patient's reflection on their state, potentially improving ASERT responses. This is not a problem when the ASERT questionnaires preceded the interview at the clinical scale. However, our analysis included both ASERT self-reports that preceded the interview and followed the clinical interview. To assess the effect of the sequence (ASERT before and after the scaling interview), we performed separate post hoc sensitivity analyses for these two cases, and the results did not differ substantially from the main analysis on all the data (Tables S3 and S4 in Multimedia Appendix 1).

Third, the definition of relapse was based only on the threshold of the total score of the MADRS or YMRS [7] and other events and factors indicating a relapse (hospitalizations, functional disability, aggression, or suicidal ideations) were not considered. The reason for this is that this information was collected retrospectively from the psychiatrists of the patients in the study, and that this information was not reliably complete for all patients at the time of the analysis, as the AKTIBIPO400 study is still ongoing. This omission likely introduces a negative bias, reducing the specificity (ie, increasing the false alarm rate) by not marking some of the weeks with low values of the scales as a relapse. The effect on sensitivity cannot be easily deduced. As the scaling is relatively frequent and regular, and the hospitalizations and other events are in good agreement with the partial data set that we already have, we consider that this issue should not represent a major source of bias for the results.

Fourth, an important issue is linked to the validation procedure: by design, data from the same patients are present in the training and testing data sets. This situation is due to the necessity to estimate the random effects of the mixed-effects model using each patient's data, as well as the need for both remission and relapse data in both data sets for a proper evaluation. Although the MADRS and YMRS data are about a month apart, this fact may represent a positive bias to the results, and the only solution is to validate the relapse detection capabilities of the model on an independent prospective data set. In real-world deployment, the proposed approach corresponds to training a model that includes pilot data from each patient (eg, the first several months of follow-up) and evaluating the model on future data.

Figure 3. Correlation between the clinical scales and the corresponding ASERT questionnaire subscores for (A) the sum of questions about depression and nonspecific questions subscore compared with the sum of MADRS scores and (B) the questions about mania subscore compared with the sum of YMRS scores. For this visualization only, the plotted values include random jitter in both the ASERT questionnaire and MADRS or YMRS, and the points are made partially transparent. The solid red line represents the fixed-effects part of the relevant mixed-effects model, and the blue dashed line represents a smooth trend (locally estimated scatterplot smoothing). ASERT: Aktibipo Self-rating; MADRS: Montgomery-Åsberg Depression Rating Scale; YMRS: Young Mania Rating Scale.



Fifth, the course of BD may consist of long (depressive) episodes. In such cases, many consecutive self-reports and clinical scales correspond to one event. These repeated measurements may have inflated the performance measures in relapse detection. From the point of view of statistical modeling, it is important to examine whether the models exploit the smoothness of the data relying only on the data history and self-similarity or whether the model properly treats the information in the data. To explore these properties of the data in relation to the relapse detection problem, we trained the relapse detection models on a subset of new relapses (relapses preceded by nonrelapse clinical scales). The results shown in Table S5 of Multimedia Appendix 1 indicate that the relapse detection models are not dependent on the data self-similarity and are able to detect new relapses with comparable performance to the detection of all relapses.

Finally, although the longitudinal study design, frequent self-report, and clinical assessments provided a rich data set for the analyses, the relapse detection models were based on a relatively small subset of patients with depressive or manic relapse.

Remote Mood Monitoring

Despite the rapid trend in mobile and digital health worldwide, concerns have been raised about the effect of smartphone-based monitoring systems on the risk of relapse of affective episodes. An exploratory analysis within the MONARCA I randomized clinical trial of patients with BD suggested that smartphone-based monitoring may reduce the risk of relapse of mania but that it can possibly increase the risk of a relapse of depression [50]. However, a recent study by Faurholt-Jepsen et al [51] on the effect of a smartphone-based monitoring system on illness activity in BD (MONARCA II) showed no significant difference between the intervention and control groups in terms of the risk of relapse. In contrast, there were differences in perceived stress and quality of life in favor of the intervention group. Some efficacy studies are yet to be evaluated [16,52]. Although a validation in a randomized controlled trial (RCT) provides a high level of evidence, few mobile solutions have been validated and the overall methodological quality of the RCTs in this field is rather low [53]. Moreover, existing RCTs of monitoring systems for BD with a clinical feedback loop (ie, providing the automatic system's output to the patient or physician) have shown mixed or unfavorable results [50,51].

There is a growing concern worldwide about privacy and user rights in the health informatics domain [54], which poses a challenge for legislators adapting to fast-developing technologies. This also poses a challenge for the patients themselves, understanding their own rights and data security policies, as well as understanding which mobile solutions in a myriad of options are of high quality and clinically validated [55]. Despite these relevant challenges in the regulatory, implementation, and educational domains, mobile apps represent a great opportunity to improve clinical practice by providing timely and widely available digital solutions.

These findings suggest that although the use of smartphone-based patient monitoring for direct decision making about further treatment may be a long and difficult journey, the feedback provided to the patients, and also to their psychiatrist, may improve the patient's experience and quality of life. Therefore, smartphone-based monitoring systems in patients with BD could provide support for both patients and health care providers. Continuous monitoring of the patient's condition, together with support provided by the clinician with their detailed insight into the patient's condition could lead to increased treatment adherence in patients with BD [51] and a fine-grained and measurement-based adaptation of treatment choices by treating clinicians [19].

Conclusions

This study presented a novel questionnaire for remote mood monitoring of patients with BD via ASERT through a mobile app. The internal structure and external validity of the ASERT were verified, and the ASERT-based model achieved high accuracy in relapse detection when used as an input into a relapse detection system. Therefore, the questionnaire provides a strong tool for remote mood monitoring in BD, calling for the preparation of a well-designed feedback system that would further extend its utility in clinical care. The simplicity, brevity, scalability, and clinical validity of ASERT make it an appropriate questionnaire for use in observational studies and, after further evaluation and validation, potentially also in routine clinical practice.

Acknowledgments

The authors thank all patients and clinical staff for their participation in the AKTIBIPO400 study. The authors are grateful to Pavel Vostatek for data preparation and Jakub Schneider for the critical review of the ASERT structure. The ASERT questionnaire and its assessment of mood development and relapse detection are (or will be) part of the Mindpax system and the related intellectual property is owned by the company. The study was sponsored by Mindpax (Mindpax s.r.o., Prague, Czech Republic).

Conflicts of Interest

JA, EB, MK, DN, and JH worked for Mindpax at different times during the study preparation and evaluation. FS, CUC, and DN are Mindpax advisory board members. CUC has been a consultant and advisor to or received honoraria from Acadia, Alkermes, Allergan, Angelini, Axsome, Gedeon Richter, Gerson Lehrman Group, Indivior, IntraCellular Therapies, Janssen (Johnson & Johnson), Karuna, LB Pharmaceuticals Inc, Lundbeck, MedAvante-ProPhase, MedInCell, Medscape, Merck, Mylan, Neurocrine, Noven, Otsuka, Pfizer, Recordati, Rovi, Servier, Sumitomo Dainippon, Sunovion, Supernus, Takeda, and Teva. CUC provided expert testimony to Janssen and Otsuka; served on a data safety monitoring board for Lundbeck, Rovi, Supernus, and Teva; received grant support from Janssen and Takeda, and is also a stock option holder for LB Pharmaceuticals Inc.

Multimedia Appendix 1

Supplementary information to the main text, including supplementary tables for the model selection procedure, the full table of medications used by the patients during the AKTIBIPO400 study, and results of post hoc analyses of the effect of the order of clinical and self-assessment on the main results.

[DOCX File , 77 KB - mental_v8i8e26348_app1.docx]

References

1. Vieta E, Berk M, Schulze TG, Carvalho AF, Suppes T, Calabrese JR, et al. Bipolar disorders. Nat Rev Dis Primers 2018 Dec 8;4:18008. [doi: 10.1038/nrdp.2018.8] [Medline: 29516993]
2. The Global Burden of Disease : 2004 Update. Geneva, Switzerland: World Health Organization; 2008.
3. Schouby Bock M, Nørgaard Van Achter O, Dines D, Simonsen Speed M, Correll CU, Mors O, et al. Clinical validation of the self-reported Glasgow Antipsychotic Side-effect Scale using the clinician-rated UKU side-effect scale as gold standard reference. J Psychopharmacol 2020 Aug;34(8):820-828. [doi: 10.1177/0269881120916122] [Medline: 32500804]
4. Correll CU, Kishimoto T, Nielsen J, Kane JM. Quantifying clinical relevance in the treatment of schizophrenia. Clin Ther 2011 Dec;33(12):B16-B39 [FREE Full text] [doi: 10.1016/j.clinthera.2011.11.016] [Medline: 22177377]
5. Guo T, Xiang Y, Xiao L, Hu C, Chiu HF, Ungvari GS, et al. Measurement-based care versus standard care for major depression: a randomized controlled trial with blind raters. Am J Psychiatry 2015 Oct;172(10):1004-1013. [doi: 10.1176/appi.ajp.2015.14050652] [Medline: 26315978]
6. Østergaard SD, Opler MG, Correll CU. Bridging the measurement gap between research and clinical care in schizophrenia: positive and negative syndrome scale-6 (panss-6) and other assessments based on the simplified negative and positive symptoms interview (SNAPSI). Innov Clin Neurosci 2017 Dec 1;14(11-12):68-72 [FREE Full text] [Medline: 29410939]
7. Tohen M, Frank E, Bowden CL, Colom F, Ghaemi SN, Yatham LN, et al. The International Society for Bipolar Disorders (ISBD) Task Force report on the nomenclature of course and outcome in bipolar disorders. Bipolar Disord 2009 Aug;11(5):453-473. [doi: 10.1111/j.1399-5618.2009.00726.x] [Medline: 19624385]
8. De Crescenzo F, Economou A, Sharpley AL, Gormez A, Quested DJ. Actigraphic features of bipolar disorder: a systematic review and meta-analysis. Sleep Med Rev 2017 Jun;33:58-69. [doi: 10.1016/j.smrv.2016.05.003] [Medline: 28185811]
9. Schneider J, Bakštein E, Kolenič M, Vostatek P, Correll CU, Novák D, et al. Motor activity patterns can distinguish between interepisode bipolar disorder patients and healthy controls. CNS Spectr 2020 Sep 4:1-11. [doi: 10.1017/S1092852920001777] [Medline: 32883376]
10. Cho C, Lee T, Kim M, In HP, Kim L, Lee H. Mood prediction of patients with mood disorders by machine learning using passive digital phenotypes based on the circadian rhythm: prospective observational cohort study. J Med Internet Res 2019 Apr 17;21(4):e11029 [FREE Full text] [doi: 10.2196/11029] [Medline: 30994461]
11. Kim H, Lee S, Lee S, Hong S, Kang H, Kim N. Depression prediction by using ecological momentary assessment, actiwatch data, and machine learning: observational study on older adults living alone. JMIR Mhealth Uhealth 2019 Oct 16;7(10):e14149 [FREE Full text] [doi: 10.2196/14149] [Medline: 31621642]
12. Cerimele JM, Goldberg SB, Miller CJ, Gabrielson SW, Fortney JC. Systematic review of symptom assessment measures for use in measurement-based care of bipolar disorders. Psychiatr Serv 2019 May 1;70(5):396-408. [doi: 10.1176/appi.ps.201800383] [Medline: 30717645]
13. mHealth: new horizons for health through mobile technologies: second global survey on eHealth. Geneva, Switzerland: World Health Organization; 2011.
14. Taylor K, Silver L. Smartphone Ownership Is Growing Rapidly Around the World, but Not Always Equally. Pew Research Center. 2019. URL: https://www.pewresearch.org/global/2019/02/05/smartphone-ownership-is-growing-rapidly-around-the-world-but-not-always-equally/ [accessed 2020-10-20]
15. Rajagopalan A, Shah P, Zhang MW, Ho RC. Digital platforms in the assessment and monitoring of patients with bipolar disorder. Brain Sci 2017 Nov 12;7(11):150 [FREE Full text] [doi: 10.3390/brainsci7110150] [Medline: 29137156]
16. Saunders KE, Bilderbeck AC, Panchal P, Atkinson LZ, Geddes JR, Goodwin GM. Experiences of remote mood and activity monitoring in bipolar disorder: a qualitative study. Eur Psychiatry 2017 Mar;41:115-121. [doi: 10.1016/j.eurpsy.2016.11.005] [Medline: 28135594]
17. Beiwinkel T, Kindermann S, Maier A, Kerl C, Moock J, Barbian G, et al. Using smartphones to monitor bipolar disorder symptoms: a pilot study. JMIR Ment Health 2016;3(1):e2 [FREE Full text] [doi: 10.2196/mental.4560] [Medline: 26740354]
18. Huang C, Yang Y, Chen M, Lee I, Yeh T, Yang M. Patient- and family-rated scale for bipolar disorder symptoms: Internal State Scale. Kaohsiung J Med Sci 2003 Apr;19(4):170-176 [FREE Full text] [doi: 10.1016/S1607-551X(09)70467-X] [Medline: 12795346]
19. Henry C, M'Bailara K, Mathieu F, Poinsot R, Falissard B. Construction and validation of a dimensional scale exploring mood disorders: MAThyS (Multidimensional Assessment of Thymic States). BMC Psychiatry 2008;8:82 [FREE Full text] [doi: 10.1186/1471-244X-8-82] [Medline: 18803831]
20. Adler M, Liberg B, Andersson S, Isacsson G, Hetta J. Development and validation of the Affective Self Rating Scale for manic, depressive, and mixed affective states. Nord J Psychiatry 2008;62(2):130-135. [doi: 10.1080/08039480801960354] [Medline: 18569776]
21. Denicoff KD, Leverich GS, Nolen WA, Rush AJ, McElroy SL, Keck PE, et al. Validation of the prospective NIMH-Life-Chart Method (NIMH-LCM-p) for longitudinal assessment of bipolar illness. Psychol Med 2000 Nov;30(6):1391-1397. [Medline: 11097079]
22. Whybrow PC, Grof P, Gyulai L, Rasgon N, Glenn T, Bauer M. The electronic assessment of the longitudinal course of bipolar disorder: the ChronoRecord software. Pharmacopsychiatry 2003 Nov;36 Suppl 3:S244-S249. [doi: 10.1055/s-2003-45137] [Medline: 14677086]

23. Tsanas A, Saunders KE, Bilderbeck AC, Palmius N, Osipov M, Clifford GD, Goodwin, et al. Daily longitudinal self-monitoring of mood variability in bipolar disorder and borderline personality disorder. J Affect Disord 2016 Nov 15;205:225-233 [FREE Full text] [doi: 10.1016/j.jad.2016.06.065] [Medline: 27449555]
24. Hidalgo-Mazzei D, Mateu A, Reinares M, Murru A, Del MB, Varo C, et al. Psychoeducation in bipolar disorder with a SIMPLe smartphone application: feasibility, acceptability and satisfaction. J Affect Disord 2016 Aug;200:58-66. [doi: 10.1016/j.jad.2016.04.042] [Medline: 27128358]
25. Hidalgo-Mazzei D, Reinares M, Mateu A, Nikolova VL, Bonnín CD, Samalin L, et al. OpenSIMPLe: a real-world implementation feasibility study of a smartphone-based psychoeducation programme for bipolar disorder. J Affect Disord 2018 Dec 1;241:436-445. [doi: 10.1016/j.jad.2018.08.048] [Medline: 30145515]
26. Aktibipo SElf-RaTing (ASERT) Questionnaire. Mindpax. 2017. URL: https://www.mindpax.me/assets/docs/ASERT-questionnaireCZ-EN-DE.pdf [accessed 2020-10-20]
27. Mindpax. URL: www.mindpax.me [accessed 2020-10-20]
28. Montgomery SA, Asberg M. A new depression scale designed to be sensitive to change. Br J Psychiatry 1979 Apr;134:382-389. [Medline: 444788]
29. Young RC, Biggs JT, Ziegler VE, Meyer DA. A rating scale for mania: reliability, validity and sensitivity. Br J Psychiatry 1978 Nov;133:429-435. [Medline: 728692]
30. AKTIBIPO400 study. URL: https://bipo.nudz.cz [accessed 2020-11-26]
31. Sachs GS. Strategies for improving treatment of bipolar disorder: integration of measurement and management. Acta Psychiatr Scand Suppl 2004(422):7-17. [doi: 10.1111/j.1600-0447.2004.00409.x] [Medline: 15330934]
32. Aiken CB, Weisler RH, Sachs GS. The Bipolarity Index: a clinician-rated measure of diagnostic confidence. J Affect Disord 2015 May 15;177:59-64. [doi: 10.1016/j.jad.2015.02.004] [Medline: 25745836]
33. Jolliffe IT. Principal Component Analysis. New York, USA: Springer-Verlag; 2002:488.
34. Cronbach LJ. Coefficient alpha and the internal structure of tests. Psychometrika 1951 Sep;16(3):297-334. [doi: 10.1007/bf02310555]
35. Pinheiro J, Bates D. Mixed-Effects Models in S and S-PLUS. New York, USA: Springer; 2000:528.
36. Craighead WE, Evans DD. Factor analysis of the montgomery-asberg depression rating scale. Depression 1996;4(1):31-33. [doi: 10.1002/(SICI)1522-7162(1996)4:1<31::AID-DEPR3>3.0.CO;2-I] [Medline: 9160651]
37. McIntyre RS, Weiller E, Zhang P, Weiss C. Brexpiprazole as adjunctive treatment of major depressive disorder with anxious distress: results from a post-hoc analysis of two randomised controlled trials. J Affect Disord 2016 Sep 1;201:116-123. [doi: 10.1016/j.jad.2016.05.013] [Medline: 27208498]
38. Cassano GB, Rucci P, Benvenuti A, Miniati M, Calugi S, Maggi L, et al. The role of psychomotor activation in discriminating unipolar from bipolar disorders: a classification-tree analysis. J Clin Psychiatry 2012 Jan;73(1):22-28 [FREE Full text] [doi: 10.4088/JCP.11m06946] [Medline: 22316575]
39. MacKinnon DF, Zandi PP, Gershon ES, Nurnberger JI, DePaulo JR. Association of rapid mood switching with panic disorder and familial panic risk in familial bipolar disorder. Am J Psychiatry 2003 Sep;160(9):1696-1698. [doi: 10.1176/appi.ajp.160.9.1696] [Medline: 12944349]
40. Perlis RH, Ostacher MJ, Patel JK, Marangell LB, Zhang H, Wisniewski SR, et al. Predictors of recurrence in bipolar disorder: primary outcomes from the Systematic Treatment Enhancement Program for Bipolar Disorder (STEP-BD). Am J Psychiatry 2006 Feb;163(2):217-224. [doi: 10.1176/appi.ajp.163.2.217] [Medline: 16449474]
41. Simon NM, Otto MW, Wisniewski SR, Fossey M, Sagduyu K, Frank E, et al. Anxiety disorder comorbidity in bipolar disorder patients: data from the first 500 participants in the Systematic Treatment Enhancement Program for Bipolar Disorder (STEP-BD). Am J Psychiatry 2004 Dec;161(12):2222-2229. [doi: 10.1176/appi.ajp.161.12.2222] [Medline: 15569893]
42. Ghaemi SN, Rosenquist KJ. Is insight in mania state-dependent?: a meta-analysis. J Nerv Ment Dis 2004 Nov;192(11):771-775. [doi: 10.1097/01.nmd.0000145036.76435.c3] [Medline: 15505521]
43. Meyer TD, Crist N, La Rosa N, Ye B, Soares JC, Bauer IE. Are existing self-ratings of acute manic symptoms in adults reliable and valid?-a systematic review. Bipolar Disord 2020 Sep;22(6):558-568. [doi: 10.1111/bdi.12906] [Medline: 32232950]
44. Eisner E, Drake R, Lobban F, Bucci S, Emsley R, Barrowclough C. Comparing early signs and basic symptoms as methods for predicting psychotic relapse in clinical practice. Schizophr Res 2018 Feb;192:124-130 [FREE Full text] [doi: 10.1016/j.schres.2017.04.050] [Medline: 28499766]
45. Španiel F, Vohlídka P, Kožený J, Novák T, Hrdlička J, Motlová L, et al. The information technology aided relapse prevention programme in schizophrenia: an extension of a mirror-design follow-up. Int J Clin Pract 2008 Dec;62(12):1943-1946 [FREE Full text] [doi: 10.1111/j.1742-1241.2008.01903.x] [Medline: 18795967]
46. Španiel F, Hrdlička J, Novák T, Kožený J, Höschl C, Mohr P, et al. Effectiveness of the information technology-aided program of relapse prevention in schizophrenia (ITAREPS): a randomized, controlled, double-blind study. J Psychiatr Pract 2012 Jul;18(4):269-280. [doi: 10.1097/01.pra.0000416017.45591.c1] [Medline: 22805901]
47. Barnett I, Torous J, Staples P, Sandoval L, Keshavan M, Onnela J. Relapse prediction in schizophrenia through digital phenotyping: a pilot study. Neuropsychopharmacology 2018 Jul;43(8):1660-1666. [doi: 10.1038/s41386-018-0030-z] [Medline: 29511333]

48. Cochran AL, Schultz A, McInnis MG, Forger DB. A Comparison of Mathematical Models of Mood in Bipolar Disorder. Cham, UK: Springer; 2017:315-341.
49. Chang S, Chou T. A dynamical bifurcation model of bipolar disorder based on learned expectation and asymmetry in mood sensitivity. Comput Psychiatr 2018 Dec;2:205-222 [FREE Full text] [doi: 10.1162/cpsy_a_00021] [Medline: 30627671]
50. Faurholt-Jepsen M, Frost M, Ritz C, Christensen EM, Jacoby AS, Mikkelsen RL, et al. Daily electronic self-monitoring in bipolar disorder using smartphones - the MONARCA I trial: a randomized, placebo-controlled, single-blind, parallel group trial. Psychol Med 2015 Oct;45(13):2691-2704. [doi: 10.1017/S0033291715000410] [Medline: 26220802]
51. Faurholt-Jepsen M, Frost M, Christensen EM, Bardram JE, Vinberg M, Kessing LV. The effect of smartphone-based monitoring on illness activity in bipolar disorder: the MONARCA II randomized controlled single-blinded trial. Psychol Med 2020 Apr 4;50(5):838-848. [doi: 10.1017/S0033291719000710] [Medline: 30944054]
52. Mühlbauer E, Bauer M, Ebner-Priemer U, Ritter P, Hill H, Beier F, et al. Effectiveness of smartphone-based ambulatory assessment (SBAA-BD) including a predicting system for upcoming episodes in the long-term treatment of patients with bipolar disorders: study protocol for a randomized controlled single-blind trial. BMC Psychiatry 2018 Oct 26;18(1):349 [FREE Full text] [doi: 10.1186/s12888-018-1929-y] [Medline: 30367608]
53. Tønning ML, Kessing LV, Bardram JE, Faurholt-Jepsen M. Methodological challenges in randomized controlled trials on smartphone-based treatment in psychiatry: systematic review. J Med Internet Res 2019 Oct 27;21(10):e15362 [FREE Full text] [doi: 10.2196/15362] [Medline: 31663859]
54. Galvin HK, DeMuro PR. Developments in privacy and data ownership in mobile health technologies, 2016-2019. Yearb Med Inform 2020 Aug;29(1):32-43 [FREE Full text] [doi: 10.1055/s-0040-1701987] [Medline: 32823298]
55. Bauer M, Glenn T, Geddes J, Gitlin M, Grof P, Kessing LV, et al. Smartphones in mental health: a critical review of background issues, current status and future concerns. Int J Bipolar Disord 2020 Jan 10;8(1):19. [doi: 10.1186/s40345-019-0164-x]

Abbreviations

ASERT: Aktibipo Self-rating
AUROC: area under the receiver operating characteristic curve
BD: bipolar disorder
LME: linear mixed-effects
MADRS: Montgomery-Åsberg Depression Rating Scale
NIMH: National Institute of Mental Health
RCT: randomized controlled trial
YMRS: Young Mania Rating Scale

Edited by J Torous; submitted 09.12.20; peer-reviewed by B Johnson, P Ossola; comments to author 04.03.21; revised version received 23.04.21; accepted 10.05.21; published 09.08.21.

Please cite as:
Anýž J, Bakštein E, Dally A, Kolenič M, Hlinka J, Hartmannová T, Urbanová K, Correll CU, Novák D, Španiel F
Validity of the Aktibipo Self-rating Questionnaire for the Digital Self-assessment of Mood and Relapse Detection in Patients With Bipolar Disorder: Instrument Validation Study
JMIR Ment Health 2021;8(8):e26348
URL: https://mental.jmir.org/2021/8/e26348
doi:10.2196/26348
PMID:34383689

Original Paper

Toward a Mobile Platform for Real-world Digital Measurement of Depression: User-Centered Design, Data Quality, and Behavioral and Clinical Modeling

Stefanie Nickels[1], PhD; Matthew D Edwards[1], PhD; Sarah F Poole[1], PhD; Dale Winter[1], PhD; Jessica Gronsbell[1], PhD; Bella Rozenkrants[1], PhD; David P Miller[1], MSc; Mathias Fleck[1], PhD; Alan McLean[1]; Bret Peterson[1], PhD; Yuanwei Chen[1], PhD; Alan Hwang[1], ScB; David Rust-Smith[1], BCS; Arthur Brant[1], MD; Andrew Campbell[2], PhD; Chen Chen[1], MS; Collin Walter[1], MBA; Patricia A Arean[3], PhD; Honor Hsin[1], MD-PhD; Lance J Myers[1], PhD; William J Marks Jr[1], MS-HCM, MD; Jessica L Mega[1], MPH, MD; Danielle A Schlosser[1], PhD; Andrew J Conrad[1], PhD; Robert M Califf[1], MD; Menachem Fromer[1], PhD

[1]Verily Life Sciences, South San Francisco, CA, United States

[2]Department of Computer Science, Dartmouth College, Hanover, NH, United States

[3]Department of Psychiatry & Behavioral Sciences, University of Washington, Seattle, WA, United States

Corresponding Author:
Stefanie Nickels, PhD
Verily Life Sciences
269 E Grand Ave
South San Francisco, CA, 94080
United States
Phone: 1 833 391 8633
Email: stefnickels@verily.com

Abstract

Background: Although effective mental health treatments exist, the ability to match individuals to optimal treatments is poor, and timely assessment of response is difficult. One reason for these challenges is the lack of objective measurement of psychiatric symptoms. Sensors and active tasks recorded by smartphones provide a low-burden, low-cost, and scalable way to capture real-world data from patients that could augment clinical decision-making and move the field of mental health closer to measurement-based care.

Objective: This study tests the feasibility of a fully remote study on individuals with self-reported depression using an Android-based smartphone app to collect subjective and objective measures associated with depression severity. The goals of this pilot study are to develop an engaging user interface for high task adherence through user-centered design; test the quality of collected data from passive sensors; start building clinically relevant behavioral measures (features) from passive sensors and active inputs; and preliminarily explore connections between these features and depression severity.

Methods: A total of 600 participants were asked to download the study app to join this fully remote, observational 12-week study. The app passively collected 20 sensor data streams (eg, ambient audio level, location, and inertial measurement units), and participants were asked to complete daily survey tasks, weekly voice diaries, and the clinically validated Patient Health Questionnaire (PHQ-9) self-survey. Pairwise correlations between derived behavioral features (eg, weekly minutes spent at home) and PHQ-9 were computed. Using these behavioral features, we also constructed an elastic net penalized multivariate logistic regression model predicting depressed versus nondepressed PHQ-9 scores (ie, dichotomized PHQ-9).

Results: A total of 415 individuals logged into the app. Over the course of the 12-week study, these participants completed 83.35% (4151/4980) of the PHQ-9s. Applying data sufficiency rules for minimally necessary daily and weekly data resulted in 3779 participant-weeks of data across 384 participants. Using a subset of 34 behavioral features, we found that 11 features showed a significant (*P*<.001 Benjamini-Hochberg adjusted) Spearman correlation with weekly PHQ-9, including voice diary–derived word sentiment and ambient audio levels. Restricting the data to those cases in which all 34 behavioral features were present, we had available 1013 participant-weeks from 186 participants. The logistic regression model predicting depression status resulted in a 10-fold cross-validated mean area under the curve of 0.656 (SD 0.079).

Conclusions: This study finds a strong proof of concept for the use of a smartphone-based assessment of depression outcomes. Behavioral features derived from passive sensors and active tasks show promising correlations with a validated clinical measure of depression (PHQ-9). Future work is needed to increase scale that may permit the construction of more complex (eg, nonlinear) predictive models and better handle data missingness.

(JMIR Ment Health 2021;8(8):e27589) doi:10.2196/27589

KEYWORDS

mental health; mobile sensing; mobile phone; mHealth; depression; location; GPS; app usage; voice diaries; adherence; engagement; mobility; sleep; physical activity; digital phenotyping; user-centered design

Introduction

Background

Mental health disorders are the leading cause of years lost to illness, disability, or premature death, as represented by disability-adjusted life years (DALYs) [1]. Perhaps, even more disquieting is the fact that we have made no progress in reducing these lost years over the past decades. From 1990 to 2016, years lived with disability (YLDs) for major depressive disorder, anxiety disorders, schizophrenia, dysthymia, and bipolar disorder have not changed, and the YLDs for opioid use disorders have even increased [2].

Although there are several reasons for this failure to improve outcomes, one area that has received increased attention is a relative lack of objective measurement in mental health [3]. Other fields have advanced by leveraging biomarkers and measurement tools to empower patients and clinicians to improve clinical decision-making. In the field of mental health, the absence of objective biomarkers leads to reliance on subjective information derived from clinical interviews and self-report surveys. Although the use of systematic self-report measures and subsequent algorithm-based treatment adjustments have shown significant promise [4,5], there is a substantial gap between the clinical trials that demonstrate this promise and what is deployed in routine practice [6]. A survey in the United Kingdom found that 80% of clinicians do not routinely use any outcome measures because they find them to be clinically irrelevant and too time consuming [7].

To address this shortcoming, the idea of taking advantage of the ubiquitous availability of smartphones has been proposed [8]. It stands to reason that novel data streams from digital sensors may provide a low-burden and scalable option to continuously quantify out-of-clinic real-world behaviors, which, in addition to traditional assessments in the clinic, could be used to bring psychiatry closer to realizing the consistently improved outcomes achieved by measurement-based care [9].

A number of studies have since explored the feasibility of linking smartphone or wearable data to clinical symptoms. In addition, passive sensing platforms have been built to collect smartphone data and examine the prediction of mental health clinical phenomena (Monsenso Monarca platform [10], Harvard Beiwe platform [11], Mindstrong [12], Dartmouth StudentLife platform [13], and Ginger.io [14], among others). A recent systematic meta-analysis review on the relationship between digital behavioral features and depressive mood symptoms concluded that this approach is promising [15]. However, most studies included in that review are not readily generalizable because of small sample sizes, where of the 26 studies of participants with mood disorders, all but one had fewer than 61 participants, and the median number of participants was 18. Furthermore, these studies were often designed as proof-of-concept investigations of single phone features (eg, actigraphy or GPS only). Even if meaningful correlations are found at this scale, the potential for identifying false-positive associations in situations where signals greatly outnumber clinical events cannot be ignored. Therefore, large-scale studies are necessary to adequately evaluate the associations and develop predictive models of clinical significance. A recent study examining the relationship coefficients between passive sensors and Patient Health Questionnaire (PHQ-2) for 271 individuals, for example, found that personalized models of behavioral features may be predictive of mood, with an area under the curve (AUC) of >0.5 for 81% of participants [16].

Objectives

Therefore, we seek to create a scalable smartphone platform that can acquire robust, multisensor streams of passively collected signals from individuals with depression, in addition to longitudinal self-report survey measures. Here, we present the results of our first feasibility study with this platform app, where we aim to (1) deploy a user-friendly interface to achieve high study engagement among participants with depression while making enrollment and all study procedures fully remote and automated; (2) test the quality of collected passive and active data, including adherence to weekly PHQ-9 tasks; (3) demonstrate the ability to construct clinically relevant behavioral measures (*features*) from passive sensors and active inputs; and (4) preliminarily explore relationships between these features and symptoms of depression.

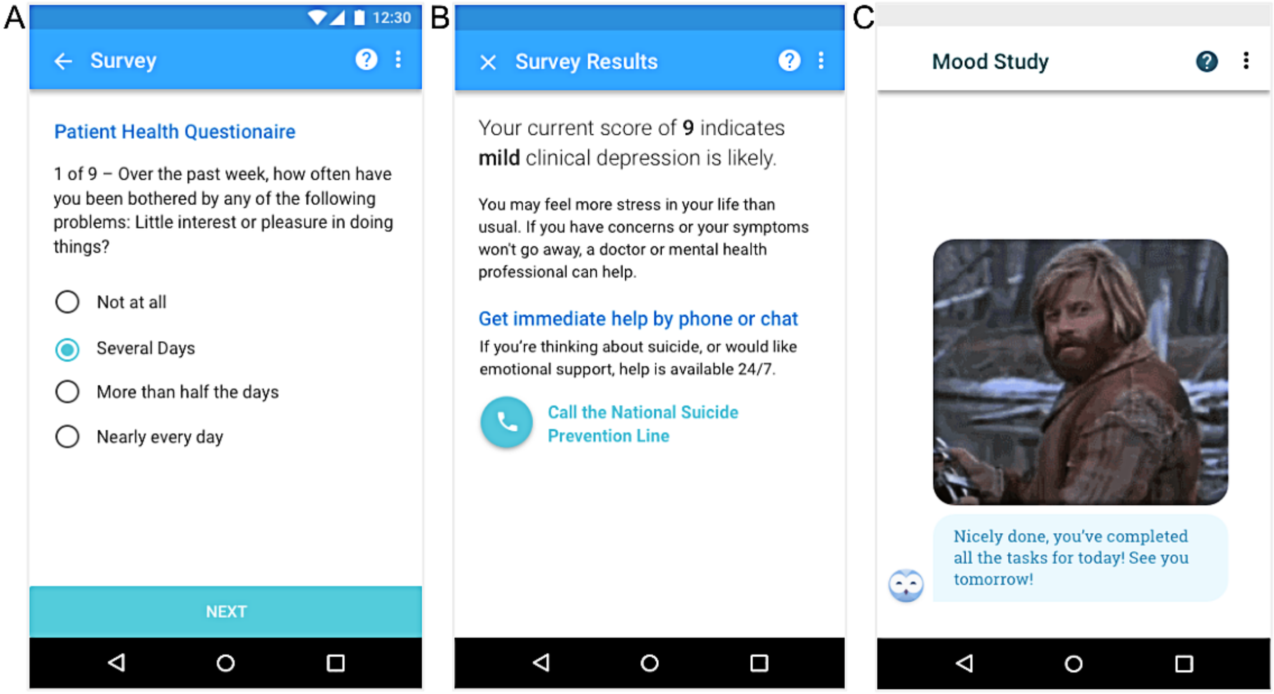
Methods

Overview

To achieve these objectives, we conducted a remote feasibility study with 600 participants. Of those participants, 80% (480/600) were treatment-seeking individuals with self-reported depression. The focus on treatment-seeking individuals was motivated by the assumption that recent treatment may lead to behavior and mood changes that can be captured during the course of the study. The remaining participants were nondepressed controls. Participants were also randomly assigned to one of the three compensation cohorts (Multimedia Appendix 1), allowing us to empirically test the effect of financial incentives on study adherence. We hypothesized that participants in the highest compensation group would show higher study adherence than those in the lower compensation groups.

The mobile app recorded data from smartphone-based sensors and provided a chatbot user interface (Figure 1) designed to engage participants and to collect self-report measures. The feasibility of launching a comprehensive sensor-based tool was evaluated via an examination of data quality and missingness of sensor and survey data as well as whether we would succeed in deriving behavioral features from the recorded raw sensor data. For an exploratory analysis, we hypothesized that location-based and social features would be related to mood symptoms, given previous literature citing such relationships. Another exploratory aim of the study was to build a multivariate model to discriminate between participants with and without depression using the derived behavioral features.

Figure 1. User interface of the study app. (A) Mobile rendering of the first question of the self-administered Patient Health Questionnaire-9 survey. (B) The score for the Patient Health Questionnaire-9 survey was tabulated and put into context for the study participant. (C) An owl avatar was used to deliver messages and occasional celebratory gif images.



Informed Consent

This study was approved by the Western Institutional Review Board Inc. In an effort to maximize transparency and mitigate privacy concerns, the informed consent form (ICF) included detailed descriptions of all sensors and the type of information that could be derived from them (Multimedia Appendix 2 provides a full copy of the ICF). It also included information about the relationship between Verily Life Sciences and Google and stated that data would not be shared with Google for advertising purposes. The app further featured a persistent notification (ie, it was visible at all times) prominently reminding the user that the app was continuously collecting data.

Participants

Participants were recruited through social media and search engine advertisements (Facebook and Google AdWords). Our goal was to enroll 600 participants, 480 of which would be treatment-seeking individuals with depression and 120 nondepressed controls. In the cohort with 480 participants, participants had to have a PHQ-9 score of ≥10 (classified as *depressed* at enrollment). For the control group, participants were only included if their PHQ-9 score was <10 (classified as *not depressed* at enrollment). Participants were excluded when a plan to inflict self-harm was reported. The full list of inclusion and exclusion criteria can be found in the ICF in Multimedia Appendix 2.

Study Design

Surveys

During the course of the 12-week study, participants provided active survey data in the form of daily questions about their sleep, functioning, and activity level. Participants were further asked to complete the PHQ-9, a quality of life questionnaire, and a voice diary on a weekly basis. Once a month, participants completed the Study Life Satisfaction scale [17] and the National Institute of Health Toolbox Loneliness scale [18], and participants completed a user feedback questionnaire after 6 weeks and 12 weeks in the study.

One subaim of this study was to determine an appropriate level of compensation that would sufficiently motivate participants to adhere to this intense survey protocol. To that end, tasks were compensated with different amounts depending on a random assignment to three compensation cohorts of US $135, US $265, or US $530 as the maximum study compensation if all tasks were completed. Participants were mailed a physical gift card, which was reloaded every 4 weeks with the accrued sum of their cohort-specific compensation. Further details of this subexperiment and the results can be found in Multimedia Appendix 1.

Sensors

We developed a mobile phone app for the Android operating system containing a broad range of *sensors* recording information from and about the following: accelerometer, ambient audio, Android device information, barometric air pressure, battery charge, Bluetooth, gyroscope, light level, magnetometer, network, phone calls, physical activity level, ping (a data point sampled every minute when the app was running), proximity (measures whether any object is within a given distance), screen state, step count, text messages, volume, and Wi-Fi networks. The ICF in Multimedia Appendix 2 provides more information and privacy statements about these sensors. Note that these sensors ranged from actual hardware sensors (eg, ambient audio) to measurements that rely on algorithmic processing of raw data (eg, inferring physical activity from raw accelerometer and gyroscope data).

Data Filtering and Aggregation

Quality control for the data consisted of a stepwise process that considered a day's worth of data for each participant and then considered a week's worth of data. In particular, participant-days were included in the analysis if the following two conditions were met: (1) there were at least 720 minutes (ie, half a day's worth) of available data for at least one of the sensors programmed to be sampled at regular 1-minute intervals (ie, ambient audio, ambient light, physical activity, location, pressure, ping, and proximity) and (2) the app had been active at least once during 18 different hours of the day. Subsequently, participant-weeks were included if there were 3 or more such days available to aggregate. In that case, the data for both regularly sampled features (eg, location) and user-driven features (eg, text messages and phone calls) were aggregated to a weekly mean. Note that using this rule, it is possible that a *specific feature* has fewer than 3 days or even no data at all available to aggregate (refer to the *Results* section).

Feature Engineering

The next step in our data processing pipeline consisted of deriving clinically interpretable weekly summaries of the available underlying data, designated as weekly *behavioral features*. For some features (eg, ambient audio, ambient light, app use, and battery charge), aggregating average daily levels to average weekly levels was the only step of feature engineering. For other types of features, additional preprocessing was required. For example, to extract physical activity minutes, an algorithm determined the type of activity from a predefined set (eg, running, walking, and biking) from the accelerometer and gyroscope sensors. For voice diaries, algorithms were developed to extract features of interest from the audio file (eg, average pause duration and duration of diary) as well as from transcribed text (eg, spoken words per minute and sentiment score). Text messages (outgoing messages only) were also analyzed for their sentiment score, word and emoji count, and daily message count. Location data, originally collected as longitude and latitude every minute, were passed through a density-based spatial clustering of applications with noise (DBSCAN) clustering algorithm to arrive at location clusters. These clusters were further augmented using time-bounded heuristics to define "home" (the single place where participants spent the most time between 11 PM and 4 AM; this was required to be the same place over the 12 weeks of the study), "work" (any place where participants spent at least 15% of their time between 10 AM and 3 PM and which is not the "home" cluster; this could be multiple places over the course of the study), and "commute" (time spent traveling between home and work). The weighted average of latitudes and longitudes for each cluster was then used to retrieve semantic information from Google Maps about the type of location (eg, hospital or doctor's office and place to exercise). Using this method, it is possible that a given central point is tagged with multiple types of location.

Feature Selection

Feature engineering, as described earlier, yielded approximately 600 possible features. With a data set of only 384 participants, we decided a priori to use only 34 of those features, based on previous literature as well as clinical usefulness and interpretability.

Statistical Analysis

Univariate Correlation Analysis

Pairwise correlations between the behavioral features and PHQ-9 were calculated on all weeks that had both the feature of interest and a weekly PHQ-9 survey available after applying the daily and weekly filtering rules. A total of 34 univariate tests were conducted, and to account for multiplicity, we used the Benjamini-Hochberg (nonnegative) procedure to adjust these 34 *P* values. The Python function *statsmodels.stats.multitest.multipletests* was used.

Multivariate Logistic Regression Model

Preprocessing for the multivariate model included standardizing all features using the *sklearn.preprocessing.StandardScaler* function. To predict the binarized PHQ-9 score (<10 or ≥10) from the 34 behavioral features, we fit an elastic net penalized logistic regression model with the function *sklearn.linear_model.LogisticRegression*. The inverse regularization parameter C was set to the default value of 1.0, and the L1 ratio (the elastic net mixing parameter) was set to 0.5. This model was passed through a 10-fold cross-validation procedure in which each participant's data for all participant-weeks were placed in either the training or test set in each fold to prevent any label leakage.

Results

User-Centered Design

Previous studies examining the feasibility of digital measurement tools have been hampered by relatively high dropout rates. For example, in a randomized controlled trial on digital sensing and cognitive behavioral therapy on 1098 depressed participants, adherence dropped to 41% by the 12-week mark [19]. A similar study reported that only 10.5% of the initial 126 participants were still active at 8 weeks [20]. Therefore, another aim of this study was to design the user experience such that interaction with the app would be self-explanatory, engaging, and enjoyable.

We conducted a number of user experience studies on depressed participants and clinical professionals to design a user interface that would be appropriate in tone for participants with depression yet maximally engaging (for a detailed description of these studies, refer to Multimedia Appendix 3).

On the basis of this research, it was ultimately decided that our app would feature a chatbot interface that guides the user through onboarding and their daily survey tasks with a series of static, predefined messages and provides unexpected moments of delight using celebratory gif images (Figure 1). The chatbot was kept deliberately lighthearted, where appropriate, and conveyed gratitude for contributing to the overall mission. Task completion was immediately followed by a celebratory moment (text encouragement plus occasional gif image) and reward (information about how much compensation was earned for completing those tasks).

Enrollment Funnel and Participant Characteristics

Enrollment started in November 2017, and the last participant finished their 12 weeks in May 2018. Figure 2 summarizes the participant drop-off at various stages. Of the 2360 participants who completed the enrollment survey, 612 (25.93%) were eligible based on the criteria reported earlier. Of those, 465 downloaded the study app, and 415 "good faith" users remained after we performed a conservative audit against potentially fraudulent activity (eg, users who seemed to be enrolled multiple times and users living in the same home as another user). The sample was further reduced to 384 users who had minimally sufficient data (ie, completed at least one in-app PHQ-9 and had at least 8 days of sufficient sensor data as per the daily filtering rule; refer to the *Methods* section).

In line with our 80:20 depressed to nondepressed recruitment ratio, the 384 participants with minimum data consisted of 313 depressed (81.5%; mean PHQ-9 at baseline 17.4, SD 4.1) and 71 nondepressed individuals (18.5%; mean PHQ-9 5.4, SD 2.9). The demographic data of these 384 participants are summarized in Table 1. Multimedia Appendix 4 provides additional demographic variables for both the 2360 individuals who completed the enrollment survey and the 384 with minimally usable data.

Figure 2. Participant flow diagram.

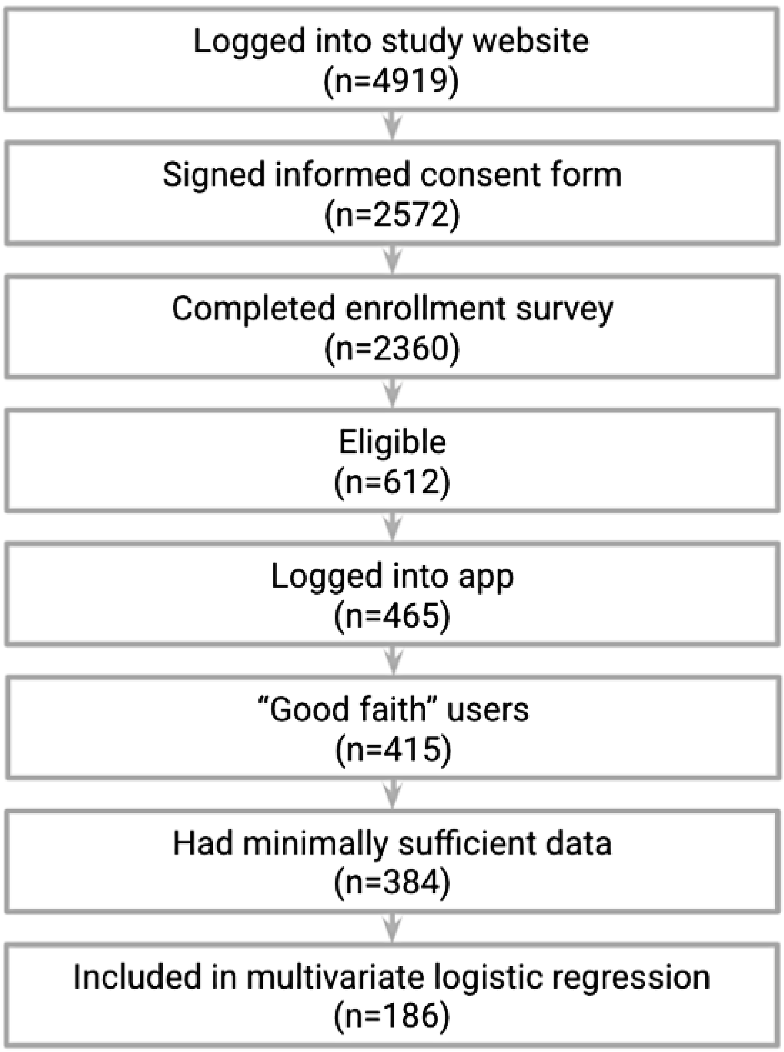


Table 1. Key demographics of the 384 participants with minimally sufficient data.

Characteristic	Participants with minimally sufficient data (n=384)	
	Depressed (n=313)	Not depressed (n=71)
Age (years), n (%)		
18-29	123 (39.3)	26 (36.6)
30-39	90 (28.8)	22 (30.9)
40-49	56 (17.9)	14 (19.7)
50-59	36 (11.5)	8 (11.2)
60-69	7 (2.2)	0 (0)
70-79	1 (0.3)	1 (1.4)
Race, n (%)		
American Indian or Alaska Native	12 (3.8)	1 (1.4)
Asian	5 (1.6)	4 (5.6)
Black or African American	29 (9.3)	12 (16.9)
Native Hawaiian or other Pacific Islander	2 (0.6)	1 (1.4)
Other	14 (4.5)	1 (1.4)
White	251 (80.2)	52 (73.2)
Sex at birth, n (%)		
Female	285 (91.1)	51 (71.8)
Male	28 (8.9)	20 (28.2)

Data Quality

Survey Data Availability

From the enrolled, good faith participants (n=415), we received 83.35% (4151/4980) of all possible weekly PHQ-9 surveys over the course of the 12-week study. Analyses regarding the effects of different compensation cohorts on PHQ-9 adherence revealed no statistically significant effect of compensation group (Multimedia Appendix 1).

Sensor Data Availability

No data could be collected if the phone was turned off; the user force-closed the study app; the app crashed unexpectedly; or if the user revoked the persistent notification permission for the study app, as Android will eventually remove background services; in this case, data collection may resume if the user navigates back into the app. When no data were collected, it was impossible to determine the cause of data missingness from the possibilities listed earlier.

Data from specific sensors may be missing if the user permanently or temporarily revoked permission for those sensors that need specific permissions (ie, Android activity recognition, app usage, Google Fit activity recognition, location (GPS), call logs, microphone for audio diaries—not needed for ambient audio, not used for phone calls as only metadata were recorded, and text logs). A later version of the study app included a sensor that records permission status for these, but for this study, it was not possible to know whether sensor data were missing because of revoked permissions or other reasons.

After applying the daily and weekly filtering rules (refer to the *Methods* section), we had available a quality control–positive data set of 3779 participant-weeks on 384 participants (Figure 2).

To visually characterize the patterns of missingness that result in removed days and weeks in the filtering rules, we consider for illustration purposes the daily availability of the minutely sampled sensors of ambient audio and location (Figure 3). These two sensors yielded some features with significant univariate correlations (Table 2) and were therefore used as examples. It is apparent that data missingness varies greatly among participants and their phone types (a detailed description of the phone types present in this study is presented in Multimedia Appendix 5). For example, the ambient audio sensor data are missing entirely for certain participants (see the bottom gray rows in the daily ambient audio panel in Figure 3) possibly because the hardware sensor is not present in certain phone types. Similarly, although location (GPS) capabilities are available on most phones, a special permission must be given to the app to record this sensor. The daily location panel shows that there are some participants who never have any location data (bottom gray rows) and many participants with intermittently missing location data. It is also apparent that, even on days where some data were available, the number of sampling points varies considerably on a day-to-day basis (yellow to blue gradient). We note that there are many participants whose ambient audio data or location are sufficient for analysis, but because of *other* missing features, they are excluded from the multivariate analysis (ie, they are among those participants below the horizontal dotted lines in Figure 3), as we required complete case data (refer to the *Methods* section). However, for the univariate correlation analysis, data from those participants were included if they met the daily and weekly sufficiency criteria.

Figure 3. Heat maps of daily data availability. For ambient audio (A) and location (B) sensors, the color scale denotes the number of data samples on a particular participant day (between 0 and 1440, ie, the number of minutes in a day) across the 415 good faith participants (heat map rows). Gray indicates that the day was excluded altogether as per daily filtering rules. Participants are ordered in the same way in both panels, and participants that fall below the horizontal dotted line were not included in the multivariate analysis because of missingness in these or other features.

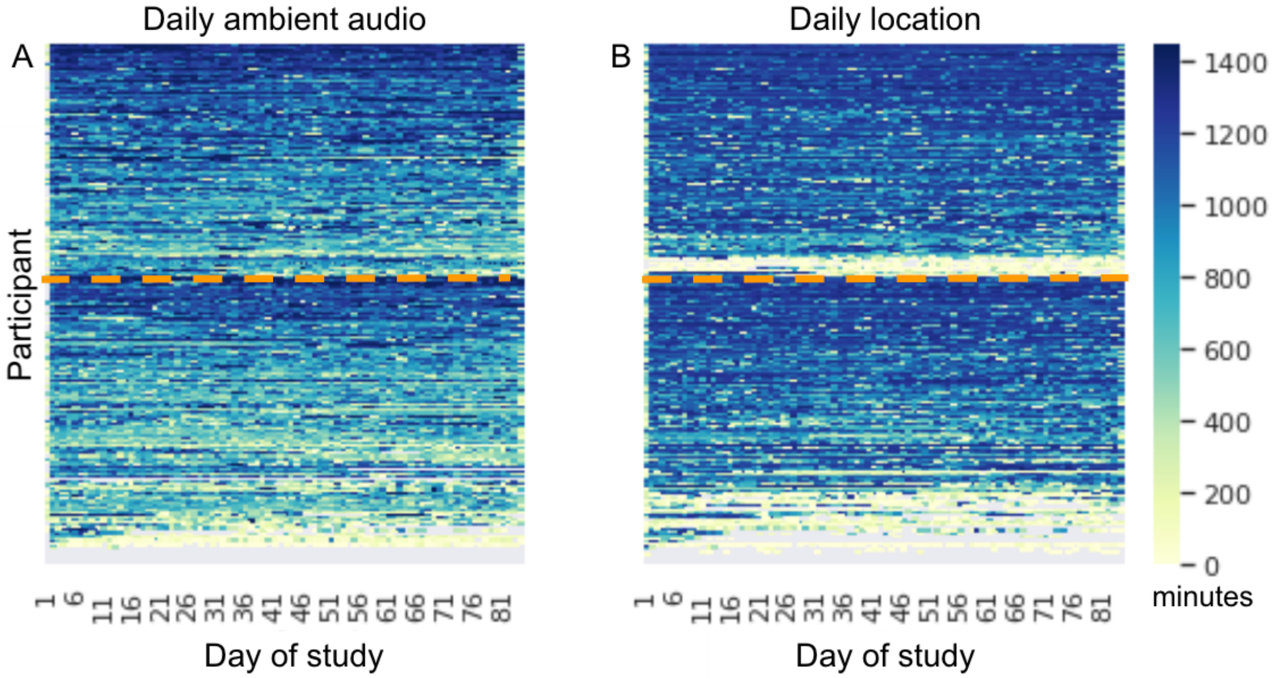


Table 2. Univariate correlations between 34 features and Patient Health Questionnaire-9 score ordered by absolute Spearman rank correlation coefficient[a].

Correlation with Patient Health Questionnaire-9 sum	Spearman *r*	Spearman *P* value	Benjamini-Hochberg-adjusted *P* value	Participant-weeks (n=3779), n (%)	Participants (n=384), n (%)
Voice diary sentiment	−0.26	<.001	<.001	2992 (79.2)	376 (97.9)
Reported sleep duration	−0.12	<.001	<.001	2797 (74)	308 (80.2)
Ambient audio level	0.10	<.001	<.001	3618 (95.7)	368 (95.8)
Phone call ringing until missed minutes	0.10	<.001	<.001	3334 (88.2)	360 (93.8)
Unique location clusters	−0.09	<.001	<.001	3623 (95.9)	379 (98.7)
Voice diary words per minute	−0.09	<.001	<.001	2992 (79.2)	376 (97.9)
Voice diary duration	0.08	<.001	<.001	3145 (83.2)	377 (98.2)
Battery percentage	−0.07	<.001	<.001	3778 (99.9)	384 (100)
Text message emoji count	−0.07	<.001	<.001	3460 (91.6)	366 (95.3)
Phone call count	0.07	<.001	<.001	3534 (93.5)	370 (96.4)
Voice diary pauses duration	0.06	<.001	.001	3038 (80.4)	377 (98.2)
Location entropy	−0.06	<.001	<.001	3623 (95.9)	379 (98.7)
Ambient light level	−0.06	.001	.003	3122 (82.6)	325 (84.6)
Outgoing text message sentiment score	−0.06	.001	.003	3140 (83.1)	351 (91.4)
Location variance	−0.06	<.001	.002	3623 (95.9)	379 (98.7)
Phone screen on minutes	0.05	.002	.003	3776 (99.9)	384 (100)
Audio system volume	−0.05	.002	.003	3779 (100)	384 (100)
Charging minutes	0.04	.009	.02	3773 (99.8)	384 (100)
Social apps usage	−0.04	.02	.03	3346 (88.5)	357 (93)
Time spent at home	0.04	.03	.049	3623 (95.9)	379 (98.7)
App usage missing	−0.03	.047	.08	3779 (100)	384 (100)
Outgoing phone call duration	−0.03	.08	.12	3295 (87.2)	355 (92.4)
Wellness apps usage	0.03	.08	.12	3346 (88.5)	357 (93)
Time spent at hospital	−0.03	.09	.12	3623 (95.9)	379 (98.7)
Number of Wi-Fi networks	−0.03	.08	.12	3778 (99.9)	384 (100)
Reported sleep duration missing	−0.03	.12	.16	3779 (100)	384 (100)
Communication apps usage	0.02	.24	.30	3346 (88.5)	357 (93)
Text message body size	0.02	.26	.31	3460 (91.6)	366 (95.3)
Ring volume	0.02	.32	.37	3779 (100)	384 (100)
Physically active minutes	−0.01	.38	.43	3612 (95.6)	382 (99.5)
Audio notification volume	−0.01	.42	.46	3779 (100)	384 (100)
Text message count	0.01	.66	.71	3460 (91.6)	366 (95.3)
Nearby Wi-Fi networks count	0.01	.76	.78	3364 (89)	378 (98.4)
Incoming phone call duration	0.00	.97	.97	3217 (85.1)	354 (92.2)

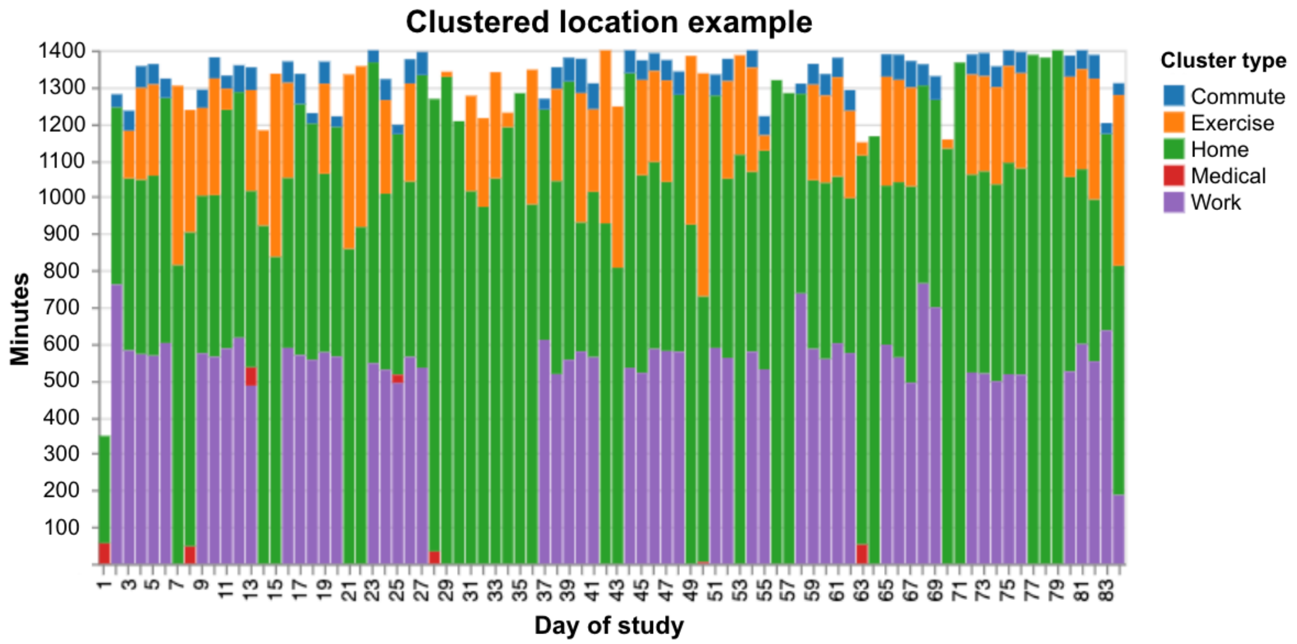
[a]The univariate correlations here are typically based on a subset of the filtered data set depending on the pairwise availability of the feature of interest and the weekly Patient Health Questionnaire-9 score, with data count as noted in the last two columns.

Behavioral and Clinical Modeling

Case Study: Location-Derived Behavioral Features

To anecdotally illustrate the quality of insights into real-world patient behavior that can be obtained from passive sensors, Figure 4 depicts a graphical summary of 1 participant's daily clustered location data. It is apparent that this participant had a strong weekly routine, going to work for approximately 550 minutes (approximately 9 hours) for 5 days, followed by no time spent at work for 2 days. The participant also regularly spends time at exercise locations (eg, a gym or a sports center) and has some medical appointments. An interesting disruption in this routine can be observed starting on day 28 of the study, during which the participant had a medical appointment, followed by 8 days without any time at work or commuting.

Figure 4. Example of clustered location data for 1 participant for the duration of the study. The total number of minutes (vertical axis) with categorized locations (denoted by various colors in the legend) are plotted as stacked bars for each day of the study on the horizontal axis. Note the week-long increased homestay starting on day 28.



Univariate Correlations Analysis With PHQ-9

To move beyond anecdotal evidence that smartphone sensors may be clinically relevant, we started with a simple question asking which, if any, behavioral features show correlations with the PHQ-9. To this end, we considered the 3779 participant-weeks (across 384 participants) available after filtering. Among the 34 behavioral features selected for analysis (refer to the *Methods* section), 11 showed a significant (P<.001) Benjamini-Hochberg-adjusted correlation with weekly PHQ-9 (Table 2). Visualizations of the spread of the PHQ-9 values and the associations between the significant behavioral features and the PHQ-9 values can be found in Multimedia Appendix 6.

In descending order of the absolute strength of the relationship, we found that a more negative sentiment of the voice diary—a measure derived from a sentiment classification algorithm [21] returning measures from −1 (extremely negative) to 1 (extremely positive)—is associated with a higher PHQ-9 score. Self-reported sleep was negatively correlated, meaning that the less sleep the participant reported that week, the higher their PHQ-9 score was. A higher ambient audio level, that is, more noise registered by the phone, was associated with higher depression severity. Letting the phone ring for longer periods until the call was missed was also associated with higher depression severity. Unique location clusters or how many different places (eg, home, work, or unlabeled clusters) the participant visited that week was negatively correlated with PHQ-9, meaning that the fewer different locations visited in a given week, the higher the PHQ-9 score. Two other measures from the voice diary, words spoken per minute (speaking rate) and the recorded duration of the diary, also showed a significant relationship: the fewer words spoken per minute and the longer the duration of voice diary, the higher the PHQ-9. The battery percentage (between 0% and 100%) was negatively associated, meaning that if the weekly mean battery percentage was lower, PHQ-9 was higher. The number of emojis in outgoing text messages showed a negative correlation, that is, a lower number of emojis was associated with a higher PHQ-9 score. Receiving or making more phone calls per week was associated with higher PHQ-9 scores. Finally, location entropy—measuring the variability in how much time a participant spent at different location clusters—showed a negative association: the less variability in where participants spent time, the higher the PHQ-9.

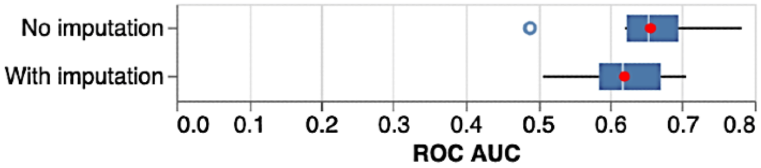
Multivariate Logistic Regression

Recall that, after filtering, there were 3779 participant-weeks available for 384 participants. Of those, 1013 weeks from 186 participants were complete cases for the 34 features and were therefore included in an initial multivariate analysis (the correlation matrix of the 34 features is presented in Multimedia Appendix 7). For these participant-weeks, 60.91% (617/1013) were classified as "depressed" (PHQ-9≥10), whereas 39.09% (396/1013) were "nondepressed" weeks (PHQ-9 score <10). We used an elastic net penalized multivariate logistic regression for predicting weekly depression status and evaluated performance with the receiver operating characteristic AUC of 10 cross-validated folds. The mean AUC of the 10 folds was 0.656 (SD 0.079; Figure 5).

We also attempted to recover some participant-weeks that were dropped for not being complete cases by performing within-participant mean imputation. This resulted in a data set of 2016 participant-weeks for 197 participants, with 1173 (58.18%) depressed and 843 (41.82%) nondepressed weeks.

Some weeks could not be recovered because a feature was missing for all 12 weeks and could therefore not be imputed using this simple within-participant procedure. With a mean AUC of 0.620 (SD 0.062; Figure 5), models trained using these imputed data performed worse on average than the models trained without imputed data, although there was less spread in the 10 AUC values.

Figure 5. Box plots of areas under the curve from 10-fold cross-validation of logistic regression predicting depression status, for the nonimputed and imputed data sets. The white line represents the median, and the red dot represents the mean. AUC: area under the curve; ROC: receiver operating characteristic.



Discussion

Principal Findings

The results of this feasibility study showed that it is possible to remotely engage a depressed study population quickly and with high adherence (>80%), data collection from a broad range of Android smartphone sensors is possible, and multiple features from multiple feature families can be extracted. We further provided a proof-of-concept demonstration of multivariate modeling of depression status.

It is also clear that passive sensor data collection using the "bring-your-own-device" approach is complex, especially across the diverse ecosystem of Android devices (Multimedia Appendix 5). The data exhibit high levels of missingness and noise related to the characteristics of the phone and the user. Furthermore, our choice of minimally required daily and weekly data amounts was arbitrary. Future studies on a scaled-up sample should develop more sophisticated models of missingness imputation that account for differences in hardware sensors and phone type quality. Furthermore, a large-scale sample would allow for more complex (eg, nonlinear) modeling of clinical outcomes, particularly considering interactions with demographics and other covariates. Therefore, to interpret this report, it is important to bear in mind that this sample skews heavily female compared with the population of individuals with depression in the United States. It also only consists of individuals who own Android phones (refer to the *Limitations* section).

Univariate correlations of 34 behavioral features showed promising correlations with the clinically validated PHQ-9 scores. Some of those replicate previous findings on similar features, for example, the negative associations between location entropy and location variance with PHQ-9 [22], the positive relationship between the duration of a diary and PHQ-9, and the negative relationship between words per minute and PHQ-9 (compare the findings herein with the total recording length [r=0.20] and speaking rate as syllables per second [r=−0.23] association with the Hamilton Depression Rating Scale in Mundt et al [23]). Of note, the univariate correlation coefficients reported in our study are, in many cases, smaller than those reported for similar behavioral features in other studies despite our larger sample (refer to the systematic review by Rohani et al [15], specifically the Multimedia Appendix 5 for a summary of the correlation coefficients from studies on patients with unipolar and bipolar depression). We posit that these inconsistencies might be because of the winner's curse—a statistical phenomenon discussed often in genome-wide association studies [24]—where associations that are close to a decision threshold are more likely to be an overestimation of the true association in the generating sample, leading to poorer replication in a separate sample (refer to the paper by Pratap et al [16] for a larger scale replication study that found no association for phone usage and a small association for GPS mobility). We also hypothesize that, for the prediction of mood scores, it may be more beneficial to define participant-level models focusing on changes against oneself over time rather than population-level models focusing on deviations from the population mean [16].

Another aim of this study was to demonstrate the potential clinical utility of behavioral features derived from passive sensors to help inform clinical practice. The location example (Figure 3) shows one area in which a clinician could gain insights about a patient's routine that could augment clinical decision-making and strengthen feedback-informed care [25]. This and other passive features such as physical activity minutes, social app usage, or how noisy a patient's environment is could perhaps give clinicians a more holistic and continuous picture of their patient's circumstances than would likely be available from momentary in-clinic visits alone. However, although the statistically significant correlations suggest some clinical relevance, future validation work is needed to establish whether some of the behavioral features alone or a multivariate model provide a sufficiently reliable signal to warrant clinical actionability.

Notably, active tasks such as voice diaries yielded rich insights into the patient's previous week, resulting in the strongest correlation with the PHQ-9 (r=0.26). Active tasks collected via a mobile app, be it voice diaries or even standard clinical measures, such as the PHQ-9, might offer another avenue for gathering perhaps not continuous but more frequent data points about a patient's well-being. As they are completed without any burden on the clinician, such self-administered measurements could remove some of the barriers to the adoption of measurement-based care in psychiatry [7]. Furthermore, one could imagine the development of an early alert system or a case ranking system based on these digitally deployed measurements that could draw attention to deteriorating patients without them having to appear in person and without them having to directly tell their therapist the uncomfortable fact that they are not getting better [25,26].

For the active measurements in particular, we believe that this study is an encouraging example of how an intuitive, fun, and engaging user interface can make interacting with an app around a generally difficult topic, such as depression, not only less burdensome but perhaps even pleasant and beneficial. In fact, anecdotal feedback about using our app suggests as much. One patient reported “[t]he app made me feel like I had an everyday purpose. I looked forward to filling it out. I enjoyed the interaction.” Another patient stated “[i]t put my mental health into perspective and I had to answer how I was feeling, not what people expect me to feel.”

Limitations

The entirely open recruitment through advertisements in this study led to a sample that was different from the larger population of individuals with depression in the United States (compare the national survey data in Luciano and Meara [27]). Most notably, our depressed sample consisted of more females (285/313, 91.1% vs 12,182/20,313, 59.97% in the national survey [27]). Furthermore, enrollment was restricted to participants owning phones with an Android operating system. All these factors lead to limited generalizability of the findings. Future recruitment strategies should involve more targeted advertising or partnering with companies or health care systems to recruit a more representative sample. We also enrolled participants based on their self-report of symptoms and treatment history on good faith; however, future studies with strengthened clinical verification methods (eg, concurrent donation of clinical records or prescription records to the research study) may enhance the validity of clinical measures.

PHQ-9 adherence (>80%) was found to be the same for all three compensation cohorts (Multimedia Appendix 1), but future work is needed to test whether adherence with a similarly engaging app but less compensation than the lowest level (US $135) will lead to similarly high adherence levels.

The semantic location features, especially locations marked as "exercise", yielded an unexpectedly high number of false positives upon manual inspection. Thus, future validation studies are required to improve these semantic location results.

Passive sensor collection, as carried out in this study, requires the user to agree to the collection of large amounts of personal data, and it is unclear whether patients outside a research context would be willing to share these data. A mitigation for this issue could be federated learning, a technique to train machine learning models locally on devices without having to pool them in a central storage location.

Despite these limitations, this study has an important place in the field with regard to the potential of low-burden collection of surveys as well as smartphone features in detecting changes in depression over time. This information could be used to improve the monitoring of treatment success or enhance the selection of treatments for each patient.

Acknowledgments

The authors would like to acknowledge Brian Anderson for help with understanding Android sensor architecture. The authors would also like to thank Ritu Kapur and Sneh Shah for their help in managing the digital phenotyping efforts at Verily.

Conflicts of Interest

At the time of investigation and preliminary analysis, SN, MDE, SFP, DW, JG, BR, DPM, M Fleck, AM, BP, YC, AH, DRS, AB, CC, CW, HH, LJM, WJM Jr, JLM, DAS, AJC, RMC, and M Fromer were employees of and owned equity in Verily Life Sciences. At the time of initial submission of this manuscript, SN, MDE, SFP, DW, DPM, YC, AH, AB, CC, LJM, WJM Jr, JLM, DAS, AJC, RMC, and M Fromer were employees of and owned equity in Verily Life Sciences.

Multimedia Appendix 1
Compensation cohort analyses.
[PDF File (Adobe PDF File), 236 KB - mental_v8i8e27589_app1.pdf]

Multimedia Appendix 2
Informed consent form.
[PDF File (Adobe PDF File), 151 KB - mental_v8i8e27589_app2.pdf]

Multimedia Appendix 3
User experience research and user-centered design.
[PDF File (Adobe PDF File), 1362 KB - mental_v8i8e27589_app3.pdf]

Multimedia Appendix 4
Study population characteristics.
[PDF File (Adobe PDF File), 56 KB - mental_v8i8e27589_app4.pdf]

Multimedia Appendix 5

Phone types.
[PDF File (Adobe PDF File), 86 KB - mental_v8i8e27589_app5.pdf]

Multimedia Appendix 6
Patient Health Questionnaire-9 figures.
[PDF File (Adobe PDF File), 321 KB - mental_v8i8e27589_app6.pdf]

Multimedia Appendix 7
Feature correlation matrix.
[PDF File (Adobe PDF File), 69 KB - mental_v8i8e27589_app7.pdf]

References

1. Murray CJL, Atkinson C, Bhalla K, Birbeck G, Burstein R, Chou D, Foreman, Lopez, Murray, Dahodwala, Jarlais, Fahami, Murray, Jarlais, Foreman, Lopez, Murray, U.S. Burden of Disease Collaborators. The state of US health, 1990-2010: burden of diseases, injuries, and risk factors. J Am Med Assoc 2013 Aug 14;310(6):591-608. [doi: 10.1001/jama.2013.13805] [Medline: 23842577]
2. US Burden of Disease Collaborators, Mokdad AH, Ballestros K, Echko M, Glenn S, Olsen HE, et al. The State of US Health, 1990-2016: Burden of Diseases, injuries, and risk factors among US states. J Am Med Assoc 2018 Dec 10;319(14):1444-1472. [doi: 10.1001/jama.2018.0158] [Medline: 29634829]
3. Aboraya A, Nasrallah HA, Elswick DE, Ahmed E, Estephan N, Aboraya D, et al. Measurement-based care in psychiatry-past, present, and future. Innov Clin Neurosci 2018 Nov 01;15(11-12):13-26 [FREE Full text] [Medline: 30834167]
4. Scott K, Lewis CC. Using measurement-based care to enhance any treatment. Cogn Behav Pract 2015 Feb;22(1):49-59 [FREE Full text] [doi: 10.1016/j.cbpra.2014.01.010] [Medline: 27330267]
5. Guo T, Xiang Y, Xiao L, Hu C, Chiu HF, Ungvari GS, et al. Measurement-based care versus standard care for major depression: a randomized controlled trial with blind raters. Am J Psychiatry 2015 Oct;172(10):1004-1013. [doi: 10.1176/appi.ajp.2015.14050652] [Medline: 26315978]
6. Fortney JC, Unützer J, Wrenn G, Pyne JM, Smith GR, Schoenbaum M, et al. A tipping point for measurement-based care. Psychiatr Serv 2017 Feb 01;68(2):179-188. [doi: 10.1176/appi.ps.201500439] [Medline: 27582237]
7. Zimmerman M, McGlinchey JB. Why don't psychiatrists use scales to measure outcome when treating depressed patients? J Clin Psychiatry 2008 Dec;69(12):1916-1919. [doi: 10.4088/jcp.v69n1209] [Medline: 19192467]
8. Insel TR. Digital phenotyping: technology for a new science of behavior. J Am Med Assoc 2017 Oct 03;318(13):1215-1216. [doi: 10.1001/jama.2017.11295] [Medline: 28973224]
9. Hsin H, Fromer M, Peterson B, Walter C, Fleck M, Campbell A, et al. Transforming psychiatry into data-driven medicine with digital measurement tools. NPJ Digit Med 2018;1:37 [FREE Full text] [doi: 10.1038/s41746-018-0046-0] [Medline: 31304319]
10. Van Wylich-Muxoll B. Monarca II Research Project: Monitoring and predicting illness activity in bipolar disorder. Monsenso. 2017. URL: https://www.monsenso.com/monarca-ii-research-project/ [accessed 2021-02-04]
11. Beiwe Research Platform. URL: https://www.beiwe.org/ [accessed 2021-02-04]
12. Mindstrong Health. URL: https://mindstrong.com/science/ [accessed 2021-02-04]
13. Wang R, Chen F, Chen Z, Li T, Harari G, Tignor S, et al. StudentLife: assessing mental health, academic performance and behavioral trends of college students using smartphones. In: UbiComp '14: Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing. NY, United States: Association for Computing Machinery; 2014 Sep 13 Presented at: UbiComp '14: Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing; September 13-17, 2014; Seattle, Washington p. 3-14. [doi: 10.1145/2632048.2632054]
14. Ginger.io behavioral health study. ClinicalTrials.gov. URL: https://clinicaltrials.gov/ct2/show/NCT02491307 [accessed 2021-02-04]
15. Rohani DA, Faurholt-Jepsen M, Kessing LV, Bardram JE. Correlations between objective behavioral features collected from mobile and wearable devices and depressive mood symptoms in patients with affective disorders: systematic review. JMIR Mhealth Uhealth 2018 Aug 13;6(8):e165 [FREE Full text] [doi: 10.2196/mhealth.9691] [Medline: 30104184]
16. Pratap A, Atkins DC, Renn BN, Tanana MJ, Mooney SD, Anguera JA, et al. The accuracy of passive phone sensors in predicting daily mood. Depress Anxiety 2019 Jan;36(1):72-81. [doi: 10.1002/da.22822] [Medline: 30129691]
17. Huebner ES. Initial development of the student's life satisfaction scale. Sch Psychol Int 2016 Jun 29;12(3):231-240. [doi: 10.1177/0143034391123010]
18. HealthMeasures. URL: https://www.healthmeasures.net/index.php?option=com_instruments&view=measure&id=660 [accessed 2021-01-21]
19. Anguera JA, Jordan JT, Castaneda D, Gazzaley A, Areán PA. Conducting a fully mobile and randomised clinical trial for depression: access, engagement and expense. BMJ Innov 2016 Jan;2(1):14-21 [FREE Full text] [doi: 10.1136/bmjinnov-2015-000098] [Medline: 27019745]

20. Wahle F, Kowatsch T, Fleisch E, Rufer M, Weidt S. Mobile sensing and support for people with depression: a pilot trial in the wild. JMIR Mhealth Uhealth 2016 Sep 21;4(3):e111 [FREE Full text] [doi: 10.2196/mhealth.5960] [Medline: 27655245]
21. Interpreting sentiment analysis values. Natural Language API Basics. URL: https://cloud.google.com/natural-language/docs/basics#interpreting_sentiment_analysis_values [accessed 2021-02-04]
22. Saeb S, Zhang M, Karr CJ, Schueller SM, Corden ME, Kording KP, et al. Mobile phone sensor correlates of depressive symptom severity in daily-life behavior: an exploratory study. J Med Internet Res 2015;17(7):e175 [FREE Full text] [doi: 10.2196/jmir.4273] [Medline: 26180009]
23. Mundt JC, Snyder PJ, Cannizzaro MS, Chappie K, Geralts DS. Voice acoustic measures of depression severity and treatment response collected via interactive voice response (IVR) technology. J Neurolinguistics 2007 Jan;20(1):50-64 [FREE Full text] [doi: 10.1016/j.jneuroling.2006.04.001] [Medline: 21253440]
24. Palmer C, Pe'er I. Statistical correction of the Winner's Curse explains replication variability in quantitative trait genome-wide association studies. PLoS Genet 2017 Jul;13(7):e1006916 [FREE Full text] [doi: 10.1371/journal.pgen.1006916] [Medline: 28715421]
25. Delgadillo J, de Jong K, Lucock M, Lutz W, Rubel J, Gilbody S, et al. Feedback-informed treatment versus usual psychological treatment for depression and anxiety: a multisite, open-label, cluster randomised controlled trial. Lancet Psychiatry 2018 Jul;5(7):564-572. [doi: 10.1016/S2215-0366(18)30162-7] [Medline: 29937396]
26. What your therapist doesn't know. The Atlantic. URL: https://www.theatlantic.com/magazine/archive/2017/04/what-your-therapist-doesnt-know/517797/ [accessed 2020-12-23]
27. Luciano A, Meara E. Employment status of people with mental illness: national survey data from 2009 and 2010. Psychiatr Serv 2014 Oct;65(10):1201-1209 [FREE Full text] [doi: 10.1176/appi.ps.201300335] [Medline: 24933361]

Abbreviations

AUC: area under the curve
DALY: disability-adjusted life year
DBSCAN: density-based spatial clustering of applications with noise
ICF: informed consent form
PHQ: Patient Health Questionnaire
YLD: year lived with disability

Edited by J Torous; submitted 05.03.21; peer-reviewed by J Melcher, J Knights; comments to author 22.03.21; revised version received 16.04.21; accepted 29.04.21; published 10.08.21.

Please cite as:
Nickels S, Edwards MD, Poole SF, Winter D, Gronsbell J, Rozenkrants B, Miller DP, Fleck M, McLean A, Peterson B, Chen Y, Hwang A, Rust-Smith D, Brant A, Campbell A, Chen C, Walter C, Arean PA, Hsin H, Myers LJ, Marks Jr WJ, Mega JL, Schlosser DA, Conrad AJ, Califf RM, Fromer M
Toward a Mobile Platform for Real-world Digital Measurement of Depression: User-Centered Design, Data Quality, and Behavioral and Clinical Modeling
JMIR Ment Health 2021;8(8):e27589
URL: https://mental.jmir.org/2021/8/e27589
doi:10.2196/27589
PMID:34383685

Review

Telemental Health For Youth With Chronic Illnesses: Systematic Review

Nancy Lau[1,2,3], PhD; Susannah F Colt[4], MA; Shayna Waldbaum[5], BSc; Alison O'Daffer[1], BA; Kaitlyn Fladeboe[1,3,6], PhD; Joyce P Yi-Frazier[1], PhD; Elizabeth McCauley[2,7], PhD; Abby R Rosenberg[1,3,6], MA, MSc, MD

[1]Palliative Care and Resilience Lab, Center for Clinical and Translational Research, Seattle Children's Research Institute, Seattle, WA, United States
[2]Department of Psychiatry and Behavioral Sciences, University of Washington School of Medicine, Seattle, WA, United States
[3]Cambia Palliative Care Center of Excellence, University of Washington, Seattle, WA, United States
[4]Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, United States
[5]Chicago Medical School, Rosalind Franklin University of Medicine and Science, North Chicago, IL, United States
[6]Department of Pediatrics, University of Washington School of Medicine, Seattle, WA, United States
[7]Center for Child Health, Behavior, and Development, Seattle Children's Research Institute, Seattle, WA, United States

Corresponding Author:
Nancy Lau, PhD
Palliative Care and Resilience Lab
Center for Clinical and Translational Research
Seattle Children's Research Institute
1920 Terry Ave
Seattle, WA, 98101
United States
Phone: 1 206 884 0569
Email: nancy.lau@seattlechildrens.org

Abstract

Background: Children, adolescents, and young adults with chronic conditions experience difficulties coping with disease-related stressors, comorbid mental health problems, and decreased quality of life. The COVID-19 pandemic has led to a global mental health crisis, and telemental health has necessarily displaced in-person care. However, it remains unknown whether such remote interventions are feasible or efficacious. We aimed to fill this research-practice gap.

Objective: In this systematic review, we present a synthesis of studies examining the feasibility and efficacy of telemental health interventions for youth aged ≤25 years with chronic illnesses.

Methods: PubMed, Embase, Web of Science, PsycInfo, and Cochrane Database of Systematic Reviews were searched from 2008 to 2020. We included experimental, quasiexperimental, and observational studies of telemental health interventions designed for children, adolescents, and young adults aged ≤25 years with chronic illnesses, in which feasibility or efficacy outcomes were measured. Only English-language publications in peer-reviewed journals were included. We excluded studies of interventions for caregivers or health care providers, mental health problems not in the context of a chronic illness, disease and medication management, and prevention programs for healthy individuals.

Results: We screened 2154 unique study records and 109 relevant full-text articles. Twelve studies met the inclusion criteria, and they represented seven unique telemental health interventions. Five of the studies included feasibility outcomes and seven included efficacy outcomes. All but two studies were pilot studies with relatively small sample sizes. Most interventions were based on cognitive behavioral therapy and problem-solving therapy. The subset of studies examining intervention feasibility concluded that telemental health interventions were appropriate, acceptable, and satisfactory to patients and their parents. Technology did not create barriers in access to care. For the subset of efficacy studies, evidence in support of the efficacy of telemental health was mixed. Significant heterogeneity in treatment type, medical diagnoses, and outcomes precluded a meta-analysis.

Conclusions: The state of the science for telemental health interventions designed for youth with chronic illnesses is in a nascent stage. Early evidence supports the feasibility of telehealth-based delivery of traditional in-person interventions. Few studies have assessed efficacy, and current findings are mixed. Future research should continue to evaluate whether telemental health may serve as a sustainable alternative to in-person care after the COVID pandemic.

(JMIR Ment Health 2021;8(8):e30098) doi:10.2196/30098

KEYWORDS

telehealth care; mental health; psychosocial issues; psychiatry; psychology; child; chronic disease

Introduction

Children, adolescents, and young adults with chronic medical conditions experience difficulties coping with disease-related stressors and decreased quality of life [1-3]. There is a strong association between physical and mental health in children, adolescents, and young adults, such that up to 60% of those with chronic illnesses are diagnosed with comorbid mental health disorders [4,5]. Common challenges include navigating diagnosis- and treatment-related distress, disruptions to normative development, changing family and peer relationships, and worries and uncertainty about the future [4-6]. The global COVID-19 pandemic has compounded these challenges and led to an increased risk of mental health symptoms, such as anxiety, depression, substance abuse, and posttraumatic stress, in healthy populations and worsening symptoms in those with pre-existing mental health disorders [7]. Further exacerbating negative impacts on psychological health, school closures have resulted in a lack of access to nonacademic support services (eg, sports and extracurricular programs and school mental health counselors), adjustment problems, and social isolation [8-10].

Overburdened health care systems have experienced a corresponding increase in demand for mental health services, and over 200 affected countries have inadequate resources to meet this influx [11,12]. Since the early stages of the pandemic, social distancing precautions have necessitated a shift in the mental health treatment standard of care from an in-person mode of delivery to telemental health (ie, via secure web-based videoconferencing platforms) [7]. Payor policy changes have been implemented to provide much needed care while ensuring the physical safety of patients and providers alike [13]. Projections suggest that US $250 billion of health care spending in the United States could become telehealth-based after the COVID-19 pandemic [14].

The rapid adoption of telehealth in health care systems and insurance program coverage has helped ensure continuity in mental health care and availability of psychosocial services in response to escalating needs during the pandemic. A number of systematic reviews in adult populations have shown that videoconferencing psychotherapies are feasible and have comparable outcomes to in-person treatment, and for anxiety and depression in particular [15-17]. There is limited information regarding the feasibility and efficacy of telemental health services for children, adolescents, and young adults with chronic illnesses, but early findings are similarly promising to adult interventions. Existing pediatric telemental health research has been limited to case studies, single-site child psychiatry department implementation efforts, and reviews pertaining to the treatment of youth with mental health concerns not in the context of medical conditions [18-22].

In this systematic review, we aimed to answer the following research questions:

1. What is the evidence for the feasibility of telemental health in child, adolescent, and young adult chronic illness populations?
2. What is the evidence for the efficacy of telemental health in child, adolescent, and young adult chronic illness populations?
3. What types of in-person interventions and intervention components have been adapted to telemental health delivery?

Methods

Data Sources

An electronic database search was executed by a research librarian in five databases (PubMed/MEDLINE, Embase, Web of Science, PsycInfo, and Cochrane Database of Systematic Reviews) on May 29, 2020, for publications from 2008 to 2020. The list of keyword parameters was based on controlled vocabulary terms prespecified by each database. The search terms included those relevant to age demographics, telemedicine and telehealth, and chronic disease, utilizing the following Boolean [23] operators: (p?ediatric* OR child* OR youth* OR teen* OR preteen* OR preadolescent* OR adolescent* OR young adult*) AND (telehealth OR eHealth OR telemedicine OR online OR telecounsel* OR teletherapy*) AND (chronic* OR condition* OR disease* OR ill* OR sick* OR syndrom* OR chronic condition OR chronic disease OR chronic illness OR long term condition OR long term disease OR noncommunicable disease OR noncommunicable condition).

Study Selection

The inclusion criteria were as follows: (1) availability in English; (2) study published in a peer-reviewed journal; (3) experimental, quasiexperimental, or observational study in which feasibility and/or efficacy outcomes were reported; (4) telemental health interventions delivered via videoconferencing platforms; and (5) interventions designed for children, adolescents, or young adults aged ≤25 years with a chronic disease (ie, a long-term medical condition lasting 3 months or longer [24]). We excluded studies of interventions that targeted caregivers or health care providers only, interventions that targeted mental health problems not in the context of a chronic illness, prevention programs for healthy individuals, and programs that targeted disease and medication management. In addition, we excluded nonpeer-reviewed publications (eg, dissertation manuscripts and conference abstracts) and study protocols for which no outcomes were reported.

First, we screened the titles and abstracts of studies retrieved for inclusion and exclusion. We then obtained the full texts of articles designated as potentially meeting the inclusion criteria to assess eligibility. Screening of all titles, abstracts, and full-text articles was first conducted by two independent coders (NL and SW). Then, disagreements between the authors were discussed while referencing the original source material to reach consensus. For screening of titles and abstracts, interrater agreement between independent coders (NL and SW) was very good, reflected by a Cohen kappa of 0.84. For screening of full-text articles, interrater agreement between independent coders (NL and SW) was good, reflected by a Cohen kappa of 0.79. For articles meeting the inclusion criteria, we independently double coded relevant information from each study in pairs from a group of three (NL, SFC, and KF).

Data Extraction and Synthesis

Data were extracted using a shared Excel template. Relevant information from each study included study design, sample size, target illness, participant age range, type of control group (where applicable), intervention name, intervention type, facilitator credentials, parental involvement, 1:1 or group-based format, homework assignments, technological components, and results. Both significant and nonsignificant outcomes were reported; we included information on *P* values and effect sizes if reported in the original study publication. After review of the articles meeting the inclusion criteria (n=12), the team determined that heterogeneity in intervention type, study design, outcome variables, and measurement timepoints precluded a meta-analysis. Thus, we described the data in a narrative synthesis.

Quality Assessment

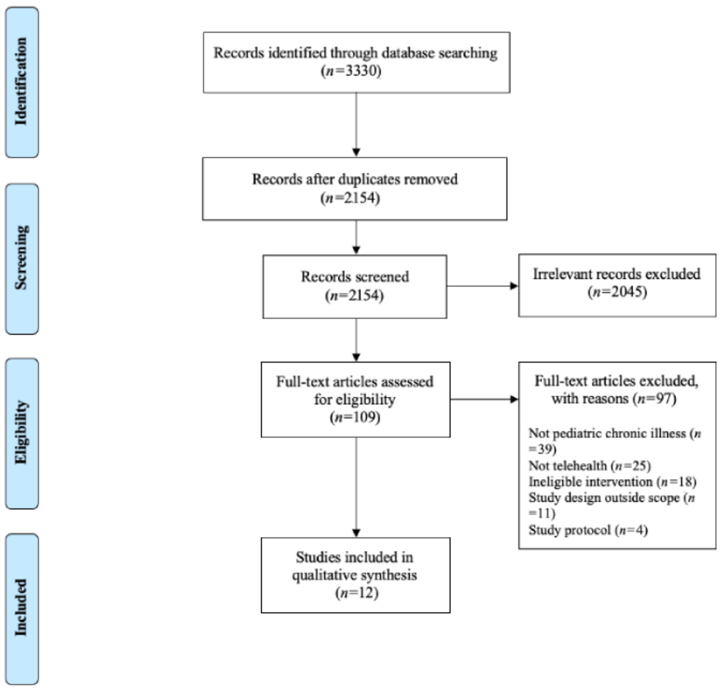
For the subset of included records that were efficacy studies, study quality was assessed by two independent coders from a group of three (NL, SFC, and KF) using the Cochrane Collaboration tool for assessing risk of bias [25]. The Cochrane tool evaluates the following seven evidence-based categories: random sequence generation (selection bias), allocation concealment (selection bias), blinding of participants and personnel (performance bias), blinding of outcome assessment (detection bias), incomplete outcome data (attrition bias), selective reporting (reporting bias), and other bias. We coded each category as low, high, or unclear risk of bias according to established standards in the *Cochrane Handbook for Systematic Reviews of Interventions* [26]. Support for judgment was directly quoted from the articles or published study protocols (where available) as relevant source materials. We utilized the well-established Grades of Recommendation, Assessment, Development, and Evaluation (GRADE) approach to evaluate the quality of evidence [27]. According to this approach, randomized trials start as high quality and are downgraded for limitations such as lack of allocation concealment; lack of blinding; attrition bias due to amount, nature, or handling of incomplete outcome data; and reporting bias. We coded a category as unclear if not enough information was available in the article to make a judgment. We resolved discrepancies in coding during regularly scheduled consensus meetings by referring back to the journal articles themselves.

Results

Overview

The search initially identified a total of 3330 articles (PubMed, 1410; Embase, 838; Web of Science, 390; PsycInfo, 607; and Cochrane Database of Systematic Reviews, 85). There were 1176 duplicates removed. This resulted in 2154 unique records screened, and 109 full-text articles were designated as potentially meeting the inclusion criteria. Review of the full-text articles resulted in 12 articles that met the criteria for inclusion. The results of the search and selection of studies are described in the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram (Figure 1).

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



Intervention Characteristics

We found seven unique telemental health interventions that were developed and tested in Australia, Canada, and the United States (Tables 1 and 2). Five (71%) were 1:1 interventions and the rest were group-based interventions. For group-based interventions, the number of facilitators ranged from 1 to 2 and the number of patients ranged from 3 to 9 per group. Four (57%) of the interventions were delivered by psychologists; other facilitators included study personnel with training in the tested intervention (n=2) and therapists at a master's level (n=1).

Table 1. Videoconferencing interventions targeted for youth chronic illness populations.

Source, year	Intervention name	Target illness	Age range, years	Country of origin	Underlying intervention theory	Facilitator credentials	Group based	Facilitator-to-patient ratio
Wade et al, 2020 [28] Moscato et al, 2019 [29]	A Survivor's Journey[a]	AYA[b] pediatric brain tumor survivors	13-24[c]	United States	Cognitive-behavioral problem-solving therapy	Licensed clinical psychologist	No	1:1
Wade et al, 2014 [30] Wade et al, 2015 [31]	Counselor Assisted Problem-Solving (CAPS) intervention	Adolescents with traumatic brain injury	12-17	United States	Family-based problem-solving therapy	Licensed clinical psychologist	Yes	1:1
Chadi et al, 2018 [32] Chadi et al, 2019 [33]	Mindful Awareness and Resilience Skills for Adolescents (MARS-A) Program[d]	Chronic medical or mental health illness	13-18	Canada	Evidence-based mindfulness	Experienced personnel with MARS-A specific training	Yes	2:9
Mcgill et al, 2017 [34] Sansom-Daly et al, 2019 [35]	Recapture Life	AYA cancer survivors	15-25	Australia	Cognitive behavioral therapy	Psychologist	Yes	1:3-5
Wade et al, 2018 [36]	Social Participation and Navigation (SPAN) Program	Adolescents with acquired brain injury	14-22	United States	Social skills, problem-solving therapy, goal setting	Trained undergraduate college student coach	No	1:1
Wade et al, 2011 [37]	Teen Online Problem-Solving (TOPS) Intervention	Adolescents with traumatic brain injury	11-18	United States	Family-based problem-solving therapy	Licensed clinical psychologist and clinical psychology PhD students	No	1:1
Ricketts et al, 2016 [38]	VoIP-delivered CBIT (CBIT-VoIP)	Youth with chronic tic disorders	8-16	United States	Comprehensive behavioral intervention for tics (CBIT)	CBIT trained therapist, master's level	No	1:1

[a]A Survivor's Journey information represents articles by Moscato et al [29] and Wade et al [28], which use the same study to report on different outcomes.

[b]AYA: adolescent and young adult.

[c]This is the age range of the study sample used in analysis. The inclusion criteria specified an age range of 13-25.

[d]MARS-A program information represents articles by Chadi et al [32,33], which are two studies of the same intervention.

Table 2. Technology, intervention, and group components.

Source, year	Intervention name	Technology components			Intervention components		
		Video platform used	Provided device	Tech support	Parents involved	Homework assignments	Self-guided
Wade et al, 2020 [28] Moscato et al, 2019 [29]	A Survivor's Journey[a]	Skype	Yes[b]	Yes[c]	No	Yes	Yes
Wade et al, 2014 [30] Wade et al, 2015 [31]	CAPS intervention	Skype	Yes	Yes[c]	Yes	Yes	Yes
Chadi et al, 2018 [32] Chadi et al, 2019 [33]	Mindful Awareness and Resilience Skills for Adolescents (MARS-A) Program[d]	Zoom	No	Yes	No	Yes	No
McGill et al, 2017 [34] Sansom-Daly et al, 2019 [35]	Recapture Life	WebEx	Yes[b]	No	Yes	Yes	No
Wade et al, 2018 [36]	Social Participation and Navigation (SPAN) Program	Skype	No	No	No	Yes	Yes
Wade et al, 2011 [37]	Teen Online Problem Solving (TOPS) intervention	Not specified	Yes[b,e]	Yes[c]	Yes	Yes	Yes
Ricketts et al, 2016 [38]	VoIP-delivered CBIT (CBIT-VoIP)	Skype	Yes[b]	Yes	Yes	Yes	No

[a]A Survivor's Journey information represents articles by Moscato et al [29] and Wade et al [28], which report on different outcomes from the same parent study.

[b]Provided to those who did not have one of their own.

[c]Tutorial provided before the start of the intervention; no ongoing tech support.

[d]MARS-A program information represents articles by Chadi et al [32,33], which are two studies of the same intervention.

[e]High speed internet access provided to everyone.

All psychosocial interventions were adapted specifically for chronic medical conditions and designed to teach coping skills to facilitate adjustment to illness via treatment manuals. Interventions were primarily based on evidence-based cognitive behavioral therapy and problem-solving therapy. The specific coping skills targeted included cognitive restructuring, mindfulness, behavioral activation, social participation, goal setting, and problem solving to facilitate adjustment to illness. Manualized intervention content (ie, standardized treatment manuals documenting session-by-session objectives and procedures to ensure intervention fidelity across facilitators) provided a systematic approach for developing adaptive coping strategies and is the standard of practice for empirically supported psychotherapies. All interventions assigned homework to facilitate skills practice. Four (57%) interventions contained self-guided online modules with disease-relevant resources in addition to regularly scheduled videoconference-based therapy sessions. For the subset of interventions with web content, participants met with a facilitator for weekly telemental health sessions. Online self-guided modules consisted of interactive didactic content, videos of patients discussing use of coping skills, and animated videos providing examples of how to directly apply coping skills to day-to-day life. Parents were active participants in family-based videoconferencing therapy sessions for 4 (57%) of the interventions.

The video platforms used were Skype (n=4), Cisco WebEx (n=1), Zoom (n=1), and unspecified (n=1). Five (71%) of the interventions provided devices to connect to video platforms for participants who needed them, and one intervention provided both devices and high-speed internet access for participants. Three (43%) of the interventions provided an introductory tutorial on how to access and use the video platform program, and 2 (29%) provided ongoing technological support.

Participants

Across all included studies, study sample sizes ranged from 14 to 132. The age range of participants was from 8 to 25 years. The targeted chronic illnesses included brain tumor, cancer, traumatic brain injury, chronic tic disorder, and chronic illness (nondisease-specific intervention). The key study characteristics are summarized in Tables 3 and 4.

Table 3. Original research publications reporting on feasibility outcomes.

Source, year	Intervention name	Study type	Sample size[a]	Age range, years	How constructs were defined and measured	Results
Moscato et al, 2019 [29]	A Survivor's Journey	Pilot feasibility study[b]	17	13-24	Feasibility: Enrollment and completion rates Acceptability: Internally developed satisfaction survey of the Teen Online Problem Solving (TOPS) intervention on a 4-point Likert scale. Overall satisfaction ratings included whether the program met expectations, what content was most and least helpful, whether the website was easy to use, understand, and navigate. System Usability Scale (5-point Likert scale) was used to measure ease of use.	Feasibility: 50% enrollment rate (which met researchers' enrollment aim), and 95% completed core sessions, which exceeded the goal of a 75% completion rate. Acceptability: Exceeded the goal of 75% of participants reporting satisfaction on most items of the satisfaction survey (eg, every participant reported that they would recommend the program to others, website was easy to use and navigate, and content was relevant to them). Did not meet the goal of 75% of participants rating the intervention above average in usability on the System Usability Scale.
Chadi et al, 2018 [32]	Mindful Awareness and Resilience Skills for Adolescents (MARS-A) Program	Qualitative portion of randomized mixed methods trial	18	13-18	Acceptability/feasibility: Program exit interviews used to foster personal reflections about participants' experiences of the MARS-A program and qualitative analysis identified four themes from interview data.	Acceptability/feasibility: Themes describing experiences for both in-person and eHealth groups were as follows: creating a safe space; fostering peer support and connection; integrating mindfulness skills into daily life; and improving well-being through mindfulness. Based on qualitative results, they concluded that eHealth delivery of a mindfulness-based intervention may be an acceptable and feasible mode of delivery for adolescents with chronic illnesses.
McGill et al, 2017 [34]	Recapture Life	Qualitative analysis of three-arm pilot randomized controlled trial (RCT)	39[d]	15-25	Acceptability: Authors stated that therapeutic alliance (collaborative element of the patient-therapist dyad) and group cohesion (quality of interpersonal processes between group members and between group members and the therapist) are important to determine the acceptability of online models of psychosocial care.	Acceptability-group cohesion: All participants endorsed at least moderate group cohesion on all group cohesion items after the last session (that they shared important things, felt accepted and respected, and the program was the best way to get help, and it helped them gain a deeper understanding). Acceptability-therapeutic alliance: All participants endorsed strong therapeutic alliance on all therapeutic alliance items, and it remained high over time from the first to last session (understanding, confidence, appreciation, and working correctly). Therapists endorsed strong therapeutic alliance on all therapeutic alliance items (participant comfort, rapport, openness, trust, peer to peer, motivation, and engagement); endorsed that participants had significantly increased openness, trust, and motivation from the first to last session (*P*<.05); and endorsed that the items pertaining to participant comfort, rapport, peer-to-peer discussion, and engagement were strong and unchanged over time from the first to last session.

Source, year	Intervention name	Study type	Sample size[a]	Age range, years	How constructs were defined and measured	Results
Sansom-Daly et al, 2019 [35][c]	Recapture Life	Three-arm RCT reporting on feasibility and acceptability	45[d]	15-25	Feasibility: Recruitment rates across sites; mean days to group commencement; median time for session commencement; proportion of eligible, interested adolescents and young adults (AYAs) who had the technological equipment and internet access required to participate; number and type of technological difficulties experienced across sessions and perceived impact on content delivery; time taken to check participants' between-session emotional safety using email/text inquiries; total number of additional catch up sessions conducted for AYAs who missed their group session, rescheduled group sessions, and scheduled group sessions outside of office hours. Acceptability: Opt-in, enrollment, and retention rates, and participant engagement (total group sessions attended) and homework completion rates. Responded to two internally developed items "Was participation in this study beneficial to you in any way?" and "Was participation in this study burdensome for you in any way?" Qualitative analysis of open-ended questionnaire responses used to explore participants' experiences with the program.	Results: Recruitment rate of 30.41% (45 randomized/148 approached), and 80% of participants had access to all required technologies. Individuals waited on average for 40 days (range 5-107) from completing the baseline questionnaire to commencing an online group with a sufficient number of peers. Sessions took a median of 4 minutes to commence, and 74% of sessions had all participants log on within 5 minutes of the scheduled start time. Six "catch up" sessions were delivered for participants unable to attend the scheduled group. Overall, 10 of 12 groups required sessions to be scheduled out-of-hours, representing 60 online sessions across the trial (approximately 90 hours). Technological difficulties were common, being experienced at least once in 71% of sessions, and 38% experienced two or more technological difficulties, but difficulties were rated as having a low impact on intervention delivery. The most common technological difficulties were poor quality audio and dropouts (43% of sessions) and webcam freezing (43% of sessions). Of participants whose scores triggered a between-session telephone call for safety, all were telephoned within 48 hours. An average of 1.8 (range 1-4) email, text, and/or phone calls was required to confirm safety. Authors concluded that the findings support program feasibility. Acceptability: Opt-in rate of 30%. Enrollment rate of 87% of those who completed baseline, and completion rate of 92%. High level of engagement with majority of participants attending at least 74% (5/6) of sessions. AYAs reported a high benefit and low burden of participation on open-ended questionnaire responses. Participants reported a completed average of 51% of program homework. Authors concluded that the findings support program acceptability.
Wade et al, 2018 [36]	Social Participation and Navigation (SPAN) Program	Nonrandomized pilot trial[e]	15	14-22	Feasibility: Number of sessions completed, and number of social participation goals achieved during the intervention. Satisfaction: Internally developed measure to assess satisfaction with the program for participants and their parents	Feasibility: Participants completed an average of 80% of sessions (range 3-10) and achieved an average of three social participation goals (range 1-7). Authors concluded that the findings support program feasibility. Satisfaction: All participants "agreed" or "strongly agreed" that the program was useful, were glad to do the program, and would recommend the program to others. All parents "agreed" or "strongly agreed" that they were glad to do the program and would recommend it to others. Authors concluded that the findings support program satisfaction.

[a]Sample size used in the analyses.

[b]Provides feasibility outcomes for the parent study; see also the study by Wade et al [28] in Table 4.

[c]All technical problems were reported as quickly correctable without major impacts on overall sessions.

[d]Although both Recapture Life studies are based on the same parent study, a subset of 39 participants was represented in the qualitative study (McGill et al [34]) and all participants were represented in the pilot randomized controlled trial (Sansom-Daly et al [35]).

[e]Wade et al [36] is represented in Tables 3 and 4 due to reporting both feasibility and efficacy outcomes of interest.

Table 4. Original research publications reporting on efficacy outcomes.

Source, year	Intervention name	Study type	RCT[a]	Control group	Sample size[b]	Age range, years	Outcomes[c]
Wade et al, 2020 [28]	A Survivor's Journey	Pilot feasibility study	No	N/A	17	13-24	Improved self-reported overall (*d*=0.58, *P*=.01) and physical quality of life (*d*=0.55, *P*<.01) at posttreatment. Improved parent-reported emotional quality of life (*d*=0.43, *P*=.03) at posttreatment.
Wade et al, 2014 [30] Wade et al, 2015 [31]	Counselor Assisted Problem-Solving (CAPS) intervention	Original RCT RCT long-term follow-up	Yes	Internet resource	132	12-17	No differences between groups in self-reported internalizing or externalizing symptoms at posttreatment. Posttreatment between groups: gains not sustained for internalizing or externalizing symptoms at 12- and 18-month follow-ups. At the 18-month follow-up: In the CAPS group, internalizing problems improved for high school–age participants only (*P*=.03).
Chadi et al, 2019 [33]	Mindful Awareness and Resilience Skills for Adolescents (MARS-A) Program	Pilot RCT	Yes	In-person MARS-A	14	13-18	No differences between groups at posttreatment in self-reported anxiety and depression. Reduced pre-post anxiety and depression for the eHealth group (Cohen *d*=0.934, *P*=.048); improvements not sustained at a 2-month follow-up. Similar frequency and duration of the mindfulness practice between groups at posttreatment.
Wade et al, 2018 [36]	Social Participation and Navigation (SPAN) Program	Pilot trial	No	N/A[d]	15	14-22	Increase in parent-reported frequency of social participation (*d*=1.11, *P*=.01), but not for teens at posttreatment. Increase in teen-reported confidence in social participation (*d*=1.45, *P*<.01) but not for parents at posttreatment. Decline in parent-reported total problems (*d*=0.96, *P*<.01), internalizing problems (*d*=0.73, *P*=.05), externalizing problems (*d*=0.79, *P*=.02), and social problems (*d*=0.82, *P*=.02), but no differences for adolescent-reported problems at posttreatment. No significant differences in the levels of social competence and confidence in the ability to manage emotions reported by teens or parents at posttreatment.
Wade et al, 2011 [37]	Teen Online Problem-Solving (TOPS) Intervention	RCT	Yes	Internet resource	35	11-18	No differences between groups in adolescent-reported parent-teen conflict, and adolescent- and parent-reported internalizing and externalizing symptoms at posttreatment. TOPS reduced parent-reported conflict compared to controls (*P*=.002) at posttreatment.
Ricketts et al, 2016 [38]	VoIP-delivered CBIT (CBIT-VoIP)	Pilot RCT	Yes	Waitlist control	20	8-16	Improved the Yale Global Tic Severity Scale score relative to controls (*d*=0.90, *P*<.01; partial η^2=0.15, *P*<.05) at posttreatment. Higher response in the Clinical Global Impression-Improvement Scale in the treatment group (χ^2=0.33, *P*<.05, Φ=0.41) at posttreatment. Improved Parent Tic Questionnaire score (*d*=1.38, *P*<.001; and partial η^2=0.26, *P*<.05) at posttreatment.

[a]RCT: randomized controlled trial.

[b]Sample size used in study analyses.

[c]Statistically significant outcomes are reported with *P*<.05. *d* represents Cohen *d*. η^2 is eta-squared. The two statistics are measures of effect size. Effect sizes are reported if information was included in the original study publication.

[d]N/A: not applicable.

Feasibility Outcomes

Based on the seminal work of Bowen et al, there is no consensus regarding how feasibility is defined and measured [39]. Due to few published standards, guides, and thresholds upon which to test and establish feasibility of an intervention, study teams when designing feasibility studies create their own internal thresholds regarding what they determine to be important and appropriate for their specific intervention and target population. Across the five studies examining intervention feasibility included in our review, some authors used and operationalized acceptability and feasibility interchangeably, others used and operationalized acceptability and satisfaction interchangeably, and still others used and operationalized all three as distinct constructs. In Table 3, we provide detailed information on how constructs of feasibility, acceptability, and satisfaction were defined and measured by the original authors of the included studies, and their respective results. Although, as expected, we did not observe consensus across all studies in their internal thresholds, common benchmarks included enrollment, adherence, and completion rates established a priori in order to determine feasibility (ie, addressing the question of "can this be done?") and use of internally developed measures or qualitative exit interviews to measure program acceptability and satisfaction (eg, whether participants enjoyed the program, whether participants benefited from the program, whether the program met their goals, and whether they would recommend the program to others). All studies reported high feasibility, acceptability, and satisfaction of telemental health interventions based on their own a priori internally established thresholds [29,32,34-36].

Sample sizes across studies ranged from 15 to 45. Authors reported that enrollment and program completion rates met prespecified target goals, with >30% enrollment and >70% program completion rates across studies. Patients endorsed the high benefit and low burden of participation and satisfaction with the program, and mentioned they would recommend the program to others on Likert-scale measures (ie, "agreed" or "strongly agreed" with all item measures) and/or program exit interviews. In group-based telemental health interventions, patients endorsed experiencing a sense of support, trust, rapport, and connection with the facilitator and other group members on therapeutic alliance measures and/or qualitative interviews. For the Social Participation and Navigation (SPAN) program, although patients and parents reported high levels of satisfaction and enjoyment with participation, parents were more likely than their children to report that the program was useful [36]. Across all studies, technological difficulties were reported to have a low impact on intervention delivery and treatment satisfaction. However, of note, all studies examining feasibility were conducted with participants aged 13 years or older.

Efficacy Outcomes

There were seven efficacy studies with sample sizes ranging from 14 to 132; five had randomized designs [30,31,33,37,38] and two were prospective cohort studies [28,36] (Table 4). Of the randomized trials, three were compared to internet resource comparison groups [30,31,37], one to an in-person version of the videoconference-based intervention [33], and one to a waitlist control group [38]. Five (71%) of the efficacy studies were pilot studies with small sample sizes (ie, ≤35 participants) [28,29,33,36,38]. Four of the studies collected both patient- and parent-reported outcome measures [28,36-38]. In our review, there was a significant amount of heterogeneity among outcomes targeted by specific interventions and corresponding treatment effects. All efficacy studies had a primary focus on psychosocial/mental health outcomes. One of the prospective cohort studies also measured physical outcomes (ie, overall and physical quality of life) [28]. Outcome measures across the five randomized trial studies were variable (ie, anxiety, depression, internalizing problems, externalizing problems, behavioral symptoms, and parent-teen conflict) [30,31,33,37,38]. Treatment outcome assessment timeframes were generally assessed immediately after treatment [28,30,33,36-38]. One trial reported on outcomes at a 2-month follow-up [33], and another trial reported on outcomes at a 12-month/18-month follow-up [31].

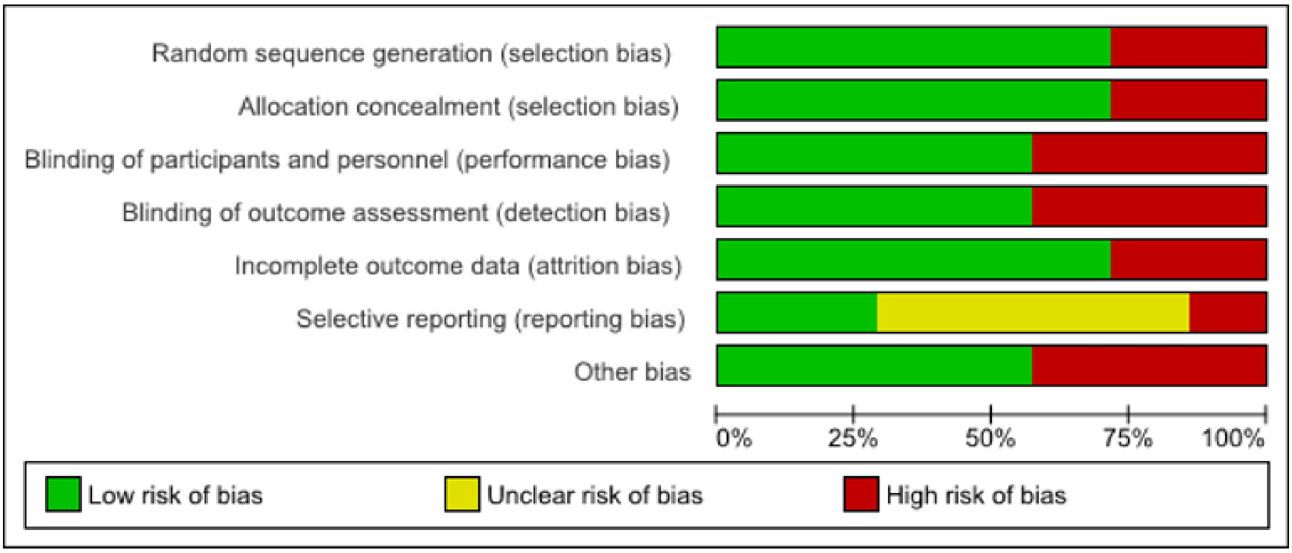
For the two prospective cohort studies, youth and parents reported some improvements in patient mental health well-being and functioning at posttreatment [28,36]. For the first study of a cognitive behavioral problem-solving intervention (sample size=17), patients reported significant improvements in emotional, physical, and overall quality of life (medium to large effects), whereas parents reported significant improvements in patient emotional quality of life only (medium effect) [28]. For the second study of a social skills and problem-solving intervention (sample size=15), patients reported significant improvements in confidence in social participation only (large effect), whereas parents reported significant improvements in patient frequency of social participation, internalizing problems, externalizing problems, social problems, and total problems (large effects) [36].

Of the randomized trial studies, four compared the intervention of interest with an active comparison group [30,31,33,37] and one with a waitlist control [38]. In the waitlist control trial (sample size=20), a cognitive behavioral intervention was associated with significant improvements in patient- and parent-reported behavioral symptoms (large effect) [38]. In a second trial testing a problem-solving intervention compared to an active internet resource comparison group (sample size=35), there were no significant group differences in patient-reported parent-teen conflict or patient- and parent-reported internalizing and externalizing symptoms; the problem-solving intervention demonstrated significant improvements in parent-reported parent-teen conflict only [37]. In the only trial to utilize a comparison group of face-to-face delivery of the same mindfulness intervention (sample size=14), both modes of delivery resulted in improvements in patient-reported posttreatment anxiety and depression with no significant differences between groups; improvements were not sustained at a 2-month follow-up [33]. Two studies of a problem-solving intervention compared to an internet resource comparison group (sample size=132) found no differences between groups in patient-reported internalizing and externalizing symptoms at posttreatment [30] and longer-term follow-up (12-month and 18-month follow-ups) [31].

Risk of Bias

Risk of bias was evaluated for all efficacy studies (Figure 2). Of the seven studies, five reported random sequence generation and allocation concealment (ie, the randomized trials). For the blinding of participants and personnel, and blinding of outcome assessment domains, four were low risk and three were high risk; the high-risk studies consisted of study designs with no control group or a waitlist control group. For attrition bias, five were low risk and two were high risk. For selective reporting bias, two were low risk, one was high risk, and four were unclear; studies were rated as unclear due to a lack of clinical trial registration or a published protocol. For other bias, four were considered low risk and three were high risk.

Figure 2. Risk of bias.



Discussion

Principal Findings

Chronic illnesses are commonly associated with comorbid mental health problems/disorders, including anxiety, stress, depression, maladjustment, and poor coping skills [2,4]. Further exacerbating baseline disease-related stressors, the COVID-19 pandemic has perpetuated a global mental health crisis and a corresponding increase in the demand for services [11]. Telemental health has become the standard of care since March of 2020, and may provide a cost-effective, scalable, and sustainable means of remote health care delivery [40]. This is the first systematic review to summarize the research evidence in support of the feasibility and efficacy of telemental health interventions for children, adolescents, and young adults with chronic illnesses.

In this article, we identified 12 studies on telemental health interventions. The interventions focused on evidence-based treatment strategies, including mindfulness, cognitive behavioral, and problem-solving strategies, across a broad range of target illnesses and psychosocial outcomes. Five (42%) of the studies included feasibility outcomes and 7 (58%) included efficacy outcomes. All but two studies were pilot studies with relatively small sample sizes. Across the small number of identified studies, telemental health interventions seemed to be appealing and acceptable for patients and parents alike. Although navigating videoconferencing platforms did not present technological barriers to treatment attendance or engagement, it is important to note that such findings were reported for studies with teenagers and young adults only who are understandably more tech savvy than younger cohorts. Single-cohort or waitlist control pilot studies examining the efficacy of telemental health interventions showed early promise and medium to large treatment effects. In randomized trials with active comparison conditions, there was little evidence of significant treatment effects across a range of mental health symptoms, and few studies included long-term outcome assessments. Only one trial compared face-to-face and telemental health delivery of the same intervention; both modes of delivery were similarly efficacious and improvements were not sustained at longer-term follow-up.

Together, this set of preliminary studies examining feasibility and efficacy provides early evidence that (1) telemental health interventions may be appropriate and acceptable to patients and their parents; (2) videoconferencing platforms may not present technological barriers to engagement and use; and (3) there is some modest early evidence in support of the efficacy of telemental health interventions, but preliminary findings are mixed. Our findings are consistent with previous reviews suggesting that telehealth may be appropriate and efficacious for adults with chronic conditions and for the delivery of mental health care [41,42].

Limitations

Some limitations need to be considered. First, relatively few papers have been published on telemental health feasibility and efficacy among children, adolescents, and young adults with chronic illnesses. This suggests that the science lags behind its rapid rate of adoption in clinical settings. Second, most papers were pilot studies with small sample sizes and were underpowered to detect clinically or statistically significant treatment effects. Third, given the heterogeneity of treatment type, disease target, measurement timepoints, and mental health outcome measures, we were unable to perform a meta-analysis to quantify treatment effects. Fourth, approximately one-third of studies did not report on the racial or ethnic distribution of their sample, and those that did reported that the majority of participants were white. Similarly, none of the studies reported on the rurality of their sample. Fifth, we limited our study to English-language publications. Thus, the generalizability of the findings remains unclear. Sixth, it was not possible to examine age and developmental differences in treatment effects due to the small number of included studies and specific interventions developed and tested with a wide age range of patients. This limitation is consistent with a previous review that found little research assessing age and developmental patterns in coping with chronic illnesses [6]. Finally, our study was limited to peer-reviewed published articles and did not consider unpublished gray literature, such as conference abstracts and dissertations, which may have led to the identification of additional studies.

Future Directions

Future research should extend beyond feasibility and early efficacy pilot studies to assess how telemental health delivery compares to in-person care in therapeutic alliance and rapport building, treatment engagement and treatment drop-out rates, and efficacy/effectiveness in large-scale randomized trials. Although some preliminary evidence suggests positive effects associated with telehealth delivery, it is important to examine the relative benefits and costs associated with remote interventions. Telehealth may not be an appropriate delivery format for all patients. Some patients may respond better to or prefer in-person care, experience higher homework compliance, have lower dropout rates, and/or establish a stronger therapeutic alliance with a provider in a face-to-face meeting. Important future research directions include the development of the best screening processes to match patient characteristics to care delivery preferences in order to optimize outcomes. Patient characteristics with the potential to impact care delivery that warrant further exploration include age and developmental stage, acuity of mental health needs, chronic illness diagnosis and illness narrative, and medical treatment. Evidence-based strategies proven to be effective when deployed through traditional in-person care may require iterative intervention adaptations to successfully translate treatment effects to telehealth modes of delivery. It may not be as straightforward as simply delivering the same manualized protocols via Skype, Zoom, or WebEx, and the optimal balance between traditional face-to-face care and remote delivery should be further examined. Equity in access to telemental health should also be examined, as some research teams supplied videoconferencing capable devices and high-speed internet access to participants who needed them, and such an approach may not be scalable as the standard of care. Ultimately, this study and future studies will help inform whether, for whom, and under what treatment conditions telemental health has the potential to serve as a sustainable long-term alternative to in-person care after the COVID-19 pandemic.

Conclusions

The strengths of this paper include the systematic approach to synthesizing the breadth of literature across chronic illness populations and the timely focus on telehealth, which has displaced in-person treatment as the standard of care in the context of the COVID-19 pandemic. Our findings suggest that although COVID-19 has necessitated remote treatment delivery, and patients and families may find this mode of delivery to be engaging and satisfactory, the state of the science is in a nascent stage and there is much to be learned about whether such interventions work, for whom they work, and in what contexts they work, as well as how they compare to in-person treatments.

Acknowledgments

NL is funded as an Implementation Science Scholar through the National Heart, Lung, and Blood Institute of the National Institutes of Health (grant number: 5K12 HL137940-02). The opinions herein represent those of the authors and not necessarily those of the funders.

Conflicts of Interest

None declared.

References

1. Warschburger P, Hänig J, Friedt M, Posovszky C, Schier M, Calvano C. Health-related quality of life in children with abdominal pain due to functional or organic gastrointestinal disorders. J Pediatr Psychol 2014;39(1):45-54. [doi: 10.1093/jpepsy/jst070] [Medline: 24055816]
2. Wallander JL, Varni JW. Effects of pediatric chronic physical disorders on child and family adjustment. J Child Psychol Psychiatry 1998 Jan;39(1):29-46. [Medline: 9534085]
3. Yeo M, Sawyer S. Chronic illness and disability. BMJ 2005 Mar 26;330(7493):721-723 [FREE Full text] [doi: 10.1136/bmj.330.7493.721] [Medline: 15790645]
4. Butler A, Van Lieshout RJ, Lipman EL, MacMillan HL, Gonzalez A, Gorter JW, et al. Mental disorder in children with physical conditions: a pilot study. BMJ Open 2018 Jan 03;8(1):e019011 [FREE Full text] [doi: 10.1136/bmjopen-2017-019011] [Medline: 29301763]
5. Barker MM, Beresford B, Bland M, Fraser LK. Prevalence and Incidence of Anxiety and Depression Among Children, Adolescents, and Young Adults With Life-Limiting Conditions: A Systematic Review and Meta-analysis. JAMA Pediatr 2019 Sep 01;173(9):835-844 [FREE Full text] [doi: 10.1001/jamapediatrics.2019.1712] [Medline: 31282938]
6. Compas BE, Jaser SS, Dunn MJ, Rodriguez EM. Coping with chronic illness in childhood and adolescence. Annu Rev Clin Psychol 2012 Apr 27;8(1):455-480 [FREE Full text] [doi: 10.1146/annurev-clinpsy-032511-143108] [Medline: 22224836]

7. Moreno C, Wykes T, Galderisi S, Nordentoft M, Crossley N, Jones N, et al. How mental health care should change as a consequence of the COVID-19 pandemic. Lancet Psychiatry 2020 Sep;7(9):813-824 [FREE Full text] [doi: 10.1016/S2215-0366(20)30307-2] [Medline: 32682460]
8. Lee J. Mental health effects of school closures during COVID-19. Lancet Child Adolesc Health 2020 Jun;4(6):421 [FREE Full text] [doi: 10.1016/S2352-4642(20)30109-7] [Medline: 32302537]
9. Razai MS, Oakeshott P, Kankam H, Galea S, Stokes-Lampard H. Mitigating the psychological effects of social isolation during the covid-19 pandemic. BMJ 2020 May 21;369:m1904. [doi: 10.1136/bmj.m1904] [Medline: 32439691]
10. Hoffman JA, Miller EA. Addressing the Consequences of School Closure Due to COVID-19 on Children's Physical and Mental Well-Being. World Med Health Policy 2020 Aug 20;12(3):300-310 [FREE Full text] [doi: 10.1002/wmh3.365] [Medline: 32904951]
11. Anjum S, Ullah R, Rana MS, Khan HA, Memon FS, Ahmed Y, et al. COVID-19 Pandemic: A Serious Threat for Public Mental Health Globally. Psychiatr Danub 2020 Aug 12;32(2):245-250 [FREE Full text] [doi: 10.24869/psyd.2020.245] [Medline: 32796793]
12. Ćosić K, Popović S, Šarlija M, Kesedžić I. Impact of Human Disasters and COVID-19 Pandemic on Mental Health: Potential of Digital Psychiatry. Psychiatr Danub 2020 Apr 15;32(1):25-31 [FREE Full text] [doi: 10.24869/psyd.2020.25] [Medline: 32303026]
13. Ramtekkar U, Bridge JA, Thomas G, Butter E, Reese J, Logan E, et al. Pediatric Telebehavioral Health: A Transformational Shift in Care Delivery in the Era of COVID-19. JMIR Ment Health 2020 Sep 18;7(9):e20157 [FREE Full text] [doi: 10.2196/20157] [Medline: 32525485]
14. Bestsennyy O, Gilbert G, Harris A, Rost J. Telehealth: a quarter-trillion-dollar post-COVID-19 reality? McKinsey & Company. 2021. URL: https://www.mckinsey.com/industries/healthcare-systems-and-services/our-insights/telehealth-a-quarter-trillion-dollar-post-covid-19-reality [accessed 2021-02-01]
15. Backhaus A, Agha Z, Maglione ML, Repp A, Ross B, Zuest D, et al. Videoconferencing psychotherapy: a systematic review. Psychol Serv 2012 May;9(2):111-131. [doi: 10.1037/a0027924] [Medline: 22662727]
16. Rees CS, Maclaine E. A Systematic Review of Videoconference ‐ Delivered Psychological Treatment for Anxiety Disorders. Australian Psychologist 2020 Nov 12;50(4):259-264. [doi: 10.1111/ap.12122]
17. García-Lizana F, Muñoz-Mayorga I. Telemedicine for depression: a systematic review. Perspect Psychiatr Care 2010 Apr;46(2):119-126. [doi: 10.1111/j.1744-6163.2010.00247.x] [Medline: 20377799]
18. Sharma A, Sasser T, Schoenfelder Gonzalez E, Vander Stoep A, Myers K. Implementation of Home-Based Telemental Health in a Large Child Psychiatry Department During the COVID-19 Crisis. J Child Adolesc Psychopharmacol 2020 Sep 01;30(7):404-413. [doi: 10.1089/cap.2020.0062] [Medline: 32639849]
19. Duncan AB, Velasquez SE, Nelson E. Using videoconferencing to provide psychological services to rural children and adolescents: a review and case example. J Clin Child Adolesc Psychol 2014 Sep 30;43(1):115-127. [doi: 10.1080/15374416.2013.836452] [Medline: 24079653]
20. Gloff NE, LeNoue SR, Novins DK, Myers K. Telemental health for children and adolescents. Int Rev Psychiatry 2015 Nov 05;27(6):513-524. [doi: 10.3109/09540261.2015.1086322] [Medline: 26540584]
21. Nelson E, Patton S. Using Videoconferencing to Deliver Individual Therapy and Pediatric Psychology Interventions with Children and Adolescents. J Child Adolesc Psychopharmacol 2016 Apr;26(3):212-220 [FREE Full text] [doi: 10.1089/cap.2015.0021] [Medline: 26745607]
22. Pesämaa L, Ebeling H, Kuusimäki ML, Winblad I, Isohanni M, Moilanen I. Videoconferencing in child and adolescent telepsychiatry: a systematic review of the literature. J Telemed Telecare 2004 Jun 24;10(4):187-192. [doi: 10.1258/1357633041424458] [Medline: 15273027]
23. Rousseau GK, Jamieson BA, Rogers WA, Mead SE, Sit RA. Assessing the usability of on-line library systems. Behaviour & Information Technology 1998 Jan;17(5):274-281. [doi: 10.1080/014492998119346]
24. Bernell S, Howard SW. Use Your Words Carefully: What Is a Chronic Disease? Front Public Health 2016 Aug 02;4:159 [FREE Full text] [doi: 10.3389/fpubh.2016.00159] [Medline: 27532034]
25. Higgins JPT, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, Cochrane Bias Methods Group, Cochrane Statistical Methods Group. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ 2011 Oct 18;343(oct18 2):d5928-d5928 [FREE Full text] [doi: 10.1136/bmj.d5928] [Medline: 22008217]
26. Higgins J, Thomas J, Chandler J, Cumpston M, Li T, Page M, et al. Cochrane Handbook for Systematic Reviews of Interventions. New York, NY: John Wiley & Sons; 2019.
27. Guyatt GH, Oxman AD, Vist G, Kunz R, Brozek J, Alonso-Coello P, et al. GRADE guidelines: 4. Rating the quality of evidence--study limitations (risk of bias). J Clin Epidemiol 2011 Apr;64(4):407-415. [doi: 10.1016/j.jclinepi.2010.07.017] [Medline: 21247734]
28. Wade SL, Narad ME, Moscato EL, LeBlond EI, King JA, Raj SP, et al. A Survivor's Journey: Preliminary efficacy of an online problem-solving therapy for survivors of pediatric brain tumor. Pediatr Blood Cancer 2020 Feb 13;67(2):e28043. [doi: 10.1002/pbc.28043] [Medline: 31724307]

29. Moscato EL, Miley AE, LeBlond EI, King JA, Raj SP, Narad ME, et al. Feasibility and acceptability of an online problem-solving therapy intervention for adolescent and young adult brain tumor survivors. Clinical Practice in Pediatric Psychology 2019 Mar;7(1):68-78. [doi: 10.1037/cpp0000265]
30. Wade S, Stancin T, Kirkwood M, Brown T, McMullen K, Taylor H. Counselor-assisted problem solving (CAPS) improves behavioral outcomes in older adolescents with complicated mild to severe TBI. J Head Trauma Rehabil 2014;29(3):198-207 [FREE Full text] [doi: 10.1097/HTR.0b013e31828f9fe8] [Medline: 23640543]
31. Wade SL, Taylor HG, Cassedy A, Zhang N, Kirkwood MW, Brown TM, et al. Long-Term Behavioral Outcomes after a Randomized, Clinical Trial of Counselor-Assisted Problem Solving for Adolescents with Complicated Mild-to-Severe Traumatic Brain Injury. J Neurotrauma 2015 Jul 01;32(13):967-975 [FREE Full text] [doi: 10.1089/neu.2014.3684] [Medline: 25738891]
32. Chadi N, Weisbaum E, Malboeuf-Hurtubise C, Ahola Kohut S, Viner C, Kaufman M, et al. Can the Mindful Awareness and Resilience Skills for Adolescents (MARS-A) Program Be Provided Online? Voices from the Youth. Children (Basel) 2018 Aug 28;5(9) [FREE Full text] [doi: 10.3390/children5090115] [Medline: 30154368]
33. Chadi N, Weisbaum E, Malboeuf-Hurtubise C, Kohut SA, Viner C, Palaniyar N, et al. In-person vs. eHealth Mindfulness-based Intervention for Adolescents with Chronic Illnesses: A Pilot Randomized Trial. APS 2019 May 31;9(1):11-23. [doi: 10.2174/2210676608666181031102235]
34. McGill BC, Sansom-Daly UM, Wakefield CE, Ellis SJ, Robertson EG, Cohn RJ. Therapeutic Alliance and Group Cohesion in an Online Support Program for Adolescent and Young Adult Cancer Survivors: Lessons from "Recapture Life". J Adolesc Young Adult Oncol 2017 Dec;6(4):568-572. [doi: 10.1089/jayao.2017.0001] [Medline: 28581346]
35. Sansom-Daly UM, Wakefield CE, Bryant RA, Patterson P, Anazodo A, Butow P, Recapture Life Working Party. Feasibility, acceptability, and safety of the Recapture Life videoconferencing intervention for adolescent and young adult cancer survivors. Psychooncology 2019 Feb 28;28(2):284-292. [doi: 10.1002/pon.4938] [Medline: 30414219]
36. Wade SL, Bedell G, King JA, Jacquin M, Turkstra LS, Haarbauer-Krupa J, et al. Social Participation and Navigation (SPAN) program for adolescents with acquired brain injury: Pilot findings. Rehabil Psychol 2018 Aug;63(3):327-337 [FREE Full text] [doi: 10.1037/rep0000187] [Medline: 30024207]
37. Wade SL, Walz NC, Carey J, McMullen KM, Cass J, Mark E, et al. Effect on behavior problems of teen online problem-solving for adolescent traumatic brain injury. Pediatrics 2011 Oct 02;128(4):e947-e953 [FREE Full text] [doi: 10.1542/peds.2010-3721] [Medline: 21890828]
38. Ricketts EJ, Goetz AR, Capriotti MR, Bauer CC, Brei NG, Himle MB, Snorrason, et al. A randomized waitlist-controlled pilot trial of voice over Internet protocol-delivered behavior therapy for youth with chronic tic disorders. J Telemed Telecare 2016 Apr;22(3):153-162 [FREE Full text] [doi: 10.1177/1357633X15593192] [Medline: 26169350]
39. Bowen DJ, Kreuter M, Spring B, Cofta-Woerpel L, Linnan L, Weiner D, et al. How we design feasibility studies. Am J Prev Med 2009 May;36(5):452-457 [FREE Full text] [doi: 10.1016/j.amepre.2009.02.002] [Medline: 19362699]
40. Whaibeh E, Mahmoud H, Naal H. Telemental Health in the Context of a Pandemic: the COVID-19 Experience. Curr Treat Options Psychiatry 2020 Apr 02;7(2):1-5 [FREE Full text] [doi: 10.1007/s40501-020-00210-2] [Medline: 32292687]
41. Mallow JA, Petitte T, Narsavage G, Barnes E, Theeke E, Mallow BK, et al. The Use of Video Conferencing for Persons with Chronic Conditions: A Systematic Review. Ehealth Telecommun Syst Netw 2016 Jun;5(2):39-56 [FREE Full text] [doi: 10.4236/etsn.2016.52005] [Medline: 29104830]
42. Langarizadeh M, Tabatabaei M, Tavakol K, Naghipour M, Rostami A, Moghbeli F. Telemental Health Care, an Effective Alternative to Conventional Mental Care: a Systematic Review. Acta Inform Med 2017 Dec;25(4):240-246 [FREE Full text] [doi: 10.5455/aim.2017.25.240-246] [Medline: 29284913]

Edited by J Torous, G Eysenbach; submitted 30.04.21; peer-reviewed by A Wickersham, C Son; comments to author 20.05.21; revised version received 23.06.21; accepted 06.07.21; published 27.08.21.

Please cite as:
Lau N, Colt SF, Waldbaum S, O'Daffer A, Fladeboe K, Yi-Frazier JP, McCauley E, Rosenberg AR
Telemental Health For Youth With Chronic Illnesses: Systematic Review
JMIR Ment Health 2021;8(8):e30098
URL: https://mental.jmir.org/2021/8/e30098
doi:10.2196/30098
PMID:34448724

Original Paper

A Compassion-Focused Ecological Momentary Intervention for Enhancing Resilience in Help-Seeking Youth: Uncontrolled Pilot Study

Christian Rauschenberg[1,2], MSc; Benjamin Boecking[3], Dipl-Psych, PhD, DClinPsy; Isabell Paetzold[1], MSc; Koen Schruers[2,4], PhD, MD; Anita Schick[1], Dipl-Psych, PhD; Thérèse van Amelsvoort[2,4], PhD, MD; Ulrich Reininghaus[1,5,6], Dipl-Psych, MSc, PhD

[1]Department of Public Mental Health, Central Institute of Mental Health, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany

[2]Department of Psychiatry and Neuropsychology, School for Mental Health and Neuroscience, Maastricht University, Maastricht, Netherlands

[3]Charité, Universitätsmedizin Berlin, Berlin, Germany

[4]Mondriaan Mental Health Center, Maastricht, Netherlands

[5]ESRC Centre for Society and Mental Health, King's College London, London, United Kingdom

[6]Centre for Epidemiology and Public Health, Health Service and Population Research Department, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, United Kingdom

Corresponding Author:
Ulrich Reininghaus, Dipl-Psych, MSc, PhD
Department of Public Mental Health
Central Institute of Mental Health
Medical Faculty Mannheim, Heidelberg University
J5
Mannheim, 68159
Germany
Phone: 49 621 1703 1930
Email: ulrich.reininghaus@zi-mannheim.de

Abstract

Background: Digital interventions offer new avenues for low-threshold prevention and treatment in young people. Ecological momentary interventions (EMIs) represent a powerful approach that allows for adaptive, real-time, and real-world delivery of intervention components in daily life by real-time processing of ecological momentary assessment (EMA) data. Compassion-focused interventions (CFIs) may be particularly amenable to translation into an EMI to strengthen emotional resilience and modify putative risk mechanisms, such as stress sensitivity, in the daily lives of young help-seeking individuals.

Objective: This study aims to investigate the feasibility, safety, and initial therapeutic effects of a novel, accessible, transdiagnostic, ecological momentary CFI for improving emotional resilience to stress (*EMIcompass*).

Methods: In this uncontrolled pilot study, help-seeking youth with psychotic, depressive, or anxiety symptoms were offered the EMIcompass intervention in addition to treatment as usual. The EMIcompass intervention consisted of a 3-week EMI (including enhancing, consolidating, and EMA-informed interactive tasks) administered through a mobile health app and three face-to-face sessions with a trained psychologist intended to provide guidance and training on the CFI exercises presented in the app (ie, training session, follow-up booster session, and review session).

Results: In total, 10 individuals (mean age 20.3 years, SD 3.8; range 14-25) were included in the study. Most (8/10, 80%) participants were satisfied and reported a low burden of app usage. No adverse events were observed. In approximately one-third of all EMAs, individuals scored high on stress, negative affect, or threat anticipation during the intervention period, resulting in real-time, interactive delivery of the CFI intervention components in addition to weekly enhancing and daily consolidating tasks. Although the findings should be interpreted with caution because of the small sample size, reduced stress sensitivity, momentary negative affect, and psychotic experiences, along with increased positive affect, were found at postintervention and the 4-week follow-up. Furthermore, reductions in psychotic, anxiety, and depressive symptoms were found (r=0.30-0.65).

Conclusions: Our findings provide evidence on the feasibility and safety of the EMIcompass intervention for help-seeking youth and lend initial support to beneficial effects on stress sensitivity and mental health outcomes. An exploratory randomized controlled trial is warranted to establish the feasibility and preliminary evidence of its efficacy.

(JMIR Ment Health 2021;8(8):e25650) doi:10.2196/25650

KEYWORDS

mental health; adolescent psychopathology; digital interventions; mobile health; self-compassion; ecological momentary assessment; mobile phone

Introduction

Background

Most mental disorders first emerge in adolescence and young adulthood (three-fourths by the age of 24 years [1]), with an estimated lifetime prevalence of approximately 50% of any mental disorder in young age groups [1-5]. Furthermore, the Global Burden of Disease study has reported that mental and substance use disorders in children and youth aged 10 to 24 years were the leading cause of overall disease burden in high-income countries [6-8]. Evidence further suggests that most mental disorders are continuous—phenomenologically and temporarily—and, in their early stages, are nonspecific in nature, often evolving into transdiagnostic phenotypes associated with a range of exit psychopathologies [9-16]. Consequently, clinical staging models as an adjunct to formal diagnoses have been introduced [17-19], highlighting the importance of transdiagnostic (indicated) prevention and early intervention [20-24].

Recent transformations in our understanding of the phenomenology, etiology, and early course of mental disorders have contributed to a move toward early detection and prevention [10-13,20,25-31]. Although conventional mental health services offer a range of therapeutic options, it has been widely documented that psychological help remains difficult to access, especially for young individuals in the early stages of mental health problems [21,22,32,33]. Furthermore, tailoring therapeutic options to specific needs and preferences of youth remains a challenge [32-36] and likely contributes to the problem that only a fraction of young people in need of help access any mental health service. Hence, young individuals often experience a long duration of untreated mental health problems, which has been identified as an important marker of poor course and outcome [32].

There is increasing interest in using digital tools to deliver mental health services [37], which may help extend access to and personalization of mental health care [38,39]. This shift has driven the development of novel mobile health (mHealth) interventions for various mental health problems [40-42], of which ecological momentary interventions (EMIs) [23,34,38,39,43], such as the Acceptance and Commitment Therapy in Daily Life [34-36,44], represent a very powerful approach. EMIs allow for adaptive, real-time, and real-world transfer of intervention components in individuals' daily lives. Thus, EMIs provide a unique opportunity to deliver personalized, precision interventions tailored to what young individuals need in a given moment and context through interactive sampling in real time and the real world. They are based on fine-grained ecological momentary assessment (EMA) data acquired through cutting-edge digital technology [21,23,24,38,39,45,46]. More recently, some authors have started to use the term just-in-time adaptive interventions, which emphasize EMI's capability of adapting the delivery of intervention components to person and context based on experience sampling or other, for example, sensing data [47,48].

One tangible prevention and early intervention strategy using digital tools is to identify and target transdiagnostic psychological mechanisms in daily life, which have been shown to be involved in the development of mental health problems [23,38]. In recent years, research using EMA—a structured diary technique, also known as experience sampling methodology [43]—has contributed to a better understanding of putative mechanisms likely to impact different stages and increase the intensity of mental health problems in individuals' daily lives, in real time and outside the research laboratory [21-23,29,43,49,50]. To date, the psychological mechanism most widely studied in daily life is elevated stress sensitivity, characterized by more intense negative affective and psychotic experiences in response to minor stressors and routine daily hassles [22,24,29,43]. Previous studies have suggested that stress sensitivity is elevated in individuals with (1) higher familial or psychometric risk, (2) an ultra–high risk state for psychosis, (3) other early mental health problems, (4) first-episode psychosis, (5) severe and enduring psychosis, and (6) depressive disorders [21,22,24,28,50-58]. In addition, heightened interpersonal sensitivity and threat anticipation have previously been reported to represent further candidate mechanisms in individuals with ultra–high risk state for psychosis, paranoia, and psychotic disorders [24,29,30,59-62] and individuals with depression and anxiety [63-66]. These transdiagnostic mechanisms reflect candidate targets to be modified by EMIs [21,22,24,29].

Compassion-focused interventions (CFIs) are considered an important strand of transdiagnostic interventions for modifying emotion regulation systems [67,68]. CFIs are part of third-wave cognitive behavioral therapy (CBT) and previous meta-analytic evidence on third-wave CBT, including CFIs [69-73], suggest that these types of interventions may yield improvements in mental health outcomes of moderate-to-large effect size. CFIs have been successfully administered to and appraised positively by help-seeking individuals, including individuals with depression, anxiety, and psychosis [74-77]. Furthermore, CFIs have been shown to induce reductions in negative affect and paranoia in moments of high stress in previous research lab experimental work [78,79]. In addition, positive imagery, an important component of CFIs, has been found effective in reducing various mental health problems, including depression, anxiety, and psychosis [76,80,81] and increasing positive affect, optimism, and behavioral activation [79,82-84]. Thus, CFIs are particularly well placed to be administered as an EMI to strengthen emotional resilience and modify putative risk mechanisms of poor mental health in young individuals with psychological distress [72,78,85], including stress sensitivity and threat anticipation [21,22]. However, the use of conventional CFIs under real-world conditions remains limited [86].

As young individuals are *digital natives*, translating CFI components into an EMI administered through an mHealth app may be a particularly promising approach, offering entirely new avenues for low-threshold prevention and intervention in youth. EMIs are fundamentally translational as they directly build on evidence of underlying momentary mechanisms in daily life and translate these into the development and evaluation of novel digital interventions by targeting these mechanisms in real time and the real world, outside the research lab or clinic [23,39,43]. However, it remains to be established whether evidence on reductions in negative affect and paranoia in moments of high stress—observed in the research laboratory—and effects on other mental health outcomes can indeed be translated to real-world and real-time delivery of EMIs that harness CFI techniques, especially in young help-seeking individuals, where accessible, youth-friendly translation of prevention and early intervention principles reflects a particular challenge.

This Study

The current study aims to establish the clinical feasibility, safety, and initial therapeutic effects of a novel, accessible, transdiagnostic, ecological momentary CFI for improving emotional resilience to stress (*EMIcompass*) in an uncontrolled phase 1 pilot study in help-seeking youth with psychotic, depressive, or anxiety symptoms. The EMIcompass intervention consisted of a 3-week EMI and three face-to-face sessions with a trained psychologist (ie, training session, follow-up booster session, and review session). Specifically, the intervention offered widely used CFI techniques (eg, compassionate and positive imagery, compassionate writing, and emotion as a wave). To facilitate the interactive, real-time, and real-world translation of the therapeutic content and techniques used in the initial training and booster sessions into individuals' daily lives, the EMI was administered through an mHealth app on a smartphone. The EMI consisted of (1) enhancing tasks, (2) consolidating tasks, and (3) EMA-informed interactive tasks that aim at an ecological translation of CFI principles and techniques to daily life. Participants were required to complete one *enhancing task* per week, which allowed them to practice new compassion-focused exercises that were then extended throughout the study period. In addition, they were required to practice the learned CFI components once a day by completing the *consolidating tasks*. Each time an enhancing task was presented, the intervention components covered by consolidating tasks were expanded. Participants were also offered *interactive tasks* if they scored high on stress, negative affect, or threat anticipation in daily EMA. The face-to-face sessions were designed to provide guidance and training on the CFI exercises and how to use the app, background information on the strategies presented, and discussions of open questions and challenges participants encountered while using the app.

The primary objective of this study is to (1) assess the clinical feasibility of delivering the EMIcompass intervention to help-seeking youth based on successful recruitment, assessment of outcomes, compliance, satisfaction, and acceptability and safety by carefully documenting any serious adverse events throughout the study period. The secondary objectives were to examine (2) initial therapeutic effects of EMIcompass on reducing stress sensitivity, negative affect, and psychotic experiences, and increasing positive affect in daily life at the end of the 3-week intervention period (*postintervention*), and after a 4-week follow-up period (*follow-up*), along with (3) the initial therapeutic effects of EMIcompass on reducing threat anticipation, psychotic, depressive, and anxiety symptoms as well as general psychopathology.

Methods

Study Design

In an uncontrolled phase 1 pilot study, help-seeking individuals with psychotic, depressive, or anxiety symptoms aged between 14 and 25 years were referred to secondary mental health services in the Netherlands (ie, Mondriaan Mental Health Trust and Virenze Mental Health Care) and received the EMIcompass intervention in addition to treatment as usual. Data were collected before the intervention (*baseline*), at the end of the 3-week intervention period (*postintervention*), and after a 4-week follow-up period (*follow-up*). Close attention was paid to establishing the clinical feasibility (eg, pragmatic inclusion and exclusion criteria based on routine assessments) and safety (ie, documentation of any serious adverse events) of this study. Our recruitment strategy drew on our previous and ongoing work with youth [22,24,29,34-36,44] and guidance for pragmatic randomized controlled trials (RCTs) [87] and hence was geared to reflect the heterogeneity of the population commonly encountered in routine care.

Sample

We recruited young individuals with psychotic, depressive, and/or anxiety symptoms who sought help from two secondary mental services (ie, Mondriaan Mental Health Center and Virenze Mental Health Care). The inclusion and exclusion criteria were equivalent in principle across the two services but were purposefully selected to be pragmatic and hence based on routine assessments for screening, diagnosis, formulation, and outcome measurement, which differed between the two services (Textbox 1). This approach was adopted to ensure that the aim of establishing feasibility reflected the population actually encountered in clinical practice (rather than imposed by researchers) while keeping the assessment burden at a minimum. The study was approved by the Ethics Review Committee of Mondriaan Mental Health Center and the Ethics Review Committee of Psychology and Neuroscience, Maastricht University. A flowchart of the study is shown in Figure 1.

The prodromal questionnaire (PQ) [88,89], which has been reported to be a very good screening measure in routine mental health services [89,90], was used to screen for psychotic symptoms. In addition, the Brief Symptom Inventory (BSI) [91,92] was used to screen for anxiety, depressive, and psychotic symptoms, and the Symptom Questionnaire-48 [93] was used in addition to the PQ to screen for anxiety and depressive symptoms.

Textbox 1. Inclusion and exclusion criteria by participating mental health services.

Inclusion Criteria

Mondriaan

- Aged between 18 and 25 years
- Prodromal questionnaire score of 6 or above
- Symptom questionnaire-48 score of 9 or above on the social phobia subscale, or score of 8 or above on the depression subscale, or score of 11 or above on the anxiety subscale
- Willingness to participate in the compassion-focused ecological momentary intervention
- Ability to give written informed consent independently, without help from others

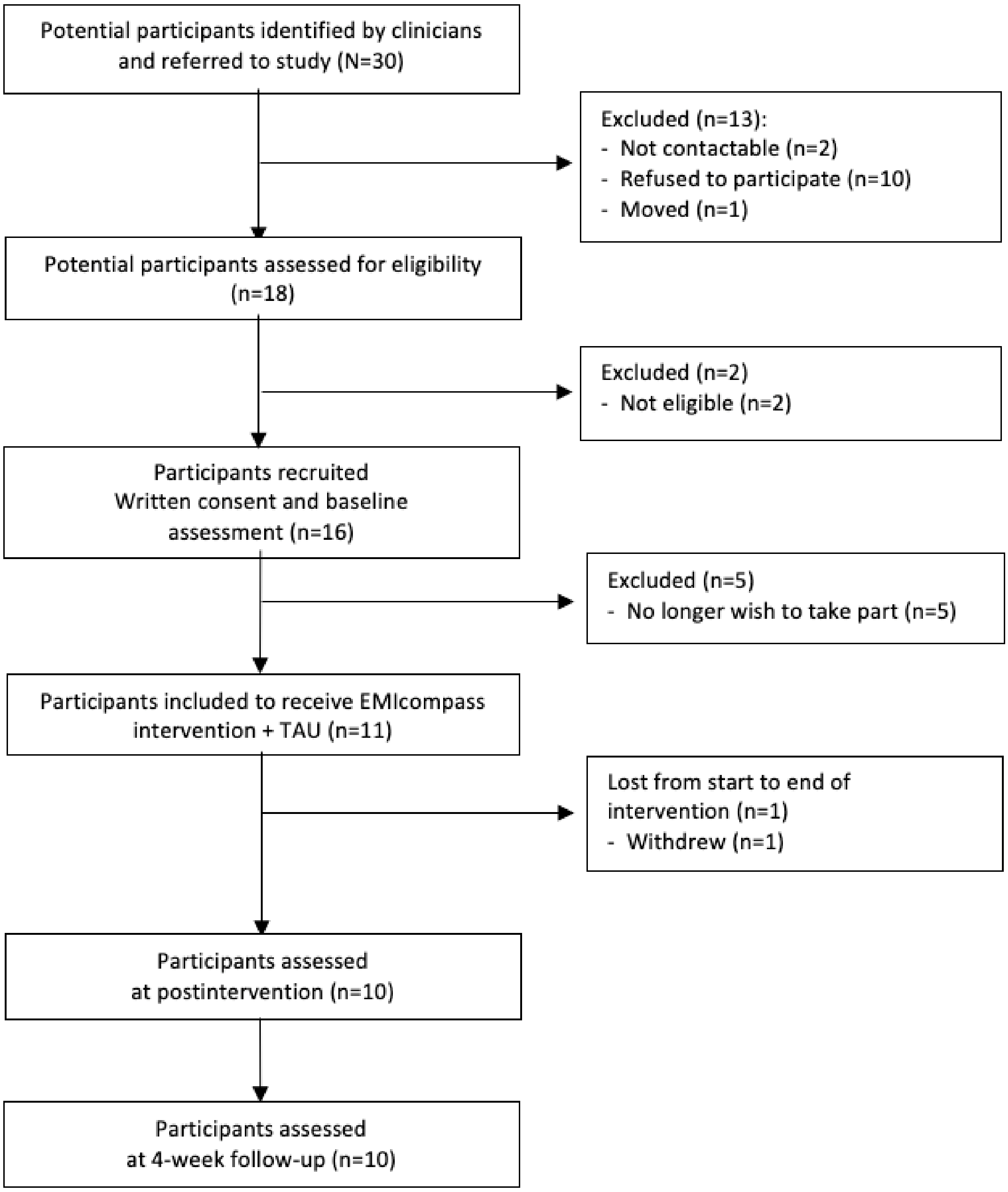
Virenze

- Aged between 14 and 25 years
- Prodromal questionnaire score of 6 or above
- Brief Symptom Inventory *t* score of 63 or above
- Willingness to participate in the compassion-focused ecological momentary intervention
- Ability to give written informed consent independently, without help from others

Exclusion Criteria

- Insufficient command of Dutch, primary clinical diagnosis of alcohol or substance dependency, severe endocrine, cardiovascular, or organic brain disease

Figure 1. Study flowchart. TAU: treatment as usual.



The EMIcompass Intervention

Development of the Manual

The intervention was structured and manualized to ensure consistent delivery. The manual was based on widely used CFI techniques (eg, compassionate and positive imagery, compassionate writing, and emotion as a wave) and developed following a process of reviewing existing manuals and the extant CFI literature [67,68,73,74,78,80] through the team's clinical experience of working with these approaches with clients and through consultation with local experts in CFI and the wider research team. The intervention was designed based on the principles of EMIs [23,34-36,39,43,44].

EMIcompass Intervention and Treatment as Usual

In this study, participants were offered the EMIcompass intervention in addition to treatment as usual, which included all the treatment they received before the start of the study (ie, good standard care delivered according to local and national guidelines by their general practitioner, psychiatrist, and other health care professionals), including CBT, third-wave CBT, dialectical behavior therapy, and other psychological interventions. The EMIcompass intervention consisted of three face-to-face sessions (one training session, one follow-up booster session, and one review session) given by a trained psychologist, who was supervised by an expert clinical psychologist in compassion-focused therapy, and a 3-week EMI administered through an mHealth app on a smartphone (PsyMate; Psymate BV). In addition, participants were offered on-demand email and/or phone contact during the intervention period.

At the beginning of the 3-week intervention period, an initial face-to-face training session was offered to participants. This session was fully manualized based on previous research that used CFIs [67,68,74,78,94]. The goal of the first session was to train individuals to cope with negative emotions by applying a personal, compassionate image that conveys compassion, care, and warmth to them based on the description of Gilbert [68], as applied by Lincoln et al [78]. This was followed by inducing negative emotions using in-sensu exposure to a personally relevant social situation that participants remember having experienced as distressing. This method has been safely applied to individuals with mental health problems [74,78] without any adverse consequences or health-related risks. Following the induction of negative emotions, participants were asked to practice a 5-minute application of the compassionate image they were trained in at the beginning of the session [67,68,78]. This step of actively using compassionate imagery after inducing negative emotions is considered essential for compassion-focused therapy to be efficacious in reducing stress sensitivity, threat anticipation, and psychotic, depressive, or anxiety symptoms in daily life [67,68]. Training the use of compassionate imagery was repeated and extended to imagery involving a *compassionate self* [68] and *emotion as a wave* [94] in the following booster session 2 weeks after the initial training session. In the review session at the end of the 3-week intervention period, the smartphone was returned, and progress and satisfaction with and acceptability of the intervention were reviewed and assessed.

To allow for interactive, real-time, and real-world translation of the therapeutic content and techniques of initial and booster sessions into individuals' daily lives, participants were offered a 3-week EMI delivered through an mHealth app. During the 3-week intervention period, the smartphone prompted a signaling sound from the smartphone seven times per day on 6 consecutive days per week to reduce the burden associated with app usage. At each beep, participants were asked to complete a brief EMA on momentary stress, positive and negative affect, and threat anticipation in daily life (see the section on EMA measures used). The EMA was scheduled at random within set blocks of time. The EMI consisted of 3 different types of tasks (Table 1): participants were asked to complete one *enhancing task* per week, allowing them to practice new compassion-focused exercises, which were subsequently extended during the study period (eg, discovering their own compassionate self and experiencing emotions as a wave). In addition, they were asked to practice the learned CFI components once a day by completing the *consolidating tasks* at a predefined time. The components covered by consolidating tasks were extended each time an enhancing task was presented. Furthermore, *interactive tasks* were offered if participants scored high on stress, negative affect, or threat anticipation in the EMA (ie, scores higher than 4 on a 7-point Likert scale). As an essential element of compassion-focused therapy is the use of compassionate imagery in moments of high stress, negative affect, or threat anticipation, these interactive tasks are thought to reflect a core active component of the 3-week compassion-focused EMI.

Table 1. Components of the EMIcompass intervention.

	Week 1	Week 2	Week 3
Compassion-focused training sessions	• Training session (compassionate image)	• Booster session (day 11-15; compassionate self-training, "emotion as a wave")	• Review session (after day 20)
Enhancing tasks	• Task 1 (day 3 or 4): compassionate self-validation	• Task 2 (day 9 or 10): "emotion as a wave"	• Task 3 (day 15 or 16): self-compassionate writing
Consolidating tasks	• Compassionate self-validation (from day 5, following enhancing EMIa task 1)	• Compassionate self-validation • "Emotion as a wave" (from day 11, following enhancing EMI task 2)	• Compassionate self-validation • "Emotion as a wave" • Self-compassionate writing (from day 17, following enhancing EMI task 3)
Interactive tasks	• Compassionate image • Compassionate self-validation (from day 5, following enhancing EMI task 1)	• Compassionate image • Compassionate self-validation • "Emotion as a wave" (from day 11, following enhancing EMI task 2)	• Compassionate image • Compassionate self-validation • "Emotion as a wave" • Self-compassionate writing (from day 17, following enhancing EMI task 3)

[a]EMI: ecological momentary intervention.

Measures

Sociodemographic Characteristics

A sociodemographic schedule was used to assess age, gender, occupation, and level of education.

Clinical Feasibility and Safety

Feasibility was assessed based on successful recruitment, assessment of outcomes, compliance with the manual, satisfaction, and acceptability. For some of the feasibility domains, a debriefing scale was used. The reasons participants declined to participate in the study were carefully recorded, and the completeness of outcomes at each time point was documented. Acceptability was assessed in the review session of the EMIcompass intervention together with the trained psychologist by asking participants to complete a feedback form about the EMI tasks and sessions and rate the extent to which they felt they benefited from and were satisfied with the intervention [74,78]. In addition, the trained psychologist asked participants in the review session to report whether they perceived the face-to-face sessions, compassion-focused exercises, and EMI tasks as helpful. App usability was assessed by asking participants to rate the readability of the text shown on the screen, any difficulties in operating the app or technical problems, the clarity of provided instructions, and whether the app was perceived as burdensome. All items were rated on a 7-point Likert scale ranging from *not at all* (rating of 1) to *moderate* (rating of 4) and *very* (rating of 7), which were subsequently grouped into three categories of *not* (rating of 3 or lower), *moderate* (rating of 4 or 5), and *very* (rating of 6 or 7) for the sake of interpretability of findings (given small numbers in each cell). Safety was assessed by carefully documenting any serious adverse events throughout the entire study period and the potential negative effects of app usage on mental health in participants.

Stress Sensitivity, Negative and Positive Affect, and Psychotic Experiences in Daily Life

EMA was used to assess stress sensitivity, negative and positive affect, psychotic experiences, and threat anticipation in daily life. For this, the same app was used as for the EMIcompass intervention (PsyMate), and assessments were completed at baseline, postintervention, and 4-week follow-up for 6 consecutive days, following the protocol from previous EMA studies [22,24,29,46,49]. Stress was operationalized as minor disturbances and distinctive unpleasant events, activities, and social situations that occur in the flow of daily life. Event-related stress was measured with an item asking participants to rate the most important event that had happened since the last beep on a 7-point Likert scale ranging from *very unpleasant* (rating of −3) to *very pleasant* (rating of 3) [54]. The item was recoded, such as higher ratings indicated higher levels of stress (with ratings of −3 coded as 7 and ratings of 3 coded as 1). Activity-related stress was measured by asking participants first to specify their current activity (eg, resting and watching TV), which was followed by asking them to rate the pleasantness of this activity on a 7-point Likert scale (1=*very unpleasant*; 7=*very pleasant*). Social stress was measured by asking participants to specify categorically with whom they were spending time (eg, nobody, partner, or family) and appraise the current social context using the items "I find being with these people pleasant" (reversed), "I feel accepted" (reversed), and "I feel excluded (if with someone)" or "I find it pleasant to be alone" (reversed) and "I would prefer to have company" (if alone) ranging from *not at all* (rating of 1) to *very much* (rating of 7). The good concurrent validity of these EMA stress measures has been reported [54,55]. Furthermore, a composite stress score was calculated using the mean score of all seven stress items [21,95]. Negative affect was assessed using five items asking participants to rate the extent to which they felt anxious, down, insecure, uncomfortable, and guilty at each entry point [54]. Positive affect was assessed by asking participants to rate the extent to which they felt cheerful and relaxed, all rated on a 7-point Likert scale ranging from *not at all* (rating of 1) to *very much* (rating of 7) [54,55,96]. Psychotic experiences were assessed using seven items ("I see things that aren't really there," "I hear things that aren't really there," "I feel suspicious/paranoid," "I feel unreal," "My thoughts are influenced by other," "I can't get these thoughts out of my head," and "I feel like I am losing control") rated on a 7-point Likert scale ranging from 1 (*not at all*) to 7 (*very much*) [55,96]. Threat anticipation was assessed by asking participants to think of what might happen in the next few hours and rate the item "I think that something unpleasant will happen" on a 7-point Likert scale (ranging from 1=*not at all* to 7=*very much*) [24,29]. Negative and positive affect, psychotic experiences, and threat anticipation scores were assessed by computing the mean scores. In line with earlier studies [22,24,29,46,49], items on stress, negative affect, and psychotic experiences were used as a proxy for individuals' stress sensitivity in daily life by modeling the association between stress and (1) negative affect and (2) psychotic experiences. Thus, we conceptualized stress sensitivity in daily life as individuals' affective and psychotic reactivity to minor daily stressors.

Psychotic, Depressive, and Anxiety Symptoms and General Psychopathology

We used non-EMA outcome measures to assess psychotic, depressive, and anxiety symptoms and general psychopathology. First, the BSI was used to assess depressive and anxiety symptoms (based on the respective BSI subscales) and general psychopathology by computing the Global Severity Index (based on 53 BSI items). Participants rated each item on a 5-point scale ranging from 0 (*not at all*) to 4 (*extremely*) [91,92]. Second, the Green et al, Paranoid Thoughts Scale, a reliable and valid scale, was used to assess psychosis [97]. The Green et al, Paranoid Thoughts Scale was modified to ask participants about paranoid ideation during the past week rather than the past month, given that the intervention period was only 3 weeks. A total score was computed using all 32 items (both with a 5-point scale: 1=*not at all*, 3=*somewhat*, and 5=*totally*). Third, the threat anticipation measure [98] was used to measure threat anticipation by asking participants to estimate the future likelihood of a list of threatening, neutral, and positive events happening to themselves and other people [62,98,99]. Items for threatening and neutral events were used to compute the total scores. Each event was rated separately for the likelihood that it will happen to oneself and another person on a 7-point scale (1=*not at all*; 7=*very likely*), resulting in four total sum scores (ie, threat anticipation-self, threat anticipation-other, neutral anticipation-self, and neutral anticipation-other), where higher scores indicate higher probability estimates. Finally, the PQ [88,89] was used to assess the presence of prodromal and attenuated psychotic symptoms (ie, positive symptoms, disorganized symptoms, negative symptoms, and general symptoms). This measure consists of 16 items that assess the presence of psychotic symptoms (0=false and 1=true), which were used to compute a total score (range 0-16). Good psychometric properties have been reported for these measures [88,97,98,100,101].

Statistical Analysis

STATA 15.1 (StataCorp) was used to analyze the data. First, descriptive statistics were used, and CIs were constructed, as appropriate, to summarize the findings on feasibility and safety. Second, as EMA data have a multilevel structure, such that multiple observations (level 1) are nested within subjects (level 2), linear mixed models were used to control for within-subject clustering of multiple observations using the *mixed* command in STATA. Thus, to examine the effects of the EMIcompass intervention on reducing stress sensitivity, EMA stress variables and time points were included as independent variables and negative affect and psychotic experiences as the outcome variable in linear mixed models, which were fitted separately for each outcome variable. We then added two-way interaction terms for stress×time and used likelihood ratio tests (*lrtest* command) to evaluate improvement in model fit and the *lincom* command to compute linear combinations of coefficients to test our hypotheses on whether stress sensitivity was reduced at postintervention and the 4-week follow-up. We standardized continuous ESM (experience sampling method) variables (mean 0, SD 1) to interpret significant interaction terms. Family-wise error-corrected *P* values were computed to control for multiple testing by multiplying the unadjusted *P* values of the two-way interaction effects by the total number of tests (N=4) for each outcome. Third, to examine the effects of the EMIcompass intervention on other EMA outcome measures, time points were included as independent variables and negative affect, positive affect, psychotic experiences, and threat anticipation as the outcome variable in separate linear mixed models. All models were controlled for potential confounders (ie, age, gender, and level of education). Finally, we used Wilcoxon signed-rank tests to examine the effects of EMIcompass on non-EMA outcome measures of threat anticipation, psychotic, depressive, and anxiety symptoms and general psychopathology at postintervention and 4-week follow-up. The resulting z scores were used to calculate the effect sizes displayed in *r* as described by Rosenthal and DiMatteo [102].

Results

Sociodemographic and Clinical Characteristics

A flowchart of the study is shown in Figure 1 and basic sample characteristics in Table 2. In total, 30 potential participants aged between 14 and 25 years were referred to the study by clinicians from the two participating mental health services. Of these, 16 provided written informed consent and were eligible, of whom 11 completed the baseline assessment and were included in the EMIcompass intervention. A participant was lost during the 3-week intervention period, whereas 10 participants (mean age 20.3 years, SD 3.8; range 14-24) completed the EMIcompass intervention and both postintervention and 4-week follow-up assessments. Most participants were women (7/10, 70%) and were currently at school/university (6/10, 60%). Half of the participants had a clinical diagnosis of major depressive disorder (5/10, 50%) and met the criteria for a comorbid mental health condition. Most participants were of White Dutch ethnicity, and some reported having used cannabis during the previous 12 months (3/10, 30%).

Table 2. Basic sample characteristics of service users (N=10).

Characteristic	Value
Age (years), mean (SD; range)	20.3 (3.8; 14-25)
Sex, n (%)	
Female	7 (70)
Male	3(30)
Ethnicity, n (%)	
White Dutch	6 (60)
Other	1 (10)
Missing value	3 (30)
Level of education, n (%)[a]	
School	2 (20)
Further	4 (40)
Higher	4 (40)
Occupation, n (%)	
School or education	6 (60)
Employed (full- or part-time)	3 (30)
Unstructured activities	1 (10)
Cannabis use[b], n (%)	
12 months	3 (30)
Lifetime	4 (40)
DSM-IV[c] diagnosis, n (%)	
Major depressive disorder	5 (50)
Attention-deficit/hyperactivity disorder	1 (10)
Reactive attachment disorder	2 (20)
None	2 (20)
Comorbid condition[d]	5 (50)

[a]Categories defined as school (elementary school), further (voorbereidend middelbaar beroepsonderwijs [VMBO]; hoger algemeen voortgezet onderwijs [HAVO], and voorbereidend wetenschappelijk onderwijs [VWO]), and higher (hoger beroepsonderwijs [HBO], and wetenschappelijk onderwijs [WO]) of the Dutch educational system.

[b]On the basis of Composite International Diagnostic Interview section of Illegal Substance Use and defined as having used cannabis more than five times on its own initiative during the previous 12 months or lifetime.

[c]DSM-IV: Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition.

[d]Consisting of the following diagnostic categories: panic disorder, attention-deficit/hyperactivity disorder, intermittent explosive disorder, borderline personality disorder, and parent-child relational problem.

Clinical Feasibility and Safety

The clinical feasibility and safety findings are shown in Table 3. Almost all individuals (9/10, 90%) reported that participating in the study did not interfere with their daily activities. Most individuals reported being very (40%-50%) or moderately satisfied (40%-50%) with tasks delivered through the EMIcompass app and moderately (20%-30%) or very (60%) satisfied across face-to-face sessions. Most participants were also very (5/10, 50%) or moderately (2/10, 20%) successful in imagining a compassionate image. Some individuals reported that the intervention positively influenced social contacts (3/10, 30%; ratings of *moderate* and *very* combined) and levels of activity (4/10, 40%). All individuals were very satisfied with the face-to-face contact sessions and felt trained psychologists understood them. Although all participants reported that they were able to follow the instructions shown on the screen, observer ratings by trained psychologists, who also delivered the face-to-face sessions, indicated that some individuals might have had problems with this (1/10, 10% in session 1 and 2/20, 20% in session 3). Findings on app usability were satisfactory, and the burden associated with app usage was perceived to be low or very low across all time points (70%-90%), although some individuals (3/10, 30%) found the number of signals per day to be moderately burdensome. In addition, some individuals perceived the items used in the PsyMate app as difficult or unclear (2/10, 20%). No severe adverse events were observed during the study period.

In-app usage data during the intervention period suggest high completion rates of EMA assessments. Specifically, the EMIcompass app triggered 1260 signals asking participants to complete brief EMA assessments (126 for each person). Of these 1260 signals, individuals reacted to 467 (37.06%), although high variability between individuals was found (range 214/1260, 16.9% to 844/1260, 66.9%). Individuals scored high on stress, negative affect, or threat anticipation in 32.1% (150/467) of EMA assessments, resulting in real-time delivery of CFI intervention components in approximately 1 out of 3 of all completed EMA assessments. When considering the assessment of outcomes at baseline, postintervention, and follow-up, we found satisfactory compliance rates (no missing data for outcome measures filled in person and at least 18/60, 30% of all EMA assessments). Thus, when combining self-reports and in-app usage data, assessing outcomes and compliance with the manual was considered satisfactory. Furthermore, the conversion rate of recruitment was 3:1 (ie, from identified to included individuals; Figure 1), which is in line with previous research and considered successful recruitment.

Table 3. Findings on safety, feasibility, and app usability of the EMIcompass intervention.

	Ratings[a]		
	Very	Moderate	Not
Safety and feasibility, n (%)			
Interference of study participation with daily activities	0 (0)	1 (10)	9 (90)
Satisfaction with face-to-face sessions	6 (60)	2 (20)	2 (20)
Session 1: compassionate image; inducing negative emotions	6 (60)	3 (30)	1 (10)
Session 2: compassionate self; emotion as a wave	6 (60)	3 (30)	1 (10)
Session 3: review session	6 (60)	3 (30)	1 (10)
Satisfaction with tasks, n (%)			
Task 1: compassionate self-validation	4 (40)	5 (50)	1 (10)
Task 2: emotion as a wave	5 (50)	3 (30)	2 (20)
Task 3: self-compassionate writing	5 (50)	3 (30)	2 (20)
Self-reported success in making a compassionate image	5 (50)	3 (30)	2 (20)
Taking part in the study positively affected activities[b]	2 (20)	2 (20)	5 (50)
Taking part in the study affected social contacts, n (%)			
Positively	1 (10)	2 (20)	7 (70)
Negatively	0 (0)	0 (0)	10 (100)
Satisfaction with contact with trained psychologist[b]	9 (100)	0 (0)	0 (0)
Participant felt understood by trained psychologist[b]	9 (100)	0 (0)	0 (0)
Self-reported level of understanding of instructions provided by trained psychologist[b]	9 (100)	0 (0)	0 (0)
Observer-rating by trained psychologists, n (%)			
Compliance in session 1	7 (70)	2 (20)	1 (10)
Compliance in session 2	7 (70)	3 (30)	0 (0)
Compliance in session 3	6 (60)	2 (20)	2 (20)
EMIcompass app usability, n (%)			
Readability of text on screen	10 (100)	0 (0)	0 (0)
Difficulties in operating the app	0 (0)	0 (0)	10 (100)
Clarity of instructions given on screen	10 (100)	0 (0)	0 (0)
Difficulties understanding used items	0 (0)	2 (20)	8 (80)
EMIcompass app perceived as burdensome, n (%)			
In terms of the number of signals per day	0 (0)	3 (30)	7 (70)
In terms of the number of items asked per signal	0 (0)	1 (10)	9 (90)
In terms of the signal sound	1 (10)	1 (10)	8 (80)
Technical problems	0 (0)	1 (10)	9 (90)

[a]Items were rated on a 7-point Likert scale ranging from not at all (rating of 1) to moderate (rating of 4) and very (rating of 7). Trained psychologists noted the answers. The answers were grouped into three categories of not (rating of 3 or lower), moderate (rating of 4 or 5), and very (rating of 6 or 7) for the sake of interpretability (given small numbers in each cell).

[b]Missing value for 1 participant.

Initial Therapeutic Effects

Stress Sensitivity, Negative and Positive Affect, and Psychotic Experiences in Daily Life

The findings on the initial therapeutic effects of the EMIcompass intervention on stress sensitivity are provided in Table 4. We found preliminary evidence that participants experienced less intense negative affect in response to event-related and activity-related stress at postintervention and in response to overall, event-related, activity-related, and social stress at follow-up than at baseline, as indicated by statistically significant two-way interaction effects for stress×time point. Furthermore, participants reported less intense psychotic experiences in response to minor stressors in daily life (ie, overall and specific types of stressors) at postintervention and follow-up than at baseline.

Table 4. Initial therapeutic effects of EMIcompass on stress sensitivity in daily life.

Outcome	Postintervention versus baseline		Follow-up versus baseline		Follow-up versus postintervention		Likelihood ratio test for interaction[a]	
	Adjusted β[b] (95% CI)	*P* value	Adjusted β (95% CI)	*P* value	Adjusted β (95% CI)	*P* value	Chi-square (*df*)	PFWE[c]
Negative affect								
Stress								
Overall	−0.12 (−0.27 to 0.03)	.11	−0.51 (−0.63 to −0.40)	<.001	−0.39 (−0.55 to −0.23)	<.001	72.6 (2)	<.001
Event-related	−0.41 (−0.56 to −0.25)	<.001	−0.39 (−0.51 to −0.27)	<.001	0.02 (−0.14 to 0.18)	.83	51.6 (2)	<.001
Activity-related	−0.25 (−0.40 to −0.09)	.002	−0.35 (−0.47 to −0.23)	<.001	−0.10 (−0.27 to 0.06)	.22	32.5 (2)	<.001
Social	0.05 (−0.10 to 0.20)	.50	−0.41 (−0.53 to −0.28)	<.001	−0.46 (−0.62 to −0.29)	<.001	47.6 (2)	<.001
Psychotic experiences								
Stress								
Overall	−0.15 (−0.25 to −0.04)	.005	−0.28 (−0.36 to −0.20)	<.001	−0.14 (−0.25 to −0.03)	.01	48.7 (2)	<.001
Event-related	−0.29 (−0.39 to −0.19)	<.001	−0.19 (−0.27 to −0.11)	<.001	0.10 (−0.01 to 0.20)	.08	40.6 (2)	<.001
Activity-related	−0.25 (−0.35 to −0.14)	<.001	−0.20 (−0.28 to −0.12)	<.001	0.05 (−0.06 to 0.16)	.40	33.3 (2)	<.001
Social	−0.01 (−0.11 to 0.09)	.86	−0.24 (−0.32 to −0.16)	<.001	−0.23 (−0.34 to −0.12)	<.001	36.3 (2)	<.001

[a]Likelihood ratio test for stress×time interaction after inclusion in the following model: (for y_{ij} negative affect, psychotic experiences or positive affect as outcome variable): $y_{ij}=\beta_0+\beta_1(STRESS_{ij})+\beta_2(TIME_j)+\beta_3(STRESS_{ij}\times TIME_j)+\varepsilon_{ij}$.

[b]Adjusted β: standardized regression coefficients (continuous independent variables were standardized [mean 0, SD 1] for interpreting interaction terms).

[c]PFWE: family-wise error-corrected *P* values were computed by multiplying the unadjusted *P* value by the total number of tests for each outcome (N=4) to adjust significance levels of likelihood ratio tests for two-way interactions.

Furthermore, Table 5 shows the findings of the initial effects of EMIcompass on momentary negative affect, psychotic experiences, and positive affect. There was preliminary evidence that participants experienced less intense negative affect and psychotic experiences and more intense positive affect in daily life at postintervention and the 4-week follow-up than at baseline. There was also evidence that individuals anticipated fewer threatening events in their daily lives at postintervention and the 4-week follow-up than at baseline.

Table 5. Initial therapeutic effects of EMIcompass on individuals' momentary stress, negative affect, psychotic experiences, positive affect, and threat anticipation.

	Baseline, mean (SD)	Postintervention, mean (SD)	Follow-up, mean (SD)	Postintervention versus baseline		Follow-up versus baseline	
				β (95% CI)	*P* value	β (95% CI)	*P* value
Positive affect	3.9 (1.8)	4.5 (1.5)	4.3 (1.6)	0.39 (0.16 to 0.62)	.001	0.31 (0.10 to 0.52)	.004
Negative affect	2.2 (1.3)	1.8 (1.1)	1.4 (0.7)	−0.44 (−0.59 to −0.30)	<.001	−0.59 (−0.72 to −0.46)	<.001
Psychotic experiences	1.7 (0.8)	1.4 (0.9)	1.3 (0.6)	−0.25 (−0.34 to −0.16)	<.001	−0.36 (−0.44 to −0.28)	<.001
Threat anticipation	2.7 (1.9)	2.2 (1.3)	1.6 (1.1)	−0.61 (−0.83 to −0.39)	<.001	−0.96 (−1.15 to −0.76)	<.001

Psychotic, Depressive, Anxiety Symptoms, and General Psychopathology

The findings on the initial therapeutic effects of EMIcompass on non-EMA outcome measures are presented in Table 6. Overall, reductions in threat anticipation, psychotic, depressive, and anxiety symptoms and general psychopathology (as indexed by the Global Severity Index) of moderate-to-large effect sizes were found at the end of the 3-week intervention period (*postintervention*) and after a 4-week follow-up period (r=0.30-0.65). There was initial evidence, despite the small sample size and, hence, limited statistical power, that these reductions were beyond what would be expected by chance alone for psychotic symptoms at postintervention and 4-week follow-up and, at trend level, for anxiety symptoms (postintervention, 4-week follow-up) and anticipation of a positive future self (4-week follow-up). The intervention effects on depressive symptoms and general psychopathology were also of medium-to-large effect size but fell short of statistical significance. Reductions in threat anticipation (self or other) were only of small-to-moderate effect size and did not reach conventional levels of statistical significance.

Table 6. Initial therapeutic effects of EMIcompass intervention on psychotic, depressive, and anxiety symptoms, general psychopathology, and threat anticipation.

	Scores, median (range)			Paired Wilcoxon signed-rank test (N=10)					
	Baseline	Postintervention	Follow-up	Postintervention versus baseline		Follow-up versus baseline		Follow-up versus postintervention	
				z	Effect size (r)[a]	z	Effect size (r)[a]	z	Effect size (r)[a]
BSI[b]									
Global Severity Index	81 (22-146)	68.5 (5-158)	51 (7-142)	−1.02	−0.32	−1.17	−0.37	−1.53	−0.48
Depression	13.5 (1-23)	12 (0-23)	7 (1-21)	−1.02	−0.33	−1.03	−0.33	−1.38	−0.44
Anxiety	11.5 (4-16)	9.5 (0-17)	7 (2-14)	−1.74	−0.55[c]	−1.79	−0.57[c]	−0.82	−0.26
GPTS[d]									
Total score	41 (32-73)	46.5 (32-83)	38 (32-70)	1.94	0.61[e]	−1.74	−0.55[c]	−2.50	−0.79[e]
Prodromal questionnaire									
Total score	5 (1-10)	5 (0-9)	2 (0-10)	−1.32	−0.42	−2.05	−0.65[e]	−1.34	−0.42
TAM[f]									
Future self (positive)	26.5 (17-37)	27 (16-37)	33 (7-42)	0.41	0.13	1.89	0.60[c]	1.79	0.57[c]
Future self (threatening)	15.5 (11-25)	16.5 (7-24)	13 (7-34)	−0.46	−0.15	−1.28	−0.40	−0.52	−0.16
Future others (positive)	31.5 (19-45)	31 (27-42)	33.5 (22-44)	0.21	0.07	1.33	0.42	1.74	0.55[c]
Future others (threatening)	15.5 (7-37)	14 (8-36)	13.5 (7-32)	−0.78	−0.25	−0.77	−0.24	−0.21	−0.07

[a]Effect size estimates are based on r described by Rosenthal and DiMatteo [102] using the following formula: r=Z/√number of pairs.

[b]BSI: Brief Symptom Inventory.

[c]$P<.10$.

[d]GPTS: Green et al, Paranoid Thoughts Scale.

[e]$P<.05$.

[f]TAM: threat anticipation measure.

Discussion

Principal Findings

The findings of this uncontrolled phase 1 pilot study suggest initial results on the feasibility, safety, and preliminary therapeutic effects of a compassion-focused ecological momentary transdiagnostic intervention designed to improve emotional resilience to stress (*EMIcompass*) in help-seeking youth with psychotic, depressive, or anxiety symptoms. First, individuals were satisfied with face-to-face and app-based intervention components, interference with daily activities was low, and observer-rated compliance with the treatment was high. The indicators of app usability were satisfactory. In addition, no adverse effects were observed. Second, there was preliminary evidence of decreased stress sensitivity, negative affect, and psychotic experiences and increased positive affect in daily life at the end of the 3-week intervention period (*postintervention*) and after a 4-week follow-up period (*follow-up*) as compared with baseline. Third, there was initial evidence, despite the small sample size and limited statistical power, of reductions in threat anticipation, psychotic, anxiety, and depressive symptoms of medium-to-large effect size (r=0.30-0.65). Overall, this reflects promising preliminary evidence of clinical feasibility and safety of the EMIcompass intervention in help-seeking youth and some evidence on initial therapeutic effects. However, findings on clinical outcomes should be interpreted with caution, considering the small sample size of this pilot study.

Strengths and Limitations

The strength of this study is that the principles of CFIs were, for the first time, translated into an EMI administered through an mHealth app as a new avenue for real-world and real-time prevention and intervention in youth. Furthermore, EMIcompass transforms evidence on putative underlying mechanisms into an intervention that directly targets these mechanisms in daily life and hence is translational. However, there are a number of limitations that must be considered when interpreting our findings. First, in line with state-of-the-art guidance on developing and evaluating complex interventions [103], mHealth interventions in particular [104], the sample size (N=10) of this pilot study was selected to be small. Thus, the primary focus of this study was to investigate feasibility and safety and estimate the effect size of initial therapeutic effects rather than statistical significance to provide the basis for a feasibility RCT [105]. Nonetheless, while considering the low statistical power and limitations associated with a small sample size, we found preliminary evidence (in terms of statistical significance) on the effects of the EMIcompass intervention on stress sensitivity. These are promising findings, as stress sensitivity is the primary target of this emotion regulation–focused intervention. Second, data on feasibility and acceptability were assessed together with or by a trained psychologist and not an independent person. Thus, we cannot rule out biases and underreporting of unhelpful experiences. Third, we used a modified version of an established debriefing scale already used for a decade in EMA studies and, more recently, in other EMIs [34,35] to assess satisfaction, engagement, and other domains of feasibility. However, the convergent validity of this measure with other established measures (eg, Mobile App Rating Scale) and other psychometric properties remain to be established. Fourth, because of the absence of a waiting list or active control group, we cannot rule out that there may be no additive therapeutic effects of the EMIcompass intervention to the therapeutic effects of the face-to-face sessions with the trained psychologists or other therapeutic interventions participants received during the intervention period in the form of treatment as usual. However, the primary aim of this pragmatic phase 1 pilot study was to provide the basis for a feasibility RCT by investigating feasibility and safety, generating initial effect sizes. Further examination of the efficacy of EMIcompass intervention is urgently warranted. Fifth, most participants were women, and half of the participants had depression, which may limit the generalizability of findings, as selection bias may have operated on our sampling procedure. Sixth, after written informed consent was obtained and baseline assessments were completed, 5 individuals decided not to participate in the study. The reasons for exclusion were not assessed, which limited our findings on feasibility. Finally, the complex nature of the investigated constructs, sample size, and study design exclude any form of causal inference.

Ideas for Future Work

The EMIcompass intervention aimed to augment current treatment options for young individuals seeking help for mental health problems. Most individuals reported being satisfied with the intervention. Although the small sample size has to be considered when interpreting findings, the preliminary therapeutic effects on candidate psychological mechanisms, including stress sensitivity and other psychopathological outcomes, were promising. Importantly, no adverse effects have been reported, and participating in the study did not hinder individuals in their daily activities. Thus, overall, findings on feasibility, safety, and initial therapeutic effects may be considered encouraging.

This is one of the first studies to develop and pilot an EMI that incorporates an adaptive and context-dependent delivery scheme of intervention components in youth with mental health problems. The *interactive tasks* were triggered in approximately 1 out of 3 of all EMA assessments when individuals experienced elevated levels of negative affect (eg, feeling anxious, insecure, down; ie, scores higher than 4 on a 7-point Likert scale) or momentary stress. Thus, real-time data processing was successfully applied based on EMA data to determine the delivery of CFI components. This may represent not only an important step toward ecologically more valid and accessible psychological interventions in youth but also a more personalized and contextualized clinical and preventive approach. In other words, the principles of EMIs allow not only to translate intervention components targeting candidate momentary mechanisms and contexts to individuals' daily lives but also take a personalized, adaptive approach informed by fine-grained real-time EMA data to produce sustainable change in the real world. Although a feasibility RCT is needed as a significant next step to investigate the efficacy of the intervention and feasibility as a basis for a confirmatory RCT [23,34], this pilot study of this novel EMI reflects an important stepping stone toward more personalized and accessible youth mental health care. Furthermore, in-app data analytics revealed high variability in compliance among individuals. This suggests that for some individuals, the number of signals per day was too high (ie, seven times per day on 6 consecutive days per week).

These findings hint toward potential avenues for the improvement of the EMIcompass intervention to be iteratively incorporated. First, future versions of the EMIcompass intervention may offer adaptive intervention trajectories that vary in the type of exercise depending on individual needs and preferences. Importantly, in doing so, potentially influencing factors (eg, educational level, language skills, cultural peculiarities, and subjective preferences) should be considered at an early stage of the design process and considered in optimizing EMIs further. Coproduction with young service users is essential during these developmental processes [106]. Second, sustained engagement in using digital tools remains a significant challenge [107], which may be addressed through the use of gamification elements, especially in youth [108,109]. However, in this study, the burden associated with app usage was low, and problems with engagement have mainly been reported for stand-alone mHealth apps without components of blended care [110]. Third, in working toward more personalized mHealth apps, more sophisticated methods may be used to inform the timing and context of when intervention components are offered (eg, by using mobile sensing data). A broader range of intervention components delivered for a longer intervention period may help enhance the effects of EMIcompass further and achieve sustainable change in individuals' daily lives. Fourth, the type of intervention components may be personalized further by assessing the effects of specific intervention components on individuals' mental health at the individual level. Fifth, it should be further examined whether and, if so, how the therapeutic alliance can be strengthened in light of a limited number of face-to-face sessions [111]. Finally, the number of signals per day triggered by the smartphone was perceived as burdensome by some participants. Thus, future versions of the EMIcompass app may lower the number of signals per day or shorten the number of items per signal [112].

Conclusions

Evidence on feasibility and safety and preliminary evidence on the therapeutic effects of the EMIcompass intervention suggest that translating CFI components into individuals' daily life through an EMI delivered by an mHealth app may be a promising novel, accessible, and transdiagnostic treatment approach in help-seeking youth by strengthening emotional resilience and directly targeting candidate psychological mechanisms. As an important next step, an exploratory RCT is warranted to demonstrate the feasibility and preliminary evidence of the efficacy of the EMIcompass intervention.

Acknowledgments

This work was supported by the Netherlands Organisation for Scientific Research (grant 451-13-022) and the German Research Foundation (grant 389624707). These funding sources had no further role in the study design; in the collection, analysis, and interpretation of data; in the writing of the report; or in the decision to submit the paper for publication. The authors would like to thank Inge Heunen, Danny Deckers, Nele Soons, Christiane Schittek, Shandery Rosalina, and Truda Driesen who helped with recruitment and data collection.

Authors' Contributions

CR was involved in developing the methodology, formal analysis, data curation, visualization, and writing—original draft, review, and editing. BB was involved in conceptualization, resources, and writing—review and editing. IP was involved in writing—review and editing. KS was involved in obtaining resources, investigation, writing—review and editing—and funding acquisition. AS involved in writing—review and editing. TVA was involved in obtaining resources, investigation, writing—review and editing—and funding acquisition. UR was involved in conceptualization, methodology, formal analysis, writing of the original draft, investigation, resources, further writing—review and editing—supervision, project administration, and funding acquisition.

Conflicts of Interest

None declared.

References

1. Kessler RC, Berglund P, Demler O, Jin R, Merikangas KR, Walters EE. Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. Arch Gen Psychiatry 2005 Jun;62(6):593-602. [doi: 10.1001/archpsyc.62.6.593] [Medline: 15939837]
2. Beesdo-Baum K, Knappe S, Asselmann E, Zimmermann P, Brückl T, Höfler M, et al. The 'Early Developmental Stages of Psychopathology (EDSP) study': a 20-year review of methods and findings. Soc Psychiatry Psychiatr Epidemiol 2015 Jun;50(6):851-866. [doi: 10.1007/s00127-015-1062-x] [Medline: 25982479]
3. Perälä J, Suvisaari J, Saarni SI, Kuoppasalmi K, Isometsä E, Pirkola S, et al. Lifetime prevalence of psychotic and bipolar I disorders in a general population. Arch Gen Psychiatry 2007 Jan;64(1):19-28. [doi: 10.1001/archpsyc.64.1.19] [Medline: 17199051]
4. Merikangas KR, He J, Burstein M, Swanson SA, Avenevoli S, Cui L, et al. Lifetime prevalence of mental disorders in U.S. adolescents: results from the National Comorbidity Survey Replication - Adolescent Supplement (NCS-A). J Am Acad Child Adolesc Psychiatry 2010 Oct;49(10):980-989 [FREE Full text] [doi: 10.1016/j.jaac.2010.05.017] [Medline: 20855043]
5. Ormel J, Raven D, van Oort F, Hartman CA, Reijneveld SA, Veenstra R, et al. Mental health in Dutch adolescents: a TRAILS report on prevalence, severity, age of onset, continuity and co-morbidity of DSM disorders. Psychol Med 2015 Jan;45(2):345-360. [doi: 10.1017/S0033291714001469] [Medline: 25066533]
6. Harhay MO, King CH. Global burden of disease in young people aged 10-24 years. Lancet 2012 Jan 07;379(9810):27-28. [doi: 10.1016/S0140-6736(12)60019-1] [Medline: 22225664]
7. Erskine HE, Moffitt TE, Copeland WE, Costello EJ, Ferrari AJ, Patton G, et al. A heavy burden on young minds: the global burden of mental and substance use disorders in children and youth. Psychol Med 2015 May;45(7):1551-1563. [doi: 10.1017/S0033291714002888] [Medline: 25534496]
8. Gore FM, Bloem PJ, Patton GC, Ferguson J, Joseph V, Coffey C, et al. Global burden of disease in young people aged 10-24 years: a systematic analysis. Lancet 2011 Jun 18;377(9783):2093-2102. [doi: 10.1016/S0140-6736(11)60512-6] [Medline: 21652063]
9. van Os J. The dynamics of subthreshold psychopathology: implications for diagnosis and treatment. Am J Psychiatry 2013 Jul;170(7):695-698. [doi: 10.1176/appi.ajp.2013.13040474] [Medline: 23820827]
10. van Os J, Reininghaus U. Psychosis as a transdiagnostic and extended phenotype in the general population. World Psychiatry 2016 Jun;15(2):118-124 [FREE Full text] [doi: 10.1002/wps.20310] [Medline: 27265696]
11. Reininghaus U, Priebe S, Bentall RP. Testing the psychopathology of psychosis: evidence for a general psychosis dimension. Schizophr Bull 2013 Jul;39(4):884-895 [FREE Full text] [doi: 10.1093/schbul/sbr182] [Medline: 22258881]
12. Reininghaus U, Böhnke JR, Hosang G, Farmer A, Burns T, McGuffin P, et al. Evaluation of the validity and utility of a transdiagnostic psychosis dimension encompassing schizophrenia and bipolar disorder. Br J Psychiatry 2016 Aug;209(2):107-113. [doi: 10.1192/bjp.bp.115.167882] [Medline: 26989099]

13. Reininghaus U, Böhnke JR, Chavez-Baldini U, Gibbons R, Ivleva E, Clementz BA, et al. Transdiagnostic dimensions of psychosis in the Bipolar-Schizophrenia Network on Intermediate Phenotypes (B-SNIP). World Psychiatry 2019 Feb;18(1):67-76 [FREE Full text] [doi: 10.1002/wps.20607] [Medline: 30600629]
14. Kotov R, Jonas KG, Carpenter WT, Dretsch MN, Eaton NR, Forbes MK, HiTOP Utility Workgroup. Validity and utility of Hierarchical Taxonomy of Psychopathology (HiTOP): I. Psychosis superspectrum. World Psychiatry 2020 Jun;19(2):151-172 [FREE Full text] [doi: 10.1002/wps.20730] [Medline: 32394571]
15. Krueger RF, Kotov R, Watson D, Forbes MK, Eaton NR, Ruggero CJ, et al. Progress in achieving quantitative classification of psychopathology. World Psychiatry 2018 Oct;17(3):282-293 [FREE Full text] [doi: 10.1002/wps.20566] [Medline: 30229571]
16. Rutigliano G, Valmaggia L, Landi P, Frascarelli M, Cappucciati M, Sear V, et al. Persistence or recurrence of non-psychotic comorbid mental disorders associated with 6-year poor functional outcomes in patients at ultra high risk for psychosis. J Affect Disord 2016 Oct;203:101-110. [doi: 10.1016/j.jad.2016.05.053] [Medline: 27285723]
17. Iorfino F, Scott EM, Carpenter JS, Cross SP, Hermens DF, Killedar M, et al. Clinical stage transitions in persons aged 12 to 25 years presenting to early intervention mental health services with anxiety, mood, and psychotic disorders. JAMA Psychiatry 2019 Nov 01;76(11):1167-1175 [FREE Full text] [doi: 10.1001/jamapsychiatry.2019.2360] [Medline: 31461129]
18. Hickie IB, Scott EM, Hermens DF, Naismith SL, Guastella AJ, Kaur M, et al. Applying clinical staging to young people who present for mental health care. Early Interv Psychiatry 2013 Feb;7(1):31-43. [doi: 10.1111/j.1751-7893.2012.00366.x] [Medline: 22672533]
19. Hartmann JA, Nelson B, Spooner R, Amminger GP, Chanen A, Davey CG, et al. Broad clinical high-risk mental state (CHARMS): Methodology of a cohort study validating criteria for pluripotent risk. Early Interv Psychiatry 2019 Jun;13(3):379-386. [doi: 10.1111/eip.12483] [Medline: 28984077]
20. McGorry PD, Hartmann JA, Spooner R, Nelson B. Beyond the "at risk mental state" concept: transitioning to transdiagnostic psychiatry. World Psychiatry 2018 Jun;17(2):133-142 [FREE Full text] [doi: 10.1002/wps.20514] [Medline: 29856558]
21. Rauschenberg C, van Os J, Goedhart M, Schieveld JN, Reininghaus U. Bullying victimization and stress sensitivity in help-seeking youth: findings from an experience sampling study. Eur Child Adolesc Psychiatry 2020 May 13:591-605. [doi: 10.1007/s00787-020-01540-5] [Medline: 32405792]
22. Rauschenberg C, van Os J, Cremers D, Goedhart M, Schieveld JN, Reininghaus U. Stress sensitivity as a putative mechanism linking childhood trauma and psychopathology in youth's daily life. Acta Psychiatr Scand 2017 Oct;136(4):373-388. [doi: 10.1111/acps.12775] [Medline: 28758672]
23. Reininghaus U, Depp CA, Myin-Germeys I. Ecological interventionist causal models in psychosis: targeting psychological mechanisms in daily life. Schizophr Bull 2016 Mar;42(2):264-269 [FREE Full text] [doi: 10.1093/schbul/sbv193] [Medline: 26707864]
24. Reininghaus U, Kempton MJ, Valmaggia L, Craig TK, Garety P, Onyejiaka A, et al. Stress sensitivity, aberrant salience, and threat anticipation in early psychosis: an experience sampling study. Schizophr Bull 2016 May;42(3):712-722 [FREE Full text] [doi: 10.1093/schbul/sbv190] [Medline: 26834027]
25. Quattrone D, Di Forti M, Gayer-Anderson C, Ferraro L, Jongsma HE, Tripoli G, EU-GEI WP2 Group, et al. Transdiagnostic dimensions of psychopathology at first episode psychosis: findings from the multinational EU-GEI study. Psychol Med 2019 Jun;49(8):1378-1391 [FREE Full text] [doi: 10.1017/S0033291718002131] [Medline: 30282569]
26. European Network of National Networks studying Gene-Environment Interactions in Schizophrenia (EU-GEI), van Os J, Rutten BP, Myin-Germeys I, Delespaul P, Viechtbauer W, et al. Identifying gene-environment interactions in schizophrenia: contemporary challenges for integrated, large-scale investigations. Schizophr Bull 2014 Jul;40(4):729-736 [FREE Full text] [doi: 10.1093/schbul/sbu069] [Medline: 24860087]
27. Morgan C, Reininghaus U, Fearon P, Hutchinson G, Morgan K, Dazzan P, et al. Modelling the interplay between childhood and adult adversity in pathways to psychosis: initial evidence from the AESOP study. Psychol Med 2014 Jan;44(2):407-419 [FREE Full text] [doi: 10.1017/S0033291713000767] [Medline: 23590972]
28. Klippel A, Myin-Germeys I, Chavez-Baldini U, Preacher KJ, Kempton M, Valmaggia L, et al. Modeling the interplay between psychological processes and adverse, stressful contexts and experiences in pathways to psychosis: an experience sampling study. Schizophr Bull 2017 Dec 01;43(2):302-315 [FREE Full text] [doi: 10.1093/schbul/sbw185] [Medline: 28204708]
29. Reininghaus U, Gayer-Anderson C, Valmaggia L, Kempton MJ, Calem M, Onyejiaka A, et al. Psychological processes underlying the association between childhood trauma and psychosis in daily life: an experience sampling study. Psychol Med 2016 Dec;46(13):2799-2813 [FREE Full text] [doi: 10.1017/S003329171600146X] [Medline: 27400863]
30. Reininghaus U, Oorschot M, Moritz S, Gayer-Anderson C, Kempton MJ, Valmaggia L, et al. Liberal acceptance bias, momentary aberrant salience, and psychosis: an experimental experience sampling study. Schizophr Bull 2019 Jun 18;45(4):871-882 [FREE Full text] [doi: 10.1093/schbul/sby116] [Medline: 30189093]
31. Hermans KS, Myin-Germeys I, Gayer-Anderson C, Kempton MJ, Valmaggia L, McGuire P, et al. Elucidating negative symptoms in the daily life of individuals in the early stages of psychosis. Psychol Med 2020 May 22:1-11. [doi: 10.1017/S0033291720001154] [Medline: 32438944]

32. Malla A, Iyer S, McGorry P, Cannon M, Coughlan H, Singh S, et al. From early intervention in psychosis to youth mental health reform: a review of the evolution and transformation of mental health services for young people. Soc Psychiatry Psychiatr Epidemiol 2016 Mar;51(3):319-326. [doi: 10.1007/s00127-015-1165-4] [Medline: 26687237]
33. McGorry P, Bates T, Birchwood M. Designing youth mental health services for the 21st century: examples from Australia, Ireland and the UK. Br J Psychiatry Suppl 2013 Jan;54:30-35. [doi: 10.1192/bjp.bp.112.119214] [Medline: 23288499]
34. Reininghaus U, Klippel A, Steinhart H, Vaessen T, van Nierop M, Viechtbauer W, et al. Efficacy of Acceptance and Commitment Therapy in Daily Life (ACT-DL) in early psychosis: study protocol for a multi-centre randomized controlled trial. Trials 2019 Dec 26;20(1):769 [FREE Full text] [doi: 10.1186/s13063-019-3912-4] [Medline: 31878966]
35. van Aubel E, Bakker JM, Batink T, Michielse S, Goossens L, Lange I, et al. Blended care in the treatment of subthreshold symptoms of depression and psychosis in emerging adults: a randomised controlled trial of Acceptance and Commitment Therapy in Daily-Life (ACT-DL). Behav Res Ther 2020 May;128:103592. [doi: 10.1016/j.brat.2020.103592] [Medline: 32146218]
36. Vaessen T, Steinhart H, Batink T, Klippel A, Van Nierop M, Reininghaus U, et al. ACT in daily life in early psychosis: an ecological momentary intervention approach. Psychosis 2019 Mar 19;11(2):93-104. [doi: 10.1080/17522439.2019.1578401]
37. Bhugra D, Tasman A, Pathare S, Priebe S, Smith S, Torous J, et al. The WPA-Lancet Psychiatry Commission on the future of psychiatry. Lancet Psychiatry 2017 Oct;4(10):775-818. [doi: 10.1016/S2215-0366(17)30333-4] [Medline: 28946952]
38. Reininghaus U. [Ecological Momentary Interventions in Psychiatry: the momentum for change in daily social context]. Psychiatr Prax 2018 Mar;45(2):59-61. [doi: 10.1055/s-0044-101986] [Medline: 29495051]
39. Myin-Germeys I, Klippel A, Steinhart H, Reininghaus U. Ecological momentary interventions in psychiatry. Curr Opin Psychiatry 2016 Jul;29(4):258-263. [doi: 10.1097/YCO.0000000000000255] [Medline: 27153125]
40. Miralles I, Granell C, Díaz-Sanahuja L, Van Woensel W, Bretón-López J, Mira A, et al. Smartphone apps for the treatment of mental disorders: systematic review. JMIR Mhealth Uhealth 2020 Apr 02;8(4):e14897 [FREE Full text] [doi: 10.2196/14897] [Medline: 32238332]
41. McKay FH, Wright A, Shill J, Stephens H, Uccellini M. Using health and well-being apps for behavior change: a systematic search and rating of apps. JMIR Mhealth Uhealth 2019 Jul 04;7(7):e11926 [FREE Full text] [doi: 10.2196/11926] [Medline: 31274112]
42. Seppälä J, De Vita I, Jämsä T, Miettunen J, Isohanni M, Rubinstein K, M-RESIST Group, et al. Mobile phone and wearable sensor-based mhealth approaches for psychiatric disorders and symptoms: systematic review. JMIR Ment Health 2019 Feb 20;6(2):e9819 [FREE Full text] [doi: 10.2196/mental.9819] [Medline: 30785404]
43. Myin-Germeys I, Kasanova Z, Vaessen T, Vachon H, Kirtley O, Viechtbauer W, et al. Experience sampling methodology in mental health research: new insights and technical developments. World Psychiatry 2018 Jun;17(2):123-132 [FREE Full text] [doi: 10.1002/wps.20513] [Medline: 29856567]
44. Batink T, Bakker J, Vaessen T, Kasanova Z, Collip D, van Os J, et al. Acceptance and commitment therapy in daily life training: a feasibility study of an mHealth intervention. JMIR Mhealth Uhealth 2016 Sep 15;4(3):e103 [FREE Full text] [doi: 10.2196/mhealth.5437] [Medline: 27634747]
45. Heron KE, Smyth JM. Ecological momentary interventions: incorporating mobile technology into psychosocial and health behaviour treatments. Br J Health Psychol 2010 Feb;15(Pt 1):1-39 [FREE Full text] [doi: 10.1348/135910709X466063] [Medline: 19646331]
46. Myin-Germeys I, Birchwood M, Kwapil T. From environment to therapy in psychosis: a real-world momentary assessment approach. Schizophr Bull 2011 Mar;37(2):244-247 [FREE Full text] [doi: 10.1093/schbul/sbq164] [Medline: 21224492]
47. Nahum-Shani I, Smith SN, Spring BJ, Collins LM, Witkiewitz K, Tewari A, et al. Just-in-Time Adaptive Interventions (JITAIs) in mobile health: key components and design principles for ongoing health behavior support. Ann Behav Med 2016 Sep 23;52(6):446-462. [doi: 10.1007/s12160-016-9830-8] [Medline: 27663578]
48. Bidargaddi N, Schrader G, Klasnja P, Licinio J, Murphy S. Designing m-Health interventions for precision mental health support. Transl Psychiatry 2020 Jul 07;10(1):222 [FREE Full text] [doi: 10.1038/s41398-020-00895-2] [Medline: 32636358]
49. Myin-Germeys I, Oorschot M, Collip D, Lataster J, Delespaul P, van Os J. Experience sampling research in psychopathology: opening the black box of daily life. Psychol Med 2009 Sep;39(9):1533-1547. [doi: 10.1017/S0033291708004947] [Medline: 19215626]
50. Oorschot M, Kwapil T, Delespaul P, Myin-Germeys I. Momentary assessment research in psychosis. Psychol Assess 2009 Dec;21(4):498-505. [doi: 10.1037/a0017077] [Medline: 19947784]
51. Collip D, Nicolson NA, Lardinois M, Lataster T, van Os J, Myin-Germeys I, G.R.O.U.P. Daily cortisol, stress reactivity and psychotic experiences in individuals at above average genetic risk for psychosis. Psychol Med 2011 Nov;41(11):2305-2315. [doi: 10.1017/S0033291711000602] [Medline: 21733219]
52. Lataster T, Wichers M, Jacobs N, Mengelers R, Derom C, Thiery E, et al. Does reactivity to stress cosegregate with subclinical psychosis? A general population twin study. Acta Psychiatr Scand 2009 Jan;119(1):45-53. [doi: 10.1111/j.1600-0447.2008.01263.x] [Medline: 18822092]
53. Myin-Germeys I, van Os J. Stress-reactivity in psychosis: evidence for an affective pathway to psychosis. Clin Psychol Rev 2007 May;27(4):409-424. [doi: 10.1016/j.cpr.2006.09.005] [Medline: 17222489]

54. Myin-Germeys I, van Os J, Schwartz JE, Stone AA, Delespaul PA. Emotional reactivity to daily life stress in psychosis. Arch Gen Psychiatry 2001 Dec;58(12):1137-1144. [doi: 10.1001/archpsyc.58.12.1137] [Medline: 11735842]
55. Palmier-Claus JE, Dunn G, Lewis SW. Emotional and symptomatic reactivity to stress in individuals at ultra-high risk of developing psychosis. Psychol Med 2012 May;42(5):1003-1012. [doi: 10.1017/S0033291711001929] [Medline: 22067414]
56. Wichers M, Schrijvers D, Geschwind N, Jacobs N, Myin-Germeys I, Thiery E, et al. Mechanisms of gene-environment interactions in depression: evidence that genes potentiate multiple sources of adversity. Psychol Med 2009 Jul;39(7):1077-1086. [doi: 10.1017/S0033291708004388] [Medline: 18834553]
57. Vaessen T, Kasanova Z, Hernaus D, Lataster J, Collip D, van Nierop M, et al. Overall cortisol, diurnal slope, and stress reactivity in psychosis: an experience sampling approach. Psychoneuroendocrinology 2018 Oct;96:61-68. [doi: 10.1016/j.psyneuen.2018.06.007] [Medline: 29906787]
58. Klippel A, Viechtbauer W, Reininghaus U, Wigman J, van Borkulo C, MERGE, et al. The Cascade of Stress: A network approach to explore differential dynamics in populations varying in risk for psychosis. Schizophr Bull 2018 Feb 15;44(2):328-337 [FREE Full text] [doi: 10.1093/schbul/sbx037] [Medline: 28338969]
59. Bell V, Freeman D. A pilot trial of cognitive behavioural therapy for interpersonal sensitivity in individuals with persecutory delusions. J Behav Ther Exp Psychiatry 2014 Dec;45(4):441-446. [doi: 10.1016/j.jbtep.2014.06.001] [Medline: 24974215]
60. Freeman D, Pugh K, Vorontsova N, Antley A, Slater M. Testing the continuum of delusional beliefs: an experimental study using virtual reality. J Abnorm Psychol 2010 Feb;119(1):83-92 [FREE Full text] [doi: 10.1037/a0017514] [Medline: 20141245]
61. Masillo A, Day F, Laing J, Howes O, Fusar-Poli P, Byrne M, et al. Interpersonal sensitivity in the at-risk mental state for psychosis. Psychol Med 2012 Sep;42(9):1835-1845. [doi: 10.1017/S0033291711002996] [Medline: 22225783]
62. Freeman D, Dunn G, Fowler D, Bebbington P, Kuipers E, Emsley R, et al. Current paranoid thinking in patients with delusions: the presence of cognitive-affective biases. Schizophr Bull 2013 Nov;39(6):1281-1287 [FREE Full text] [doi: 10.1093/schbul/sbs145] [Medline: 23223342]
63. Harb GC, Heimberg RG, Fresco DM, Schneier FR, Liebowitz MR. The psychometric properties of the Interpersonal Sensitivity Measure in social anxiety disorder. Behav Res Ther 2002 Aug;40(8):961-979. [doi: 10.1016/s0005-7967(01)00125-5] [Medline: 12186358]
64. Wilhelm K, Boyce P, Brownhill S. The relationship between interpersonal sensitivity, anxiety disorders and major depression. J Affect Disord 2004 Apr;79(1-3):33-41. [doi: 10.1016/S0165-0327(02)00069-1] [Medline: 15023478]
65. Bentall RP, de Sousa P, Varese F, Wickham S, Sitko K, Haarmans M, et al. From adversity to psychosis: pathways and mechanisms from specific adversities to specific symptoms. Soc Psychiatry Psychiatr Epidemiol 2014 Jul;49(7):1011-1022. [doi: 10.1007/s00127-014-0914-0] [Medline: 24919446]
66. Freeman D. Persecutory delusions: a cognitive perspective on understanding and treatment. Lancet Psychiatry 2016 Jul;3(7):685-692. [doi: 10.1016/S2215-0366(16)00066-3] [Medline: 27371990]
67. Gilbert P. Introducing compassion-focused therapy. Adv Psychiatr Treat 2018 Jan 02;15(3):199-208. [doi: 10.1192/apt.bp.107.005264]
68. Gilbert P. Compassion Focused Therapy (CBT Distinctive Features). London, United Kingdom: Routledge; 2010:1-248.
69. Hunot V, Moore THM, Caldwell DM, Furukawa TA, Davies P, Jones H, et al. 'Third wave' cognitive and behavioural therapies versus other psychological therapies for depression. Cochrane Database Syst Rev 2013;10:CD008704. [doi: 10.1002/14651858.CD008704.pub2] [Medline: 24142844]
70. Johns LC, Oliver JE, Khondoker M, Byrne M, Jolley S, Wykes T, et al. The feasibility and acceptability of a brief Acceptance and Commitment Therapy (ACT) group intervention for people with psychosis: the 'ACT for life' study. J Behav Ther Exp Psychiatry 2016 Mar;50:257-263. [doi: 10.1016/j.jbtep.2015.10.001] [Medline: 26480469]
71. Shawyer F, Farhall J, Thomas N, Hayes SC, Gallop R, Copolov D, et al. Acceptance and commitment therapy for psychosis: randomised controlled trial. Br J Psychiatry 2017 Feb;210(2):140-148. [doi: 10.1192/bjp.bp.116.182865] [Medline: 27979820]
72. Tai S, Turkington D. The evolution of cognitive behavior therapy for schizophrenia: current practice and recent developments. Schizophr Bull 2009 Sep;35(5):865-873 [FREE Full text] [doi: 10.1093/schbul/sbp080] [Medline: 19661198]
73. Kirby JN, Tellegen CL, Steindl SR. A meta-analysis of compassion-based interventions: current state of knowledge and future directions. Behav Ther 2017 Nov;48(6):778-792. [doi: 10.1016/j.beth.2017.06.003] [Medline: 29029675]
74. Ascone L, Sundag J, Schlier B, Lincoln TM. Feasibility and effects of a brief compassion-focused imagery intervention in psychotic patients with paranoid ideation: a randomized experimental pilot study. Clin Psychol Psychother 2017 Mar;24(2):348-358. [doi: 10.1002/cpp.2003] [Medline: 26888312]
75. Heriot-Maitland C, McCarthy-Jones S, Longden E, Gilbert P. Compassion focused approaches to working with distressing voices. Front Psychol 2019;10:152 [FREE Full text] [doi: 10.3389/fpsyg.2019.00152] [Medline: 30774614]
76. Leaviss J, Uttley L. Psychotherapeutic benefits of compassion-focused therapy: an early systematic review. Psychol Med 2015 Apr;45(5):927-945 [FREE Full text] [doi: 10.1017/S0033291714002141] [Medline: 25215860]
77. Cuppage J, Baird K, Gibson J, Booth R, Hevey D. Compassion focused therapy: exploring the effectiveness with a transdiagnostic group and potential processes of change. Br J Clin Psychol 2018 Jun;57(2):240-254. [doi: 10.1111/bjc.12162] [Medline: 29044607]

78. Lincoln TM, Hohenhaus F, Hartmann M. Can paranoid thoughts be reduced by targeting negative emotions and self-esteem? An experimental investigation of a brief compassion-focused intervention. Cogn Ther Res 2012 Aug 14;37(2):390-402. [doi: 10.1007/s10608-012-9470-7]
79. Pearson J, Naselaris T, Holmes EA, Kosslyn SM. Mental Imagery: Functional mechanisms and clinical applications. Trends Cogn Sci 2015 Oct;19(10):590-602 [FREE Full text] [doi: 10.1016/j.tics.2015.08.003] [Medline: 26412097]
80. Holmes EA, Mathews A. Mental imagery in emotion and emotional disorders. Clin Psychol Rev 2010 Apr;30(3):349-362. [doi: 10.1016/j.cpr.2010.01.001] [Medline: 20116915]
81. Braehler C, Gumley A, Harper J, Wallace S, Norrie J, Gilbert P. Exploring change processes in compassion focused therapy in psychosis: results of a feasibility randomized controlled trial. Br J Clin Psychol 2013 Jun;52(2):199-214. [doi: 10.1111/bjc.12009] [Medline: 24215148]
82. Holmes EA, Blackwell SE, Burnett Heyes S, Renner F, Raes F. Mental imagery in depression: phenomenology, potential mechanisms, and treatment implications. Annu Rev Clin Psychol 2016;12:249-280. [doi: 10.1146/annurev-clinpsy-021815-092925] [Medline: 26772205]
83. Blackwell SE, Holmes EA. Brightening the day with flashes of positive mental imagery: a case study of an individual with depression. J Clin Psychol 2017 May;73(5):579-589 [FREE Full text] [doi: 10.1002/jclp.22455] [Medline: 28152198]
84. Renner F, Ji JL, Pictet A, Holmes EA, Blackwell SE. Effects of engaging in repeated mental imagery of future positive events on behavioural activation in individuals with major depressive disorder. Cognit Ther Res 2017;41(3):369-380 [FREE Full text] [doi: 10.1007/s10608-016-9776-y] [Medline: 28515538]
85. Galante J, Dufour G, Vainre M, Wagner AP, Stochl J, Benton A, et al. A mindfulness-based intervention to increase resilience to stress in university students (the Mindful Student Study): a pragmatic randomised controlled trial. Lancet Public Health 2018 Dec;3(2):72-81 [FREE Full text] [doi: 10.1016/S2468-2667(17)30231-1] [Medline: 29422189]
86. Linardon J. Can acceptance, mindfulness, and self-compassion be learned by smartphone apps? A systematic and meta-analytic review of randomized controlled trials. Behav Ther 2020 Jul;51(4):646-658. [doi: 10.1016/j.beth.2019.10.002] [Medline: 32586436]
87. Hotopf M. The pragmatic randomised controlled trial. Adv Psychiatr Treat 2018 Jan 02;8(5):326-333. [doi: 10.1192/apt.8.5.326]
88. Loewy RL, Bearden CE, Johnson JK, Raine A, Cannon TD. The prodromal questionnaire (PQ): preliminary validation of a self-report screening measure for prodromal and psychotic syndromes. Schizophr Res 2005 Nov 01;79(1):117-125. [Medline: 16276559]
89. Ising HK, Veling W, Loewy RL, Rietveld MW, Rietdijk J, Dragt S, et al. The validity of the 16-item version of the Prodromal Questionnaire (PQ-16) to screen for ultra high risk of developing psychosis in the general help-seeking population. Schizophr Bull 2012 Nov;38(6):1288-1296 [FREE Full text] [doi: 10.1093/schbul/sbs068] [Medline: 22516147]
90. Savill M, D'Ambrosio J, Cannon TD, Loewy RL. Psychosis risk screening in different populations using the Prodromal Questionnaire: a systematic review. Early Interv Psychiatry 2018 Feb;12(1):3-14 [FREE Full text] [doi: 10.1111/eip.12446] [Medline: 28782283]
91. Derogatis L. BSI, Brief Symptom Inventory: Administration, Scoring & Procedures Manual. Minneapolis, MN: National Computer Systems; 1993:A.
92. Derogatis LR, Melisaratos N. The Brief Symptom Inventory: an introductory report. Psychol Med 1983 Aug;13(3):595-605. [Medline: 6622612]
93. Carlier I, Schulte-Van Maaren Y, Wardenaar K, Giltay E, Van Noorden M, Vergeer P, et al. Development and validation of the 48-item Symptom Questionnaire (SQ-48) in patients with depressive, anxiety and somatoform disorders. Psychiatry Res 2012 Dec 30;200(2-3):904-910. [doi: 10.1016/j.psychres.2012.07.035] [Medline: 22884307]
94. Leahy R, Tirch D, Napolitano L, Linehan M, McKay D. Emotion Regulation in Psychotherapy: A Practitioner's Guide. New York, United States: Guilford Press; 2011:1-304.
95. Pries L, Klingenberg B, Menne-Lothmann C, Decoster J, van Winkel R, Collip D, et al. Polygenic liability for schizophrenia and childhood adversity influences daily-life emotion dysregulation and psychosis proneness. Acta Psychiatr Scand 2020 May;141(5):465-475 [FREE Full text] [doi: 10.1111/acps.13158] [Medline: 32027017]
96. Myin-Germeys I, Marcelis M, Krabbendam L, Delespaul P, van Os J. Subtle fluctuations in psychotic phenomena as functional states of abnormal dopamine reactivity in individuals at risk. Biol Psychiatry 2005 Jul 15;58(2):105-110. [doi: 10.1016/j.biopsych.2005.02.012] [Medline: 16038680]
97. Green CE, Freeman D, Kuipers E, Bebbington P, Fowler D, Dunn G, et al. Measuring ideas of persecution and social reference: the Green et al. Paranoid Thought Scales (GPTS). Psychol Med 2008 Jan;38(1):101-111. [doi: 10.1017/S0033291707001638] [Medline: 17903336]
98. Corcoran R, Cummins S, Rowse G, Moore R, Blackwood N, Howard R, et al. Reasoning under uncertainty: heuristic judgments in patients with persecutory delusions or depression. Psychol Med 2006 Aug;36(8):1109-1118. [doi: 10.1017/S003329170600794X] [Medline: 16734944]
99. Bentall RP, Rowse G, Shryane N, Kinderman P, Howard R, Blackwood N, et al. The cognitive and affective structure of paranoid delusions: a transdiagnostic investigation of patients with schizophrenia spectrum disorders and depression. Arch Gen Psychiatry 2009 Mar;66(3):236-247. [doi: 10.1001/archgenpsychiatry.2009.1] [Medline: 19255373]

100. Boulet J, Boss MW. Reliability and validity of the Brief Symptom Inventory. Psychol Assess: J Consult Clin Psychol 1991 Sep;3(3):433-437. [doi: 10.1037/1040-3590.3.3.433]
101. Freeman D, Loe BS, Kingdon D, Startup H, Molodynski A, Rosebrock L, et al. The revised Green, Paranoid Thoughts Scale (R-GPTS): psychometric properties, severity ranges, and clinical cut-offs. Psychol Med 2019 Nov 20:244-253. [doi: 10.1017/S0033291719003155] [Medline: 31744588]
102. Rosenthal R, DiMatteo MR. Meta-analysis: recent developments in quantitative methods for literature reviews. Annu Rev Psychol 2001;52:59-82. [doi: 10.1146/annurev.psych.52.1.59] [Medline: 11148299]
103. Anderson R. New MRC guidance on evaluating complex interventions. Br Med J 2008 Oct 22;337:a1937. [doi: 10.1136/bmj.a1937] [Medline: 18945728]
104. Collins LM, Murphy SA, Strecher V. The multiphase optimization strategy (MOST) and the sequential multiple assignment randomized trial (SMART): new methods for more potent eHealth interventions. Am J Prev Med 2007 May;32(5 Suppl):112-118 [FREE Full text] [doi: 10.1016/j.amepre.2007.01.022] [Medline: 17466815]
105. Eldridge SM, Lancaster GA, Campbell MJ, Thabane L, Hopewell S, Coleman CL, et al. Defining feasibility and pilot studies in preparation for randomised controlled trials: development of a conceptual framework. PLoS One 2016;11(3):e0150205 [FREE Full text] [doi: 10.1371/journal.pone.0150205] [Medline: 26978655]
106. Bucci S, Schwannauer M, Berry N. The digital revolution and its impact on mental health care. Psychol Psychother 2019 Jun;92(2):277-297. [doi: 10.1111/papt.12222] [Medline: 30924316]
107. Torous J, Lipschitz J, Ng M, Firth J. Dropout rates in clinical trials of smartphone apps for depressive symptoms: a systematic review and meta-analysis. J Affect Disord 2020 Feb 15;263:413-419. [doi: 10.1016/j.jad.2019.11.167] [Medline: 31969272]
108. Cheng VW, Davenport T, Johnson D, Vella K, Hickie IB. Gamification in apps and technologies for improving mental health and well-being: systematic review. JMIR Ment Health 2019 Jun 26;6(6):e13717 [FREE Full text] [doi: 10.2196/13717] [Medline: 31244479]
109. Christie GI, Shepherd M, Merry SN, Hopkins S, Knightly S, Stasiak K. Gamifying CBT to deliver emotional health treatment to young people on smartphones. Internet Interv 2019 Dec;18:100286 [FREE Full text] [doi: 10.1016/j.invent.2019.100286] [Medline: 31890633]
110. Lattie EG, Adkins EC, Winquist N, Stiles-Shields C, Wafford QE, Graham AK. Digital mental health interventions for depression, anxiety, and enhancement of psychological well-being among college students: systematic review. J Med Internet Res 2019 Jul 22;21(7):e12869 [FREE Full text] [doi: 10.2196/12869] [Medline: 31333198]
111. Hollis C, Sampson S, Simons L, Davies EB, Churchill R, Betton V, et al. Identifying research priorities for digital technology in mental health care: results of the James Lind Alliance Priority Setting Partnership. Lancet Psychiatry 2018 Oct;5(10):845-854. [doi: 10.1016/S2215-0366(18)30296-7] [Medline: 30170964]
112. Eisele G, Vachon H, Lafit G, Kuppens P, Houben M, Myin-Germeys I, et al. The effects of sampling frequency and questionnaire length on perceived burden, compliance, and careless responding in experience sampling data in a student population. Assessment 2020 Sep 10:A. [doi: 10.1177/1073191120957102] [Medline: 32909448]

Abbreviations

BSI: Brief Symptom Inventory
CBT: cognitive behavioral therapy
CFI: compassion-focused intervention
EMA: ecological momentary assessment
EMI: ecological momentary intervention
ESM: experience sampling method
mHealth: mobile health
PQ: prodromal questionnaire
RCT: randomized controlled trial

Edited by J Torous; submitted 10.11.20; peer-reviewed by A Beck, E Kleiman; comments to author 21.12.20; revised version received 27.04.21; accepted 25.05.21; published 05.08.21.

Please cite as:
Rauschenberg C, Boecking B, Paetzold I, Schruers K, Schick A, van Amelsvoort T, Reininghaus U
A Compassion-Focused Ecological Momentary Intervention for Enhancing Resilience in Help-Seeking Youth: Uncontrolled Pilot Study
JMIR Ment Health 2021;8(8):e25650
URL: https://mental.jmir.org/2021/8/e25650
doi:10.2196/25650
PMID:34383687

Original Paper

Mental Health and the Perceived Usability of Digital Mental Health Tools Among Essential Workers and People Unemployed Due to COVID-19: Cross-sectional Survey Study

Felicia Mata-Greve[1], PhD; Morgan Johnson[1], MS; Michael D Pullmann[1,2], PhD; Emily C Friedman[1,2], CPE, MID; Isabell Griffith Fillipo[1,3], BA; Katherine A Comtois[1,4], MPH, PhD; Patricia Arean[1,2,3], PhD

[1]Department of Psychiatry and Behavioral Sciences, University of Washington, Seattle, WA, United States

[2]Advanced Laboratories for Accelerating the Reach and Impact of Treatments for Youth and Adults with Mental Illness (ALACRITY) Center, University of Washington, Seattle, WA, United States

[3]Creating Technological Innovations (CREATIV) Lab, University of Washington, Seattle, WA, United States

[4]Center for Suicide Prevention and Recovery, University of Washington, Seattle, WA, United States

Corresponding Author:
Patricia Arean, PhD
Department of Psychiatry and Behavioral Sciences
University of Washington
1959 NE Pacific Street
Seattle, WA, 98195
United States
Phone: 1 206 616 2129
Email: parean@uw.edu

Abstract

Background: COVID-19 has created serious mental health consequences for essential workers or people who have become unemployed as a result of the pandemic. Digital mental health tools have the potential to address this problem in a timely and efficient manner.

Objective: The purpose of this study was to document the extent of digital mental health tool (DMHT) use by essential workers and those unemployed due to COVID-19, including asking participants to rate the usability and user burden of the DMHT they used most to cope. We also explored which aspects and features of DMHTs were seen as necessary for managing stress during a pandemic by having participants design their own ideal DMHT.

Methods: A total of 2000 people were recruited from an online research community (Prolific) to complete a one-time survey about mental health symptoms, DMHT use, and preferred digital mental health features.

Results: The final sample included 1987 US residents that identified as either an essential worker or someone who was unemployed due to COVID-19. Almost three-quarters of the sample (1479/1987, 74.8%) reported clinically significant emotional distress. Only 14.2% (277/1957) of the sample used a DMHT to cope with stress associated with COVID-19. Of those who used DMHTs to cope with COVID-19, meditation apps were the most common (119/261, 45.6%). Usability was broadly in the acceptable range, although participants unemployed due to COVID-19 were less likely to report user burden with DMHTs than essential workers ($t_{198.1}$=–3.89, P<.001). Individuals with emotional distress reported higher financial burden for their DMHT than nondistressed individuals ($t_{69.0}$=–3.21, P=.01). When the sample was provided the option to build their own DMHT, the most desired features were a combination of mindfulness/meditation (1271/1987, 64.0%), information or education (1254/1987, 63.1%), distraction tools (1170/1987, 58.9%), symptom tracking for mood and sleep (1160/1987, 58.4%), link to mental health resources (1140/1987, 57.4%), and positive psychology (1131/1986, 56.9%). Subgroups by employment, distress, and previous DMHT use status had varied preferences. Of those who did not use a DMHT to cope with COVID-19, most indicated that they did not consider looking for such a tool to help with coping (1179/1710, 68.9%).

Conclusions: Despite the potential need for DMHTs, this study found that the use of such tools remains similar to prepandemic levels. This study also found that regardless of the level of distress or even past experience using an app to cope with COVID-19, it is possible to develop a COVID-19 coping app that would appeal to a majority of essential workers and unemployed persons.

(JMIR Ment Health 2021;8(8):e28360) doi:10.2196/28360

KEYWORDS

digital health; COVID-19; essential worker; unemployed; usability; user burden; mental health; e-mental health; survey; distress

Introduction

Background

The COVID-19 pandemic has led to necessary public health mandates, such as physical distancing and stay-at-home orders. While these orders are important to contain the outbreak, they have led to concerns about increased isolation and loneliness among the general population, and prolonged exposure to stress among essential workers (eg, those working in food distribution, construction, mail delivery, etc) and those who are unemployed or furloughed owing to the pandemic [1-4]. Rates of negative mental health outcomes, especially fear, anxiety, and stress, in the general population during this pandemic are higher compared to prepandemic times [1,5].

Individuals struggling financially are reporting challenges with job security (ie, being laid off), housing costs, and making enough money to make ends meet [6]. Essential workers and those unemployed due to COVID-19 have many unique stressors, including but not limited to, concern about COVID-19 exposure, caring for family while working or searching for work, uncertainty about their job security, financial stress, guilt about not contributing to frontline COVID-19 efforts, under- or uninsured status, and access to no or nonmedical grade personal protective equipment [1-4]. While both groups have shared concerns, recent studies have shown that half of all essential workers are likely experiencing at least one adverse mental health symptom and increased anxiety or fatigue due to work demands in high stress or changing settings [3,7]. For the unemployed, there is concern about higher rates of suicidality and suicide attempts. Previous pandemics, such as the Spanish flu of 1918 and the 2003 SARS (severe acute respiratory syndrome) epidemic, led to an increase in suicide, and loss of employment and financial stress are risk factors for suicide [4,8]. Although the recent availability of vaccines and the eventual reopening of services mean that these concerns will eventually resolve, the need to understand how to best support essential workers and unemployed people emotionally during this time is still important, as future pandemics are predicted to be likely [9], and the long-term emotional impact of the current pandemic is still unknown [10].

In response to these mental health concerns, public health systems and digital mental health companies responded by increasing access to existing technology-based care (ie, telemedicine) or modifying digital mental health tools (DMHTs), such as online resources or mobile phone apps to address perceived concerns specific to COVID-19. For example, in the United States, Medicare restrictions on telemedicine were lifted to allow for better access to health care [11]. DMHTs are also available as potential solutions to decrease stress and mental health symptoms and address the mental health care shortage during COVID-19 [8,12]. In anticipation of the need for low- or no-cost care, organizations such as the Veterans Affairs Health Care System created a free mobile app to help veterans cope with COVID-19. A report from March 2020, as physical distancing began in the United States, found that there was an increased volume of people using these tools [13]. In addition, many organizations and tech companies are turning to DMHTs to support the emotional well-being of frontline health care workers [14].

These recent events lend an important opportunity to learn about the utility of digital mental health to support populations impacted by prolonged pandemic conditions. No research has evaluated the use of DMHTs by two of the most affected populations outside of frontline health care workers and older adults or adults with disability: essential workers and those unemployed due to COVID-19. As identified in several studies, the use of DMHTs tends to be poor, with most people downloading then discontinuing use of these tools in quick succession [15,16]. As Mohr and colleagues [17] have noted, digital mental health service use could be improved if intervention developers better understood what features people felt were important to have, the usability of these tools, and what role these services should have in the context of mental wellness [18-20].

This Study

Considering the need to better understand the mental health challenges faced by essential workers and those unemployed due to COVID-19, the potential long-term effects of the societal challenges imposed by the pandemic, the potential for future pandemics, and the limited information we have on the usability and user burden of DMHTs to cope with the stress of COVID-19, we conducted a study with the following aims:

- Aim 1: Document psychological distress through clinically validated measures by the total sample, employment status (ie, unemployed due to COVID-19 and essential workers), and DMHT use (ie, reported using DMHTs to cope with COVID-19, reported not using DMHTs to cope with COVID-19);
- Aim 2: Explore DMHT use in response to COVID-19–related stress and differences by employment status and psychological distress (ie, distressed, not distressed);
- Aim 3: Assess usability and user burden ratings of DMHTs by total sample, employment status, and psychological distress;
- Aim 4: Understand the needs of these at-risk populations by identifying what DMHT features were ranked as most important by employment status, psychological distress, and DMHT use during this time.

Methods

Recruitment

A total of 2000 adults (≥18 years old) were recruited from Prolific Research Platform [21]. Using online research platforms is becoming increasingly popular in behavioral health research due its affordability, efficiency, access, and reliability [22]. Recent studies highlight that participants recruited from Prolific are more diverse and honest as well as provide higher data quality compared to other popular platforms, such as Amazon Mechanical Turk [22,23]. This national, cross-sectional study collected responses from October 26, 2020, to December 14, 2020. Participants were screened and invited to consent for participation in the anonymous, confidential survey online. Each participant was paid $3. The research was approved by the University of Washington's institutional review board.

Measures

Measures were selected and created to maximize participant engagement and reduce respondent burden. The investigative team reviewed brief measures of constructs of interest and gave preference to longer measures where no reliable or valid brief measure was available.

Inclusion Screening

Participants must have been ≥18 years old, speak English, and self-reported as either an essential worker during COVID-19 or unemployed or furloughed due to COVID-19. They also had the opportunity to indicate their current job (if an essential worker) or past job (if an unemployed worker). Participants were excluded if they were under 18 years of age, did not speak English, had no access to a mobile device (eg, smartphone or tablet), did not report being an essential worker or unemployed due to COVID-19, or lived outside of the United States.

Bad-Actor Screening

Even with the best safeguards in place, online recruitment can sometimes result in the accidental inclusion of individuals who participate in bad faith to accumulate monetary incentives ("bad actors") [24]. We instituted the procedures explained below to identify potential bad actors.

The first was to use research platforms (described above) that conduct their own extensive participant vetting. These procedures include but are not limited to: (1) every account needing a unique non-VOIP (voice over IP) phone number to verify, (2) restricting signups based on IP address and internet service provider, (3) limiting the number of accounts that can use the same IP address and machine to prevent duplicate accounts, (4) limiting the number of unique IP addresses per study, and (5) unique payment accounts (eg, PayPal) for each participant account. For example, in order to have 2 participant accounts that receive payment from Prolific, a participant would need to have 2 PayPal accounts. Payment accounts, such as PayPal, have steps to prevent duplicate accounts, such as analyzing internal data to monitor for patterns of unusual use [25].

The second method involved the use of an attention check built into our survey [26]. This method consisted of one question where participants were given this instruction: "To confirm you are paying attention, please select 'strongly disagree'" and then choices between strongly agree to strongly disagree were provided.

The third method involved the review of open-ended responses to screen out bot-like communication, repetitious, and nonsensical responses. Each of these methods confirmed that the final sample in this study could be qualified as comprising "good actors."

Demographics

Participants completed a questionnaire about demographics, which collected information about age, race, ethnicity, gender identity, sexual orientation, marital status, education, employment status, income, and living situation.

Mental Health and Possible Substance Use Disorder

Participants completed the 2-item Patient Health Questionnaire (PHQ-2) [27], the 2-item Generalized Anxiety Disorder (GAD-2) [28], and the Cut-Annoyed-Guilty-Eye Adapted to Include Drugs (CAGE-AID) [29]. The PHQ-2 and GAD-2 have good sensitivity and specificity with sensitivity to change over time in comparison to the PHQ-9 and GAD-7 [28-30]. The CAGE-AID demonstrates good sensitivity and poor specificity for substance use disorders. As a result, individuals who scored beyond the cut-off on the CAGE-AID (≥1) were categorized as a possible case of substance use disorder, in accordance with the National HIV Curriculum [29,31].

Suicidal Behaviors

Suicidal behaviors were measured using the Suicide Behaviors Questionnaire–Revised (SBQ-R) [32], a 4-item self-report measure that assesses suicide attempts, ideation, communication, and intent in one's lifetime. If the total score is greater than or equal to 7, the score is deemed to have good sensitivity and specificity for identifying individuals at risk for suicidal behaviors in a nonpsychiatric general adult population. Given some limitations of the SBQ-R, a single validated item (ie, "Have you attempted to kill yourself?") was added. The addition of this item provides further accuracy and classification of individuals at risk of suicide [33].

Psychological Distress

Participants were placed in the "distressed" category if they endorsed one or more of the clinical cut-offs, which included ≥3 on the PHQ-2 [27], ≥3 on the GAD-2 [28], ≥1 on the CAGE-AID, ≥7 [29,31] on SBQ-R [32], or reported a history of a suicide attempt [33].

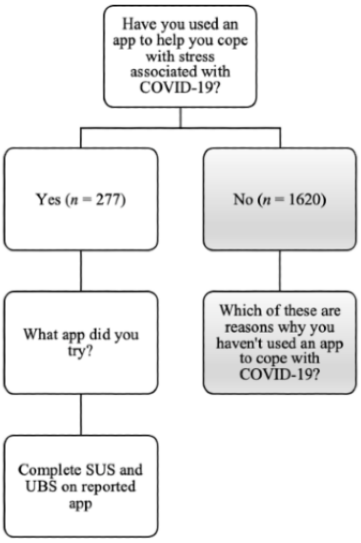
DMHT Questionnaire

This questionnaire was developed by the research team with expertise in digital mental health (author PA). The measure was tested for face validity, understandability, and respondent burden among the internal group. The questionnaire consisted of three distinct tasks: use of DMHTs during COVID-19, usability and burden of DMHTs during COVID-19, and design of an ideal DMHT for COVID-19, which are described below.

Use of DMHTs

All participants were asked whether they have used an app to cope with stress associated with COVID-19. If the participant responded yes, they were asked to list which apps they used, and if they used more than one, to list the app they used the most to cope with COVID-19. Participants were then asked to rate the app that they used most frequently in terms of features they liked, features they did not like, and then on the app's usability and user burden. If participants did not report using an app to cope with COVID-19 stress, they were asked to provide reasons for why they did not use an app (Figure 1).

Figure 1. Respondent pathway. SUS: System Usability Scale, UBS: Use Burden Scale.



Usability

Usability was measured with the System Usability Scale (SUS) [34], a 10-item measure that examines the usability of a particular intervention. The scale assesses a system's likability, learnability, complexity, need for technical support, system integration, and efficiency. The SUS is the industry standard for measuring the usability of a variety of digital tools and systems and has normative data to allow for cross system and app comparisons, even between those that are outwardly very dissimilar to one another [35].

User Burden

User burden was measured using the 20-item Use Burden Scale (UBS) [36]. This scale creates five subscales to assess different types of user burden: difficulty of use ("this app demands too much mental effort"), physical demands ("use of this app is too physically demanding"), time and social burden ("I spend too much time using this app"; "using this app has a negative impact on my social life"), mental and emotional burden ("this app presents too much information at once"), and privacy and financial burden ("the value of the app is not worth the cost for me"). This measure was developed in order to assess the adoption, retention, and experience of various technologies with the ability to compare and calibrate burden across different tools. User burden is linked to app retention and has been used in the context of mobile app research [37].

Design of a COVID-19 App

All participants, regardless of whether they reported app use for stress associated with COVID-19, were asked which features they thought would be helpful to include in an app for coping with COVID-19 (ie, information or education, meditation/mindfulness, symptom tracking, brain games, distraction tools, gratitude exercises, links to resources, chatbot, or tips to cope with COVID-19) on a scale from 0="not at all important" to 9="very important." This method of asking opinions of those who do and do not use digital technology, particularly when the needs of a given population are unknown, is commonly used in app development. The opinions of people familiar and unfamiliar with apps are needed to design a digital tool with the broadest reach [38].

After indicating which features participants preferred in an app to cope with COVID-19, they were then asked to build their own app, by selecting from a preset list of features and then adding their own desired features that were not previously listed. The app feature list was created using premade categories from One Mind Psyberguide [39], a nonprofit tool that reviews digital mental health tools for consumers, and M-Health Index and Navigation Database (MIND) [40] (see Multimedia Appendix 1 for the full survey).

Statistical Analysis

To describe the sample, we ran crosstabulations (with chi-square tests or Fisher exact tests) and independent samples *t* tests to examine possible differences in the demographic and descriptive variables by employment status (ie, unemployed vs essential worker groups) and DMHT use (ie, DMHT user vs non–DMHT user). For variables with multiple discrete categories (eg, education), if these analyses indicated a significant omnibus chi-square test, we examined standardized residuals to identify which categories were responsible for the omnibus significant difference, and reported on all categories with absolute value standardized residuals greater than 2.

For the first aim, descriptive statistics were used to document the frequencies and means of the psychological distress composite among the entire sample and stratified by employment status. We also compared those who reported using an app to cope with COVID-19 to those who reported not using an app to cope with COVID-19. Specific reports on depression, anxiety, possible substance use disorder, suicidal behavior, and history of suicide attempt may be found in Multimedia Appendix 2.

For the second aim, we calculated frequencies and differences in DMHT use for the whole sample, between essential workers and those unemployed and between those reporting distress and no distress.

For the third aim, we computed means and SDs to examine DMHT ratings from the SUS and the UBS only for those who reported using a DMHT to cope with COVID-19. Differences across the top 3 apps were assessed using an ANOVA (analysis of variance). For the sample that did not report using a DMHT to cope during COVID-19, we provided the reasons for not using a DMHT and the frequency by which those reasons were endorsed in the sample.

For the fourth aim, we computed frequencies and central tendencies of the data to assess preferred DMHT components for the whole sample and compared these findings first between essential workers and those unemployed, then between distressed and nondistressed subsamples, and finally between those who reported having used a DMHT and those who did not.

The aims described above that examined significant differences by employment, distress, or DMHT use status were assessed using chi-square tests, Fisher exact tests, or independent samples *t* tests. All statistical analyses were performed with SAS version 9.4 (SAS Institute Inc). To adjust for increased type 1 error rates due to multiple tests, we applied the Benjamini-Hochberg procedure, which applies the acceptable fraction of tests that may be erroneously statistically significant, deemed the "false discovery rate" [41,42]. We applied a false discovery rate (Q) of 10% to 119 statistical tests.

Open-ended responses from the DMHT survey for categories (ie, "What app did you try? If you tried more than one app, please pick the one you liked the most") and app features listed during the create-your-own-app survey were qualitatively coded. Like Rubanovich et al [43], the first author (FM-G) referenced the Apple App Store and Google Play to verify spelling and DMHT titles. As an example, *Calm*, *CALM*, *Calm App*, *Calm*, and *Camh* were all coded as "Calm." If a DMHT was unable to be identified via Google Play, Apple App Store, or an internet search, or the participant response was undecipherable (eg, "IDK," "NA"), it was categorized as missing (n=18).

Categorization of DMHTs was completed by authors FM-G and MJ. Informed by a modified grounded theory approach [44], each response was reviewed in order to identify meaningful units of information. Responses were compared with one another and grouped based on common responses until categories were identified. If the authors were unfamiliar with a DMHT, they read descriptions and reviews of the DMHT to determine its main feature. Some participants described DMHTs instead of names. In these cases, the response was coded for a DMHT category, but not for a specific DMHT title. As an example, the following responses, "I used a few meditation apps and one about CBT," "mindfulness app," and "meditation app" were coded into the mindfulness/meditation category. Categories and definitions were informed by Psyberguide, MIND, and experience working with digital mental health researchers. An identical process was conducted to code desired app features.

Data Exclusion and Cleaning

Duplicate cases were identified and removed. Missingness accounted for less than 5% of the data evaluated item by item. Measures were scored unless all items were missing. As an exception, PHQ-2, GAD-2, and CAGE-AID required all items to be answered to attain a final score.

Results

Sample Description

A total of 2485 participants completed the initial screener. Of this, 598 (23.7%) observations were deleted due to missing IDs, duplicate responses, "bad actors," or not meeting inclusion criteria. The final analytic sample (Table 1) consisted of 1987 adults with 1013 (50.9%) participants reporting unemployment due to COVID-19 and 974 (49.0%) identifying as an essential worker during COVID-19. The most common open-ended responses for jobs among essential workers included education, customer service or retail, management, information technology (IT), health care, pharmacy, delivery or postal work, and food service (eg, cashiers, servers, restaurant workers, grocery store workers). Although we sampled throughout the United States, compared to the US census, the majority of the overall sample was European American (1538/1987, 77.4%, compared to the US census figure of 60%), with a somewhat higher representation of Asian Americans (238/1987, 12.0% vs 5% US census) and a lower representation of African Americans (172/1987, 8.7% vs 13% US census) and Latinx Americans (212/1956, 10.8% vs 18% US census) [45]. The sample was almost split evenly between male and female (female: 1027/1987, 52.2%).

Compared to the essential workers, the unemployed group had significantly more people who identified as being: Hispanic or Latinx, or an unlisted race; younger; any gender other than male; any sexuality other than straight; and never married. The group comprised significantly less White individuals. Of note, there were almost twice as many in the "single or never married" category than what would be expected compared to the US census data [46]; however, our sample was relatively young (ie, early 30s) compared to the US population [47]. Additionally, there were socioeconomic differences across groups. Compared to the essential workers, the unemployed group had significantly more individuals with lower education, less income, and lived somewhere other than a house or apartment.

Compared to participants that did not use a DMHT to cope with COVID-19 stress, DMHT users had a significantly higher proportion of individuals who identified as transgender and a lower proportion of individuals who identified as women or men. DMHT users were more likely to be married compared to non–DMHT users. In terms of socioeconomic differences, DMHT users had a significantly smaller proportion of individuals with lower levels of education and a higher percentage of individuals with higher education compared to non–DMHT users. Finally, compared to non–DMHT users, DMHT users were less likely to live in a house and more likely to live in an apartment.

Table 1. Sample characteristics.

Characteristic	Unemployed (n=1013)	Essential worker (n=974)	*P* value	Non–DMHT user (n=1680)	DMHT user (n=277)	*P* value	Total (N=1987)
Race (not mutually exclusive), n (%)							
Asian American	129 (12.7)	109 (11.2)	.29[a]	207 (12.3)	29 (10.5)	.38[a]	238 (12.0)
European American/White	763 (75.3)	775 (79.6)	.02[a]	1300 (77.4)	225 (81.2)	.15[a]	1538 (77.4)
African American/Black	95 (9.4)	77 (7.9)	.24[a]	153 (9.1)	16 (5.8)	.07[a]	172 (8.7)
Hawaiian/Pacific Islander	8 (0.8)	3 (0.3)	.23[b]	11 (0.7)	0 (0)	.38[b]	11 (0.6)
American Indian/Alaska Native	24 (2.4)	24 (2.5)	.89[a]	36 (2.1)	9 (3.2)	.26[a]	48 (2.4)
Unlisted	52 (5.1)	22 (2.3)	<.001[a]	63 (3.8)	11 (4.0)	.86[a]	74 (3.7)
Ethnicity, (%)			.003[a]			.91[a]	
Hispanic/Latinx	128 (12.9)	84 (8.7)		179 (10.8)	30 (11.0)		212 (10.8)
Not Hispanic/Latinx	863 (87.1)	881 (91.3)		1483 (89.2)	243 (89.0)		1744 (89.2)
Age (years)			<.001[c]			.92[c]	
Mean (SD)	30.4 (11.1)	33.3 (9.9)		31.8 (10.8)	31.9 (9.7)		31.9 (10.6)
Range	18.0-73.0	18.0-78.0		18.0-78.0	18.0-73.0		18.0-78.0
Gender, n (%)			<.001[b]			<.001[b]	
Women	573 (57.4)	454 (46.9)		853 (51.1)	166 (60.1)		1027 (52.2)
Men	384 (38.4)	499 (51.5)		774 (46.3)	96 (34.8)		883 (44.9)
Nonbinary	35 (3.5)	12 (1.2)		38 (2.3)	9 (3.3)		47 (2.4)
Transgender	3 (0.3)	2 (0.2)		2 (0.1)	3 (1.1)		5 (0.3)
Unlisted	4 (0.4)	1 (0.1)		3 (0.2)	2 (0.7)		5 (0.3)
Sexuality, n (%)			<.001[a]			.30[a]	
Heterosexual/straight	681 (69.0)	802 (82.9)		1276 (76.6)	192 (71.6)		1483 (75.9)
Gay/lesbian/homosexual	69 (7.0)	41 (4.2)		89 (5.3)	20 (7.5)		110 (5.6)
Bisexual	189 (19.1)	104 (10.8)		243 (14.6)	45 (16.8)		293 (15.0)
Unlisted	48 (4.9)	20 (2.1)		57 (3.4)	11 (4.1)		68 (3.5)
Marital status, n (%)			<.001[a]			.02[a]	
Never married	737 (73.8)	500 (51.9)		1065 (63.8)	156 (57.6)		1237 (63.0)
Widowed	8 (0.8)	5 (0.5)		13 (0.8)	0 (0)		13 (0.7)
Married	177 (17.7)	402 (41.7)		473 (28.3)	101 (37.3)		579 (29.5)
Separated	15 (1.5)	7 (0.7)		21 (1.3)	1 (0.4)		22 (1.1)
Divorced	61 (6.1)	50 (5.2)		98 (5.9)	13 (4.8)		111 (5.7)
Education, n (%)			<.001[a]			<.001[a]	
High school graduate (or equivalent) or less	154 (15.3)	77 (7.9)		215 (12.8)	11 (4.0)		231 (11.7)
Some college	367 (36.5)	192 (19.7)		480 (28.6)	74 (26.7)		559 (28.3)
Trade/technical/vocational training/associate degree	125 (12.4)	108 (11.1)		209 (12.4)	22 (7.9)		233 (11.8)
Bachelor's degree	283 (28.2)	353 (36.3)		540 (32.1)	92 (33.2)		636 (32.2)
Higher education (master's, professional, or doctorate degree)	76 (7.6)	243 (25.0)		236 (14.0)	78 (28.2)		319 (16.1)

Characteristic	Unemployed (n=1013)	Essential worker (n=974)	*P* value	Non–DMHT user (n=1680)	DMHT user (n=277)	*P* value	Total (N=1987)
Income ($US), n (%)			<.001[a]			.28[a]	
<$10K	245 (25.2)	42 (4.4)		247 (15.0)	37 (13.7)		287 (14.8)
$10,000-$31,199	305 (31.3)	180 (18.7)		421 (25.6)	58 (21.4)		485 (25.1)
$31,200-$33,280	62 (6.4)	37 (3.9)		87 (5.3)	10 (3.7)		99 (5.1)
$33,281-$49,999	134 (13.8)	169 (17.6)		257 (15.6)	43 (15.9)		303 (15.7)
$50,000-$59,999	58 (6.0)	93 (9.7)		127 (7.7)	23 (8.5)		151 (7.8)
$60,000-$69,999	46 (4.7)	74 (7.7)		104 (6.3)	16 (5.9)		120 (6.2)
$70,000-$99,999	73 (7.5)	165 (17.2)		200 (12.2)	36 (13.3)		238 (12.3)
$100,000-$149,999	38 (3.9)	147 (15.3)		144 (8.8)	37 (13.7)		185 (9.6)
≥$150,000	13 (1.3)	54 (5.6)		56 (3.4)	11 (4.1)		67 (3.5)
Living situation, n (%)			<.001[a]			.004[a]	
House	611 (61.2)	624 (64.4)		1071 (64.0)	150 (54.7)		1235 (62.8)
Apartment	347 (34.7)	335 (34.6)		557 (33.3)	119 (43.4)		682 (34.7)
Other	41 (4.1)	10 (1.0)		45 (2.7)	5 (1.8)		51 (2.6)

[a]Chi-square test.

[b]Fisher exact test.

[c]Unequal variance two-sample *t* test.

Aim 1: Document Psychological Distress Among the Sample

Table 2 reports psychological distress (see the *Measures* section for calculation of the composite score) for the whole sample with stratification by employment status and DMHT-use status. We found that almost three-quarters of the sample fell into the "distressed" category (1479/1976, 74.8%), meaning they had scores at or above the clinical cut-off for at least one of the following: depression (PHQ-2), anxiety (GAD-2), risk for substance use disorder (CAGE-AID), risk for suicidal behaviors (SBQ-R), and history of suicide attempt. The unemployed group was more likely to be distressed than the essential worker group (815/1013, 81.2% vs 664/974, 68.3%; χ^2_1=43.40, *P*<.001; Table 2). DMHT users were significantly more likely to be distressed compared to non–DMHT users (236/277, 85.2% vs 1234/1680, 73.5%; χ^2_1=17.55, *P*<.001; Table 2). Table S1 in Multimedia Appendix 3 provides a further breakdown of depression, anxiety, risk for substance use disorder, risk for suicidal behaviors, and history of suicide attempt by total sample, employment status, and DMHT-use status.

Table 2. Psychological distress stratified by employment status and digital mental health tool (DMHT) use.

Variable	Unemployed (n=1013)	Essential worker (n=974)	*P* value	Non–DMHT user (n=1680)	DMHT user (n=277)	*P* value	Total (N=1987)
Psychological distress, n (%)			<.001[a,b]			<.001[a,b]	
Nondistressed	189 (18.8)	308 (31.7)		446 (26.5)	41 (14.8)		497 (25.2)
Distressed	815 (81.2)	664 (68.3)		1234 (73.5)	236 (85.2)		1479 (74.8)

[a]Chi-square test.

[b]*P* values <.05 and less than the Benjamini-Hochberg critical value were considered to be statistically significant.

Aim 2: Explore DMHT Use in Response to COVID-19

Of the 1957 participants who responded, 277 (14.2%) reported using a DMHT to cope with stress associated with COVID-19. There was no significant difference in the proportion of participants who used a DMHT in the unemployed (137/1013, 13.5%) and essential worker (140/974, 14.4%) groups (χ^2_1=0.25, *P*=.62). Distressed individuals (236/1470, 16.1%) were significantly more likely to use a DMHT app compared to nondistressed individuals (41/487, 8.4%; χ^2_1=17.55, *P*<.001).

Most Used DMHTs

Total Sample

Among the total sample, which included 261 responses, the most used DMHTs were 2 meditation apps, Calm (41/261, 15.7%) and Headspace (38/261, 14.6%), followed by BetterHelp (11/261, 4.2%). A total of 119 participants (45.6%) reported using meditation apps, 25 (9.6%) reported using virtual therapy or DMHTs that facilitated contact with a virtual provider, and 21 (8.1%) used DMHTs with a chat feature (Table 3).

Table 3. Categories of digital mental health tools (DMHTs).

Category	Definition	Participants, n (%)[a]
Meditation/mindfulness	A DMHT offering primarily meditation or mindfulness (eg, Calm, Headspace)	119 (45.6)
Virtual therapy or contact with a virtual provider	A DMHT offering primarily virtual therapy via text, phone, or video, or appointments with a physician (eg, BetterHelp, Sanvello)	25 (9.6)
Chat feature	The main feature was a chat function for one-on-one chats with a peer or chatbot, group chats, or connecting with others in an organized forum (eg, Woebot, Wysa)	21 (8.1)
Health	Tools that offer education or tips to promote healthy habits with exercise, nutrition, physical health, or sleep (eg, Downdog)	20 (7.7)
COVID-19 contact tracing	A DMHT with information related to local COVID-19 cases, rates of infection, and information about symptoms or testing (eg, Contact Tracing)	13 (5.0)
Entertainment and distraction	A DMHT with entertainment, which may include movies, music, games, GIFs, memes, or other forms of entertainment (eg, Among Us, Music app)	12 (4.6)
Social media	A social media platform (eg, TikTok, Reddit)	10 (3.8)
Symptom tracking	A DMHT that allows users to monitor symptoms or daily activities (eg, eMoods, The Pattern)	10 (3.8)
COVID-19 coping	A DMHT providing emotional coping skills and education in the context of COVID-19 stressors (eg, COVID Coach)	8 (3.1)
Positive psychology	A DMHT with gratitude exercises or methods to promote positivity, such as daily verses, positive thoughts, uplifting stories, or uplifting quotes (eg, InnerHour)	7 (2.7)
Finance	A DMHT with resources for financial decisions, financial decision-making, or spending tips (eg, Yes, Pacific)	7 (2.7)
Journal	A DMHT with primarily writing or journaling features (eg, Day One, Iona)	4 (1.5)
News	Information about international or national occurrences (eg, WHO Info)	3 (1.2)
Crisis	Using a DMHT to manage crisis or safety (eg, suicide)	1 (0.4)
Language learning	Using a DMHT in order to practice or learn a new language	1 (0.4)

[a]A total of 18 responses were coded as "missing" due to being indecipherable or unidentifiable; percentages do not reflect missingness.

Employment Status

The leading entries by the unemployed sample were 3 meditation apps: Calm (26/131, 19.8%), Headspace (22/131, 16.8%), and Insight Timer (7/131, 5.3%). The most common DMHT categories among individuals unemployed due to COVID-19 were meditation (70/131, 53.4%), virtual therapy or DMHTs that facilitated virtual contact with a mental health provider (11/131, 8.4%), and DMHTs with a chatbot (11/131, 8.4%). The most frequently reported DMHTs by the essential worker sample were Headspace (16/130, 12.3%), Calm (15/130, 11.5%), and COVID Coach (8/130, 6.2%). By category, essential workers reported using mostly meditation (49/130, 37.7%), DMHTs with virtual therapy or contact with a virtual provider (14/130, 10.8%), health DMHTs (12/130, 9.4%), and COVID-19 contact tracing (12/130, 9.4%).

Distress Status

Similarly, the leading entries by the distressed sample were 2 meditation apps, Calm (33/223, 14.8%) and Headspace (32/223, 14.3%), followed by BetterHelp (10/223, 4.5%). Most of the distressed sample used meditation (100/223, 44.8%), virtual therapy or contact with a virtual provider (24/223, 10.8%), and DMHTs with a chat feature (19/223, 8.5%). The most frequently reported DMHTs by the nondistressed group were Calm, (8/38, 21.1%), Headspace (6/38, 15.8%), and COVID Coach (2/38 5.3%). Among the individuals in the nondistressed group, the most frequently used app categories were meditation (9/38, 50%), COVID-19 contact tracing (4/38, 10.5%), and social media (3/38, 7.9%).

Further comparisons of app categories by employment and distress statuses may be found in Table S2 (Multimedia Appendix 3).

Reasons for Lack of Use

Most of the sample (1710/1957, 85.9%) reported that they did not use a DMHT to cope with COVID-19. The primary reasons for not using a DMHT to cope with COVID-19 were (1) not thinking to look for an app (1179/1710, 68.9%), (2) not thinking apps would help them (605/1710, 35.4%), and (3) having other ways of coping (421/1710, 24.6%). Table S3 in Multimedia Appendix 3 lists all reasons for lack of use. These top 3 responses were endorsed by all subgroups.

There were differences that emerged by employment status and distress status. Compared to essential workers, those who were unemployed due to COVID-19 were more likely to report not thinking to look for a DMHT (629/876, 71.8% vs 550/834, 65.9%; χ^2_1=6.84, *P*=.009) and not having money to spend on a data plan to use a DMHT (112/876, 12.8% vs 54/834, 6.5%; χ^2_1=19.41, *P*<.001).

Compared to the nondistressed group, distressed individuals were more likely to not think to look for an app (293/456, 64.3% vs 886/1243, 71.3%; χ^2_1=7.75, *P*=.005), to not think apps would help them (142/456, 31.1% vs 463/1243, 37.2%; χ^2_1=5.43, *P*=.02), to prefer working with a professional (32/456, 7.0% vs 191/1243, 15.4%; χ^2_1=20.39, *P*<.001), to not have money to spend on a data plan to use apps (25/456, 5.5% vs 141/1243, 11.3%; χ^2_1=13.00, *P*<.001), and to not find an app that was relevant to their needs (19/456, 4.2% vs 103/1243 8.3%; χ^2_1=8.50, *P*=.004). However, compared to nondistressed individuals, distressed workers were less likely to state that having another way of coping was the reason for why they did not use a DMHT (281/1243, 22.6% vs 140/456, 30.7%; χ^2_1=11.73, *P*<.001).

Aim 3: Assess DMHT Usability and User Burden

Data for the following analyses were taken from the 277 participants who reported using a DMHT to cope with COVID-19. Individuals who did not report using a DMHT to cope with COVID-19 did not complete the SUS or UBS (Figure 1).

Employment Status

As shown in Table 4, compared to the essential workers, those who were unemployed due to COVID-19 reported significantly less user burden when using DMHTs (mean 13.69, SD 17.76 vs mean 7.23, SD 8.24; $t_{198.1}$=–3.89, *P*<.001). Specifically, those who were unemployed rated their selected DMHT as being significantly less difficult to use (mean 2.77, SD 4.02 vs mean 1.53, SD 2.15; $t_{214.5}$=3.20, *P*=.002), and having less physical burden (mean 1.54, SD 2.89 vs mean 0.43, SD 1.37; $t_{198.3}$=4.06, *P*<.001), time and social burden (mean 2.60, SD 4.00 vs mean 1.07, SD 2.15; $t_{215.1}$=3.95, *P*<.001), mental and emotional burden (mean 2.46, SD 3.92 vs mean 1.07, SD 2.08; $t_{213.4}$=3.69, *P*<.001), and privacy burden (mean 2.30, SD 3.16 vs mean 1.25, SD 2.14; $t_{245.1}$=3.25, *P*=.001). The conditions did not differ for reports of financial burden (mean 2.04, SD 2.33 vs mean 1.88, SD 2.47; t_{272}=–0.53 *P*=.59). In addition, there was no significant difference in ratings of usability between unemployed individuals (mean 76.96, SD 16.21) and essential workers (mean 74.32, SD 17.01; t_{271}=–1.31, *P*=.19).

Table 4. User burden and system usability stratified by workers and psychological distress.

Variable	Unemployed (n=137)	Essential worker (n=140)	*P* value	Nondistressed (n=41)	Distressed (n=236)	*P* value	Total (N=277)
Overall burden			<.001[b,c]			.30[a]	
Count, n	134	140		41	233		274
Mean (SD)	7.2 (8.2)	13.7 (17.8)		8.4 (13.8)	10.9 (14.4)		10.5 (14.3)
Difficulty of use			.002[b,c]			.42[a]	
Count, n	134	140		41	233		274
Mean (SD)	1.5 (2.2)	2.8 (4.0)		1.8 (3.8)	2.2 (3.2)		2.2 (3.3)
Physical burden			<.001[b,c]			.52[a]	
Count, n	134	139		40	233		273
Mean (SD)	0.4 (1.4)	1.5 (2.9)		0.8 (2.0)	1.0 (2.4)		1.0 (2.3)
Social and time burden			<.001[b,c]			.99[a]	
Count, n	134	140		41	233		274
Mean (SD)	1.1 (2.2)	2.6 (4.0)		1.9 (3.3)	1.9 (3.3)		1.9 (3.3)
Mental and emotional burden			<.001[b,c]			.80[a]	
Count, n	134	140		41	233		274
Mean (SD)	1.1 (2.1)	2.5 (3.9)		1.7 (3.3)	1.8 (3.2)		1.8 (3.2)
Privacy burden			.001[b,c]			.22[a]	
Count, n	134	140		41	233		274
Mean (SD)	1.2 (2.1)	2.3 (3.2)		1.3 (2.6)	1.9 (2.8)		1.8 (2.8)
Financial burden			.59[a]			.002[b,c]	
Count, n	134	140		41	233		274
Mean (SD)	1.9 (2.5)	2.0 (2.3)		1.1 (1.8)	2.1 (2.5)		2.0 (2.4)
System Usability Score			.19[a]			.96[a]	
Count, n	134	139		40	233		273
Mean (SD)	77.0 (16.2)	74.3 (17.0)		75.5 (17.6)	75.6 (16.5)		75.6 (16.6)

[a]Equal variance two-sample *t* test.

[b]Unequal variance two-sample *t* test.

[c]*P* values <.05 and less than the Benjamini-Hochberg critical value were considered to be statistically significant.

Distress Status

As shown in Table 4, there was no difference in reported DMHT burden between the distressed and nondistressed subsamples (mean 10.91, SD 14.37 vs mean 8.41, SD 13.82; t_{272}=–1.03, *P*=.30) or in overall usability (mean 75.63, SD 16.52 vs mean 75.50, SD 17.59; t_{271}=–0.05, *P*=.96). Likewise, we found no difference between groups in types of burden (Table 4). The one exception was that distressed individuals reported higher financial burden for their selected DMHT than nondistressed individuals (mean 2.12, SD 2.46 vs mean 1.07, SD 1.81; $t_{69.0}$=–3.21, *P*=.01).

Finally, we explored the user burden and usability ratings of the three most used apps (ie, Calm, Headspace, and BetterHelp; shown in Table S4 in Multimedia Appendix 3). There were no statistically significant differences among the apps in terms of the total SUS, total UBS, and UBS subscales, except for privacy burden (Calm: mean 1.54, SD 2.82 vs Headspace: mean 0.50, SD 1.03 vs BetterHelp: mean 2.00, SD 2.14; $F_{2,87}$=3.25, *P*=.04).

Aim 4: Identify Important DMHT Features

Total Sample

The sample reported the following top-rated features for DMHTs: (1) information or education (mean 6.09, SD 2.66); (2) mindfulness or meditation tools (mean 6.06, SD 2.59); (3) link to resources, counseling, or crisis support (mean 5.93, SD 2.80); and (4) tools to focus on positive events and influences in life (mean 5.88, SD 2.46).

Participants also had the option to write in what DMHT features they felt were important to include but were not provided in the list of options. The top suggested features among the 764 responses were the ability to chat with a mental health professional, support personnel, or peer (n=57); entertainment and distraction (n=39); and positive psychology (n=29). The feature “entertainment and distraction” was defined as “different forms of entertainment such as music, movies, movie clips, GIFs, memes, games, or other forms of distraction.” Additionally, participants reported wanting regularly occurring (ie, daily) gratitude exercises or activities to promote positivity, such as verses, quotes, and uplifting or hopeful stories, which we categorized as “positive psychology” features. Example responses included: “give positive messages in the morning or something like that,” “daily gratitude,” and “a good news section… I don’t want to be told COVID-19 isn’t a problem. I want to know what hope there is.”

When provided the option to build their own app, the sample most frequently endorsed the following features: mindfulness/meditation (1271/1987, 64.0%), information or education (1254/1987, 63.1%), and distraction tools (1170/1987, 58.9%) (Table 5).

Table 5. Digital mental health tool (DMHT) features stratified by worker status and psychological distress.

Feature	Unemployed (n=1013), n (%)	Essential worker (n=974), n (%)	*P* value	Nondistressed (n=497), n (%)	Distressed (n=1479), n (%)	*P* value
Mindfulness/meditation	687 (67.8)	584 (60.0)	<.001[a,b]	305 (61.4)	966 (65.3)	.11[a]
Information or education	636 (62.8)	618 (63.4)	.76[a]	327 (65.8)	927 (62.7)	.21[a]
Distraction tools (drawing, puzzles, music)	630 (62.2)	540 (55.4)	.002[a,b]	276 (55.5)	894 (60.4)	.05[a]
Symptom tracking (tracking sleep or mood)	605 (59.7)	555 (57.0)	.22[a]	270 (54.3)	890 (60.2)	.02[a,b]
Link to resources, counseling, or crisis support	604 (59.6)	536 (55.0)	.04[a,b]	276 (55.5)	864 (58.4)	.26[a]
Tools to focus on the positive events and influences in life	578 (57.1)	553 (56.8)	.90[a]	267 (53.7)	864 (58.4)	.07[a]
Brain games to improve thinking	525 (51.8)	480 (49.3)	.26[a]	257 (51.7)	748 (50.6)	.66[a]
How to cope with COVID-19	406 (40.1)	409 (42.0)	.39[a]	200 (40.2)	615 (41.6)	.60[a]
A chatbot to help you with daily stress	352 (34.7)	293 (30.1)	.03[a,b]	139 (28.0)	506 (34.2)	.01[a,b]

[a]Chi-square test.

[b]*P* values <.05 and less than the Benjamini-Hochberg critical value were considered to be statistically significant.

Employment Status

The three most important DMHT components for essential workers and unemployed individuals were information or education (essential: mean 6.09, SD 2.70; unemployed: mean 6.09, SD 2.61); mindfulness/meditation (essential: mean 6.17, SD 2.55; unemployed: mean 5.94, SD 2.62); and link to resources, counseling, or crisis support (essential: 6.00, SD 2.89; unemployed: mean 5.86, SD 2.72). Unemployed participants were more likely to rate distraction tools (drawing, puzzles, and music) (mean 5.84, SD 2.55 vs mean 5.42, SD 2.59; t_{1945}=3.59, *P*<.001) and mindfulness/meditation (mean 6.17, SD 2.55 vs mean 5.94, SD 2.61; t_{1945}=2.02, *P*=.04) as more important than essential workers.

When provided the option to build their own DMHT, the most common features listed by essential workers were information and education (618/974, 63.4%), mindfulness/meditation (584/974, 60.0%), and symptom tracking (tracking sleep or mood; 555/974, 57%). The most common features reported by unemployed persons was mindfulness/meditation (687/991, 67.8%), information or education (636/991, 62.8%), and distraction tools (eg, drawing, puzzles, music) (630/991, 62.2%). In comparing the desired features for a DMHT by employment status, unemployed participants were more likely to request that their DMHT include mindfulness/meditation (687/1013, 67.8% vs 584/974, 60.0%; χ^2_1=13.31, *P*<.001); distraction tools (drawing, puzzles, and music; 630/1013, 62.2% vs 540/974, 55.4%; χ^2_1=9.34, *P*=.002); link to resources, counseling, or crisis support (604/1013, 59.6%, vs 536/974, 55.0%; χ^2_1=4.29, *P*=.04); and a chatbot to help with daily stress (352/1013, 34.7%, vs 293/974, 30.1%; χ^2_1=4.93, *P*=.03) than the essential worker group (Table 5).

Distress Status

The most important DMHT components among distressed and nondistressed users included information or education (distressed: mean 6.01, SD 2.67; nondistressed: mean 6.32, SD 2.59); mindfulness/meditation (distressed: mean 6.09, SD 2.56; nondistressed: mean 5.96, SD 2.68); and link to resources, counseling, or crisis support (distressed: mean 5.95, SD 2.81; nondistressed: mean 5.88, SD 2.80). Distressed individuals also rated tools to focus on positive life events and influences as important (mean 5.90, SD 2.42).

When provided the option to build their own DMHT, nondistressed individuals indicated information or education (327/497, 65.8%), followed by mindfulness/meditation (305/497, 61.4%), distraction tools (276/497, 55.5%), and link to resources, counseling, or crisis support (276/497, 55.5%). Similarly, distressed individuals desired to include mindfulness/meditation (966/1479, 65.3%), followed by information or education (927/1479, 62.7%) and distraction tools (894/1479, 60.4%). Compared to nondistressed individuals, distressed participants preferred to include symptom tracking (270/497, 54.3% vs 890/1479, 60.2%; χ^2_1=5.25, *P*=.02) and a chatbot (139/497, 28.0% vs 506/1479, 34.2%; χ^2_1=6.60, *P*=.01) within their DMHT (Table 5).

DMHT Use Status

Participants who used DMHTs to cope during COVID-19 reported the following features as having the highest importance for a DMHT: (1) mindfulness/meditation (mean 7.10, SD 2.05); (2) tools to focus on the positive events and influences in life (mean 6.23, SD 2.24); (3) link to resources, counseling, or crisis support (mean 5.94, SD 2.64); and (4) symptom tracking (mean 5.90, SD 2.40). On the other hand, non–DMHT users indicated their most important features were (1) information or education (mean 6.14, SD 2.69); (2) link to resources, counseling, or crisis support (mean 5.93, SD 2.83); and (3) mindfulness/meditation tools (mean 5.89, SD 2.63).

When asked to build their own DMHT, individuals who did not use a DMHT to cope during the COVID-19 pandemic preferred to include information or education (1091/1680, 64.9%), mindfulness/meditation (1071/1680, 63.8%), and distraction tools (1031/1680, 61.4%). DMHT users preferred to include mindfulness/meditation (200/277, 72.2%), tools to focus on the positive events and influences in life (178/277, 64.3%), and symptom tracking (tracking sleep or mood; 166/277, 59.9%).

Participants who used DMHTs to cope during COVID-19 were more likely than those who did not use DMHTs to prefer mindfulness/meditation features (200/277, 72.2% vs 1071/1680, 63.8%; χ^2_1=7.46, *P*=.006), positive psychology features (178/277, 64.3% vs 953/1680, 56.7%; χ^2_1=5.53, *P*=.02), and chatbot features (108/277, 39.0% vs 537/1680, 32.0%; χ^2_1=5.31, *P*=.02). Conversely, compared to non–DMHT users, DMHT users were less likely to prefer brain games to improve thinking (124/277, 44.8% vs 881/1680, 52.4%; χ^2_1=5.61, *P*=.02), and distraction tools (139/277, 50.2% vs 1031/1680, 61.4%; χ^2_1=12.38, *P*<.001) (Table 6).

Table 6. Digital mental health tool (DMHT) features stratified by user status.

Feature	Non–DMHT user (n=1680), n (%)	DMHT user (n=277) , n (%)	*P* value	Total (N=1987), n (%)
Mindfulness/meditation	1071 (63.8)	200 (72.2)	.006[a,b]	1271 (64.0)
Information or education	1091 (64.9)	163 (58.8)	.05[a]	1254 (63.1)
Distraction tools (drawing, puzzles, music)	1031 (61.4)	139 (50.2)	<.001[a,b]	1170 (58.9)
Symptom tracking (tracking sleep or mood)	994 (59.2)	166 (59.9)	.81[a]	1160 (58.4)
Link to resources, counseling, or crisis support	986 (58.7)	154 (55.6)	.33[a]	1140 (57.4)
Tools to focus on the positive events and influences in life	953 (56.7)	178 (64.3)	.02[a,b]	1131 (56.9)
Brain games to improve thinking	881 (52.4)	124 (44.8)	.02[a,b]	1005 (50.6)
How to cope with COVID-19	689 (41.0)	126 (45.5)	.16[a]	815 (41.0)
A chatbot to help you with daily stress	537 (32.0)	108 (39.0)	.02[a,b]	645 (32.5)

[a]Chi-square test.

[b]*P* values <.05 and less than the Benjamini-Hochberg critical value were considered to be statistically significant.

Discussion

Principal Findings

This study documented DMHT use among essential workers and unemployed individuals during the COVID-19 pandemic and determined which features such users would prefer to have in a DMHT offering. DMHT use has been deemed by many in the field to be subpar, and some have suggested that poor uptake and adherence to such tools is the result of user burden and inadequate match to user needs [17]. Indeed, our findings indicate that despite reports of increased downloads [48] and user registration by digital mental health companies [13], use of DMHTs by essential workers and those unemployed due to COVID-19 is very similar to prepandemic reports (14%). Compared to our study (14%), previous studies found that 10% of outpatient psychiatric clinic patients used a DMHT [49] and only 17% of a sample with no self-reported mental health distress report downloading an app “to help relax” [50].

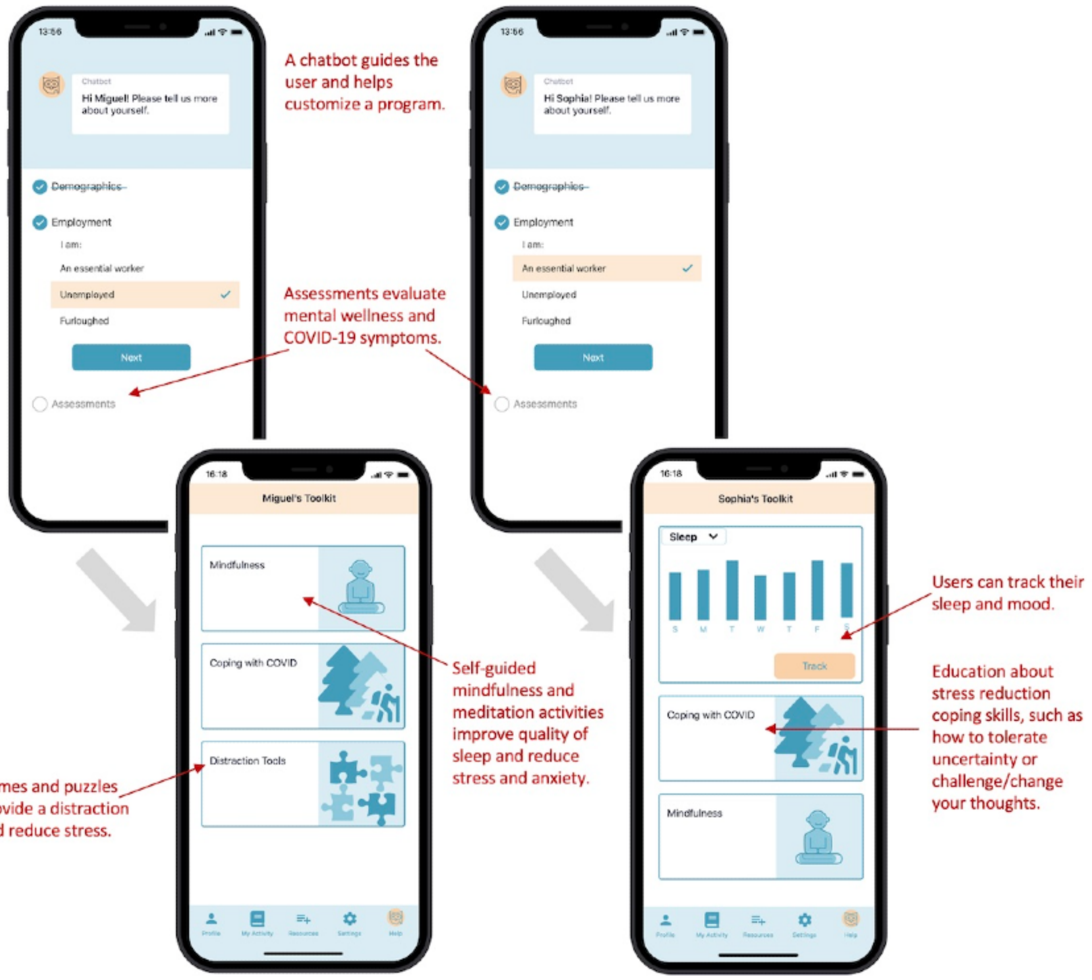
Of those who reported using a DMHT, by far the most common DMHTs were those that focused on mindfulness/meditation strategies (46%), with access to virtual therapy (10%) in second place. This finding did not vary by level of distress or employment status except among the nondistressed group using COVID-19 contact tracing (8% of this subsample). This finding is nearly identical to another recent study that found that Calm and Headspace were the top 2 downloaded apps among iPhone users during COVID-19 [48].

Additionally, when asked to rate the usability and user burden of the DMHT tool participants used the most, system usability fell in the “acceptable” range [34], and time, mental and emotional, physical, financial, and privacy burdens were seen as acceptable, with essential workers finding these tools to be more burdensome than the unemployed group. Increased perceived burdensome may be partially explained by previous findings suggesting that essential workers have increased fatigue from elevated anxiety and work demands during the ongoing pandemic [3].

Individuals with increased mental health needs (ie, the distressed group) reported more financial burden of DMHTs than the nondistressed. It is understandable that during a pandemic, where people are struggling financially, there would be concerns about the costs of DMHTs, given that many popular and widely publicized tools require a paid subscription. In the United States, those who lost their jobs during COVID-19 are faced with insufficient insurance to cover the costs of mental health care [51], and those who are struggling financially likely have additional financial concerns aside from a DMHT subscription fee, such as the cost of data plans and the technology needed to use these services. In fact, an earlier study noted that most individuals with depression and/or anxiety symptoms preferred using health apps that were free or had low cost for download (eg, <$5) [43]. As such, reimbursement is one part of the solution for increasing access to care for everyone, and until technology is more affordably available to all, the use of these services will be compromised [52].

When asked to design their own DMHT for coping with COVID-19, again mindfulness/meditation was listed as an important feature for all subgroups in this study. Interestingly, information and education about COVID-19 was also consistently listed as an important feature in all subgroups except for people who had used DMHTs during the pandemic. In addition to mindfulness/meditation, people who used DMHTs to cope with COVID-19 preferred positive psychology tools and mood and sleep tracking. Figure 2 illustrates the preferences between the unemployed and essential worker groups. This finding has important implications for DMHT development focused on pandemic response and other prolonged environmental disasters. Developers would be able to create a single tool that includes mindfulness/meditation, information and education about COVID-19 coping, and distraction tools, which would appeal to a wide group of people with different needs during COVID-19, with only a few added features for specific populations.

Figure 2. Preferred digital mental health tool features according to participants.



A final finding in this study was reasons for not turning to DMHTs to cope with COVID-19. Most of the sample indicated that they did not use a DMHT because they did not think to look for such a tool. Past reports suggest that this result may be due to a lack of information about how DMHTs might be effective [53]. This assumption is further supported by the fact that one-third of the sample did not think a DMHT would be helpful to them, and one-quarter of the sample indicated that they had other means of coping. The potential lack of confidence in DMHTs might be addressed through education to health providers on the effectiveness of DMHTs [54], the creation of reimbursement codes in the United States that would allow providers to prescribe these services [55], or the further use of a human-centered design from DMHT companies to create tools that are appropriately targeting user needs and concerns.

Comparison With Prior Work

A strength of this study is that we explicitly asked a large sample of users about their app preferences and perceived importance of various features. This survey was different from previous studies that have primarily focused on downloads and user metrics [48], insight from providers and private digital health companies [56], and self-report from individuals exclusively with mild depression or mild anxiety symptoms with exclusion of severer mental health conditions (eg, suicidality) [43]. It is also novel in its consideration of user-centered design principles (eg, ease of use and learnability) when developing and identifying DMHT features that would be most acceptable to a very large sample of potential target consumers. Consistent with emerging models that integrate community-based research, implementation science, and user-centered design principles [57,58], this is an important first step in a well-planned process of DMHT design to identify the needs and preferred features that users, both experienced and unexperienced, and preferences for what tools they would like to see in a DMHT. Previous studies that used self-report of physical health and mental health apps found that users typically only use an app for one feature [43]. It might be that future apps need to have multiple features incorporated to meet the overarching needs of similar populations. As Mohr et al [17] have noted previously, health app developers tend to create a tool based on what the developer feels is essential and historically only designs around these developer-driven features, rather than asking the end-user what role they see digital health playing in their lives, what needs they have that are unmet, and what functions they want these tools to have. By starting with understanding end-user needs and preferences, DMHT developers may see not only an increase in DMHT uptake but long-term use as well.

The findings of this study differ from findings in recent studies on the use of technology to cope with the consequences of COVID-19. According to recent research in the general population, there has been increased desire for apps or online resources that allow for fitness at home, owing to physical distancing and stay-at-home orders that have led to a shift from gyms and group fitness classes to exercise at home [59]. During prepandemic times, Rubanovich et al [43] found that people with depression and anxiety symptoms reported more frequently using health apps featuring fitness, pedometers, or heart rate monitoring apps than DMHTs. Conversely, in our study, fitness apps and tools were listed very low in the list of tools participants used for coping with COVID-19. Although studies on the use of fitness apps among essential workers and employee groups are sparse, existing research suggests that the use of such tools in practice is low [60], which may explain why these tools were not in the top group of DMHTs listed by these participants. According to past research, those who are unemployed may likewise not have resources to engage in fitness apps, and generally are less likely to engage in fitness tracking [61]. Finally, another COVID-19 study found that more contact tracing and COVID-19 informational apps were being downloaded than DMHTs in North America [62]. We note here that downloads are often not equivalent to tool use as recent research has found that many people do download such tools but rarely use them long term [20,63]. Our study specifically asked about which DMHTs people used to cope with COVID-19 stress.

Our study adds to the existing body of work by understanding how DMHTs could be made to be more accessible to those at risk for the emotional consequences of COVID-19. Many experts in digital mental health have argued for the need to better personalize such tools [54] and to include the perspectives of the intended consumer in the design of such tools [8].

Limitations

Although this study has important implications regarding the use of DMHTs from a human-centered design approach, it does have limitations. First, this is a cross-sectional study surveying the US participants' experiences and opinions at one point in time. Second, the participants of this sample are likely to be more accepting of digital tools, as they were recruited from an online research platform. As such, the information from this study is limited to those who are currently using and familiar with technology. Third, this study did not consider cross-cultural acceptance of DMHTs, which is an important caveat since a DMHT may be different in countries that already support such tools as part of their health care system. Fourth, we are unable to explicitly comment on the sample's overall experience with apps or DMHTs during prepandemic times. The focus of this paper was to explore whether users were using available, low-cost DMHTs to address COVID-19–related stress. Future studies should conduct a more thorough assessment of both current and previous DMHT use.

Conclusions

Despite the limitations, this study provides important information to the mental health care system and to those who develop and provide DMHTs during prolonged stressful events. Policy makers and providers may not be able to rely on existing DMHTs to address the emotional health of essential workers and people who are unemployed. This study points to the need to ensure DMHTs address the needs that the intended consumer feels is most important, that these tools are not burdensome under high-stress conditions, and that they are affordable to people who have limited means.

Acknowledgments

The research reported in this publication was supported by the National Institutes of Health (grant P50MH115837 and T32MH020021). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Authors' Contributions

PA, KAC, and FM-G contributed to study concepts and design. PA and KAC obtained funding. MJ and FM-G conducted or interpreted the statistical analyses. MP consulted on the analytic approach. FM-G, MJ, and PA drafted the manuscript with contributions from all the authors. All authors read and approved the final manuscript for publication.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Consent form and COVID-19 mental health apps survey.
[DOC File , 92 KB - mental_v8i8e28360_app1.doc]

Multimedia Appendix 2
Distress measures stratified by app users.
[DOC File , 74 KB - mental_v8i8e28360_app2.doc]

Multimedia Appendix 3
Additional supplementary tables.
[DOCX File , 53 KB - mental_v8i8e28360_app3.docx]

References

1. Coping with Stress. Centers for Disease Control and Prevention. 2020. URL: https://www.cdc.gov/coronavirus/2019-ncov/daily-life-coping/managing-stress-anxiety.html [accessed 2021-07-05]
2. Fitzpatrick KM, Harris C, Drawve G. Fear of COVID-19 and the mental health consequences in America. Psychol Trauma 2020 Aug;12(S1):S17-S21. [doi: 10.1037/tra0000924] [Medline: 32496100]
3. Benhamou K, Piedra A. CBT-Informed Interventions for Essential Workers During the COVID-19 Pandemic. J Contemp Psychother 2020 Jul 18;50(4):275-283. [doi: 10.1007/s10879-020-09467-3]
4. Gunnell D, Appleby L, Arensman E, Hawton K, John A, Kapur N, et al. Suicide risk and prevention during the COVID-19 pandemic. The Lancet Psychiatry 2020 Jun;7(6):468-471. [doi: 10.1016/s2215-0366(20)30171-1]
5. Mental Health and COVID-19. World Health Organization. 2021. URL: https://www.who.int/teams/mental-health-and-substance-use/covid-19 [accessed 2021-07-05]
6. Stress in America 2020: A National Mental Health Crisis. American Psychological Association. 2020. URL: https://www.apa.org/news/press/releases/stress/2020/report-october [accessed 2021-07-05]
7. Czeisler ME, Lane RI, Petrosky E, Wiley JF, Christensen A, Njai R, et al. Mental Health, Substance Use, and Suicidal Ideation During the COVID-19 Pandemic — United States, June 24–30, 2020. MMWR Morb Mortal Wkly Rep 2020 Aug 14;69(32):1049-1057. [doi: 10.15585/mmwr.mm6932a1]
8. Sher L. The impact of the COVID-19 pandemic on suicide rates. QJM: International Journal of Medicine 2020 2020 Jun 30;113(10):707-712. [doi: 10.1093/qjmed/hcaa202]
9. Dodds W. Disease now and potential future pandemics. In: The World's Worst Problems. Cham: Springer; 2019:31-44.
10. Tokic D. Long-term consequences of the 2020 coronavirus pandemics: Historical global ‐ macro context. J Corp Acct Fin 2020 Jun 08;31(3):9-14. [doi: 10.1002/jcaf.22448]
11. Telehealth: Delivering care safely during COVID-19. US Department of Health And Human Services. 2020. URL: https://www.hhs.gov/coronavirus/telehealth/index.html [accessed 2021-07-05]
12. Cénat JM, Dalexis RD, Kokou-Kpolou CK, Mukunzi JN, Rousseau C. Social inequalities and collateral damages of the COVID-19 pandemic: when basic needs challenge mental health care. Int J Public Health 2020 Jul 10;65(6):717-718 [FREE Full text] [doi: 10.1007/s00038-020-01426-y] [Medline: 32651593]
13. Leading health and fitness apps in the google play store worldwide in March 2020. Statista. 2020. URL: http://www.statista.com/statistics/695697/top-android-health-apps-in-google-play-by-revenue/ [accessed 2021-07-05]
14. Drissi N, Alhmoudi A, Al Nuaimi H, Alkhyeli M, Alsalami S, Ouhbi S. Investigating the Impact of COVID-19 Lockdown on the Psychological Health of University Students and Their Attitudes Toward Mobile Mental Health Solutions: Two-Part Questionnaire Study. JMIR Form Res 2020 Oct 20;4(10):e19876 [FREE Full text] [doi: 10.2196/19876] [Medline: 32969340]

15. Rathbone AL, Prescott J. The Use of Mobile Apps and SMS Messaging as Physical and Mental Health Interventions: Systematic Review. J Med Internet Res 2017 Aug 24;19(8):e295 [FREE Full text] [doi: 10.2196/jmir.7740] [Medline: 28838887]
16. Marshall JM, Dunstan DA, Bartik W. Clinical or gimmickal: The use and effectiveness of mobile mental health apps for treating anxiety and depression. Aust N Z J Psychiatry 2019 Sep 25;54(1):20-28. [doi: 10.1177/0004867419876700]
17. Mohr DC, Weingardt KR, Reddy M, Schueller SM. Three Problems With Current Digital Mental Health Research... and Three Things We Can Do About Them. Psychiatr Serv 2017 May 01;68(5):427-429. [doi: 10.1176/appi.ps.201600541] [Medline: 28412890]
18. Zhang R, Nicholas J, Knapp AA, Graham AK, Gray E, Kwasny MJ, et al. Clinically Meaningful Use of Mental Health Apps and its Effects on Depression: Mixed Methods Study. J Med Internet Res 2019 Dec 20;21(12):e15644 [FREE Full text] [doi: 10.2196/15644] [Medline: 31859682]
19. Torous J, Andersson G, Bertagnoli A, Christensen H, Cuijpers P, Firth J, et al. Towards a consensus around standards for smartphone apps and digital mental health. World Psychiatry 2019 Feb;18(1):97-98. [doi: 10.1002/wps.20592] [Medline: 30600619]
20. Chen AT, Wu S, Tomasino KN, Lattie EG, Mohr DC. A multi-faceted approach to characterizing user behavior and experience in a digital mental health intervention. J Biomed Inform 2019 Jun;94:103187 [FREE Full text] [doi: 10.1016/j.jbi.2019.103187] [Medline: 31026595]
21. Prolific. 2020. URL: https://www.prolific.co/ [accessed 2021-07-05]
22. Palan S, Schitter C. Prolific.ac—A subject pool for online experiments. Journal of Behavioral and Experimental Finance 2018 Mar;17:22-27. [doi: 10.1016/j.jbef.2017.12.004]
23. Peer E, Brandimarte L, Samat S, Acquisti A. Beyond the Turk: Alternative platforms for crowdsourcing behavioral research. Journal of Experimental Social Psychology 2017 May;70:153-163. [doi: 10.1016/j.jesp.2017.01.006]
24. Part 1: Cybersecurity Overview. Solutions Reservoir. URL: http://solutionsreservoir.com/resources/introduction-to-cybersecurity/part-1-cybersecurity-overview [accessed 2021-07-05]
25. Lettmann H, Lumsden J. Prolific's participant pool – the present and the future. Prolific. 2018 Sep 18. URL: https://blog.prolific.co/prolifics-participant-pool-its-present-and-its-future/ [accessed 2021-07-05]
26. Oppenheimer DM, Meyvis T, Davidenko N. Instructional manipulation checks: Detecting satisficing to increase statistical power. Journal of Experimental Social Psychology 2009 Jul;45(4):867-872. [doi: 10.1016/j.jesp.2009.03.009]
27. Kroenke K, Spitzer R. The patient health questionnaire-2: Validity of a two-item depression screener. Medical Care 2003;41(11):1284-1292 [FREE Full text] [doi: 10.1097/01.mlr.0000093487.78664.3c]
28. Kroenke K, Spitzer RL, Williams JB, Monahan PO, Löwe B. Anxiety disorders in primary care: prevalence, impairment, comorbidity, and detection. Ann Intern Med 2007 Mar 06;146(5):317-325. [doi: 10.7326/0003-4819-146-5-200703060-00004] [Medline: 17339617]
29. Brown RL, Rounds LA. Conjoint screening questionnaires for alcohol and other drug abuse: criterion validity in a primary care practice. Wis Med J 1995;94(3):135-140. [Medline: 7778330]
30. Staples LG, Dear BF, Gandy M, Fogliati V, Fogliati R, Karin E, et al. Psychometric properties and clinical utility of brief measures of depression, anxiety, and general distress: The PHQ-2, GAD-2, and K-6. Gen Hosp Psychiatry 2019 Jan;56:13-18 [FREE Full text] [doi: 10.1016/j.genhosppsych.2018.11.003] [Medline: 30508772]
31. Hinkin CH, Castellon SA, Dickson-Fuhrman E, Daum G, Jaffe J, Jarvik L. Screening for drug and alcohol abuse among older adults using a modified version of the CAGE. Am J Addict 2001;10(4):319-326. [Medline: 11783746]
32. Osman A, Bagge CL, Gutierrez PM, Konick LC, Kopper BA, Barrios FX. The Suicidal Behaviors Questionnaire-Revised (SBQ-R): validation with clinical and nonclinical samples. Assessment 2001 Dec 26;8(4):443-454. [doi: 10.1177/107319110100800409] [Medline: 11785588]
33. Millner AJ, Lee MD, Nock MK. Single-Item Measurement of Suicidal Behaviors: Validity and Consequences of Misclassification. PLoS One 2015;10(10):e0141606 [FREE Full text] [doi: 10.1371/journal.pone.0141606] [Medline: 26496707]
34. Brooke J. SUS: A "quick and dirty" usability scale. In: Jordan PW, Thomas B, Weedermeester BA, McClelland AL, editors. Usability Evaluation in Industry. London, UK: Taylor and Francis; 1996:189-194.
35. Bangor A, Kortum PT, Miller JT. An Empirical Evaluation of the System Usability Scale. International Journal of Human-Computer Interaction 2008 Jul 30;24(6):574-594. [doi: 10.1080/10447310802205776]
36. Suh H, Shahriaree N, Hekler E. Developing and validating the user burden scale: A tool for assessing user burden in computing systems. In: CHI Conference on Human Factors in Computing Systems. 2016 Presented at: Computer Human Interaction (CHI); May 7-12; San Jose, CA. [doi: 10.1145/2858036.2858448]
37. Cheng K, Lin V, Nijhawan K, Westhem. Apps with benefits: using benefits and burdens to predict mobile app usage. 2017 Presented at: Computer Human Interaction (CHI); May 6-11; Denver, CO p. 2452-2458. [doi: 10.1145/3027063.3053276]
38. Kujala S, Kauppinen M. Identifying and selecting users for user-centered design. In: Proceedings of the Third Nordic Conference on Human-Computer Interaction. 2004 Presented at: NordiCHI04; April 24-29; Tampere, Finland. [doi: 10.1145/1028014.1028060]
39. Help me find an app. OneMindPsyberguide. URL: https://onemindpsyberguide.org/apps/ [accessed 2021-07-05]

40. Explore relevant apps and reviews. M-Health Index and Navigation Database. URL: https://apps.digitalpsych.org/ [accessed 2021-07-05]
41. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society: Series B (Methodological) 2018 Dec 05;57(1):289-300. [doi: 10.1111/j.2517-6161.1995.tb02031.x]
42. Benjamini Y, Hochberg Y. On the Adaptive Control of the False Discovery Rate in Multiple Testing with Independent Statistics. Journal of Educational and Behavioral Statistics 2000;25(1):60-83. [doi: 10.2307/1165312]
43. Rubanovich CK, Mohr DC, Schueller SM. Health App Use Among Individuals With Symptoms of Depression and Anxiety: A Survey Study With Thematic Coding. JMIR Ment Health 2017 Jun 23;4(2):e22 [FREE Full text] [doi: 10.2196/mental.7603] [Medline: 28645891]
44. Corbin J, Strauss AL. Basics of Qualitative Research: Grounded Theory Procedures and Techniques, 3rd Ed. Thousand Oaks, CA: SAGE Publications; 2008.
45. United States Quick Facts. US Census. 2019. URL: https://www.census.gov/quickfacts/fact/table/US/PST045219 [accessed 2021-07-05]
46. Unmarried and Single Americans. US Census. 2020. URL: https://www.census.gov/newsroom/stories/unmarried-single-americans-week.html [accessed 2021-07-05]
47. Median age of the resident population of the United States from to 1960-2019. Statista. 2021 Jan 20. URL: https://www.statista.com/statistics/241494/median-age-of-the-us-population/ [accessed 2021-07-05]
48. Wang X, Markert C, Sasangohar F. Investigating Popular Mental Health Mobile Application Downloads and Activity During the COVID-19 Pandemic. Hum Factors 2021 Mar 07. [doi: 10.1177/0018720821998110] [Medline: 33682467]
49. Torous J, Wisniewski H, Liu G, Keshavan M. Mental Health Mobile Phone App Usage, Concerns, and Benefits Among Psychiatric Outpatients: Comparative Survey Study. JMIR Ment Health 2018 Nov 16;5(4):e11715 [FREE Full text] [doi: 10.2196/11715] [Medline: 30446484]
50. Robbins R, Krebs P, Jagannathan R, Jean-Louis G, Duncan DT. Health App Use Among US Mobile Phone Users: Analysis of Trends by Chronic Disease Status. JMIR Mhealth Uhealth 2017 Dec 19;5(12):e197 [FREE Full text] [doi: 10.2196/mhealth.7832] [Medline: 29258981]
51. Wasserman D, Iosue M, Wuestefeld A, Carli V. Adaptation of evidence-based suicide prevention strategies during and after the COVID-19 pandemic. World Psychiatry 2020 Oct;19(3):294-306 [FREE Full text] [doi: 10.1002/wps.20801] [Medline: 32931107]
52. Watts G. COVID-19 and the digital divide in the UK. The Lancet Digital Health 2020 Aug;2(8):e395-e396. [doi: 10.1016/s2589-7500(20)30169-2]
53. Renn BN, Hoeft TJ, Lee HS, Bauer AM, Areán PA. Preference for in-person psychotherapy versus digital psychotherapy options for depression: survey of adults in the U.S. NPJ Digit Med 2019 Feb 11;2(1):6-00 [FREE Full text] [doi: 10.1038/s41746-019-0077-1] [Medline: 31304356]
54. Torous J, Jän Myrick K, Rauseo-Ricupero N, Firth J. Digital Mental Health and COVID-19: Using Technology Today to Accelerate the Curve on Access and Quality Tomorrow. JMIR Ment Health 2020 Mar 26;7(3):e18848 [FREE Full text] [doi: 10.2196/18848] [Medline: 32213476]
55. Mohr DC, Azocar F, Bertagnolli A, Choudhury T, Chrisp P, Frank R, Banbury Forum on Digital Mental Health. Banbury Forum Consensus Statement on the Path Forward for Digital Mental Health Treatment. Psychiatr Serv 2021 Jun 20;72(6):677-683. [doi: 10.1176/appi.ps.202000561] [Medline: 33467872]
56. Inkster B. Early Warning Signs of a Mental Health Tsunami: A Coordinated Response to Gather Initial Data Insights From Multiple Digital Services Providers. Front Digit Health 2021 Feb 10;2:64. [doi: 10.3389/fdgth.2020.578902]
57. Lyon AR, Brewer SK, Areán PA. Leveraging human-centered design to implement modern psychological science: Return on an early investment. Am Psychol 2020 Nov;75(8):1067-1079. [doi: 10.1037/amp0000652] [Medline: 33252945]
58. Lyon AR, Munson SA, Renn BN, Atkins DC, Pullmann MD, Friedman E, et al. Use of Human-Centered Design to Improve Implementation of Evidence-Based Psychotherapies in Low-Resource Communities: Protocol for Studies Applying a Framework to Assess Usability . JMIR Res Protoc 2019 Oct 09;8(10):e14990 [FREE Full text] [doi: 10.2196/14990] [Medline: 31599736]
59. Nyenhuis SM, Greiwe J, Zeiger JS, Nanda A, Cooke A. Exercise and Fitness in the Age of Social Distancing During the COVID-19 Pandemic. J Allergy Clin Immunol Pract 2020;8(7):2152-2155 [FREE Full text] [doi: 10.1016/j.jaip.2020.04.039] [Medline: 32360185]
60. Salzwedel A, Rabe S, Zahn T, Neuwirth J, Eichler S, Haubold K, et al. User Interest in Digital Health Technologies to Encourage Physical Activity: Results of a Survey in Students and Staff of a German University. JMIR Mhealth Uhealth 2017 Apr 19;5(4):e51 [FREE Full text] [doi: 10.2196/mhealth.7192] [Medline: 28428156]
61. Régnier F, Chauvel L. Digital Inequalities in the Use of Self-Tracking Diet and Fitness Apps: Interview Study on the Influence of Social, Economic, and Cultural Factors. JMIR Mhealth Uhealth 2018 Apr 20;6(4):e101 [FREE Full text] [doi: 10.2196/mhealth.9189] [Medline: 29678807]

62. Collado-Borrell R, Escudero-Vilaplana V, Villanueva-Bueno C, Herranz-Alonso A, Sanjurjo-Saez M. Features and Functionalities of Smartphone Apps Related to COVID-19: Systematic Search in App Stores and Content Analysis. J Med Internet Res 2020 Aug 25;22(8):e20334 [FREE Full text] [doi: 10.2196/20334] [Medline: 32614777]
63. Ng MM, Firth J, Minen M, Torous J. User Engagement in Mental Health Apps: A Review of Measurement, Reporting, and Validity. Psychiatr Serv 2019 Jul 01;70(7):538-544 [FREE Full text] [doi: 10.1176/appi.ps.201800519] [Medline: 30914003]

Abbreviations

ANOVA: analysis of variance
CAGE-AID: Cut-Annoyed-Guilty-Eye Adapted to Include Drugs
DMHT: digital mental health tool
GAD-2: 2-item Generalized Anxiety Disorder
MIND: M-Health Index and Navigation Database
PHQ-2: 2-item Patient Health Questionnaire
SBQ-R: Suicide Behaviors Questionnaire–Revised
UBS: Use Burden Scale
SARS: severe acute respiratory syndrome
SUS: System Usability Scale
VOIP: voice over IP

Edited by G Eysenbach; submitted 03.03.21; peer-reviewed by S Schueller, K Van Orden, S Kunkle; comments to author 24.03.21; revised version received 28.04.21; accepted 16.05.21; published 05.08.21.

Please cite as:
Mata-Greve F, Johnson M, Pullmann MD, Friedman EC, Griffith Fillipo I, Comtois KA, Arean P
Mental Health and the Perceived Usability of Digital Mental Health Tools Among Essential Workers and People Unemployed Due to COVID-19: Cross-sectional Survey Study
JMIR Ment Health 2021;8(8):e28360
URL: https://mental.jmir.org/2021/8/e28360
doi:10.2196/28360
PMID:34081592

Original Paper

Integration of Digital Tools Into Community Mental Health Care Settings That Serve Young People: Focus Group Study

Ashley A Knapp[1], BSc, MA, PhD; Katherine Cohen[1], BSc, MA; Jennifer Nicholas[2], PhD; David C Mohr[1], PhD; Andrew D Carlo[1], MD, PhD; Joshua J Skerl[3], BA, MAEd; Emily G Lattie[1], PhD

[1]Feinberg School of Medicine, Northwestern University, Chicago, IL, United States

[2]Orygen Centre for Youth Mental Health, University of Melbourne, Melbourne, Australia

[3]Ohio Guidestone Organization, Berea, OH, United States

Corresponding Author:
Ashley A Knapp, BSc, MA, PhD
Feinberg School of Medicine
Northwestern University
750 N Lake Shore Dr
Chicago, IL, 60611
United States
Phone: 1 312 503 3751
Email: Ashley.knapp@northwestern.edu

Abstract

Background: Digital mental health tools have substantial potential to be easily integrated into people's lives and fundamentally impact public health. Such tools can extend the reach and maximize the impact of mental health interventions. Before implementing digital tools in new settings, it is critical to understand what is important to organizations and individuals who will implement and use these tools. Given that young people are highly familiar with technology and many mental health concerns emerge in childhood and adolescence, it is especially crucial to understand how digital tools can be integrated into settings that serve young people.

Objective: This study aims to learn about considerations and perspectives of community behavioral health care providers on incorporating digital tools into their clinical care for children and adolescents.

Methods: Data were analyzed from 5 focus groups conducted with clinicians (n=37) who work with young people at a large community service organization in the United States. This organization provides care to more than 27,000 people annually, most of whom are of low socioeconomic status. The transcripts were coded using thematic analysis.

Results: Clinicians first provided insight into the digital tools they were currently using in their treatment sessions with young people, such as web-based videos and mood-tracking apps. They explained that their main goals in using these tools were to help young people build skills, facilitate learning, and monitor symptoms. Benefits were expressed, such as engagement of adolescents in treatment, along with potential challenges (eg, accessibility and limited content) and developmental considerations (eg, digital devices getting taken away as punishment). Clinicians discussed their desire for a centralized digital platform that securely connects the clinician, young person, and caregivers. Finally, they offered several considerations for integrating digital tools into mental health care, such as setting up expectations with clients and the importance of human support.

Conclusions: Young people have unique considerations related to complex accessibility patterns and technology expectations that may not be observed when adults are the intended users of mental health technologies. Therefore, these findings provide critical insights to inform the development of future tools, specifically regarding connectivity, conditional restraints (eg, devices taken away as punishment and school restrictions), expectations of users from different generations, and the blended nature in which digital tools can support young people.

(JMIR Ment Health 2021;8(8):e27379) doi:10.2196/27379

KEYWORDS

digital mental health; treatment; young people; children; adolescents; community mental health care; mobile phone

Introduction

Background

As digital devices become increasingly ubiquitous, there is growing potential for digital mental health (DMH) tools to be easily integrated into people’s daily lives and have a fundamental public health impact. DMH tools may maximize the public health impact of mental health support and extend its reach to new populations while also enabling unprecedented individualization and treatment optimization. However, efficacious DMH interventions have not yet been translated into effective and sustained use in real-world care settings. This well-described concern is known as the research-to-practice gap [1,2]. Furthermore, engagement with DMH tools is low, for example, Baumel et al [3] report that just 4% of mental health apps are opened daily. These engagement and implementation challenges may arise because conditions (eg, needs, preferences, goals, and barriers) are specific to deployment settings, and stakeholders have not traditionally been considered in the design of evidence-based DMH care.

Recently, the DMH field has attempted to address the research-to-practice gap by focusing on engagement and implementation. Frameworks, such as the Accelerated Creation-to-Sustainment (ACTS) model [4], have been proposed to guide the successful implementation and sustainment of tools in real-world settings. The ACTS model comprises three iterative stages (Create, Trial, and Sustainment) and encourages evaluation and design at each stage. For example, in the *Create* stage, the framework encourages in-depth qualitative assessments (eg, interviews and design workshops) and usability testing to inform the design of the service and its implementation alongside the technology. In addition, ACTS and other frameworks focused on the swift translation of research evidence to practice, such as the Designing for Accelerated Translation [5], recommend the incorporation of user-centered design methods throughout all phases of development to better identify user needs and context, as well as relevant surrounding factors of deployment settings. These methods intend to partner with key stakeholders to better understand current needs and existing organizational structures in an effort to design targeted, pragmatic, and sustainable DMH tools that can be optimally integrated into the proposed context of use.

A recent review highlighted the variety of factors that may impact health care providers' willingness to adopt digital health tools, including technological aspects such as ease of use, compatibility, and personalization and social and organizational factors, including workflow, evidence base, and monetary considerations [6]. However, specifically with regard to mental health, research to understand the context, needs, and goals of providers’ use of DMH tools in clinical care settings is sparse. This is especially true for child and adolescent mental health providers, despite indications that the potential for success of technological tools within clinical care may be highest for young clients, given this generation’s immersion and interest in technology [7,8]. Furthermore, young people with mental health difficulties are a critical population, as most mental health conditions emerge in adolescence and early adulthood, underscoring the importance of preventive and early intervention efforts [9-11]. However, very few studies have specifically reported on the implementation and engagement considerations of DMH tools among mental health workers who primarily work with young people [12-14].

Through a quantitative survey, Cliffe et al [12] assessed child mental health professionals’ use of and attitudes toward technology in clinical care. They found that most clinicians were uncertain of which technologies were available, which led to their primary use of older technologies, such as helplines and websites. Although not as frequent, clinicians reported the use of apps focused on emotional management and mindfulness. Similarly, Orlowski et al [13] used focus groups and semistructured interviews with mental health workers serving young people to explore the acceptability of technology to engage with clients. Clinicians identified potential strengths of technology, such as tracking symptoms or increasing engagement, but expressed concerns about internet access in rural areas, maintaining confidentiality, and crisis management. To the best of the authors’ knowledge, the only other published study that elicited feedback from mental health workers serving young people was specific to the use of a single app that asked young people to chart their symptoms as an augment to their treatment sessions [14]. Some key takeaways from this exploratory study were that clinicians found the mood graphs were helpful in engaging parents in their child’s treatment and that privacy, anonymity, and connectivity were important to the success and use of this app.

Objectives

Given the importance of understanding the nuances in providing care to young people in efforts to successfully integrate DMH tools into health services for this population, additional data to build on early exploratory studies are needed. By focusing on unique stakeholder needs and contextual factors of different care settings, we can begin to cultivate a body of literature that could serve as a guide in informing the design and testing of technological tools for mental health workers who serve young people. In turn, we may be able to create tools that are effective, pragmatic, and sustainably integrated into care settings. This study aims to extend the literature on how mental health clinicians are currently using technology within their clinical care to young people and expand upon previous work by investigating the challenges of technology use in treating this critical population. In addition, we explore the types of digital tools that clinicians find helpful in their support of young people and considerations regarding integrating these tools into everyday practice.

Methods

Participants

We used data collected from clinical staff members who provide behavioral health care for children and adolescents at a large community service organization in a Midwestern state. A partnership was formed after the organization reached out to our research center and indicated an interest in incorporating technology into their behavioral health care services.

Procedures

Focus groups were divided based on the clinical services provided by the staff members and lasted approximately 1 hour. There were 7 groups in total; 2 comprised clinicians who provide in-home services for children and adolescents (labeled as *in home 1* and *in home 2*), two comprised those who provide in-school services for children and adolescents (labeled as *school 1* and *school 2*), one comprised those who provide services through an outpatient clinic (labeled as *outpatient 1*), one comprised those who provide services to adults, and one comprised clinical supervisors. As the focus of this study was on the perspectives of clinicians serving young clients, the data from the groups comprising clinicians serving adults only and the group of clinical supervisors were excluded from analyses; thus, in total, 5 focus groups were used for this study.

Participants were recruited through email. The focus groups took place in conference rooms at the organization's central office. Groups were run by an academic research clinical psychologist and the director of our center's research operations, who has a master's degree in public health and a background in community mental health. The focus groups were recorded and transcribed with the participants' consent. All participants who took part in the focus groups were given US $5 gift cards and were provided with food during the group. For the clinician groups, participants were first asked questions regarding their typical clinical encounters, client communication, use of supplemental treatment resources or tools, and interest in technology-based resources or tools. They were then asked questions regarding a specific technology-based tool that has been validated in several previous studies (IntelliCare) [15-17]. Clinicians were asked what would be needed to implement this tool and similar DMH tools and how they might fit in with their current practices, which led to a broader discussion on desired DMH digital tools, features, and integration considerations. All study procedures were approved by the authors' institutional review board before enrolling the participants.

Data Analysis

Overview

This study generated an extensive mixed methods data set that provided insight into several questions related to how technology-enabled mental health interventions could be used in community mental health care. Because of the size and richness of the data set, we conducted two sets of analyses. The first used mixed methods (including quantitative measures of implementation climate and clinical orientation in conjunction with focus groups) to understand community mental health providers and supervisor attitudes toward using a variety of technologies in their work and identify barriers to and facilitators of implementation [18]. In contrast, this study is a purely qualitative analysis of issues specific to the use of technology-enabled mental health services with young clients from the perspective of child and adolescent clinicians only, as mental health treatment with young people presents unique challenges and innovation in youth community mental health care remains understudied [12,13,19]. Code creation, codebook formation, coding, and derived themes were independently performed for each analysis. The procedure for this study is outlined below.

Thematic Coding

Focus group transcripts were coded by authors using a thematic analysis approach [20]. Coders first reviewed the transcripts for thematic content and created a codebook with the primary themes they identified. After the codebook was created, 2 coders reviewed the transcripts a second time to ensure the codebook accuracy before completing a final round of coding. The team-based approach to coding allowed for analyst triangulation, providing a check for validity and rigor within the analysis. Within this process, themes at every step of the analysis and reporting were determined by consensus among researchers [21]. In an additional verification of the coding of the data and implications derived, the research team partnered with a key stakeholder from the community behavioral health care organization to verify the credibility of the results and implications drawn. In addition, author KC, who was independent of the coding process, compared the results with those of other analyses to verify that there was no overlap between papers. The research question explored in this paper expands on the results discussed in the previously published analysis [18], and the results of that analysis are presented to provide the necessary context.

Results

Participants

The community service organization serves approximately 27,000 people annually and has offices and clinics across the state serving several different local communities, with most serving low socioeconomic status. The organization serves a broad range of behavioral health concerns in youth, ranging from issues such as attention-deficit/hyperactivity disorder to depression. The clinicians in the focus groups reported that they serve youth with a wide range of presenting concerns. In addition, 61.5% (218/354) of the staff at this organization had a master's degree, 87.0% (308/354) were female, and 72.5% (257/354) were White. A total of 37 staff members participated in the 5 focus groups included in this analysis (7 in home 1, 7 in home 2, 10 in school 1, 7 in school 2, and 6 in outpatient 1).

Overview of Themes

Clinicians' feedback regarding the opportunities, challenges, and future directions for integrating digital tools into clinical care with youth was divided into three main themes. Theme 1 identifies the potential strengths of digital tools, highlighting how digital tools can help young people build skills, facilitate learning, and monitor symptoms. Theme 2 presents the potential challenges of using digital tools in practice, such as limited accessibility, outdated programs, and limited content. Theme 3 introduces clinicians' ideas for future DMH work, including their desire for a centralized digital platform and considerations for integration into care. Table 1 summarizes the key findings across the themes.

Table 1. Summary of key findings.

Themes and subthemes	Key findings
Theme 1: Digital tools in clinical care with young people	
Build skills, facilitate learning, and monitor symptoms	• Clinicians reported using digital platforms was an excellent way to better engage young people in treatment. • There was a notable variety in the types of tools that clinicians used, and the main goals of using these tools were to acquire and practice skills, facilitate learning and discussion, and monitor symptoms. • An example was the use of mood-tracking apps to facilitate conversations around patterns and precursors to changes in young people's mental health symptoms. • Clinicians noted the importance of considering the child's developmental stage when deciding their level of involvement in digital tools.
Theme 2: Challenges of using digital tools in practice	
Accessibility	• Challenges related to young people's limited or no access to broadband and/or digital devices outside of treatment sessions were frequently mentioned, expanding the definition of accessibility. • Clinicians mentioned other restrictions caregivers and schools placed on smartphone use, such as limits on screen time and the context in which the device can or cannot be used. • These restrictions, although typically implemented by parents to promote well-being, can significantly interfere with the ability of adolescents to engage in DMH[a] programs at times when they perceive themselves most in need of the tools. • As for connectivity and smartphone access, our results imply that solutions, such as making content available offline or providing desktop versions, should be considered so that accessibility is not dependent on consistent internet and smartphone access.
Outdated programs and limited content	• Although clinicians could see the utility of digital platforms beyond outdated designs, young people were often deterred from initially engaging or maintaining sustained use with these platforms based on the outdated design. • Similarly, other digital tools clinicians used had limited content, and thus their clients would become bored quickly or tire easily of these apps because they ran out of new content quickly. • To combat the vast drop-off observed in these instances, clinicians underscored the importance of keeping the design and content of young people–facing platforms fresh and up to date.
Theme 3: Desired digital platform and integration considerations	
Centralized digital platform	• Several clinicians desired a centralized digital tool that securely connected the clinician, the young person, and the young person's family. • Clinicians were especially drawn to the possibility of this tool to help generalize skills learned in sessions to young people's day-to-day lives, track mood patterns in real time to discuss in treatment, and increase parent communication and engagement in their child's treatment. • For this to be effective, clinicians underscored the importance of building rapport and attaining buy-in from parents and other stakeholders so that teens have access to devices and services to use as therapy resources when triggered at school or home.
Considerations for integration into care	• Regarding the design of the young person–facing platform, clinicians emphasized the importance of visuals, compared with primary text, and features, such as earning badges and creating avatars, to keep young people engaged. • Clinicians underscored the importance of designing the program brand to be discreet and having other privacy features programmed to ensure confidentiality of mental health information stored on devices. • Clinicians unanimously agreed that using digital tools to augment therapy would be most effective if a human, such as a therapist, parent, or teacher, was behind this tool to check in with and guide young people. • Another consideration discussed when using digital tools to augment therapy was the importance of setting up expectations and boundaries with the young person and parents so the capabilities of the tool can be understood by them. • Clinicians also suggested working with the families so that it is understood that immediate help may not be available via this modality in crisis situations and to create a safety plan for those instances. • Another strategy mentioned to mitigate the risk of an unrecognized crisis communicated via digital devices was to program the tool to automatically detect and guide the young person to the appropriate contact and resources.

[a]DMH: digital mental health.

Theme 1: Digital Tools in Clinical Care With Young People

Overview

Given the rapid emergence of technology, with young people at the forefront, clinicians discussed how using digital tools in practice was a good way to connect with this population and meet them where they are. As a clinician stated:

> *The world we live in is technology driven, and the kids are way more technology savvier than we are.* [In home 2]

Clinicians discussed that the overwhelming majority of their clients preferred digital platforms to nondigital platforms:

> *And even a lot of clients, whenever I bring up anything about journaling or anything like that, they're completely against writing things on paper. But they'll do their notes on their phone.* [In home 2]

As such, clinicians reported that using this modality was an excellent way to better engage young people in treatment.

In the first report, Lattie et al [18] reported on clinicians using technological resources to support skill building and empowerment of adults and young people. The following section explores the particular types of digital tools clinicians reported using exclusively with young people, along with how they incorporated these tools in their service of this population.

Build Skills and Facilitate Learning

Most frequently, clinicians mentioned using digital tools to help children and adolescents acquire and practice skills in treatment sessions, with the intent of young people using these skills in their daily lives. Several clinicians, for example, used digital tools, such as apps and web-based videos, to help guide their child and adolescent clients through practice sessions focused on skills such as meditation, relaxation, and mindfulness. Clinicians modeled and practiced these techniques with children and adolescents in the session while also encouraging them to practice these skills in everyday life between sessions. As a clinician mentioned:

> *I've been doing a lot of imagery work and progressive muscle relaxation prompts. So, I've had some of my kids record on their phone their own voice using those prompts, so that they can use it before bed or when they get up in the morning. And then kind of a way to empower them too.* [School 1]

This clinician empowered young people by having them lead the relaxation exercise in a digital recording on their personal devices, which reinforced their knowledge and practice of that skill and also led to a newfound familiarity with tailored tools that they could use in their daily lives. Using gamification to support skills such as communication was also mentioned. Although not designed as a DMH tool, a clinician reported using the app *Heads Up*, where the device was placed on the forehead of one person, and the other person used word cues to have that person guess the word or phrase presented on the device screen. This clinician commented:

> *I've seen some kid clients come alive because they're excited because they wanna beat their score. And just helping them like, "How do you have to communicate? You have to keep talking. You have to keep going." It's helped with that.* [Outpatient 1]

Apps such as these and others specifically designed as DMH tools were used not only to engage young people in the treatment session but also to prompt practice of particular skills in real time.

Digital tools have also been used to demonstrate the value of therapeutic techniques. When wanting to provide tangible evidence of the mind-body connection to the young person, a clinician mentioned:

> *...since I got my Apple Watch, I've been letting kids put it on and then having them do jumping jacks and showing them their heart rate and then using the breathe app, so they see it. And they're, "Oh, my gosh. My heartbeat went down when I took deep breaths." And I was like, "Yeah. Your body's calming down."* [School 1]

The clinician used biofeedback via the Apple Watch to show the child how deep breathing helped slow down their heart rate. Similarly, other clinicians mentioned using different digital tools such as Fitbit to direct the child and adolescent clients through breathing exercises, whereas others mentioned apps, such as a drawing app or mediation app, to help the client downregulate emotions at the moment. After teaching and practicing the skills with young people in a session, clinicians mentioned encouraging them to practice the learned skills during the week (with or without the support of the demonstration technology) and equally encouraged parents to support their children in these efforts.

Clinicians also discussed using several different types of digital tools, such as web-based videos, websites, and apps, to facilitate learning and discussions during sessions. In particular, digital tools are frequently used to support psychoeducation around skills, such as mindfulness, or mental health conditions such as attention-deficit/hyperactivity disorder, which then prompted discussions with their clients. For example, several clinicians mentioned viewing YouTube videos tailored to their client's needs together in session and then discussing the contents of the video, asking questions such as, "Did you understand this? What did you think about this?" [Outpatient 1].

Other clinicians mentioned using apps, such as interactive story apps where the user *chooses their own adventure* to provide tangible and engaging examples of situations to young people and then talking through the actions the child or adolescent chose in these apps. Overall, clinicians emphasized that digital tools provided a platform that initiated learning in a way that was engaging and understood by young people and also elicited rich discussions around key topics of interest. Clinicians mentioned that a unique benefit of using digital tools in practice was the ability to tailor the lessons to the specific mental health needs of the child or adolescent, as reported in Lattie et al [18] and their developmental stage (eg, Mind Yeti for adolescents with attention-deficit/hyperactivity disorder and GoNoodle for children with attention-deficit/hyperactivity disorder).

Monitor Symptoms

Other common digital tools that clinicians described using with young people were tracking tools, such as apps like Daylio, to chart symptoms between sessions. A clinician mentioned:

> *I use Moodpath, which helps clients especially–it's for depression. So, tracking the symptoms. They can use the smiley faces throughout the day to track where they're at.* [Outpatient 1]

The clinician further described that this app generated a chart that depicted the daily and weekly patterns of symptoms, which they found very valuable to reference periodically in session with the client. This resonated with other clinicians in the group, especially the utility of the tracking apps in engaging young people in real time throughout the week in language young people could understand and relate to (eg, smiley faces), compared with the complex, and often difficult to understand, clinical terms used to describe anxiety and depression symptoms. A clinician described how the mood-tracking app was used in session:

> *They [the adolescent] can bring it up on their phone...and we look at just is she daily fluctuating? If so, what happened during that day? What happened during that week? Is this a cycle? Is it three days anxiety, two days depression? Was it five days she had anxiety?* [Outpatient 1]

As depicted in this example, clinicians described using digital tracking tools as a platform to facilitate and tailor the conversation with young people by discussing patterns, precursors, and the context of their symptoms.

Theme 2: Challenges of Using Digital Tools in Practice

Overview

Along with the benefits of using digital tools with young people, clinicians have also reported on their concerns. In the first report, Lattie et al [18] underscored that the major concerns clinicians had were around the confidentiality and privacy of digital tools. For example, concerns were expressed regarding confidentiality breaches if others, such as a friend or family member, used the client's device or in situations in which the client used a communal digital device (eg, shared device among family or within a public setting). Additional barriers to the use of digital tools in clinical care are discussed as follows.

Accessibility

Several clinicians expressed accessibility concerns about adolescents' use of digital tools outside of treatment. One such concern was broadband access, for example, a clinician said:

> *Because some kids don't have data plans. They just have a phone. That they found. Their data goes fast so they're bouncing off of Wi-Fi.* [School 2]

Indeed, challenges related to limited or no access to broadband and/or digital devices have been frequently mentioned. Another barrier was the frequent report of parents removing devices and internet access from their teenagers as a form of punishment. This was outlined by a clinician:

> *But then I'm thinking about the fact that it [young person's smartphone] is normally one of the first things that get taken away if they do have a bad day. So, this is the thing you can use when you're having a bad day to calm down, but then mom and dad won't let you use it because you had a bad day, so you're even more frustrated because you don't get to use the thing you're supposed to use when you have a bad day.* [School 2]

Clinicians mentioned other restrictions caregivers and schools placed on device use, such as limits on screen time and the context in which the device can or cannot be used. For example, a clinician described:

> *My kids, they do guided meditation at home. Right before bed, to calm themselves or something like that. But that's–a lot of parents have rules, like not past ten.* [School 2]

These restrictions, although typically implemented by parents to promote well-being, can significantly interfere with the ability of adolescents to engage in DMH programs when they perceive themselves most in need of the tools.

Outdated Programs and Limited Content

Other salient concerns of incorporating digital tools in their care for young people were the outdated design of many DMH tools and the limited content of other tools. Referring to an app designed to help maintain healthy habits, a clinician mentioned the following:

> *The clients that I got to try it, they would try it for a day and then they would be like, 'Okay, I'm done with this.' I wish that–it has a really great usage to it and everything, but it just needs to be updated.* [In home 1]

Although clinicians could see the utility of the app beyond its current outdated design, young people were often deterred from initially engaging or maintaining sustained use with these apps based on outdated design and content. Similarly, other digital tools that clinicians used had limited content, and thus their clients would become bored quickly or tire easily of these apps because they ran out of new content quickly.

Theme 3: Desired Digital Platform and Integration Considerations

Centralized Digital Platform

Several clinicians desired a centralized digital tool that securely connected the clinician, young person, and young person's family. As a clinician underscored:

> *I'd say I would find it really cool–again just thinking about my couple teenagers that use a mood tracking app–if there was some sort of way we could link accounts securely so that I could even login...Or even if I say, "Hey, do this activity on this app before I see you next week." I can check and see when–like some sort of interaction base would be awesome.* [In home 2]

Clinicians viewed this type of tool to meet several goals. For example, they imagined that clinicians could use this tool to suggest apps or digital tools to young people and their families for use outside of sessions, remind young people to complete activities or practice skills between sessions, securely track their clients' mood patterns and progress in real time, securely touch base with clients outside of treatment broadly and in specific situations (eg, weeks when the in-person sessions are canceled), and connect directly and securely with parents to check in and provide resources. They also mentioned that this tool could be used to better engage the parents in their child's mental health care, with a clinician describing the following:

> *[I] feel like sometimes I'll give parents follow up things to do while I'm not there, and they'll forget about it throughout the week, but because they're on their phone or whatever so much throughout the week, I feel like we could send them reminders or this is what we need to do before the next week. I think that that would encourage them to be more engaged, at least in the process.* [In home 1]

Clinicians also brainstormed their ideas for the young people–facing platform of this centralized tool and other young people–facing platforms. They discussed that this type of service would fit well within schools and other settings teens frequent and that they believe this modality would be preferable to a teen compared with paper worksheets that are easily lost, not interactive, or discreet. The teenager could, for instance, inconspicuously use their smartphone or school tablet when they are anxious or angry to practice learned strategies (eg, relaxation and cognitive restructuring). For example, a clinician suggested it would be:

> *some sort of app they could have on their phone that could help them. I think sometimes we teach them things, and they don't remember to do them at home when they're feeling upset. So, maybe it could be something positive that our kids could use when we're not with them and we can't review coping skills with them...* [School 2]

For this to be effective, however, clinicians underscored the importance of building rapport and attaining buy-in from parents and other stakeholders (eg, teachers) so that teens have access to these digital devices and services (eg, smartphones and internet access) to use as therapy resources when triggered at school or home.

Regarding the design of the young people–facing platform, clinicians emphasized the importance of visuals, compared with primary text and features, such as earning badges, to keep young people engaged. As a clinician reasoned:

> *They love badges. And decorating their avatars, like getting a new hat...So, they're very motivated to get through their modules when they get to earn something at the end.* [School 2]

In addition to engagement, clinicians discussed the importance of privacy and being inconspicuous. For example, if the platform is on the young person's device, having a password or facial recognition to get in and a subtle design if a peer or friend uses the teen's device, they will not be immediately aware of the platform's function, nor will they be able to easily gain access. As a clinician outlined:

> *The app doesn't read as something like, My Personal Diary...it reads as something that you might just pass by if you don't know what its intention is, which can be good for teenagers who are afraid of people looking into their stuff.* [School 1]

Considerations for Integration Into Care

Clinicians unanimously agreed that such a tool would be most effective if a human, such as a therapist, parent, or guidance counselor, were *behind* this tool to check in with and guide young people. They talked about who would be the best person to support the teen via this service and emphasized training that person. Specifically, training that person not only on the technical side of the mental health platform but also around safeguards and ethical considerations when using this platform with young people and families. They also mentioned the importance of setting up expectations and boundaries with their clients and parents in the context of this tool, so the capabilities of the tool can be understood by them. As a clinician stated:

> *So, you set up the boundaries at the beginning...“This is what this can help you with. This is what it can't. This is when we really need to have that face to face.”* [School 1]

By doing so, young people and their families can fully use all the service's functions and also understand that immediate help may not be available via this modality in a crisis situation. This safeguard was discussed as a means to mitigate the risk of an unrecognized crisis, in addition to other strategies, such as programming the tool to automatically detect and guide the teen to the appropriate contact and resources.

Discussion

Principal Findings

Given the rapid emergence of technology use among young people [7,8,22], clinicians have expressed great interest in more effectively incorporating technology into their clinical care with young people. Similar to previous research [13], clinicians noted that, in general, using this modality was an excellent way to better engage young people in treatment, and in particular, offered a novel way to build skills, facilitate learning, and monitor symptoms. There was a notable variety of the types of tools that were used—from the use of a heart rate feature to provide *in vivo* biofeedback during in-session mindfulness exercises to the use of web-based videos to facilitate learning and discussions around the child's specific mental health need in language that is relatable and age-appropriate to the use of mood-tracking apps to facilitate conversations around patterns and precursors to changes in teen's mental health symptoms. Our findings extend the quantitative results of Cliffe et al [12] by providing a more nuanced view on how clinicians are using digital tools with young people.

At the same time, clinicians expressed concerns about young people's limited or complete lack of access to digital devices and connectivity outside of treatment sessions. Their feedback provided a different view from recent reports indicating nearly ubiquitous teen access to technology [7]. Their feedback suggests that accessibility [13] is much more complex, including considerations around inconsistent access (eg, shared device with family), limited connectivity or data plans, and conditional restraints specific to young people (eg, device taken away as punishment and school restrictions). Barriers such as these pose a clear threat to the utility of digital tools to provide young people with mental health support in real time when triggered. It is critical to include caregivers and others who play a prominent role in young people's lives, such as teachers, in the treatment plan so that barriers, such as technology restrictions, can be pre-emptively addressed. As for connectivity and device access, our results imply that solutions, such as making content available offline or providing desktop versions, should be considered so that accessibility is not dependent on consistent internet and device access. Similarly, DMH tools could be offered via more accessible technology, such as native apps that do not require a consistent Wi-Fi connection to function, or through SMS text messaging to allow for interventions to be delivered to individuals without smartphones. Finally, it is important to highlight that these accessibility concerns were reported by clinicians who primarily work with families of lower socioeconomic status. If we are to strive for health equity and digital inclusivity through ubiquitous access to DMH tools, it is imperative that we understand and address the barriers, such as those mentioned, faced by already underserved and marginalized populations [23]. Without such considerations, the introduction of inaccessible DMH tools could lead to an exacerbation of existing disparities and inequities, as opposed to mitigation [24].

There was high interest among clinicians in a centralized digital tool connecting therapists, young people, and their caregivers. Clinicians were especially attracted to the possibility of a centralized tool to help generalize skills learned in sessions to young people's daily lives and increase parental communication and engagement in their child's treatment. The involvement of a young person in a centralized tool will most likely depend on their age. For example, the clinician could connect with a teenager and parent via the tool; however, the clinician would likely only be interacting with the parent if the client was of a younger age. Furthermore, clinicians had several design recommendations for young people–facing platforms. To increase initial engagement and maintain sustained use, clinicians underscored the importance of keeping the design and content of young people–facing platforms fresh and up to date [25]. Features such as badges and avatars were underscored to combat the vast drop-off typically observed in DMH tool use [26]. Visuals were also recommended in contrast to lengthy texts and readings to capture and maintain children's and adolescents' interest in the tool. Finally, clinicians emphasized the importance of designing the program brand, such as an app icon on a device, to be discreet and have other privacy features programmed (eg, facial recognition or fingerprint) to ensure confidentiality of mental health information stored on devices. These privacy preferences are consistent with previous studies in which teens and their therapists prioritized ambiguous branding of an app (ie, MD vs Mood Diary) to keep curious siblings or friends from accessing their mental health information [14].

Strikingly, when discussing the use of DMH tools, clinicians overwhelmingly outlined how they used tools within their face-to-face sessions. In this way, DMH tools were integrated into care in bespoke ways, and clinicians, therefore, surfaced several considerations for broader implementation. First, they underscored the importance of including some level of human support with digital tools. The importance of human support aligns well with frameworks around technology-enabled services [4,27,28] and research that suggests digital tools are more effective for some users when supported by coaches than standalone tools [29]. This is also consistent with other research in which clinicians who work with adults recommend that DMH tools should be used to enhance face-to-face treatment, not as a replacement for it [30,31]. It is therefore critical that clinical training and continuing education keep pace with the increased interest and expectation to integrate digital tools in routine care by upskilling clinicians' DMH literacy, providing training for digitally enhanced models of care. Second, clinicians emphasized the importance of setting boundaries with young people and families when augmenting therapy with digital tools, similar to previous literature [13]. They suggested working with families so that it is understood that immediate help may not be available via this modality in crisis situations and to create a safety plan for those instances. Clinicians also brainstormed that a response to this challenge could involve algorithms built into tools that detect crisis-related words or phrases and automatically connect users to resources and services [32].

Most of these recommendations can be easily incorporated into the design of future DMH tools; however, the challenge of successful DMH implementation is still significant. Achieving successful and sustainable integration of digital tools into a predigital health system will likely require collaboration between specialists in DMH and implementation [33-37] and a shift from randomized controlled trials to effectiveness-implementation hybrid trials [38]. Only then can the field build evidence on how DMH tools can be successfully embedded into the daily work of clinical settings and have continued success without research support. Critically, such research and testing of different implementation strategies have a high potential to help fill the research-to-practice gap and to create sustainable tools within care settings that fulfill the promise of DMH [38].

Limitations and Future Directions

It is important to understand our findings in the context of their limitations. This study elicited feedback from clinicians and supervisors practicing in a large community behavioral health care organization. Although this is the first step in designing DMH tools that can be incorporated into mental health care for young people, it is critical that these findings are complemented by feedback from young people undergoing treatment. Young people provide unique feedback based on their lived experiences and their own use of DMH tools within care settings [14,39,40]. In addition, gathering feedback from other stakeholders (eg, caregivers and teachers) who play a prominent role in the lives of young people will also be significant in creating DMH tools that can be used and seamlessly incorporated into young people's lives. Given the paucity of research on DMH tools for child and adolescent mental health providers [12-14], it will also be important to replicate or elicit feedback from clinicians from other health care settings serving diverse populations of young people. We can then begin to understand which features of DMH tools are universal and which features are most effective within particular settings and populations of young people and clinicians. Furthermore, it is important to note that two sets of research questions were pursued with the large data set garnered from the focus groups. The research team took particular precautions to verify the integrity of the data, such as partnering with a key stakeholder from the community behavioral health care organization to verify the credibility and an independent check of result overlap. As qualitative data collection and analysis gain traction in the field of DMH, it is important to establish standards of practice for the field to ensure rigor and credibility. Finally, a representative from the community behavioral health care organization reached out directly to the research group to learn more about incorporating technology into their organization, and there was interest expressed from organizational leadership that led to this series of focus groups. Thus, this particular group of clinicians may be more interested in DMH tools than other clinicians. It will be important to include feedback from additional behavioral health care settings and providers with varying interests in incorporating DMH into practice.

Conclusions

This study examined feedback from child and adolescent mental health care providers from a large community mental health organization on the use of digital tools used in care settings. Clinicians discussed how they incorporated digital tools into their clinical care to enhance skill building, facilitate learning, and monitor symptoms. Clinicians also provided insight into accessibility, suggesting that considerations should include consistency in access, connectivity, and conditional restraints specific to young people. Finally, clinicians expressed high interest in a centralized digital tool to help consolidate learned skills in daily life and increase communication with parents. Future studies are needed, especially those that elicit feedback from young people and other stakeholders, to form a body of literature that guides the design and implementation of sustainable DMH tools that support the mental health of children and adolescents.

Acknowledgments

This research was supported by the National Institute of Mental Health (K01 MH121854, K08 MH112878, P50 MH119029, and T32 MH115882). ADC is a senior fellow at the Meadows Mental Health Policy Institute.

Conflicts of Interest

DCM, PhD, has accepted honoraria and consulting fees from Apple Inc, Otsuka Pharmaceuticals, Pear Therapeutics, and the One Mind Foundation; has royalties from Oxford Press; and has an ownership interest in Adaptive Health Inc. JJS is employed by OhioGuidestone.

References

1. Kazdin AE. Evidence-based treatment and practice: new opportunities to bridge clinical research and practice, enhance the knowledge base, and improve patient care. Am Psychol 2008 Apr;63(3):146-159. [doi: 10.1037/0003-066X.63.3.146] [Medline: 18377105]
2. McHugh RK, Barlow DH. The dissemination and implementation of evidence-based psychological treatments. A review of current efforts. Am Psychol 2010;65(2):73-84. [doi: 10.1037/a0018121] [Medline: 20141263]
3. Baumel A, Muench F, Edan S, Kane JM. Objective user engagement with mental health apps: systematic search and panel-based usage analysis. J Med Internet Res 2019 Sep 25;21(9):e14567 [FREE Full text] [doi: 10.2196/14567] [Medline: 31573916]
4. Mohr DC, Lyon AR, Lattie EG, Reddy M, Schueller SM. Accelerating digital mental health research from early design and creation to successful implementation and sustainment. J Med Internet Res 2017 May 10;19(5):e153 [FREE Full text] [doi: 10.2196/jmir.7725] [Medline: 28490417]
5. Ramsey AT, Proctor EK, Chambers DA, Garbutt JM, Malone S, Powderly WG, et al. Designing for Accelerated Translation (DART) of emerging innovations in health. J Clin Transl Sci 2019 Jun;3(2-3):53-58. [doi: 10.1017/cts.2019.386] [Medline: 31528365]
6. Jacob C, Sanchez-Vazquez A, Ivory C. Social, organizational, and technological factors impacting clinicians' adoption of mobile health tools: systematic literature review. JMIR Mhealth Uhealth 2020 Feb 20;8(2):e15935 [FREE Full text] [doi: 10.2196/15935] [Medline: 32130167]
7. Anderson M, Jiang J. Teens, social media and technology 2018. Pew Research Center. 2018. URL: https://www.pewresearch.org/internet/2018/05/31/teens-social-media-technology-2018/ [accessed 2021-07-16]
8. Odgers C, Robb M. Tweens, teens, tech, and mental health: coming of age in an increasingly digital, uncertain, and unequal world, 2020. Common Sense Media. 2020. URL: https://www.commonsensemedia.org/sites/default/files/uploads/pdfs/tweens-teens-tech-and-mental-health-full-report-final-for-web1.pdf [accessed 2021-07-16]
9. Dahl RE. Adolescent brain development: a period of vulnerabilities and opportunities. Keynote address. Ann N Y Acad Sci 2004 Jun;1021:1-22. [doi: 10.1196/annals.1308.001] [Medline: 15251869]

10. Beardslee WR, Chien PL, Bell CC. Prevention of mental disorders, substance abuse, and problem behaviors: a developmental perspective. Psychiatr Serv 2011 Mar;62(3):247-254. [doi: 10.1176/ps.62.3.pss6203_0247] [Medline: 21363895]
11. Kessler RC, Avenevoli S, Costello EJ, Georgiades K, Green JG, Gruber MJ, et al. Prevalence, persistence, and sociodemographic correlates of DSM-IV disorders in the National Comorbidity Survey Replication Adolescent Supplement. Arch Gen Psychiatry 2012 Apr;69(4):372-380 [FREE Full text] [doi: 10.1001/archgenpsychiatry.2011.160] [Medline: 22147808]
12. Cliffe B, Croker A, Denne M, Stallard P. Clinicians' use of and attitudes towards technology to provide and support interventions in child and adolescent mental health services. Child Adolesc Ment Health 2020 May;25(2):95-101. [doi: 10.1111/camh.12362] [Medline: 32307835]
13. Orlowski S, Lawn S, Matthews B, Venning A, Wyld K, Jones G, et al. The promise and the reality: a mental health workforce perspective on technology-enhanced youth mental health service delivery. BMC Health Serv Res 2016 Oct 10;16(1):562 [FREE Full text] [doi: 10.1186/s12913-016-1790-y] [Medline: 27724951]
14. Matthews M, Doherty G. In the moodngaging teenagers in psychotherapy using mobile phones. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2011 Presented at: CHI '11: CHI Conference on Human Factors in Computing Systems; May 7 - 12, 2011; Vancouver BC Canada p. 2947-2956. [doi: 10.1145/1978942.1979379]
15. Lattie EG, Schueller SM, Sargent E, Stiles-Shields C, Tomasino KN, Corden ME, et al. Uptake and usage of intellicare: a publicly available suite of mental health and well-being apps. Internet Interv 2016 May;4(2):152-158 [FREE Full text] [doi: 10.1016/j.invent.2016.06.003] [Medline: 27398319]
16. Graham AK, Greene CJ, Kwasny MJ, Kaiser SM, Lieponis P, Powell T, et al. Coached mobile app platform for the treatment of depression and anxiety among primary care patients: a randomized clinical trial. JAMA Psychiatry 2020 Sep 01;77(9):906-914. [doi: 10.1001/jamapsychiatry.2020.1011] [Medline: 32432695]
17. Mohr DC, Schueller SM, Tomasino KN, Kaiser SM, Alam N, Karr C, et al. Comparison of the effects of coaching and receipt of app recommendations on depression, anxiety, and engagement in the IntelliCare platform: factorial randomized controlled trial. J Med Internet Res 2019 Aug 28;21(8):e13609 [FREE Full text] [doi: 10.2196/13609] [Medline: 31464192]
18. Lattie EG, Nicholas J, Knapp AA, Skerl JJ, Kaiser SM, Mohr DC. Opportunities for and tensions surrounding the use of technology-enabled mental health services in community mental health care. Adm Policy Ment Health 2020 Jan 18;47(1):138-149 [FREE Full text] [doi: 10.1007/s10488-019-00979-2] [Medline: 31535235]
19. Garland AF, Haine-Schlagel R, Brookman-Frazee L, Baker-Ericzen M, Trask E, Fawley-King K. Improving community-based mental health care for children: translating knowledge into action. Adm Policy Ment Health 2013 Jan;40(1):6-22 [FREE Full text] [doi: 10.1007/s10488-012-0450-8] [Medline: 23212902]
20. Braun V, Clarke V, Hayfield N, Terry G. Thematic analysis. In: Liamputtong P, editor. Handbook of Research Methods in Health Social Sciences. Singapore: Springer; 2019:843-860.
21. Tracy SJ. Qualitative Quality: Eight “Big-Tent” criteria for excellent qualitative research. Qual Inq 2010 Oct 01;16(10):837-851 [FREE Full text] [doi: 10.1177/1077800410383121]
22. Rideout V, Robb M. Social media, social life: Teens reveal their experiences. Common Sense Media. 2018. URL: https://www.commonsensemedia.org/sites/default/files/uploads/research/2018_cs_socialmediasociallife_executivesummary-final-release_3_lowres.pdf [accessed 2021-07-16]
23. Definitions: Digital inclusion. National Digital Inclusion Alliance. URL: https://www.digitalinclusion.org/definitions/ [accessed 2021-05-08]
24. Brewer LC, Fortuna KL, Jones C, Walker R, Hayes SN, Patten CA, et al. Back to the future: achieving health equity through health informatics and digital health. JMIR Mhealth Uhealth 2020 Jan 14;8(1):e14512 [FREE Full text] [doi: 10.2196/14512] [Medline: 31934874]
25. Carlo AD, Ghomi RH, Renn BN, Strong MA, Areán PA. Assessment of real-world use of behavioral health mobile applications by a novel stickiness metric. JAMA Netw Open 2020 Aug 03;3(8):e2011978 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.11978] [Medline: 32744628]
26. Carlo AD, Ghomi RH, Renn BN, Areán PA. By the numbers: ratings and utilization of behavioral health mobile applications. NPJ Digit Med 2019;2:54 [FREE Full text] [doi: 10.1038/s41746-019-0129-6] [Medline: 31304400]
27. Schueller S, Tomasino K, Mohr D. Integrating human support into behavioral intervention technologies: the efficiency model of support. Clin Psychol Sci Pract 2016 Nov 17;24(1):27-45 [FREE Full text] [doi: 10.1111/cpsp.12173]
28. Richards D, Richardson T. Computer-based psychological treatments for depression: a systematic review and meta-analysis. Clin Psychol Rev 2012 Jun;32(4):329-342. [doi: 10.1016/j.cpr.2012.02.004] [Medline: 22466510]
29. Baumeister H, Reichler L, Munzinger M, Lin J. The impact of guidance on Internet-based mental health interventions — a systematic review. Internet Interv 2014 Oct;1(4):205-215. [doi: 10.1016/j.invent.2014.08.003] [Medline: 25615425]
30. Berry N, Bucci S, Lobban F. Use of the internet and mobile phones for self-management of severe mental health problems: qualitative study of staff views. JMIR Ment Health 2017 Nov 01;4(4):e52 [FREE Full text] [doi: 10.2196/mental.8311] [Medline: 29092809]
31. Sinclair C, Holloway K, Riley G, Auret K. Online mental health resources in rural Australia: clinician perceptions of acceptability. J Med Internet Res 2013;15(9):e193 [FREE Full text] [doi: 10.2196/jmir.2772] [Medline: 24007949]

32. Bauer AM, Hodsdon S, Hunter S, Choi Y, Bechtel J, Fortney JC. Lessons from the deployment of the SPIRIT App to support collaborative care for rural patients with complex psychiatric conditions. Proc ACM Int Conf Ubiquitous Comput 2017 Sep;2017:772-780 [FREE Full text] [doi: 10.1145/3123024.3125610] [Medline: 29075683]
33. Dopp AR, Parisi KE, Munson SA, Lyon AR. Integrating implementation and user-centred design strategies to enhance the impact of health services: protocol from a concept mapping study. Health Res Policy Syst 2019 Jan 08;17(1):1 [FREE Full text] [doi: 10.1186/s12961-018-0403-0] [Medline: 30621706]
34. Dopp AR, Parisi KE, Munson SA, Lyon AR. A glossary of user-centered design strategies for implementation experts. Transl Behav Med 2018 Dec 07;9(6):1057-1064. [doi: 10.1093/tbm/iby119] [Medline: 30535343]
35. Graham AK, Lattie EG, Powell BJ, Lyon AR, Smith JD, Schueller SM, et al. Implementation strategies for digital mental health interventions in health care settings. Am Psychol 2020 Nov;75(8):1080-1092. [doi: 10.1037/amp0000686] [Medline: 33252946]
36. Schueller SM, Torous J. Scaling evidence-based treatments through digital mental health. Am Psychol 2020 Nov;75(8):1093-1104. [doi: 10.1037/amp0000654] [Medline: 33252947]
37. Lyon AR, Brewer SK, Areán PA. Leveraging human-centered design to implement modern psychological science: return on an early investment. Am Psychol 2020 Nov;75(8):1067-1079. [doi: 10.1037/amp0000652] [Medline: 33252945]
38. Curran GM, Bauer M, Mittman B, Pyne JM, Stetler C. Effectiveness-implementation hybrid designs: combining elements of clinical effectiveness and implementation research to enhance public health impact. Med Care 2012 Mar;50(3):217-226 [FREE Full text] [doi: 10.1097/MLR.0b013e3182408812] [Medline: 22310560]
39. Bhattacharya A, Liang C, Zeng EY, Shukla K, Wong ME, Munson SA, et al. Engaging teenagers in asynchronous online groups to design for stress management. Proc 18th ACM Int Conf Interact Des Child (2019) 2019 Jun;2019:26-37 [FREE Full text] [doi: 10.1145/3311927.3323140] [Medline: 32954383]
40. Grist R, Porter J, Stallard P. Acceptability, use, and safety of a mobile phone app (BlueIce) for young people who self-harm: qualitative study of service users' experience. JMIR Ment Health 2018 Feb 23;5(1):e16 [FREE Full text] [doi: 10.2196/mental.8779] [Medline: 29475823]

Abbreviations

ACTS: Accelerated Creation-to-Sustainment
DMH: digital mental health

Edited by J Torous; submitted 25.01.21; peer-reviewed by B Cliffe, S Markwardt; comments to author 02.03.21; revised version received 07.05.21; accepted 20.05.21; published 19.08.21.

Please cite as:
Knapp AA, Cohen K, Nicholas J, Mohr DC, Carlo AD, Skerl JJ, Lattie EG
Integration of Digital Tools Into Community Mental Health Care Settings That Serve Young People: Focus Group Study
JMIR Ment Health 2021;8(8):e27379
URL: https://mental.jmir.org/2021/8/e27379
doi:10.2196/27379
PMID:34420928

Review

Commercial Off-The-Shelf Video Games for Reducing Stress and Anxiety: Systematic Review

Federica Pallavicini[1], PhD; Alessandro Pepe[1], PhD; Fabrizia Mantovani[1], PhD

Department of Human Sciences for Education "Riccardo Massa", University of Milano Bicocca, Milano, Italy

Corresponding Author:
Federica Pallavicini, PhD
Department of Human Sciences for Education "Riccardo Massa"
University of Milano Bicocca
Piazza dell'Ateneo Nuovo 1
Milano, 20126
Italy
Phone: 39 0264484903
Email: federica.pallavicini@gmail.com

Abstract

Background: Using commercial off-the-shelf video games rather than custom-made computer games could have several advantages for reducing stress and anxiety, including their low cost, advanced graphics, and the possibility to reach millions of individuals worldwide. However, it is important to emphasize that not all commercial video games are equal, and their effects strongly depend on specific characteristics of the games.

Objective: The aim of this systematic review was to describe the literature on the use of commercial off-the-shelf video games for diminishing stress and anxiety, examining the research outcomes along with critical variables related to computer game characteristics (ie, genre, platform, time of play).

Methods: A systematic search of the literature was performed following the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analysis) guidelines. The search databases were PsycINFO, Web of Science, Medline, IEEExplore, and the Cochrane Library. The search string was: [("video game*") OR ("computer game*")] AND [("stress") OR ("anxiety") OR ("relaxation")] AND [("study") OR ("trial") OR ("training")].

Results: A total of 28 studies met the inclusion criteria for the publication period 2006-2021. The findings demonstrate the benefit of commercial off-the-shelf video games for reducing stress in children, adults, and older adults. The majority of the retrieved studies recruited young adults, and fewer studies have involved children, middle-aged adults, and older adults. In addition to exergames and casual video games, other genres of commercial off-the-shelf games helped to reduce stress and anxiety.

Conclusions: Efficacy in reducing stress and anxiety has been demonstrated not only for exergames and casual video games but also for other genres such as action games, action-adventure games, and augmented reality games. Various gaming platforms, including consoles, PCs, smartphones, mobile consoles, and virtual reality systems, have been used with positive results. Finally, even single and short sessions of play had benefits in reducing stress and anxiety.

Trial Registration: International Platform of Registered Systematic Review and Meta-analysis Protocols INPLASY202130081; https://inplasy.com/?s=INPLASY202130081

(JMIR Ment Health 2021;8(8):e28150) doi:10.2196/28150

KEYWORDS

commercial off-the-shelf video games; video games; stress; anxiety; relaxation

Introduction

Background

Since the emergence of the COVID-19 pandemic in 2020, the frequency of stress and anxiety has markedly increased worldwide [1-6], with a prevalence of 29.6% and 31.9%, respectively [6]. The fear of contracting the virus, changes in lifestyle behaviors, social isolation, boredom, and uncertainty have exacerbated stress and anxiety in populations globally, which likely has long-lasting psychological and physical consequences [6-8]. Therefore, finding age-appropriate and cost-effective ways to support individuals in managing stress, anxiety, and their effects is urgently needed [9-11].

Video games represent one of the most appealing technological interventions for developing programs to reduce stress and anxiety since they are motivating, engaging, and easily accessible [12]. In 2020, the number of gamers worldwide was estimated at approximately 2.6 billion, and the games market is expected to exceed US $200 billion by the end of 2023 [13]. In contrast to popular belief, which views male children or teenagers as typical gamers, the average age range of video game players is 35-44 years, and across all players, 59% are male and 41% are female [14].

Computer games go far beyond the boundaries of entertainment. Video games are increasingly being used in sectors such as education [15-17] and mental health [18-21]. Some studies have highlighted the potential dangers of video games in terms of problematic use [22-25] and their relation to psychological functioning [23,26,27], whereas others have emphasized that the enjoyment and intrinsic motivation often associated with computer games make them a valuable and attractive new learning method [15-17] and offer psychological support to people [18-21].

Many schools, from the elementary to university level, adopt video games. Computer games can help to stimulate individuals in all of the transversal competencies collectively defined as "soft skills" (eg, creativity and the ability to deal with problems) and in teaching specific subjects such as mathematics or history [28-30]. Concerning mental health, several studies have demonstrated the usefulness of video games for training cognitive skills, including attentional processes, memory, and cognitive flexibility, especially in the elderly and adults [20,31-33].

Besides serving as useful tools for training cognitive processes [20,31,34], video games may also have a benefit in reducing stress and anxiety [35,36]. Computer games offer various positive emotions-triggering situations [12,18,36,37]. One of the most commonly reported motives for playing modern video games is the pleasure offered by digital games: people look for and are more willing to buy games that elicit positive emotions (eg, happiness and surprise) and enjoyment [38-40]. Like other pleasurable activities, video game playing stimulates dopamine release, a neurotransmitter linked to sensations of pleasure and reward [41].

The fundamental objective of video games is to entertain the player and elicit positive emotions [12,18], which, as stated by the "broaden-and-build" theory [42,43], have positive effects on the psychological well-being of the individual [44-46]. Positive emotions are considered to form the basis for the growth and flourishing self [46], and are especially important to increase subjective well-being [46-48]. Furthermore, video games can elicit the so-called "flow" state [49-51], defined as "the optimal experience when nothing else matters" [52,53], with benefits including increased self-efficacy, a stronger sense of self, and improved overall quality of life [54-56].

Moreover, in many cases, as is true for other entertainment media, video games play a role in distraction from undesirable emotions such as anxiety and stress by providing a temporary diversion from (real-world) adverse events or emotions [19,57-61].

Finally, video games, especially multiplayer games, offer the possibility of establishing a social connection in playing with friends or with people online [62,63]. This fact has become particularly relevant since the COVID-19 pandemic broke out. Gaming for social compensation might mitigate the experienced emotional distress during pandemic-related self-isolation [64,65].

Commercial Off-The-Shelf Video Games for Relaxation

Most of the studies performed to date on video games for stress and anxiety reduction have focused mainly on custom-made games (ie, games created ad hoc by researchers to educate, train, or change behavior) [66-68]. This type of game is often defined in the literature as a "serious game" [69], as gaming features are used as the primary medium for serious purposes [66]. Several custom-made video games for relaxation integrate biofeedback techniques into the game modes, such as Deep [70,71], Nevermind [72], MindLight [73,74], Dojo [75,76], and StressJam [77]. Furthermore, studies have shown that ad hoc video games could help adults with anxiety better handle emotional and physiological responses to stressors [78] and improve behavioral performance on anxiety-related stress tasks [79].

Interestingly, in addition to custom-made video games, commercial off-the-shelf (COTS) video games have also shown potential application for improving mental health [19], including the reduction of stress and anxiety (eg, [80-82]).

Using COTS video games rather than video games created ad hoc could have several advantages, including their low cost and ready-to-use format, advanced graphic quality, and the possibility to reach millions of players worldwide. As underlined in a recent paper, COTS games may disrupt health care over the coming decade [19]. Massive corporate funding for COTS games is often much higher than the budgets available to develop custom-made games, making it possible to reach a very high quality of video games in gameplay and user experience [19]. Besides, in contrast to the limited number of players that usually have the opportunity to try a custom-made game, millions of individuals can play a COTS game.

However, it is important to emphasize that not all COTS video games are equal, and their effects strongly depend on the specific characteristics of the game itself, such as its genre [83,84]. A recent systematic review [35] indicated that casual video games (CVGs), characterized by low cognitive loads and generally short time demands such as Tetris and Angry Birds, represent a particularly useful genre for diminishing stress and anxiety. Several other genres of COTS video games appear to be promising for decreasing stress and anxiety in individuals, including exergames [85] or survival horror games [86].

Research Questions

Within this context, the aim of this systematic review was to describe the literature on the use of COTS video games for decreasing stress and anxiety. The secondary objective was to organize the research with respect to critical variables related to video game characteristics (ie, genre, platform, time of play).

Methods

Databases Searched

A systematic search of the literature was performed on March 31, 2021 by two of the authors (FP and AP) following the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) guidelines [87]. The study was preregistered (March 23, 2021) on the International Platform of Registered Systematic Review and Meta-analysis Protocols (INPLASY202130081). The search databases used were PsycINFO, Web of Science, Medline, IEEExplore, and Cochrane Library.

Inclusion Criteria

In line with the PRISMA guidelines [87], the authors (FP, AP, FM) established clear inclusion criteria to determine a paper's eligibility for inclusion in the review. Only studies meeting the following criteria were considered eligible for inclusion: (1) human participants (clinical and nonclinical populations); (2) COTS video games played on a console, mobile console, PC, smartphone/tablet, or virtual reality device; (3) the comparator group was usual care intervention, nonvideo game group, or none; (4) the outcomes measured were levels of stress, anxiety, or both; and (5) the study design was a randomized controlled trial (RCT; ie, participants are randomly assigned to an experimental group or a control group), quasiexperimental (ie, nonequivalent groups, pretest-posttest, and interrupted time series), or cross-sectional/correlational (ie, employing questionnaires and large samples).

Papers published in English in peer-reviewed journals were selected and subjected to a check for the above inclusion criteria. According to the PRISMA guidelines, the authors (FP, AP, FM) established a specific date range; studies published between January 2006 and March 2021 were selected. This period was chosen as the first reports of the effects of video games on stress and anxiety reduction appeared around the 2010s [36].

Exclusion Criteria

Studies were excluded if they: (1) did not focus on the use of COTS games for diminishing stress and/or anxiety; (2) focused on games that did not meet the definition of a COTS game (ie, "games that one can purchase on the high street" [19]) or could not be purchased in online or physical stores; (3) used a modified version in its mechanics or features of a COTS game that change a fundamental aspect of the game; (4) used custom-made games (ie, serious games); (5) did not specify the title of the game used; (6) did not specify the average age or age range of the participants.

Search Terms and Selection of Papers for Inclusion

The search string was: [("video game*") OR ("computer game*")] AND [("stress") OR ("anxiety") OR ("relaxation")] AND [("study") OR ("trial") OR ("training")]. Initially, two of the authors (FP and AP) checked the titles and abstracts of the identified articles to determine their eligibility. Subsequently, they independently reviewed the full texts of potentially eligible papers. Any disagreements between the two authors (FP and AP) were discussed until reaching a consensus. When papers provided insufficient data for inclusion in the analysis, the corresponding authors were contacted to provide additional data. No additional articles emerged via manual searching and reviewing the reference lists of relevant papers.

Data Extraction

Two authors (FP and AP) independently extracted data on study characteristics and video game characteristics.

The study characteristics included the populations included in the study (participants, mean age or age range); study design (RCT, quasiexperimental, cross-sectional/correlational study); measures used for the assessment of outcomes (eg, self-report questionnaires, physiological data, cognitive task); study outcomes (ie, stress, anxiety, or both, and differences in the outcome measures related to playing COTS games). The populations, study design, measures of outcomes, and study outcomes were considered relevant variables according to the approach adopted in previous reviews [35,37,88,89] to facilitate easily classified and comparable studies in the literature. An indication of the mean age or age range identified studies performed with children (ie, under 12 years old), adolescents (12-18 years old), young adults (18-35 years old), middle-aged adults (36-55 years old), and older adults (over 55 years old). The division of these age ranges also followed previous studies [90-92].

The video game characteristics extracted included the game genre categorized as CVGs, action, adventure, racing, sports, role-playing game (RPG), strategy, simulation, exergames, and augmented reality (AR) (see Table 1); the platform for the game (console, mobile console, PC, smartphone/tablet, virtual reality); and time spent playing (duration and the total amount of sessions). Video game genre classification was considered because not all video games are equal from many aspects, and their effects strongly depend on specific characteristics of the game itself [93,94]. Since there is no standard accepted taxonomy of genre, although one of the most commonly adopted is the system proposed by Herz [95], the above categorization was proposed to be as similar as possible to the Entertainment Software Association (ESA) classification [14,96] (see Table 1). In addition to the ESA classification genres, AR games were added since they appear to be essential to the main research questions of this review [97]. Delivery platforms were considered since they represent important information about how computer games can be accessed. Since new technologies such as mobile devices and virtual reality have recently expanded how games are played, we further considered the delivery devices. Finally, in the studies that indicated play time, this information was included in the analysis, which can offer valuable insights about how and for how long to use COTS to effectively reduce stress and anxiety.

Table 1. Definitions of the main genres of video games adopted in the systematic review.

Video game genre	Definition	Examples
Action games	Require quick action and emphasize physical challenges, including hand-eye coordination and reaction time. This genre includes many subgenres such as fighting games, shooter games, and platform games	Super Mario Bros, Doom, Call of Duty, Mortal Kombat, Street Fighter
Adventure games	Characterized by complex plots and emphasize exploration and problem-solving. Typically, pure adventure games have situational problems for the player to solve, with very little or no action. If there is action, it generally includes isolated minigames	Zork, The Walking Dead, Until Dawn, Life is Strange, Heavy Rain, Beyond: Two Souls
Action-adventure games	A hybrid genre that combines core elements from both action and adventure game genres. These games require many of the same physical skills as action games, and offer a storyline, an inventory system, and other adventure games. This genre includes the subgenre of survival horror games, typically designed to scare the players	Tomb Raider, The Last of Us, Grand Theft Auto, Uncharted, Resident Evil, Left 4 Dead, Cyberpunk 2077
Casual video games	Short games with little or no plot that can be played in short sessions; they are quick and straightforward to learn	Bejeweled, Plants vs. Zombies, Tetris, FreeCell
Racing games	Racing competition with any vehicles (from real-world racing leagues to fantastical settings)	Gran Turismo, Need for Speed, Mario Kart
Sports games	Simulate the practice of sports, including team sports, combat sports, and extreme sports	FIFA 2020, NBA 2K20, Steep
Role-playing games	Players control an avatar and develop it over a certain period of time. This genre includes the subgenre of massively multiplayer online role-playing games, which are role-playing video games played online with large numbers of players	Final Fantasy, EverQuest, World of Warcraft, Dragon Quest, Diablo
Strategy games	Emphasize strategic thinking and resource management. This genre includes the multiplayer online battle arena. Each player controls a single character with unique abilities that improve throughout a game and contribute to the team's overall strategy	Age of Empires, Civilization, Halo Wars, League of Legend, Dota II
Simulation games	Designed to closely simulate aspects of real life or fictional reality	The Sims, SimCity
Exergames	A combination of video gaming and physical exercise; these games require physical effort from the player to play the game	Just Dance, Ring Fit Adventure
Augmented reality games	Combine the use of mobile technology with physical exploration in the real world	Pokémon Go, Ingress

Study Quality and Risk of Bias Assessment

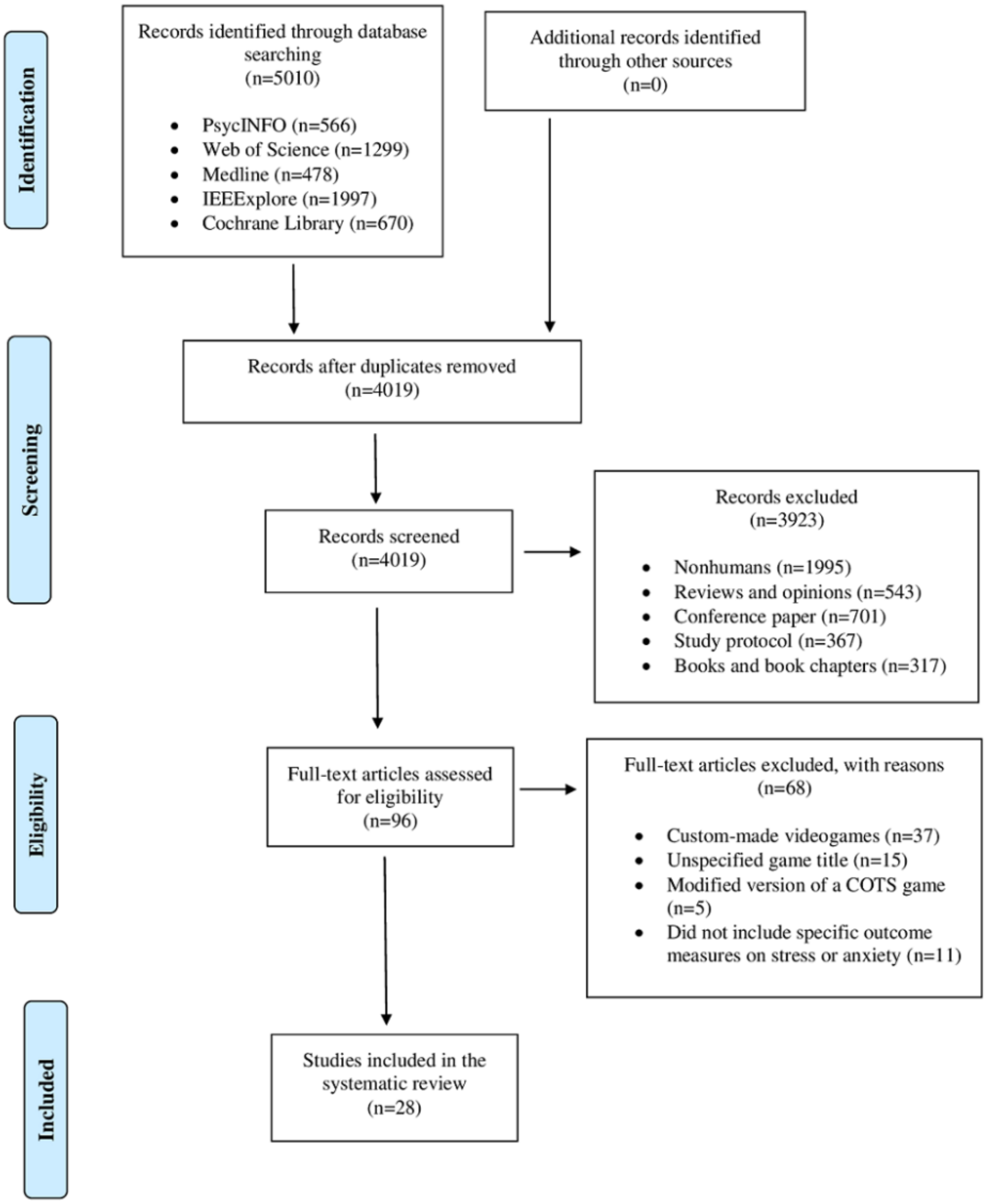
The Mixed Methods Appraisal Tool (MMAT) [98] was used to assess the methodological quality of studies included in this systematic review. The MMAT has high reliability and efficiency as a quality assessment protocol and is capable of concomitantly appraising methodological quality across various types of empirical research [99]. Two authors (FP and AP) independently assessed study quality. Interrater reliability (Cohen κ=0.816) [100], calculated using the software package SPSS, demonstrated substantial agreement [101]. Disagreements on study quality were resolved by discussion between the two authors.

Results

Retrieved Articles

The search strategy retrieved 5010 records. After deduplication and language examination, 991 studies were excluded from the review process, and 3923 studies were excluded after the first screening and title and/or abstract analysis. Ninety-six full-text copies of the remaining articles were obtained and subjected to further evaluation. After reading the full text, 68 studies were excluded from this review for the following reasons: 11 studies did not focus on the use of COTS games for diminishing stress and/or anxiety, 26 used custom-made video games, 15 did not specify the game's title, 5 adopted a modified version of a COTS game, and 11 did not include specific outcome measures on stress or anxiety. Finally, 28 studies remained for inclusion in the review (see Figure 1 and Multimedia Appendix 1).

Figure 1. PRISMA flow chart of the systematic review. COTS: commercial off-the-shelf.



Quality Assessment Outcomes

An overall quality score was assigned to each study using the MMAT scoring system [98]. Studies could be awarded a score of 0, 25, 50, 75, or 100 (with 100 indicating the highest quality). The distribution of MMAT scores among the included studies varied substantially according to study design (Table 2).

Among the 28 studies, 5 (18%) scored 100, 12 (43%) scored 75, 5 (18%) scored 50, 5 (18%) scored 25, and 1 (4%) scored 0. See Multimedia Appendix 2 for details of quality assessment for each study.

Of the 17 studies employing a quantitative RCT methodology, 11 studies did not perform randomization appropriately (ie, inserted a simple statement such as “we randomly allocated”), 10 studies did not specify if outcome assessors were blinded to the intervention provided, and 5 studies did not check or specify if the groups were comparable at baseline.

Table 2. Study design and Mixed Methods Appraisal Tool (MMAT) quality score distribution.

MMAT score distribution	References	Studies, n (%)
Quantitative randomized controlled trials (n=17)		
0	None	0 (0)
25	[36,102-105]	5 (29)
50	[106-108]	3 (18)
75	[109-113]	5 (29)
100	[85,86,114,115]	4 (24)
Quantitative nonrandomized studies (n=8)		
0	[116]	1 (13)
25	None	0 (0)
50	[80]	1 (13)
75	[117-122]	6 (75)
100	None	0 (0)
Quantitative descriptive studies (n=3)		
0	None	0 (0)
25	None	0 (0)
50	[123]	1 (33)
75	[124]	1 (33)
100	[82]	1 (33)

None of the 8 studies that used a quantitative nonrandomized methodology accounted for confounders in the design and analysis. Two of these studies did not clearly describe the target population and the sample (inclusion and exclusion criteria) [80,116]. One study did not adopt appropriate measurements and did not report complete outcome data [116].

Two of the three studies using a quantitative descriptive methodology did not report establishing an appropriate sampling strategy to address the research question [123,124] and one failed to ensure that the sample was representative of the population under study [123].

Characteristics of Included Studies

Study Design

RCT was the design of choice for 17 of the included studies. Nine studies adopted a quasiexperimental design and 2 studies [82,124] used a cross-sectional/correlational research design (see Table 2).

Populations

The number of participants ranged from 27 [109] to 337 [85,111] in RCT studies and from 1 [123] to 40 [120] in quasiexperimental studies. The two cross-sectional/correlational studies included 3915 [82] and 133 participants [124], respectively. The majority of studies (n=19) recruited young adults, while four studies involved older adults [105,109,112,117], 3 studies involved children [103,110,123], and 2 studies focused on middle-aged adults [82,102]. No study that recruited adolescents emerged in our final article list (Table 3).

Fifteen studies involved healthy participants, in most cases recruited from university staff and students, whereas three studies recruited participants among full-time workers [82], soldiers [86], and older adults in federal programs for assistance [117]. The other studies included soldiers who had posttraumatic stress disorder [113]; older adults who lived in a nursing home [112] and with Parkinson disease [109]; adults with at least minimal symptoms of depression [114], clinical depression, and comorbid anxiety [102], systemic lupus erythematosus [122], obesity [108], physical disabilities [116], and hematologic malignancies [105]; women after emergency cesarean section [115]; children with molar incisor hypomineralization–affected teeth [103], chronic wounds on the lower limbs that require active dressing changes [110], and who sustained second- and third-degree burns to the shoulders, neck, chest, bilateral forearms, and left thigh [123] (Table 3).

Table 3. Characteristics of included studies (N=28).

Characteristics	References	Studies, n (%)
Health conditions		
Healthy	[36,80,82,85,86,104,106,107,111,117-121,124]	15 (54)
Dressing change pain	[110,123]	2 (7)
Posttraumatic stress disorder	[113]	1 (4)
Parkinson disease	[109]	1 (4)
Minimal depression	[114]	1 (4)
Clinical depression and anxiety	[102]	1 (4)
Systemic lupus erythematosus	[122]	1 (4)
Obesity	[108]	1 (4)
Physical disabilities	[116]	1 (4)
Hematologic malignancies	[105]	1 (4)
Emergency cesarean section	[115]	1 (4)
Routine dental treatment	[103]	1 (4)
Institutionalized older people	[112]	1 (4)
Study outcome		
Stress	[36,80,82,85,86,104,106-108,111,118,121,123,124]	14 (50)
Anxiety	[102,103,105,109,110,112-114,117,119,120,122]	12 (43)
Both stress and anxiety	[115,116]	2 (7)
Age range		
Children (<12 years)	[103,110,123]	3 (11)
Adolescents (12-18 years)	None	0 (0)
Young adults (18-35 years)	[36,80,85,86,104,107,108,110,111,113-116,118-122,124]	19 (68)
Middle-aged adults (36-55 years)	[82,102]	2 (7)
Older adults (>55 years)	[105,109,112,117]	4 (14)
Gender		
Both male and female	[36,85,102,104-107,109-112,114,116,117,119-121,124]	18 (64)
Male only	[80,86,108,113]	4 (14)
Female only	[115,120,122,123]	4 (14)
Unspecified	[103,118]	2 (7)

Outcome Measures

All studies used self-reported quantitative measures of psychological constructs. Five studies used the State-Trait Anxiety Inventory [125], three studies [105,115,122] used the Hospital Anxiety and Depression Scale [126], and two studies [85,124] used the Perceived Stress Scale [127]. Ten studies included physiological measures, three adopted cognitive tasks [104,109,112], and four used other types of performance tasks [105,109,110,121].

Study Outcomes

Stress

Fourteen studies focused primarily on investigating COTS games for reducing stress (Table 3). Eight studies reported that COTS games were superior for reducing stress when compared with basic stress management training [86], guided relaxation or sitting quietly [104], a traditional exercise program at a moderate intensity [118], surfing the web [36], a passive video game distraction [123], a standard distraction procedure [110], and not playing games [85,111]. By contrast, a study including adult men with overweight/obesity reported increased stress levels after playing COTS games, which were higher than those recorded after watching nonviolent television [108]. Two studies compared the effects of different video games on reducing stress [106,121], showing that playing an action game elicited higher arousal and stress than playing a CVG [106,121]. Another study compared two versions (ie, cooperative vs competitive) of the same action-adventure game, showing a decrease in stress levels after both [107]. One study showed that although an action game increased stress, it also elicited happiness in players [80]. Finally, two studies investigated the relationship between stress and the use of some COTS games. Psychological stress was significantly reduced among Pokémon Go players than among nonplayers [82]. Moreover, a relationship emerged between stress levels and the use of the massively multiplayer online role-playing game (MMORPG) World of Warcraft. In particular, individuals with a low level of stress reported playing this game to enhance their offline lives. By contrast, highly stressed individuals reported that playing this game magnified rather than relieved their suffering [124].

Anxiety

Twelve studies reported outcomes for decreasing anxiety (Table 3). Nine studies reported improvement after playing a COTS game compared to an eye movement desensitization and reprocessing (EMDR) therapy alone [113], watching a film [117], surfing the web [114], anxiolytic medication [102], physiotherapy alone [105], passive video game distraction [123], and not playing a game [103,109,112]. Furthermore, in two studies that did not include a control group, exergames diminished anxiety in only one session [120] and in a more extended program including a total of 30 sessions [122]. Another study compared the efficacy of an exergame and a CVG played in virtual reality; anxiety reduction was more significant in the case of the exergame [119].

Combined Approach

Only two studies focused on both stress and anxiety in a combined manner (Table 3) [115,116]. In the first study, exergames were efficacious in reducing anxiety in a sample of individuals with physical disabilities; however, no differences emerged in stress or depression [116]. In the other study, self-reported acute stress symptoms and the frequency of intrusive traumatic memories after traumatic childbirth reduced after engaging in the brief cognitive intervention, including playing Tetris; however, no differences emerged regarding anxiety and depression [115].

Video Game Characteristics

Genre

Twenty-four studies used only one video game, whereas the other four studies adopted two video games of different genres [106,108,119,121]. Twelve studies used exergames, whereas nine studies used CVGs. Four studies adopted an action game [80,106,108,121], which was a shooter game in three studies [80,108,121] and a fighting game in the other [106]. Three studies used action-adventure games [86,107,110], which was a survival horror game in one study [86]. The other studies used an MMORPG [124], sports game [108], racing game [103], and an AR game [82] (see Table 4).

Table 4. Video game characteristics (N=28).

Characteristics	References	Studies, n (%)
Genre		
Casual video games	[36,102,104,106,113-115,119,121]	9 (28)
Exergames	[85,105,109,111,112,116-120,122,123]	12 (38)
Action games	[80,106,108,121]	4 (13)
Role-playing games	[124]	1 (3)
Action-adventure games	[86,107,110]	3 (9.4)
Sports games	[108]	1 (3)
Racing games	[103]	1 (3)
Augmented reality games	[82]	1 (3)
Platform		
PC	[36,80,86,102,104,106,114,121,124]	9 (32)
Console	[85,105,107-109,111,112,116-118,120,122,123]	13 (46)
Smartphone	[82]	1 (4)
Mobile console	[113,115]	2 (7)
Virtual reality	[103,110,119]	3 (11)
Total time of play (minutes)		
<10	[103,104,119,121]	4 (16)
11-60	[36,80,85,106-111,115-117,120,123]	14 (48)
61-180	[86,105,118]	3 (12)
>180	[102,112-114,122]	5 (16)
Unspecified	[82,124]	2 (8)

Platform

Games delivered via console were the most popular, with 13 studies using this game platform (Table 4). In particular, 7 studies used Nintendo Wii Fit, 5 studies used Microsoft Xbox 360 with Xbox Kinect, 1 study used tXbox One [107], and 1 study used Sony PlayStation 3 [108]. Nine studies used a PC, three studies

used a virtual reality viewer [103,110,119], one study used a smartphone [82], and two studies used a portable console (ie, Nintendo DS and Nintendo DS XL) [113,115].

Time of Play

In the 26 studies that measured the effect of time playing COTS games on stress and anxiety levels, there was a heterogeneous result. The mean number of sessions was 6.6, ranging from 1 session (eg, [107,116,121]) to 30 sessions [122].

The actual time spent playing video games differed among studies, ranging from about 2 minutes [103] to up to 15 hours [112,122]. Only two studies (ie, the two cross-sectional studies) did not indicate the exact playing time (Table 4).

Discussion

Principal Findings

This systematic review examined studies performed to investigate the efficacy of COTS video games for diminishing stress and anxiety. After applying the inclusion criteria, 28 papers were included for analysis. Most studies were published after 2014, with many studies (almost 40%) published after 2018. Interest in this field was crucially fueled by publication of the first study on this topic in 2009 [36]. Seventeen studies (61%) met the MMAT quality assessment score of 75% or above. This suggests that much of the research in this area is of high quality; nevertheless, the quality scores varied substantially according to the study design.

With respect to the population of focus, the majority of studies involved young adults (ie, 18-35 years). This finding also emerged in a previous systematic review on the use of video games, including COTS and custom-made games, to train cognitive skills [20]. A possible explanation of this tendency could be that many studies have enlisted college students as participants for recruitment simplicity.

Three studies involved middle-aged adults (ie, 36-55 years old) [82,102,113]. Based on emerging results, the use of COTS games can offer essential support for people of this age group, who, besides representing the most significant percentage of video game players [14], are particularly susceptible to high stress and anxiety [128,129].

Three studies recruited older adults (ie, up to 55 years old). The results of these studies suggest that the use of COTS games can be helpful for the elderly population not only to improve cognition [130-132] or to enhance physical activity [133-135] but also for relaxation [105,109,117]. This fact appears relevant since, if older adults generally do have lower stress and anxiety and better emotional regulation than younger adults [136], given the COVID-19 pandemic, this age group is currently experiencing significant adverse psychological consequences [134,135,137]. The COVID-19 pandemic has exacerbated stressors for older adults because of the risk of becoming seriously ill and the need for social isolation to mitigate this risk.

With respect to younger age groups, two studies recruited children (ie, under 12 years old) [110,123]. Playing COTS games, mainly through consoles (ie, Nintendo Wii), helped to diminish anxiety and alleviated pain in even very young children during painful or invasive medical procedures such as burn dressing changes and dental treatment. This fact appears to be important because there is a need for therapeutic alternatives within this age range with relatively limited medication options [138].

Finally, no study emerged specifically focusing on adolescents (ie, 12-18 years old). A possible explanation could be related to the intense debate in the scientific community and the general public about the effect of video games, especially those characterized by high levels of violence, on young people's mental health [139,140].

Concerning the health characteristics of the participants included in the studies, interestingly, COTS video games reduced stress and anxiety not only in healthy individuals (eg, [106,118,141]) but also in patients suffering from different mental disorders such as posttraumatic stress disorder [113]; Parkinson disease [109]; depression [102,114]; comorbid anxiety [102]; as well as physical problems such as physical disabilities [116], systemic lupus erythematosus [122], hematologic malignancies [105], or severe burns [123].

Regarding the experimental design, this systematic review showed that most studies (almost 60%) used an RCT design. Future studies should continue using this type of experimental design, representing the most reliable empirical design to prove a treatment's effectiveness, thereby minimizing the impact of confounding variables [142].

The outcome measures adopted in the studies included in this systematic review predictably primarily constituted self-administered psychological questionnaires, which were used in all studies. Nonetheless, numerous studies also included physiological measures (eg, heart rate variability, blood pressure, concentration of salivary cortisol), cognitive tests, and performance tasks, which seem to be more reliable in assessing change over time. Therefore, openness to different methods of assessment is desirable from the perspective of empirical evidence. Moreover, since many different tools are used, especially self-report questionnaires, in the future, it will be essential to define a set of standard measures for the evaluation of stress and anxiety that are specific to the different age ranges of participants.

With respect to outcomes, 14 studies included in this systematic review primarily focused on investigating COTS games for reducing stress, 12 focused on anxiety, and 2 assessed both of these conditions [115,116]. Empirical evidence emerged concerning the efficacy of COTS video games in reducing both stress and anxiety.

COTS games appear to be superior for reducing stress when compared with both control procedures (eg, sitting quietly or surfing the web) [85,110,111,123] and traditional techniques such as stress management training [86], guided relaxation [104], or a standard distraction procedure [110]. An action-adventure game (ie, Lego: Marvel Superheroes) decreased stress both in the cooperative and competitive versions [107]. In the two studies investigating the relationship between stress and the use of some COTS games, psychological stress was significantly more reduced among Pokémon Go players than among nonplayers [82]. In addition, a relationship emerged between stress levels and use of the MMORPG game World of Warcraft [124].

However, other studies reported no decreases in the levels of stress of the players. A more significant increase in stress levels emerged after playing FIFA 2013 and Call of Duty compared with the levels recorded after watching nonviolent television [108]. Playing an action game, specifically a shooter game (ie, Counter-Strike), elicited an arousal stress response but also increased happiness in players [80]. Furthermore, playing action games (ie, Mortal Kombat: Komplete Edition, Light Heroes) elicited higher levels of stress than CVGs (ie, Tetris Ultimate, Clusterz) [106,121].

Concerning anxiety, studies that emerged from this systematic review reported better improvement after playing a COTS game compared with not playing the game [103,109,112], surfing the web [114], watching a film [117], a passive video game distraction [123], EMDR therapy [113], anxiolytic medication [102], or physiotherapy alone [105]. Furthermore, two studies found a significant decrease in anxiety after a single exergame session [120] and after an exercise program with the same video game genre [122].

Regarding studies examining both stress and anxiety, in the first, anxiety, but not stress or depression, decreased after an intervention using exergames in a sample of individuals with physical disabilities [116]. In the second, playing a CVG (ie, Tetris) within a brief cognitive intervention reduced the frequency of intrusive traumatic memories after emergency cesarean section, but did not affect anxiety or depression [115].

With respect to video game characteristics, considering game genre distribution, exergames were the most frequently used, closely followed by CVGs. This result is partly surprising, as previous literature focused almost exclusively on CVGs to reduce stress and anxiety. This genre of games has proven to be able to diminish state anxiety [35,102,114] as well stress of the players [36], even to a greater extent than medical treatment [114]. In particular, playing a CVG under a prescribed condition added to an individual's medication regimen significantly reduced state anxiety symptom severity and had a medium effect on trait anxiety compared with the medication intervention alone [114].

However, based on the findings that emerged in this systematic review, CVGs are not the only promising genre for decreasing players' stress and anxiety. As noted above, many studies included in this review used different genres, especially exergames. Owing to their high level of interactivity and high-quality entertainment [143-145], exergames represent one of the most appealing video game genres for inducing positive emotions and decreasing stress and anxiety. In addition to exergames and CVGs, this systematic review showed that other genres of COTS video games could also be helpful for the reduction of stress and anxiety, including action games [80,106,108,121], particularly shooter games [80,106,121] and a fighting game [106]; action-adventure games [86,107,110], including survival horror games [86]; RPGs, in particular MMORPGs [124]; sports games [108]; racing games [103]; and AR games [82].

Concerning action games, and in particular shooter games, a study performed using Counter-Strike reported a high physiological arousal response, accompanied by the perception of a positive emotional state and decreased negative emotions [80]. Based on this result, it appears possible that shooter games activate an intense arousal response in the player while improving their emotional state, likely because they require high cognitive resources [146,147]. However, this hypothesis requires further investigation.

Two other studies included in this review showed an increase in stress levels and a physiological arousal response after playing a shooter game (ie, Call of Duty), which were higher than those measured after watching nonviolent television [108] or playing a CVG (ie, Clusterz) [121].

This systematic review also offers evidence about the efficacy of action-adventure games for reducing stress, not only in young adults but also in children [86,107,110]. This fact is interesting because this genre of games includes titles suitable for children, such as those used in the two studies that emerged from the review (ie, Ice Age 2: Meltdown and Lego: Marvel Superheroes). A subgenre of action-adventure video games that reduced stress and anxiety in young adults was survival horror.

One study performed on young adult male soldiers reported that playing a horror game (ie, Left 4 Dead) combined with biofeedback techniques reduced stress to a greater extent than training as usual [86]. Therefore, together with exergames and CVGs, horror games could represent another game type for effectively managing stress and anxiety.

One study in this review provides preliminary evidence that racing games could help decrease the players' stress and anxiety, especially in children [103]. This video game genre shares many characteristics with CVGs, such as ease of learning and short duration. For this reason, it would be interesting to further explore the use of racing games for the reduction of stress and anxiety in other age groups such as adolescents and adults who could also obtain a benefit.

Based on the results of this review, another interesting genre to help reduce stress and anxiety is AR games [82], which combines smart mobile technology with physical exploration in the real world. In particular, the players of Pokémon Go, one of the most famous titles in this category released in 2016, reported a lower level of stress than nonplayers. These findings seem very intriguing since, unlike most video games, AR games such as Pokémon Go have unique features that encourage social interaction and physical movement [148-150]. They may have a possible therapeutic role in helping stressed or anxious people deal with their everyday experiences.

Regarding the platform, 11 of the included studies delivered games via a console, especially Nintendo Wii Fit and Microsoft Xbox 360 with Xbox Kinect. Curiously, one of the most famous and popular consoles (ie, Sony PlayStation) was used in only one study in the PlayStation 3 version [108]. No study has used recent versions of this console, namely PlayStation 4 and PlayStation 5, released in 2013 and 2020. Nine studies included in the review used a PC as the platform and virtual reality systems were used in three studies [103,110,119]. Only one study adopted smartphones [82] or mobile consoles [113].

Finally, the effect of time of play was heterogeneous, both in terms of the number of sessions and the specific time spent playing video games. In particular, the number of sessions ranged from a minimum of 1 (eg, [36,110,117] to a maximum of 30 over 10 weeks [122], and the total playing time varied from a few minutes [102] to over 15 hours [122]. The fact that even single and short sessions (ie, 1 or 5 minutes) of play were effective in reducing stress and anxiety appears particularly interesting.

Potential Risks in Using COTS Video Games for Relaxation

In addition to offering data in favor of the effectiveness of COTS video games in reducing stress and anxiety, the results of this review also raise some critical reflections on the possible risks of using these games for this aim.

First, COTS games appear to be not always useful for relaxation. Some studies included in this review reported an increase in stress after playing a sports game (ie, FIFA 2013) [108], as well as a more intense stress and arousal response after an action game, in particular a fighting game (ie, Mortal Kombat: Komplete Edition), than a CVG (ie, Tetris Ultimate) [106]. In another study using an MMORPG (ie, World of Warcraft), highly stressed individuals reported that playing this game magnified rather than relieved their suffering [124].

Second, some gaming platforms are not suitable for all ages. In particular, concerning the use of video games played in virtual reality, it is important to emphasize the possible risks for children under 12 years old [151]. As indicated by all of the manufacturers of head-mounted displays, including Oculus VR, use in individuals under 12 years old is not recommended [151]. This decision connects to the fact that children are more vulnerable to virtual reality, as they are highly susceptible and can more easily confuse what is real and what is not real; thus, children may be less able or unable to distinguish between the real world and the virtual world [152,153]. To date, only one study focused on the safety of current virtual reality devices for children with respect to possible negative consequences on children's eyes [154].

Limitations

This review does not claim to be comprehensive but rather summarizes the research on COTS video games for reducing stress and anxiety based on specific keywords used in the search string, the databases searched, and the time period under analysis. Moreover, this review analyzed video games using a specific categorization of their genres, although the best approach to classify video games is an ongoing discussion. Therefore, it is essential to emphasize the specificity of the classification used, which resembles the video games' ESA classification as much as possible [14,96]. In addition, the included studies presented high heterogeneity for stress and anxiety levels and the recruited sample regarding age and health conditions. There was also considerable heterogeneity found in the COTS game genres, platforms, times of play, and methods used to assess stress and anxiety levels among the included studies. Therefore, the results from this systematic review require careful interpretation.

Future Directions

This systematic review provides several directions for future studies in this research field. First, given that COTS video games are used not only by young adults but also by people of all age groups [14], it seems necessary to further explore the use of video games to reduce stress and anxiety in diverse populations, especially in younger and older individuals. Furthermore, future studies should investigate the effectiveness of COTS video games in adolescents, a population that was not involved in any of the studies included in this review. This fact seems essential since the use of video games can favor adolescent adherence to psychological support programs more than traditional psychotherapy [155], and because young individuals often suffer from high levels of stress and anxiety, especially in this particular historical moment linked to the COVID-19 pandemic [155-157].

Second, since few studies performed to date used COTS games of a genre other than exergames or CVGs, future studies are needed to explore the efficacy of other genres in reducing stress and anxiety. In particular, based on the results that emerged from the review, the action, action-adventure, RPG, sports, racing, and AR games appear to be particularly interesting in this regard.

Third, these studies often adopted gaming platforms that are dated or not very accessible to the public. Future studies should also adopt more popular and widely used gaming devices (eg, PlayStation 4 or the newest PlayStation 5, Oculus Quest, or Oculus Quest 2). It also seems essential to investigate improved mobile gaming in the future, which could offer unique advantages over traditional tools such as a PC or console because of its potential ubiquity and real-time use.

Fourth, the most effective number of sessions and playing time required to achieve relaxation remain unclear. Future studies should address such aspects in detail, for instance by comparing shorter and longer times of play to identify the optimal playing time for reducing stress and anxiety. It also seems to be essential to verify how long the benefits of COTS games on anxiety and stress can last through follow-up studies.

Fifth, the quality assessment performed using the MMAT suggests that even if much of the research in this area is of high quality, methodological concerns are a significant issue for many studies. Researchers should follow reporting guidelines to ensure the completeness of the dissemination of research findings.

Finally, future studies are needed to explore the relationship between the effectiveness of COTS games for reducing stress and anxiety, and individual preferences concerning the genre and the gaming platform. In fact, such characteristics can have a meaningful impact on the efficacy of specific video games for diminishing stress and anxiety [58,158,159]. Furthermore, in the future, it will be necessary to investigate how other individual characteristics may influence the efficacy of COTS games in reducing stress and anxiety, including personality and cognitive ability.

Implications for Clinical Practice

The findings of this review have some practical implications for health care practitioners. The COTS games that have more experimental evidence for their effectiveness in reducing stress and anxiety are the exergames and CVGs. Even short sessions of playing (eg, 1 or 5 minutes) can be helpful for relaxation. It is possible to use COTS video games for reducing stress and anxiety not only in healthy people but also in individuals with several mental and physical health problems. Finally, when selecting the gaming platform, it is essential to consider the player's age (ie, avoiding virtual reality for children under 12 years of age).

Conclusions

To summarize, this systematic review provides evidence of the benefits of COTS video games for reducing stress in children, young adults, and older adults. Efficacy has been demonstrated not only for exergames and CVGs but also for other genres of video games including action games, action-adventure games, and AR games. Various gaming platforms (ie, consoles, PCs, smartphones, portable consoles) showed positive results, including the most innovative platforms represented by virtual reality systems. Given their low cost and popularity among millions of players worldwide, COTS games may be an important tool in reducing stress and anxiety for many individuals, diminishing the existing psychological support barriers.

Acknowledgments

FP conceived of the review and wrote the first draft of the manuscript. All authors contributed to manuscript revision, and read and approved the submitted version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Description of the studies included in this review.
[DOCX File , 49 KB - mental_v8i8e28150_app1.docx]

Multimedia Appendix 2
Quality assessment scores using the Mixed Methods Appraisal Tool (MMAT).
[DOCX File , 29 KB - mental_v8i8e28150_app2.docx]

References

1. Brooks SK, Webster RK, Smith LE, Woodland L, Wessely S, Greenberg N, et al. The psychological impact of quarantine and how to reduce it: rapid review of the evidence. Lancet 2020 Mar 14;395(10227):912-920 [FREE Full text] [doi: 10.1016/S0140-6736(20)30460-8] [Medline: 32112714]
2. Du J, Dong L, Wang T, Yuan C, Fu R, Zhang L, et al. Psychological symptoms among frontline healthcare workers during COVID-19 outbreak in Wuhan. Gen Hosp Psychiatry 2020 Apr;67:144-145 [FREE Full text] [doi: 10.1016/j.genhosppsych.2020.03.011] [Medline: 32381270]
3. Prete G, Fontanesi L, Porcelli P, Tommasi L. The psychological impact of COVID-19 in Italy: worry leads to protective behavior, but at the cost of anxiety. Front Psychol 2020 Dec 10;11:566659. [doi: 10.3389/fpsyg.2020.566659] [Medline: 33362631]
4. Odriozola-González P, Planchuelo-Gómez Á, Irurtia MJ, de Luis-García R. Psychological symptoms of the outbreak of the COVID-19 confinement in Spain. J Health Psychol 2020 Oct 30:1359105320967086. [doi: 10.1177/1359105320967086] [Medline: 33124471]
5. Wang Y, Di Y, Ye J, Wei W. Study on the public psychological states and its related factors during the outbreak of coronavirus disease 2019 (COVID-19) in some regions of China. Psychol Health Med 2021 Jan;26(1):13-22. [doi: 10.1080/13548506.2020.1746817] [Medline: 32223317]
6. Salari N, Hosseinian-Far A, Jalali R, Vaisi-Raygani A, Rasoulpoor S, Mohammadi M, et al. Prevalence of stress, anxiety, depression among the general population during the COVID-19 pandemic: a systematic review and meta-analysis. Global Health 2020 Jul 06;16(1):57 [FREE Full text] [doi: 10.1186/s12992-020-00589-w] [Medline: 32631403]

7. Conway PM, Campanini P, Sartori S, Dotti R, Costa G. Main and interactive effects of shiftwork, age and work stress on health in an Italian sample of healthcare workers. Appl Ergon 2008 Sep;39(5):630-639. [doi: 10.1016/j.apergo.2008.01.007] [Medline: 18313640]
8. Vinstrup J, Jakobsen MD, Andersen LL. Perceived stress and low-back pain among healthcare workers: a multi-center prospective cohort study. Front Public Health 2020;8:297. [doi: 10.3389/fpubh.2020.00297] [Medline: 32850571]
9. Imperatori C, Dakanalis A, Farina B, Pallavicini F, Colmegna F, Mantovani F, et al. Global storm of stress-related psychopathological symptoms: a brief overview on the usefulness of virtual reality in facing the mental health impact of COVID-19. Cyberpsychol Behav Soc Netw 2020 Nov;23(11):782-788. [doi: 10.1089/cyber.2020.0339] [Medline: 32640852]
10. Carmassi C, Foghi C, Dell'Oste V, Cordone A, Bertelloni CA, Bui E, et al. PTSD symptoms in healthcare workers facing the three coronavirus outbreaks: What can we expect after the COVID-19 pandemic. Psychiatry Res 2020 Oct;292:113312 [FREE Full text] [doi: 10.1016/j.psychres.2020.113312] [Medline: 32717711]
11. Hacimusalar Y, Kahve AC, Yasar AB, Aydin MS. Anxiety and hopelessness levels in COVID-19 pandemic: A comparative study of healthcare professionals and other community sample in Turkey. J Psychiatr Res 2020 Oct;129:181-188 [FREE Full text] [doi: 10.1016/j.jpsychires.2020.07.024] [Medline: 32758711]
12. Granic I, Lobel A, Engels RCME. The benefits of playing video games. Am Psychol 2014 Jan;69(1):66-78. [doi: 10.1037/a0034857] [Medline: 24295515]
13. Wijman T. The world's 2.7 billion gamers will spend $159.3 billion on games in 2020; the market will surpass $200 billion by 2023. Newzoo. 2020 May 08. URL: https://newzoo.com/insights/articles/newzoo-games-market-numbers-revenues-and-audience-2020-2023/ [accessed 2021-01-12]
14. 2020 Essential facts about the video game industry. Entertainment Software Association. 2020. URL: https://www.theesa.com/esa-research/2020-essential-facts-about-the-video-game-industry/ [accessed 2021-01-12]
15. Hwang G, Chen C. Influences of an inquiry-based ubiquitous gaming design on students' learning achievements, motivation, behavioral patterns, and tendency towards critical thinking and problem solving. Br J Educ Technol 2016 May 06;48(4):950-971. [doi: 10.1111/bjet.12464]
16. Gee JP. What video games have to teach us about learning and literacy. Comput Entertain 2003 Oct 01;1(1):20. [doi: 10.1145/950566.950595]
17. Vlachopoulos D, Makri A. The effect of games and simulations on higher education: a systematic literature review. Int J Educ Technol High Educ 2017 Jul 10;14(1):22. [doi: 10.1186/s41239-017-0062-1]
18. Jones CM, Scholes L, Johnson D, Katsikitis M, Carras MC. Gaming well: links between videogames and flourishing mental health. Front Psychol 2014;5:260. [doi: 10.3389/fpsyg.2014.00260] [Medline: 24744743]
19. Colder Carras M, Van Rooij AJ, Spruijt-Metz D, Kvedar J, Griffiths MD, Carabas Y, et al. Commercial video games as therapy: a new research agenda to unlock the potential of a global pastime. Front Psychiatry 2017;8:300. [doi: 10.3389/fpsyt.2017.00300] [Medline: 29403398]
20. Pallavicini F, Ferrari A, Mantovani F. Video games for well-being: a systematic review on the application of computer games for cognitive and emotional training in the adult population. Front Psychol 2018;9:2127. [doi: 10.3389/fpsyg.2018.02127] [Medline: 30464753]
21. Halbrook YJ, O'Donnell AT, Msetfi RM. When and how video games can be good: a review of the positive effects of video games on well-being. Perspect Psychol Sci 2019 Nov;14(6):1096-1104. [doi: 10.1177/1745691619863807] [Medline: 31672105]
22. Reichmuth P, Werning S. Pixel Pashas, Digital Djinns. ISIM Rev 2006;18(1):46-47 [FREE Full text]
23. Anderson CA, Shibuya A, Ihori N, Swing EL, Bushman BJ, Sakamoto A, et al. Violent video game effects on aggression, empathy, and prosocial behavior in eastern and western countries: a meta-analytic review. Psychol Bull 2010 Mar;136(2):151-173. [doi: 10.1037/a0018251] [Medline: 20192553]
24. Anderson CA, Bushman BJ. Effects of violent video games on aggressive behavior, aggressive cognition, aggressive affect, physiological arousal, and prosocial behavior: a meta-analytic review of the scientific literature. Psychol Sci 2001 Sep;12(5):353-359. [doi: 10.1111/1467-9280.00366] [Medline: 11554666]
25. Engelhardt CR, Bartholow BD, Kerr GT, Bushman BJ. This is your brain on violent video games: neural desensitization to violence predicts increased aggression following violent video game exposure. J Exp Soc Psychol 2011 Sep;47(5):1033-1036. [doi: 10.1016/j.jesp.2011.03.027]
26. Mathews CL, Morrell HER, Molle JE. Video game addiction, ADHD symptomatology, and video game reinforcement. Am J Drug Alcohol Abuse 2019;45(1):67-76. [doi: 10.1080/00952990.2018.1472269] [Medline: 29874473]
27. Müller KW, Wölfling K. Both sides of the story: Addiction is not a pastime activity. J Behav Addict 2017 Jun 01;6(2):118-120 [FREE Full text] [doi: 10.1556/2006.6.2017.038] [Medline: 28662617]
28. Hewett KJE, Zeng G, Pletcher BC. The acquisition of 21st-century skills through video games: Minecraft design process models and their web of class roles. Simulat Gaming 2020 Feb 16;51(3):336-364. [doi: 10.1177/1046878120904976]
29. Ashinoff B. The potential of video games as a pedagogical tool. Front Psychol 2014;5:1109. [doi: 10.3389/fpsyg.2014.01109] [Medline: 25324811]
30. Kenwright B. USA: Association for Computing Machinery, Inc; 2017 Presented at: SIGGRAPH Asia 2017 Symposium on Education; 2017; Bangkok, Thailand p. 1-10. [doi: 10.1145/3134368.3139220]

31. Stanmore E, Stubbs B, Vancampfort D, de Bruin ED, Firth J. The effect of active video games on cognitive functioning in clinical and non-clinical populations: A meta-analysis of randomized controlled trials. Neurosci Biobehav Rev 2017 Jul;78:34-43 [FREE Full text] [doi: 10.1016/j.neubiorev.2017.04.011] [Medline: 28442405]
32. Bediou B, Adams DM, Mayer RE, Tipton E, Green CS, Bavelier D. Meta-analysis of action video game impact on perceptual, attentional, and cognitive skills. Psychol Bull 2018 Jan;144(1):77-110. [doi: 10.1037/bul0000130] [Medline: 29172564]
33. Green CS, Bavelier D. Action video game training for cognitive enhancement. Curr Opin Behav Sci 2015 Aug;4:103-108. [doi: 10.1016/j.cobeha.2015.04.012]
34. Baniqued PL, Kranz MB, Voss MW, Lee H, Cosman JD, Severson J, et al. Corrigendum: Cognitive training with casual video games: points to consider. Front Psychol 2014;5:234. [doi: 10.3389/fpsyg.2014.00234] [Medline: 24688477]
35. Pine R, Fleming T, McCallum S, Sutcliffe K. The effects of casual videogames on anxiety, depression, stress, and low mood: a systematic review. Games Health J 2020 Aug;9(4):255-264. [doi: 10.1089/g4h.2019.0132] [Medline: 32053021]
36. Russoniello CV, O' Brien K, Parks JM. The effectiveness of casual video games in improving mood and decreasing stress. J Cyber Ther Rehab 2009;2(1):53-66.
37. Villani D, Carissoli C, Triberti S, Marchetti A, Gilli G, Riva G. Videogames for emotion regulation: a systematic review. Games Health J 2018 Apr;7(2):85-99. [doi: 10.1089/g4h.2017.0108] [Medline: 29424555]
38. Ravaja N, Salminen M, Holopainen J, Saari T, Laarni J, Järvinen A. Emotional response patterns and sense of presence during video games. New York, USA: ACM Press; 2004 Presented at: NordiCHI '04 Third Nordic Conference on Human-Computer Interactions; 2004; Tampere, Finland p. 339-347. [doi: 10.1145/1028014.1028068]
39. Olson CK. Children's motivations for video game play in the context of normal development. Rev Gen Psychol 2010 Jun 01;14(2):180-187. [doi: 10.1037/a0018984]
40. Olson C. Are electronic games health hazards or health promoters? In: Kowert R, Quandt T, editors. The video game debate: Unravelling the physical, social, and psychological effects of digital games. Oxford, UK: Taylor & Francis Group; 2015:39-53.
41. Koepp MJ, Gunn RN, Lawrence AD, Cunningham VJ, Dagher A, Jones T, et al. Evidence for striatal dopamine release during a video game. Nature 1998 May 21;393(6682):266-268. [doi: 10.1038/30498] [Medline: 9607763]
42. Fredrickson BL. The role of positive emotions in positive psychology. The broaden-and-build theory of positive emotions. Am Psychol 2001 Mar;56(3):218-226 [FREE Full text] [doi: 10.1037//0003-066x.56.3.218] [Medline: 11315248]
43. Fredrickson BL, Branigan C. Positive emotions broaden the scope of attention and thought-action repertoires. Cogn Emot 2005 May 01;19(3):313-332 [FREE Full text] [doi: 10.1080/02699930441000238] [Medline: 21852891]
44. Gloria CT, Steinhardt MA. Relationships among positive emotions, coping, resilience and mental health. Stress Health 2016 Apr;32(2):145-156. [doi: 10.1002/smi.2589] [Medline: 24962138]
45. Cohn MA, Fredrickson BL, Brown SL, Mikels JA, Conway AM. Happiness unpacked: positive emotions increase life satisfaction by building resilience. Emotion 2009 Jun;9(3):361-368 [FREE Full text] [doi: 10.1037/a0015952] [Medline: 19485613]
46. Fredrickson BL, Joiner T. Positive emotions trigger upward spirals toward emotional well-being. Psychol Sci 2002 Mar;13(2):172-175. [doi: 10.1111/1467-9280.00431] [Medline: 11934003]
47. Greenglass ER, Fiksenbaum L. Proactive coping, positive affect, and well-being. Eur Psychol 2009 Jan;14(1):29-39. [doi: 10.1027/1016-9040.14.1.29]
48. Quoidbach J, Berry EV, Hansenne M, Mikolajczak M. Positive emotion regulation and well-being: comparing the impact of eight savoring and dampening strategies. Personality Ind Diff 2010 Oct;49(5):368-373. [doi: 10.1016/j.paid.2010.03.048]
49. Cowley B, Charles D, Black M, Hickey R. Toward an understanding of flow in video games. Comput Entertain 2008 Jul;6(2):1-27. [doi: 10.1145/1371216.1371223]
50. Michailidis L, Balaguer-Ballester E, He X. Flow and immersion in video games: the aftermath of a conceptual challenge. Front Psychol 2018;9:1682. [doi: 10.3389/fpsyg.2018.01682] [Medline: 30233477]
51. Klasen M, Weber R, Kircher TTJ, Mathiak KA, Mathiak K. Neural contributions to flow experience during video game playing. Soc Cogn Affect Neurosci 2012 Apr;7(4):485-495 [FREE Full text] [doi: 10.1093/scan/nsr021] [Medline: 21596764]
52. Csikszentmihalyi M. Flow: The psychology of optimal experience. New York, NY: Harper and Row; 1990.
53. Nakamura J, Csikszentmihalyi M. The concept of flow. In: Snyder CR, Lopez SJ, editors. Handbook of positive psychology. New York: Oxford University Press; 2002:89-105.
54. Rankin K, Walsh LC, Sweeny K. A better distraction: exploring the benefits of flow during uncertain waiting periods. Emotion 2019 Aug;19(5):818-828. [doi: 10.1037/emo0000479] [Medline: 30265082]
55. Csikszentmihalyi M. Flow: the psychology of happiness. London: Ebury Publishing; 2013.
56. Seligman MEP, Csikszentmihalyi M. Positive psychology: an introduction. Am Psychol 2000;55(1):5-14. [doi: 10.1037/0003-066X.55.1.5]
57. Billieux J, Van der Linden M, Achab S, Khazaal Y, Paraskevopoulos L, Zullino D, et al. Why do you play World of Warcraft? An in-depth exploration of self-reported motivations to play online and in-game behaviours in the virtual world of Azeroth. Comput Hum Behav 2013 Jan;29(1):103-109. [doi: 10.1016/j.chb.2012.07.021]
58. Sherry J, Lucas K, Greenberg B, Lachlan K. Video game uses and gratifications as predictors of use and game preference. In: Vorderer P, Bryant J, editors. Playing video games: motives, responses, and consequences. Milton Park, Oxfordshire, UK: Taylor & Francis; 2006:213-224.

59. Yang C, Liu D. Motives matter: motives for playing Pokémon Go and implications for well-being. Cyberpsychol Behav Soc Netw 2017 Jan;20(1):52-57. [doi: 10.1089/cyber.2016.0562] [Medline: 28080150]
60. Dahlquist LM, Weiss KE, Law EF, Sil S, Herbert LJ, Horn SB, et al. Effects of videogame distraction and a virtual reality type head-mounted display helmet on cold pressor pain in young elementary school-aged children. J Pediatr Psychol 2010 Jul 28;35(6):617-625 [FREE Full text] [doi: 10.1093/jpepsy/jsp082] [Medline: 19786489]
61. Jameson E, Trevena J, Swain N. Electronic gaming as pain distraction. Pain Res Manag 2011;16(1):27-32. [doi: 10.1155/2011/856014] [Medline: 21369538]
62. Trepte S, Reinecke L, Juechems K. The social side of gaming: How playing online computer games creates online and offline social support. Comput Hum Behav 2012 May;28(3):832-839. [doi: 10.1016/j.chb.2011.12.003]
63. Arbeau K, Thorpe C, Stinson M, Budlong B, Wolff J. The meaning of the experience of being an online video game player. Comput Hum Behav Rep 2020 Aug;2:100013. [doi: 10.1016/j.chbr.2020.100013]
64. Giardina A, Di Blasi M, Schimmenti A, King D, Starcevic V, Billieux J. Online gaming and prolonged self-isolationvidence from Italian gamers during the COVID-19 outbreak. Clin Neuropsychiatry 2021;18(1):65-74. [doi: 10.36131/cnfioritieditore20210106]
65. Lewis JE, Trojovsky M, Jameson MM. New social horizons: anxiety, isolation, and animal crossing during the COVID-19 pandemic. Front Virtual Real 2021 Mar 30;2:14. [doi: 10.3389/frvir.2021.627350]
66. Fleming TM, Bavin L, Stasiak K, Hermansson-Webb E, Merry SN, Cheek C, et al. Serious games and gamification for mental health: current status and promising directions. Front Psychiatry 2016;7:215. [doi: 10.3389/fpsyt.2016.00215] [Medline: 28119636]
67. Barnes S, Prescott J. Empirical evidence for the outcomes of therapeutic video games for adolescents with anxiety disorders: systematic review. JMIR Serious Games 2018 Feb 28;6(1):e3 [FREE Full text] [doi: 10.2196/games.9530] [Medline: 29490893]
68. Horne-Moyer HL, Moyer BH, Messer DC, Messer ES. The use of electronic games in therapy: a review with clinical implications. Curr Psychiatry Rep 2014 Dec;16(12):520 [FREE Full text] [doi: 10.1007/s11920-014-0520-6] [Medline: 25312026]
69. Zyda M. From visual simulation to virtual reality to games. Computer 2005 Sep;38(9):25-32. [doi: 10.1109/mc.2005.297]
70. Weerdmeester J, Van Rooij M, Harris O, Smit N, Engels R, Granic I. Exploring the role of self-efficacy in biofeedback video games. New York, NY: Association for Computing Machinery, Inc; 2017 Presented at: The ACM SIGCHI Annual Symposium on Computer-Human Interaction in Play (CHI PLAY 2017); October 15-18, 2017; Amsterdam p. 453-461. [doi: 10.1145/3130859.3131299]
71. Bossenbroek R, Wols A, Weerdmeester J, Lichtwarck-Aschoff A, Granic I, van Rooij MMJW. Efficacy of a virtual reality biofeedback game (DEEP) to reduce anxiety and disruptive classroom behavior: single-case study. JMIR Ment Health 2020 Mar 24;7(3):e16066 [FREE Full text] [doi: 10.2196/16066] [Medline: 32207697]
72. Lobel A, Gotsis M, Reynolds E, Annetta M, Engels R, Granic I. Designingutilizing biofeedback games for emotion regulation: The case of Nevermind. USA: Association for Computing Machinery; 2016 Presented at: CHI'16: CHI Conference on Human Factors in Computing Systems; May 7-12, 2016; San Jose, CA, USA p. 1946-1951. [doi: 10.1145/2851581.2892521]
73. Schoneveld EA, Wols A, Lichtwarck-Aschoff A, Otten R, Granic I. Mental health outcomes of an applied game for children with elevated anxiety symptoms: a randomized controlled non-inferiority trial. J Child Fam Stud 2020 Jun 24;29(8):2169-2185. [doi: 10.1007/s10826-020-01728-y]
74. Schoneveld EA, Malmberg M, Lichtwarck-Aschoff A, Verheijen GP, Engels RC, Granic I. A neurofeedback video game (MindLight) to prevent anxiety in children: a randomized controlled trial. Comput Hum Behav 2016 Oct;63:321-333. [doi: 10.1016/j.chb.2016.05.005]
75. Scholten H, Malmberg M, Lobel A, Engels RCME, Granic I. A randomized controlled trial to test the effectiveness of an immersive 3D video game for anxiety prevention among adolescents. PLoS One 2016 Jan 27;11(1):e0147763 [FREE Full text] [doi: 10.1371/journal.pone.0147763] [Medline: 26816292]
76. Schuurmans AAT, Nijhof KS, Engels RCME, Granic I. Using a videogame intervention to reduce anxiety and externalizing problems among youths in residential care: an initial randomized controlled trial. J Psychopathol Behav Assess 2018 Nov 29;40(2):344-354 [FREE Full text] [doi: 10.1007/s10862-017-9638-2] [Medline: 29937625]
77. Maarsingh BM, Bos J, Van Tuijn CFJ, Renard SB. Changing stress mindset through Stressjam: a virtual reality game using biofeedback. Games Health J 2019 Oct;8(5):326-331 [FREE Full text] [doi: 10.1089/g4h.2018.0145] [Medline: 31539291]
78. Dennis TA, O'Toole L. Mental health on the go: effects of a gamified attention bias modification mobile application in trait anxious adults. Clin Psychol Sci 2014 Sep 01;2(5):576-590 [FREE Full text] [doi: 10.1177/2167702614522228] [Medline: 26029490]
79. Dennis-Tiwary TA, Egan LJ, Babkirk S, Denefrio S. For whom the bell tolls: neurocognitive individual differences in the acute stress-reduction effects of an attention bias modification game for anxiety. Behav Res Ther 2016 Feb;77:105-117 [FREE Full text] [doi: 10.1016/j.brat.2015.12.008] [Medline: 26745621]
80. Shin M, Heard R, Suo C, Chow CM. Positive emotions associated with "Counter-Strike" game playing. Games Health J 2012 Oct;1(5):342-347. [doi: 10.1089/g4h.2012.0010] [Medline: 26192000]

81. Russoniello CV, Fish M, O'Brien K. The efficacy of casual videogame play in reducing clinical depression: a randomized controlled study. Games Health J 2013 Dec;2(6):341-346. [doi: 10.1089/g4h.2013.0010] [Medline: 26197075]
82. Watanabe K, Kawakami N, Imamura K, Inoue A, Shimazu A, Yoshikawa T, et al. Pokémon GO and psychological distress, physical complaints, and work performance among adult workers: a retrospective cohort study. Sci Rep 2017 Sep 07;7(1):10758. [doi: 10.1038/s41598-017-11176-2] [Medline: 28883633]
83. Choi E, Shin S, Ryu J, Jung K, Kim S, Park M. Commercial video games and cognitive functions: video game genres and modulating factors of cognitive enhancement. Behav Brain Funct 2020 Feb 03;16(1):2 [FREE Full text] [doi: 10.1186/s12993-020-0165-z] [Medline: 32014027]
84. van Muijden J, Band GPH, Hommel B. Online games training aging brains: limited transfer to cognitive control functions. Front Hum Neurosci 2012;6:221. [doi: 10.3389/fnhum.2012.00221] [Medline: 22912609]
85. Huang H, Wong M, Yang Y, Chiu H, Teng C. Impact of playing exergames on mood states: a randomized controlled trial. Cyberpsychol Behav Soc Netw 2017 Apr;20(4):246-250. [doi: 10.1089/cyber.2016.0322] [Medline: 28394215]
86. Bouchard S, Bernier F, Boivin E, Morin B, Robillard G. Using biofeedback while immersed in a stressful videogame increases the effectiveness of stress management skills in soldiers. PLoS One 2012;7(4):e36169 [FREE Full text] [doi: 10.1371/journal.pone.0036169] [Medline: 22558370]
87. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009 Jul 21;6(7):e1000097 [FREE Full text] [doi: 10.1371/journal.pmed.1000097] [Medline: 19621072]
88. Boyle EA, Hainey T, Connolly TM, Gray G, Earp J, Ott M, et al. An update to the systematic literature review of empirical evidence of the impacts and outcomes of computer games and serious games. Comput Educ 2016 Mar;94:178-192. [doi: 10.1016/j.compedu.2015.11.003]
89. Connolly TM, Boyle EA, MacArthur E, Hainey T, Boyle JM. A systematic literature review of empirical evidence on computer games and serious games. Comput Educ 2012 Sep;59(2):661-686 [FREE Full text] [doi: 10.1016/j.compedu.2012.03.004]
90. Petry NM. A comparison of young, middle-aged, and older adult treatment-seeking pathological gamblers. Gerontologist 2002 Feb;42(1):92-99. [doi: 10.1093/geront/42.1.92] [Medline: 11815703]
91. Pigozzi LB, Pereira DD, Pattussi MP, Moret-Tatay C, Irigaray TQ, Weber JBB, et al. Quality of life in young and middle age adult temporomandibular disorders patients and asymptomatic subjects: a systematic review and meta-analysis. Health Qual Life Outcomes 2021 Mar 10;19(1):83 [FREE Full text] [doi: 10.1186/s12955-021-01727-7] [Medline: 33691709]
92. Jaworska N, MacQueen G. Adolescence as a unique developmental period. J Psychiatry Neurosci 2015 Sep;40(5):291-293 [FREE Full text] [doi: 10.1503/jpn.150268] [Medline: 26290063]
93. Achtman RL, Green CS, Bavelier D. Video games as a tool to train visual skills. Restor Neurol Neurosci 2008;26(4-5):435-446 [FREE Full text] [Medline: 18997318]
94. van Muijden J, Band GPH, Hommel B. Online games training aging brains: limited transfer to cognitive control functions. Front Hum Neurosci 2012;6:221. [doi: 10.3389/fnhum.2012.00221] [Medline: 22912609]
95. Herz JC. Joystick nation: how video games ate our quarters, stole our hearts, and rewired our minds. Boston, MA: Little, Brown, and Co; 1997.
96. 2018 Essential facts about the computer and video game industry. Entertainment Software Association. 2018. URL: https://www.theesa.com/resource/2018-essential-facts-about-the-computer-and-video-game-industry/ [accessed 2021-01-19]
97. Ellis LA, Lee MD, Ijaz K, Smith J, Braithwaite J, Yin K. COVID-19 as 'game changer' for the physical activity and mental well-being of augmented reality game players during the pandemic: mixed methods survey study. J Med Internet Res 2020 Dec 22;22(12):e25117 [FREE Full text] [doi: 10.2196/25117] [Medline: 33284781]
98. Hong QN, Fàbregues S, Bartlett G, Boardman F, Cargo M, Dagenais P, et al. The Mixed Methods Appraisal Tool (MMAT) version 2018 for information professionals and researchers. Educ Inf 2018 Dec 18;34(4):285-291. [doi: 10.3233/EFI-180221]
99. Hong QN, Gonzalez-Reyes A, Pluye P. Improving the usefulness of a tool for appraising the quality of qualitative, quantitative and mixed methods studies, the Mixed Methods Appraisal Tool (MMAT). J Eval Clin Pract 2018 Jun 21;24(3):459-467. [doi: 10.1111/jep.12884] [Medline: 29464873]
100. Cohen J. A coefficient of agreement for nominal acales. Educ Psychol Meas 2016 Jul 02;20(1):37-46. [doi: 10.1177/001316446002000104]
101. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977 Mar;33(1):159-174. [Medline: 843571]
102. Fish MT, Russoniello CV, O'Brien K. Zombies vs. anxiety: an augmentation study of prescribed video game play compared to medication in reducing anxiety symptoms. Simul Gaming 2018 May 04;49(5):553-566. [doi: 10.1177/1046878118773126]
103. Özükoç C. Reducing anxiety in children with molar incisor hypomineralization using virtual reality ? A randomized crossover study. CR Acad Bulg Sci 2020;73(11):1600-1608. [doi: 10.7546/crabs.2020.11.16]
104. Rupp MA, Sweetman R, Sosa AE, Smither JA, McConnell DS. Searching for affective and cognitive restoration: examining the restorative effects of casual video game play. Hum Factors 2017 Nov;59(7):1096-1107. [doi: 10.1177/0018720817715360] [Medline: 28636838]

105. Schumacher H, Stüwe S, Kropp P, Diedrich D, Freitag S, Greger N, et al. A prospective, randomized evaluation of the feasibility of exergaming on patients undergoing hematopoietic stem cell transplantation. Bone Marrow Transpl 2018 May;53(5):584-590. [doi: 10.1038/s41409-017-0070-8] [Medline: 29335629]
106. Porter AM, Goolkasian P. Video games and stress: how stress appraisals and game content affect cardiovascular and emotion outcomes. Front Psychol 2019;10:967. [doi: 10.3389/fpsyg.2019.00967] [Medline: 31133924]
107. Roy A, Ferguson CJ. Competitively versus cooperatively? An analysis of the effect of game play on levels of stress. Comput Hum Behav 2016 Mar;56:14-20. [doi: 10.1016/j.chb.2015.11.020]
108. Siervo M, Gan J, Fewtrell MS, Cortina-Borja M, Wells JCK. Acute effects of video-game playing versus television viewing on stress markers and food intake in overweight and obese young men: A randomised controlled trial. Appetite 2018 Jan 01;120:100-108. [doi: 10.1016/j.appet.2017.08.018] [Medline: 28843974]
109. Alves MLM, Mesquita BS, Morais WS, Leal JC, Satler CE, Dos Santos Mendes FA. Nintendo Wii™ versus Xbox Kinect™ for assisting people with Parkinson's disease. Percept Mot Skills 2018 Jun;125(3):546-565. [doi: 10.1177/0031512518769204] [Medline: 29665760]
110. Hua Y, Qiu R, Yao W, Zhang Q, Chen X. The effect of virtual reality distraction on pain relief during dressing changes in children with chronic wounds on lower limbs. Pain Manag Nurs 2015 Oct;16(5):685-691. [doi: 10.1016/j.pmn.2015.03.001] [Medline: 25972074]
111. Huang H, Nguyen HV, Cheng TCE, Wong M, Chiu H, Yang Y, et al. A randomized controlled trial on the role of enthusiasm about exergames: players' perceptions of exercise. Games Health J 2019 Jun;8(3):220-226. [doi: 10.1089/g4h.2018.0057] [Medline: 30418041]
112. Jahouh M, González-Bernal JJ, González-Santos J, Fernández-Lázaro D, Soto-Cámara R, Mielgo-Ayuso J. Impact of an intervention with Wii video games on the autonomy of activities of daily living and psychological-cognitive components in the institutionalized elderly. Int J Environ Res Public Health 2021 Feb 07;18(4):1570 [FREE Full text] [doi: 10.3390/ijerph18041570] [Medline: 33562249]
113. Butler O, Herr K, Willmund G, Gallinat J, Kühn S, Zimmermann P. Trauma, treatment and Tetris: video gaming increases hippocampal volume in male patients with combat-related posttraumatic stress disorder. J Psychiatry Neurosci 2020 Jul 01;45(4):279-287 [FREE Full text] [doi: 10.1503/jpn.190027] [Medline: 32293830]
114. Fish MT, Russoniello CV, O'Brien K. The efficacy of prescribed casual videogame play in reducing symptoms of anxiety: a randomized controlled study. Games Health J 2014 Oct;3(5):291-295. [doi: 10.1089/g4h.2013.0092] [Medline: 26192483]
115. Horsch A, Vial Y, Favrod C, Harari MM, Blackwell SE, Watson P, et al. Reducing intrusive traumatic memories after emergency caesarean section: A proof-of-principle randomized controlled study. Behav Res Ther 2017 Jul;94:36-47 [FREE Full text] [doi: 10.1016/j.brat.2017.03.018] [Medline: 28453969]
116. Singh DKA, Rahman NNA, Seffiyah R, Chang SY, Zainura AK, Aida SR, et al. Impact of virtual reality games on psychological well-being and upper limb performance in adults with physical disabilities: A pilot study. Med J Malaysia 2017 Apr;72(2):119-121 [FREE Full text] [Medline: 28473675]
117. de Morais MA, de Lima BE, Bandeira Santos LC. Acute effect of Xbox exercise on mood states in older adults. Act Adapt Aging 2019 May 28;44(2):146-156. [doi: 10.1080/01924788.2019.1618689]
118. Naugle KE, Naugle KM, Wikstrom EA. Cardiovascular and affective outcomes of active gaming: using the nintendo wii as a cardiovascular training tool. J Strength Cond Res 2014 Feb;28(2):443-451 [FREE Full text] [doi: 10.1519/JSC.0b013e31829999c3] [Medline: 23660574]
119. Pallavicini F, Pepe A. Virtual reality games and the role of body involvement in enhancing positive emotions and decreasing anxiety: within-subjects pilot study. JMIR Serious Games 2020 Jun 17;8(2):e15635 [FREE Full text] [doi: 10.2196/15635] [Medline: 32554371]
120. Viana RB, Alves CL, Vieira CA, Vancini RL, Campos MH, Gentil P, et al. Anxiolytic effects of a single session of the exergame Zumba Fitness on healthy young women. Games Health J 2017 Dec;6(6):365-370. [doi: 10.1089/g4h.2017.0085] [Medline: 29077484]
121. Yeh CS. Exploring the effects of videogame play on creativity performance and emotional responses. Comput Hum Behav 2015 Dec;53:396-407. [doi: 10.1016/j.chb.2015.07.024]
122. Yuen HK, Holthaus K, Kamen DL, Sword DO, Breland HL. Using Wii Fit to reduce fatigue among African American women with systemic lupus erythematosus: a pilot study. Lupus 2011 Oct;20(12):1293-1299 [FREE Full text] [doi: 10.1177/0961203311412098] [Medline: 21700656]
123. Sil S, Dahlquist LM, Burns AJ. Case study: videogame distraction reduces behavioral distress in a preschool-aged child undergoing repeated burn dressing changes: a single-subject design. J Pediatr Psychol 2013 Apr 17;38(3):330-341 [FREE Full text] [doi: 10.1093/jpepsy/jss128] [Medline: 23248343]
124. Snodgrass JG, Lacy MG, Dengah HF, Eisenhauer S, Batchelder G, Cookson RJ. A vacation from your mind: problematic online gaming is a stress response. Comput Hum Behav 2014 Sep;38:248-260. [doi: 10.1016/j.chb.2014.06.004]
125. Spielberger CD, Gorsuch RL, Lushene PR, Vagg PR, Jacobs AG. Manual for the State-Trait Anxiety Inventory (Form Y). Palo Alto, CA: Consulting Psychologists Press; 1983.

126. Spinhoven P, Ormel J, Sloekers PP, Kempen GI, Speckens AE, Van Hemert AM. A validation study of the Hospital Anxiety and Depression Scale (HADS) in different groups of Dutch subjects. Psychol Med 1997 Mar;27(2):363-370. [doi: 10.1017/s0033291796004382] [Medline: 9089829]
127. Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. J Health Soc Behav 1983 Dec;24(4):385-396. [doi: 10.2307/2136404]
128. Wu J, Tong H, Liu Z, Tao J, Chen L, Chan CCH, et al. Neurobiological effects of perceived stress are different between adolescents and middle-aged adults. Brain Imaging Behav 2021 Apr;15(2):846-854 [FREE Full text] [doi: 10.1007/s11682-020-00294-7] [Medline: 32737826]
129. Mirescu C, Gould E. Stress and adult neurogenesis. Hippocampus 2006;16(3):233-238. [doi: 10.1002/hipo.20155] [Medline: 16411244]
130. Bonnechère B, Langley C, Sahakian BJ. The use of commercial computerised cognitive games in older adults: a meta-analysis. Sci Rep 2020 Sep 17;10(1):15276. [doi: 10.1038/s41598-020-72281-3] [Medline: 32943742]
131. Mansor NS, Chow CM, Halaki M. Cognitive effects of video games in older adults and their moderators: a systematic review with meta-analysis and meta-regression. Aging Ment Health 2020 Jun;24(6):841-856. [doi: 10.1080/13607863.2019.1574710] [Medline: 30784296]
132. Kueider AM, Parisi JM, Gross AL, Rebok GW. Computerized cognitive training with older adults: a systematic review. PLoS One 2012;7(7):e40588 [FREE Full text] [doi: 10.1371/journal.pone.0040588] [Medline: 22792378]
133. Hall AK, Chavarria E, Maneeratana V, Chaney BH, Bernhardt JM. Health benefits of digital videogames for older adults: a systematic review of the literature. Games Health J 2012 Dec;1(6):402-410. [doi: 10.1089/g4h.2012.0046] [Medline: 26192056]
134. Maillot P, Perrot A, Hartley A. Effects of interactive physical-activity video-game training on physical and cognitive function in older adults. Psychol Aging 2012 Sep;27(3):589-600. [doi: 10.1037/a0026268] [Medline: 22122605]
135. Taylor LM, Kerse N, Frakking T, Maddison R. Active video games for improving physical performance measures in older people: a meta-analysis. J Geriatr Phys Ther 2018;41(2):108-123 [FREE Full text] [doi: 10.1519/JPT.0000000000000078] [Medline: 26974212]
136. Lee EE, Depp C, Palmer BW, Glorioso D, Daly R, Liu J, et al. High prevalence and adverse health effects of loneliness in community-dwelling adults across the lifespan: role of wisdom as a protective factor. Int Psychogeriatr 2019 Oct;31(10):1447-1462 [FREE Full text] [doi: 10.1017/S1041610218002120] [Medline: 30560747]
137. Vahia IV, Jeste DV, Reynolds CF. Older adults and the mental health effects of COVID-19. JAMA 2020 Dec 08;324(22):2253-2254. [doi: 10.1001/jama.2020.21753] [Medline: 33216114]
138. Ipser JC, Stein DJ, Hawkridge S, Hoppe L. Pharmacotherapy for anxiety disorders in children and adolescents. Cochrane Database Syst Rev 2009 Jul 08;13(3):CD005170. [doi: 10.1002/14651858.CD005170.pub2] [Medline: 19588367]
139. Porter G, Starcevic V. Are violent video games harmful? Australas Psychiatry 2007 Oct;15(5):422-426. [doi: 10.1080/10398560701463343] [Medline: 17828641]
140. Ferguson CJ. Do Angry Birds make for angry children? a meta-analysis of video game influences on children's and adolescents' aggression, mental health, prosocial behavior, and academic performance. Perspect Psychol Sci 2015 Sep;10(5):646-666. [doi: 10.1177/1745691615592234] [Medline: 26386002]
141. Huang M, Ali R, Liao J. The effect of user experience in online games on word of mouth: A pleasure-arousal-dominance (PAD) model perspective. Comput Hum Behav 2017 Oct;75:329-338. [doi: 10.1016/j.chb.2017.05.015]
142. Levin KA. Study design VII. Randomised controlled trials. Evid Based Dent 2007;8(1):22-23. [doi: 10.1038/sj.ebd.6400473] [Medline: 17380181]
143. Li J, Theng Y, Foo S. Exergames for older adults with subthreshold depression: does higher playfulness lead to better improvement in depression? Games Health J 2016 Jun;5(3):175-182. [doi: 10.1089/g4h.2015.0100] [Medline: 27135146]
144. Ho SS, Lwin MO, Sng JR, Yee AZ. Escaping through exergames: presence, enjoyment, and mood experience in predicting children's attitude toward exergames. Comput Hum Behav 2017 Jul;72:381-389. [doi: 10.1016/j.chb.2017.03.001]
145. Staiano AE, Calvert SL. Exergames for physical education courses: physical, social, and cognitive benefits. Child Dev Perspect 2011 Jun;5(2):93-98 [FREE Full text] [doi: 10.1111/j.1750-8606.2011.00162.x] [Medline: 22563349]
146. Steenbergen L, Sellaro R, Stock A, Beste C, Colzato LS. Action video gaming and cognitive control: playing first person shooter games is associated with improved action cascading but not inhibition. PLoS One 2015 Dec 10;10(12):e0144364 [FREE Full text] [doi: 10.1371/journal.pone.0144364] [Medline: 26655929]
147. Bavelier D, Green CS. Enhancing attentional control: lessons from action video games. Neuron 2019 Oct 09;104(1):147-163 [FREE Full text] [doi: 10.1016/j.neuron.2019.09.031] [Medline: 31600511]
148. Bonus JA, Peebles A, Mares M, Sarmiento IG. Look on the bright side (of media effects): Pokémon Go as a catalyst for positive life experiences. Media Psychol 2017 Apr 07;21(2):263-287. [doi: 10.1080/15213269.2017.1305280]
149. Chong Y, Sethi DK, Loh CHY, Lateef F. Going forward with Pokemon Go. J Emerg Trauma Shock 2018;11(4):243-246. [doi: 10.4103/JETS.JETS_87_17] [Medline: 30568365]
150. Xian Y, Xu H, Xu H, Liang L, Hernandez AF, Wang TY, et al. An initial evaluation of the impact of Pokémon GO on physical activity. J Am Heart Assoc 2017 May 16;6(5):e005341 [FREE Full text] [doi: 10.1161/JAHA.116.005341] [Medline: 28512111]

151. Oculus Best Practices. Oculus VR. 2017. URL: https://tinyurl.com/3sx7wc4f [accessed 2021-01-19]
152. Segovia KY, Bailenson JN. Virtually true: children's acquisition of false memories in virtual reality. Media Psychol 2009 Nov 23;12(4):371-393. [doi: 10.1080/15213260903287267]
153. Kenwright B. Virtual reality: ethical challenges and dangers [Opinion]. IEEE Technol Soc Mag 2018 Dec;37(4):20-25. [doi: 10.1109/mts.2018.2876104]
154. Tychsen L, Foeller P. Effects of immersive virtual reality headset viewing on young children: visuomotor function, postural stability, and motion sickness. Am J Ophthalmol 2020 Jan;209:151-159. [doi: 10.1016/j.ajo.2019.07.020] [Medline: 31377280]
155. Zayeni D, Raynaud J, Revet A. Therapeutic and preventive use of video games in child and adolescent psychiatry: a systematic review. Front Psychiatry 2020;11:36. [doi: 10.3389/fpsyt.2020.00036] [Medline: 32116851]
156. Zhou S, Zhang L, Wang L, Guo Z, Wang J, Chen J, et al. Prevalence and socio-demographic correlates of psychological health problems in Chinese adolescents during the outbreak of COVID-19. Eur Child Adolesc Psychiatry 2020 Jun;29(6):749-758 [FREE Full text] [doi: 10.1007/s00787-020-01541-4] [Medline: 32363492]
157. Marques de Miranda D, da Silva Athanasio B, Sena Oliveira AC, Simoes-E-Silva AC. How is COVID-19 pandemic impacting mental health of children and adolescents? Int J Disaster Risk Reduct 2020 Dec;51:101845 [FREE Full text] [doi: 10.1016/j.ijdrr.2020.101845] [Medline: 32929399]
158. Chesham A, Wyss P, Müri RM, Mosimann UP, Nef T. What older people like to play: genre preferences and acceptance of casual games. JMIR Serious Games 2017 Apr 18;5(2):e8 [FREE Full text] [doi: 10.2196/games.7025] [Medline: 28420601]
159. Gackenbach J, Bown J. Video Game Presence as a Function of Genre: A Preliminary Inquiry. Loading 2011;5(8):4-28.

Abbreviations

AR: augmented reality
COTS: commercial off-the-shelf
CVG: casual video games
EMDR: eye movement desensitization and reprocessing
ESA: Entertainment Software Association
MMAT: Mixed Methods Appraisal Tool
MMORPG: massively multiplayer online role-playing games
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analysis
RCT: randomized controlled trial
RPG: role-playing game

Edited by J Torous; submitted 23.02.21; peer-reviewed by G Dermody, A Wols; comments to author 21.03.21; revised version received 26.04.21; accepted 22.05.21; published 16.08.21.

Please cite as:
Pallavicini F, Pepe A, Mantovani F
Commercial Off-The-Shelf Video Games for Reducing Stress and Anxiety: Systematic Review
JMIR Ment Health 2021;8(8):e28150
URL: https://mental.jmir.org/2021/8/e28150
doi:10.2196/28150
PMID:34398795

Original Paper

Mental Health Service User and Worker Experiences of Psychosocial Support Via Telehealth Through the COVID-19 Pandemic: Qualitative Study

Annie Venville[1], PhD; Sarah O'Connor[2], MPH; Hannah Roeschlein[2], BHSc; Priscilla Ennals[2], PhD; Grace McLoughlan[2]; Neil Thomas[3], PhD

[1]Social Work, College of Health and Biomedicine, Victoria University, Footscray, Australia

[2]Neami National, Melbourne, Australia

[3]Centre for Mental Health, Swinburne University of Technology, Melbourne, Australia

Corresponding Author:
Annie Venville, PhD
Social Work
College of Health and Biomedicine
Victoria University
Ballarat Road
Footscray, 3011
Australia
Phone: 61 +6199194571 ext 4571
Email: annie.venville@vu.edu.au

Abstract

Background: During the COVID-19 pandemic, we saw telehealth rapidly become the primary way to receive mental health care. International research has validated many of the benefits and challenges of telehealth known beforehand for specific population groups. However, if telehealth is to assume prominence in future mental health service delivery, greater understanding of its capacity to be used to provide psychosocial support to people with complex and enduring mental health conditions is needed.

Objective: We focused on an Australian community-managed provider of psychosocial intervention and support. We aimed to understand service user and worker experiences of psychosocial support via telehealth throughout the COVID-19 pandemic.

Methods: This study was jointly developed and conducted by people with lived experience of mental ill health or distress, mental health service providers, and university-based researchers. Semistructured interviews were conducted between August and November 2020 and explored participant experiences of receiving or providing psychosocial support via telehealth, including telephone, text, and videoconferencing. Qualitative data were analyzed thematically; quantitative data were collated and analyzed using descriptive statistics.

Results: Service users (n=20) and workers (n=8) completed individual interviews via telephone or videoconferencing platform. Service users received psychosocial support services by telephone (12/20, 60%), by videoconferencing (6/20, 30%), and by both telephone and videoconferencing (2/20, 10%). Of note, 55% (11/20) of service user participants stated a future preference for in-person psychosocial support services, 30% (6/20) preferred to receive a mixture of in-person and telehealth, and 15% (3/20) elected telehealth only. Two meta-themes emerged as integral to worker and service user experience of telehealth during the pandemic: (1) creating safety and comfort and (2) a whole new way of working. The first meta-theme comprises subthemes relating to a sense of safety and comfort while using telehealth; including trusting in the relationship and having and exercising choice and control. The second meta-theme contains subthemes reflecting key challenges and opportunities associated with the shift from in-person psychosocial support to telehealth.

Conclusions: Overall, our findings highlighted that most service users experienced telehealth positively, but this was dependent on them continuing to get the support they needed in a way that was safe and comfortable. While access difficulties of a subgroup of service users should not be ignored, most service users and workers were able to adapt to telehealth by focusing on maintaining the relationship and using choice and flexibility to maintain service delivery. Although most research participants expressed a preference for a return to in-person psychosocial support or hybrid in-person and telehealth models, there was a general recognition that intentional use of telehealth could contribute to flexible and responsive service delivery. Challenges to telehealth provision of psychosocial support identified in this study are yet to be fully understood.

(JMIR Ment Health 2021;8(8):e29671) doi:10.2196/29671

KEYWORDS

telehealth; mental health; psychosocial support; COVID-19; service user; workers; qualitative; e-mental health; support; telemedicine; intervention; user experience

Introduction

COVID-19 has resulted in a dramatic increase in the use of telehealth for mental health treatment, intervention, and support. Although there had been incremental growth in the use of telehealth modalities such as videoconferencing and telephone prior to the pandemic, it had largely been seen as a substitute for in-person service provision, particularly for people unable to travel or located in remote or underserved regions; evidence showed that videoconferencing and telephone delivery was feasible with high satisfaction and equivalent outcomes, and in some cases, could reduce discomfort associated with in-person interactions [1-6]. Effective group facilitation via videoconference [7-9], and decreased feelings of anxiety and isolation in group videoconferencing settings, even for those with limited digital literacy [10], had also been reported.

Some of the most commonly cited barriers to implementing telehealth services before the pandemic included reduced opportunities for nonverbal cues and eye contact [8], reduced worker confidence in rapport-building and limited worker buy-in [11], technical issues (including poor audio quality, time lag, problems downloading software, and slow internet connection) [10]; inequitable access to internet, and reduced rapport between participants and workers [7-9]. Challenges notwithstanding, in the face of widespread international service disruption wrought by COVID-19, telehealth rapidly became, for many, the primary way to receive mental health care.

During the pandemic, extensive reporting of the benefits of telehealth consultations [12] for mental health assessment, treatment, interventions, and ongoing support largely validated the prepandemic experiences of psychiatrist- or psychologist-based services and psychoeducational programs (eg, chronic disease management, therapeutic support for anxiety and depression, and support groups). However, there has been less direct study of the impact on psychosocial support delivered to people with moderate to severe mental health problems. Psychosocial support aims to assist people make sense of what is going on for them, explore possibilities, and manage daily tasks in their natural environments—the effectiveness of which is typically reliant on strong relationships between service users and workers. These types of interventions frequently involve the provision of practical support and “walking alongside” people as they engage in work, home, and community life. The pandemic further tested the mental health of many people by creating a situation that restricted people from natural support, community connections, and the routines and rhythms of daily life that contribute to the well-being of individuals, communities, and society [13].

This project explored the provision of psychosocial support via telehealth for people with complex mental health conditions. The aims of this study were twofold—first, to understand the service user and workers experience of provision of psychosocial support via telehealth through the COVID-19 pandemic, and second, to use findings to inform planning for the incorporation of telehealth into future service delivery.

Methods

Study Setting

The study focus was an Australian community-managed provider of psychosocial intervention and support called Neami National. Neami community support workers provide tailored psychosocial interventions focused on supporting adults with moderate to severe mental health problems such as psychotic disorders and severe mood disorders, to live well in the community or navigate acute distress or ill health. Neami assists over 9000 people each year to work toward goals that may include community participation, managing daily tasks, undertaking work or study, finding housing, and strengthening social connections. These psychosocial supports and interventions are commissioned by federal or state governments or local health districts and are regularly described by service users as central to sustaining their mental health and well-being, increasing their ability to cope with the ups and downs of life, and reducing their need to access more intensive, clinical or acute mental health services.

Prior to the pandemic, Neami predominantly provided psychosocial supports for individuals or groups in face-to-face settings with infrequent telehealth offerings. During the COVID-19 response period, government lockdown restrictions affected Neami service provision across Australia, and almost overnight, in-person psychosocial support transitioned to telehealth (videoconferencing, text or instant messaging, and telephone). Nationally, staff continued to provide a range of service offerings including group and individual work via telehealth; however, the more practical face-to-face supports could not be provided in their usual way. This study took place between August and November 2020, during which Victoria was at the epicenter of Australia’s COVID-19 second wave crisis; Victoria experienced the most severe and prolonged lockdown in Australia (15 weeks). During this time, other Australian states and territories experienced fewer restrictions; however, telehealth was widely offered. Neami sought to learn from pandemic-forced changes to service delivery by understanding worker and service user experiences of these different ways of receiving psychosocial support and identifying the barriers, enablers, and opportunities for telehealth in psychosocial- or recovery-focused practice.

Design

We designed a qualitative exploratory methodology that reflected an interpretive approach [14] to understand service user and worker experience of psychosocial support via telehealth through the COVID-19 pandemic. The research was jointly developed and conducted by people with lived experience of mental ill health or distress, Neami staff, and university-based researchers.

Participants

A combination of purposive, convenience, and snowball sampling was used to recruit 2 participant groups—service users and workers—from a range of Neami sites across Australia. The study was promoted to service users (1) via information posted on a digital platform enabling direct contact with the research team, and (2) through conversation with Neami workers, who obtained permission to pass on contact details to the researchers. Workers were informed of the study through internal Neami channels and team meetings. Workers with experiences of service users being challenged by telehealth or who had become disconnected because of telehealth were intentionally sought out. Prospective participants who expressed interest in the research received detailed information about the study via an email or telephone call from a member of the research team (AV). A second contact was made with prospective participants to gain consent to participate and arrange a preferred time and modality for the interview—phone, videoconference, or a hard copy of the interview questions if they wished to respond in writing. Participants gave verbal consent at the commencement of the interview or signed and returned a written consent form prior to interview. The voluntary nature of participation was explicitly communicated, and people were advised that they could change their mind or withdraw at any time. They were also advised that their choice to not participate or to withdraw participation would not in any way affect their relationship with Neami, Neami workers, or the service they provide or receive.

Data Collection

A semistructured interview guide was used with both service users and workers. Questions explored participant experiences of receiving or providing psychosocial support via telehealth platforms including telephone, text, and videoconferencing. We were particularly interested in learning about the challenges and opportunities telehealth support offered during a global pandemic, how it differed from the types of support people had received or provided in the past, and what role they saw it playing into their future. The interviews were conducted by 5 researchers (AV, PE, SO, HR, GM) and took place between August and November 2020. They ranged in length from 30 to 60 minutes and were recorded and transcribed verbatim. Service user participants received a Aus $40 (approximately US $29.45) gift voucher as a token of appreciation.

Data Analysis

Interview data were analyzed thematically in keeping with the techniques described by Braun and Clark [15]. The process was inductive and began with 3 researchers (AV, HR, and SO) rereading the transcripts to increase familiarity with the data and noting ideas as preliminary codes. A discussion of these ideas with the whole research team then led to the generation of initial categories. Researchers (AV, HR, SO, and GM) returned to the transcripts and organized data relevant to each category. The categories were reviewed with the research team for input and feedback, and then, mapped to generate themes and subthemes. Themes were again reviewed, refined, and summarized by all researchers to ensure that they had clear parameters and fit with the coded extracts. Data extracts were checked for accuracy (SO, HR, AV). The study was approved by the University ethics committee (HRE20-115) and Neami.

Results

Table 1 provides a worked sample of the theme development.

Table 1. Subtheme development.

Theme	Subthemes	Categories	Sample quotations
Creating safety and comfort	Trusting in the relationship	Trusting workers—feeling safe	• “We did develop quite a good relationship, I would say. I feel quite comfortable being able to talk to her and, you know, let her know what's happening” [Li, Service use] • “she's always been really good and dedicates a lot of time to me“ [Christine, Service user] • “we all started off a bit clunky and then as confidence increased and as I think you said before, you know, part of that rapport building is just kind of being a bit - you know, the genuineness and if people can do that, then that encourages rapport and then that encourages trust, and then that encourages the connection. So, it's those things, isn't it? It's a package deal.” [Julia, Service user]
		Determination to keep service users engaged	• “They just found it too hard which was really disappointing because it's something I was quite excited about and they just found it too hard to do and I guess being able to sort of explain or sort of help someone navigate that process, it's kind of a bit ironic that, you know, I'm trying to do that via telehealth, sort of teach them how to use the telehealth and it just didn't work.” [Liam, Worker]
		Equalizing power—mutual learning (and we are all in this pandemic together)	• “I feel like they know more about the platforms than we do. So, I feel sometimes they've helped us out with using it than the other way around.” [Megan, Worker]

Participant Characteristics

Participants included 20 service users and 8 workers. Detailed diagnostic information was not collected from individual service users; Neami supports people with enduring mental health challenges and diagnoses—typically psychotic disorders, mood disorders, or personality disorders. Half of service user participants (10/20, 50%) and 38% of worker participants (3/8) lived in Victoria, which was subject to a prolonged period of lockdown during 2020. Most service users (17/20, 85%) had not used telehealth services for psychosocial support prior to the pandemic. Service users ranged in age from 17 to 68 years, with 55% service user participants (11/20) under 40 years of age and 45% over 40 years of age (9/20). Of the 20 service user participants, 16 (75%) self-identified as female, 4 self-identified (25%) as male, and 0 (0%) self-identified as nonbinary. All service users (20/20, 100%) described English as the main language spoken at home; 1 service user self-identified as Aboriginal and/or Torres Strait Islander, 4 self-identified cultural backgrounds including Iranian, Lebanese and Malay/Chinese. Four service users disclosed health conditions that they felt affected their ability to engage with telehealth including mild hearing impairments, back injury, migraines, chronic pain, complex PTSD, sensory issues, ear issues, and fatigue. It was not possible for us to contact service users who had disengaged from mental health support during the pandemic and impressions of their experiences of telehealth relied upon reports from workers. Of the 8 worker participants, 5 self-identified as female, and 3 self-identified as male. All worked in direct psychosocial support roles, with duration of employment at the Neami ranging from 10 months to 6 years.

Current and Preferred Use of Telehealth Services

During the study period, 60% (12/20) of service user participants received psychosocial support services from Neami by telephone, 30% of service user participants (6/20) received support services by videoconferencing, and 10% of service user participants (2/20) received support services by both telephone and videoconferencing. For those who chose to disclose information about other telehealth services with which they had engaged (for example medicine, psychiatry, psychology), 18% (3/17) received services via phone, 41% (7/17) via videoconferencing, and 41% (7/17) via a blend of approaches. Of note, 55% (11/20) of service user participants stated a future preference for in-person psychosocial support services, while 30% (6/20) preferred to receive a mixture of in-person and telehealth, and 15% (3/20) elected telehealth only.

Key Themes

Analysis of interview data revealed 2 major themes as integral to worker and service user experience of telehealth during the pandemic: (1) creating safety and comfort, and (2) a whole new way of working (Table 2).

Table 2. Themes, subthemes, and categories.

Themes and subthemes	Categories
Creating safety and comfort	
Trusting in the relationship	• Trusting workers—feeling safe • Determination to keep service users engaged • Equalizing power—mutual learning (and we are all in this pandemic together)
Matching service offering with service user need	• Getting what I need (or not) • Not getting practical support • Recognizing and responding to the nature of distress • The pros and cons of different platforms for different needs
Having and exercising choice and control	• Options made explicit; choices offered • Choices in platforms • Choices in ways and timing of engagement • Choices in content/focus
Spaces and strategies that enable privacy and safety	• Private spaces to talk freely • The value of physical therapeutic spaces • Flexibility required for privacy and safety • Digital security not really a concern • Having time creates a sense of safety
A whole new way of working	
Doing things differently	• Following the rules or making up the rules as we go • Trying things out—Learning as we go and “winging it” • Shifting assumptions about “good practice” • Service users and workers showing flexibility
Good practice in telehealth takes time, effort and organizational support	• Building skills and confidence • Developing and adjusting to new ways of working • Resourcing workers • Resourcing service users

Creating Safety and Comfort

Trusting in the Relationship

Trust was a feature of strong and safe worker and service user relationships. This, alongside worker skills and attributes, positively influenced service users' feelings of safety and comfort with telehealth. Service users expressed appreciation for workers showing patience and focus, taking the time to listen, being respectful and understanding, and adapting to reduced visual cues. All but one service user felt that telehealth sessions benefited from having a preexisting relationship with their Neami worker; this was also the case for telehealth sessions with their general practitioner, psychiatrist, or other health professional. For many service users, their sense of safety and comfort with telehealth increased with a developing therapeutic relationship, time, and practice:

> *Yeah, it was hard first, like our first few sessions because I was like, oh, new person again… and I was like, "I'm scared I'm going to have to retell everything and be like kind of opening up trauma again." It wasn't like that at all. That was just my thoughts racing but our first session on the phone was like - she was very just like, "Get to know each other, don't really talk about the mental health side of it yet." Then our second session was more comfortable. I knew her name, she knows me, and then from then we kept having phone calls.* [Carol, Service user]

Likewise, workers appreciated the trust service users placed in them and the generosity shown. For example, workers expressed some surprise that many service users

> *still rated it as a positive experience despite my misgivings about sort of like, you know, all the hiccups and interruptions and, you know, other headaches of the technology it wrought.* [Aran, Worker]

Overall, there was an acknowledgment that both workers and service users were learning together and were simultaneously adapting to the impacts of the pandemic and varying forms of lockdown and restriction. These shared experiences appeared to equalize power in some service user–worker relationships.

Matching Service Offering With Service User Need

Different forms of telehealth were perceived to be more or less helpful, at different times, for recognizing and responding to varying degrees of distress. All participants acknowledged that telehealth offered a convenience that was not always possible with in-person support. It was handy to quickly and easily obtain scripts for medication, have a brief medical consultation, and gain emotional support:

> *...picking up the phone isn't as hard as having to go out and drive all the way kind of to the other side of town to go sit in a waiting room and wait to see someone* [Christine, Service user]

In addition to savings in time and effort, some service users found that the telephone provided a safer space for psychosocial support

> *because you are not face-to-face with them in person. It feels more like a safer zone that, you know, you don't see that person* [Li, Service user]

However, for some service users, particularly those experiencing high levels of distress, telephone sessions felt far more challenging and unsafe than in-person support. One person in particular described feeling very alone and unsafe with all forms of telehealth:

> *Oh, I actually hate it. I was really surprised with hearing how it worked for other people because it just doesn't work for me. It really doesn't feel safe to me and, yeah…..like, when I'm really distressed or going into heavy things or vulnerable things, it's just I just often feel like I can't do it and I can't do it on my own and I need a lot of help and support and I just don't feel the support through telehealth.* [Anita, Service user]

Service users who found it more challenging to communicate distress over the telephone appreciated the visual cues offered by videoconferencing. This was very important for people who described having trouble expressing themselves when feeling unwell:

> *I end up losing a lot of language and so if they don't have the visual cue, I feel like possibly they don't actually pick up how not good I actually am.* [Mia, Service user]

Videoconferencing platforms added an extra challenge for some service users who reported that seeing their own image on screen made them feel uncomfortable and unsafe. For 2 participants with eating disorders, this experience was at times so distressing and distracting they reported needing to end the session earlier than scheduled.

Having and Exercising Choice and Control

> *If I had an option to pick face-to-face or telehealth, I'd probably go face-to-face but if the question was, "What's easier?" I'd say telehealth* [Carol, Service user]

Participants told us that having some capacity to control the terms of engagement with telehealth led to more positive experiences. Although the organization encouraged the use of a specific digital teleconferencing platform, workers found that many service users did not have the equipment or sufficient data required to gain access. In response to these challenges, many workers chose to adopt the platform of the service user's choice. Workers and service users reported that platform preferences were usually based on familiarity, ease of access, and confidence. In addition to choice in platform type (eg, phone, video, text messaging), workers also offered choice in the timing and availability of connection (eg, when and for how long, planned, short notice, spontaneous, asynchronous) and in the focus of the session (eg, putting some things on hold that did not feel right to focus on via telehealth, choosing matters of highest priority).

Workers told us that by offering choices, they hoped to increase service users' engagement, safety, and comfort with telehealth. However, the reality was that a significant number of service users did not have real choices or the ability to exercise control. For example, some did not have the technology or the data available to them to engage with videoconferencing or even phone contact—

> *...they literally don't have the tools to do it...* [Violet, Worker]

or could not navigate platforms on their own. Workers were concerned that a significant number of people had dropped out of contact with the service, and although some service users had indicated that they would resume contact when meetings could be in-person, others just disappeared. This seemed to be particularly the case for people engaged with group-based psychosocial interventions:

> *I saw people just fall away.* [Ava, Worker]

Spaces and Strategies That Enable Privacy and Safety

The physical space in which telehealth sessions occurred mattered to both workers and service users. A small number of service users noted that spaces in their homes did not afford the safe "therapeutic holding" they experienced in a physical office with a physically present support worker. This meant that those service users did not engage in the same therapeutic work they might have done had they attended in-person. Irrespective of the nature of the therapeutic work, feeling safe in a telehealth session for most service users meant that those they lived with could not overhear their sessions. Of equal importance for service users was the need to be reassured that their support worker was also engaging with them from a quiet and private place:

> *There's been times where it's been challenging and I was like, "Oh, am I actually - do I want to say stuff because I don't want anyone else to hear…..." I would tell her I don't feel comfortable saying some things when I can hear other people in the back.* [Carol, Service user]

Challenges to safety and privacy seemed to be exacerbated in group-based work where both service users and workers had less control over the contexts of the other group participants;

> *I've had people ask me, who is listening to this? Is this being recorded?* [Ava, Worker]

Notably, for all participants in this study, safety was also associated with the relational space afforded or restricted by telehealth;

> *I started to realize how much of my job isn't actually about what people say but picking up on other things.* [Violet, Worker]

Some service users sensed that their support workers had more time available when using telehealth; this created less time pressure during their exchanges, allowed for more conversation, and enhanced a sense of safety. Being able to control the level of anonymity in the exchange was also important for feelings of safety and comfort. This often meant that it was significant for service users to have the choice of turning their video on or off in sessions;

> *I could also have the option of turning off my camera but still being there, and they were very like accepting of that.* [Carol, Service user]

Although all workers expressed support and understanding for service users who wanted to only show a blank screen, or be the voice at the end of a phone line, some expressed doubts about their ability to connect when working with very vulnerable people they could not see—

> *Really hard to do on the phone because there's that whole disconnect.* [Ava, Worker]

Of note, few workers reported the need to address the security of online platforms used for telehealth with service users. This may be because while the security of the online platform used was important for some service users, most of the people we interviewed stated that they either had not thought about it or were not concerned. For some of the service users, trust in their worker influenced their trust in the platform used for telehealth:

> *I trust that the people that I'm having the Zoom meeting with have done their homework and they've chosen a platform which is going to be secure and not be hacked…* [Laura, Service user]

A Whole New Way of Working

Doing Things Differently

The data revealed that telehealth was not just about a different mode of connecting; it required a whole new way of working. Unable to rely on previously proven ways of providing psychosocial support, workers made great efforts to be flexible and responsive in maintaining engagement and meeting people's needs. Workers told us they needed to think and work flexibly and differently.

While the flexibility of new and different ways of working was appreciated by many service users, for some, the increased flexibility came at the cost of predictability. For one participant, changes to service delivery meant that there was

> *way too much choice*

and this left her feeling as though she

> *didn't know the rules anymore.* [Anita, Service user]

Workers told us that the loss of predictability and constant need for flexibility and adaptation left them feeling fatigued. Constantly adapting meant that changes to practice were often made in the moment. Sessions became shorter and sometimes more frequent; poor internet connections were managed by quickly switching to phone; and a lack of privacy or safety at home was accommodated by doing sessions in the car, or at a local park, or at a different time altogether. Most of the workers felt as though they were constantly "winging it"; this was the case even for those with self-described high-level technological skills.

A number of service users told us that despite workers' best efforts, the real value of psychosocial support was lost in telehealth because it did not give them the social interaction and practical support important for their well-being. As one participant said,

> *I still don't know how to handle life- for me to do this I need practice.* [Hannah, Service user]

Workers were particularly concerned that telehealth was not suitable for people experiencing financial hardship, unstable accommodation, homelessness, or in prison:

> *So, although we've got the platforms available…I think they really need that sort of practical face-to-face engagement, like either center based or outreach.* [Liam, Worker]

Good Practice in Telehealth Takes Time, Effort, and Organizational Support

Although most service users identified time savings as one of the major benefits of telehealth, workers described many aspects of telehealth as taking considerable time and effort. Being responsive to service user needs and working flexibly meant that extra time was required for planning, building technological skills, and developing supportive relationships. Workers spent time providing technical support to service users while simultaneously learning how to use new platforms themselves. They reported variable levels of organizational recognition of the time required to engage via telehealth and expressed a need for

> *having dedicated time to become proficient* [Lucas, Worker]

themselves, as well as dedicated time and resources to support service users' technology needs:

> *They [service users] just found it too hard which was really disappointing because it's something I was quite excited about and they just found it too hard to do and I guess being able to sort of explain or sort of help someone navigate that process.* [Liam, Worker]

Workers told us that it was

> *harder to make that connection...* [Violet, Worker]

with service users via telehealth, and described feeling drained and

> *stretched thin* [Lucas, Worker]

by the extra time and effort needed to build rapport. Interestingly, although most workers felt that a preexisting relationship with a service user should have made the transition to telehealth easier it was not always the case:

> *Even though we already had a prior relationship, I didn't have that sense from them that, this is a safe space, and I can talk about anything that I need to talk about in this time. They were more guarded, so I was having to work a bit harder.* [Ava, Worker]

While most research participants expressed a preference for a return to in-person psychosocial support, or hybrid in-person and telehealth models, there was a general recognition that intentional use of telehealth could contribute to flexible and responsive service delivery. Workers were hopeful that offering telehealth in future could enable better access to services for people who often find it more difficult to connect with support such as those in remote areas and people living with disabilities. Workers identified 4 key foundations necessary to build their capacity and confidence to deliver high quality psychosocial supports via telehealth. These were (1) receiving the training they need, (2) having dedicated time to develop their practice and do the behind-the-scenes work required, (3) spending time as a team to work through challenges and discover solutions collectively, and (4) having a supportive and confident manager who championed the use of telehealth. These 4 foundations were particularly important to workers with lower self-reported levels of confidence using technology.

Discussion

Principal Findings

We examined the experiences of Australian community mental health service users with the transition to telehealth service delivery during the COVID-19 pandemic. Overall, our findings highlighted that most service users experienced telehealth positively, but this was dependent on them continuing to get the support they needed in a way that was safe and comfortable. The first of 2 themes was ensuring telehealth was a safe means of interacting. Among a population with high rates of prior trauma and past experiences of involuntary or coerced treatment, trust, and safety are especially important in service delivery [16].

In common with previous telehealth implementations [5,7,17], service users indicated a need to adapt to new ways of interacting with services, and workers expressed concerns that technical difficulties and a lack of in-person presence would interfere with maintaining an effective helping relationship. We found that during this pandemic, when telehealth was typically the only way to receive psychosocial support, service users were generally willing to adapt, and adjustment to a new service delivery model could be navigated for most. Generating a sense of safety and ensuring needs could still be met appeared important to successful telehealth service delivery. Notably, safety did not seem to be related to security or usability of the telecommunication platforms used. Rather, this appeared primarily related to being able to maintain valued elements of supportive relationships with workers. However, it appeared that workers could foster this sense of safety and continuity of support by ensuring clients could exercise choices about which communication modalities to use and how to use them. This also allowed for opportunities presented by telehealth to be capitalized upon, such as convenience of access and appointments being less rushed. In line with other reports of the use of telehealth [1-7], using more remote communication channels sometimes overcame discomfort associated with in-person interaction. There also appeared to be a potential to increase the power and control the client had within the relationship through these choices, which may be especially important in this often-disempowered client group.

The service user's own environment was also a major factor in the creation of a sense of safety, with privacy being raised as an issue, consistent with the findings of some other examinations of telehealth [18,19]. Potentially, this can be navigated by problem solving—finding a suitable location for telehealth calls to take place. However, it is also important to note that workers reported disengagement from a number of service users who were unable to use telehealth and concern in meeting the needs for persons without stable accommodation, who may be most in need. This is in line with international reports of experiences, for example, a major survey of UK mental health practitioners found the majority had lost contact with some service users due to the shift to remote working [12].

The second overall theme highlighted some of the demands that the shift to telehealth resulted in for both service users and workers, including a need for flexibility, and tolerance of unpredictability. Several worker participants cited technological issues experienced by both themselves and service users as a key challenge to providing effective support via telehealth. This emphasizes the importance of services that ensure supporting systems are prioritized at an organizational level in order to allow workers to have the technical support and equipment needed, in good time [20]. In common with previous research findings [7], workers also experienced telehealth as fatiguing. This partly appeared attributable to having to adapt to a new mode of practice, but also to having to meet the additional out-of-session work required and facing greater challenges in reading nonverbal cues to maintain rapport. It is notable that randomized controlled trials have not found formal client ratings of therapeutic rapport to be lower with telehealth than with in-person delivery [5,6,21]. On the other hand, the demand on workers in maintaining rapport can be greater in the absence of nonverbal cues.

Notably, the challenges to telehealth provision of psychosocial support identified in this study, in particular creating the sense of walking alongside people in life, are yet to be fully understood. Unexpectedly, however, the need for different communication approaches also appeared to create a positive change in the way that both workers and service users thought about the nature of service interaction. This appeared to open up the possibility of new ways of practice, by embracing remote communication as a previously underutilized means of connecting with and supporting clients. This aligns with hopes that the rapid introduction of telehealth may create opportunities to transform care [22,23]. Overall, this study highlighted that, for participants, the introduction of telehealth represented a fundamental change in practice than merely a change in means of communication.

Findings highlight a range of enablers and opportunities that can inform service delivery during the immediate postpandemic period and beyond. They demonstrate that more flexible ways of providing support, including hybrid approaches combining face-to-face and telehealth options, are welcome and should be embraced. However, service changes must be accompanied by acknowledgment that telehealth involves a different way of working together and not simply a different platform. Adequate supports—resources; training; coaching and encouragement; and time to practice, build skills, and confidence—are required to enable both service users and workers in navigating the new way of working and the new rules of engagement. Positive early telehealth interactions are critical in supporting people, to persist and overcome any technical issues that arise; therefore, organizations need to foster positive attitudes and skills in staff around the potential and practice of telehealth [24] for psychosocial service delivery. Resistance to more hybrid modes of psychosocial service delivery is likely to persist from some workers and service users; however, witnessing or hearing about the benefits and being adequately supported to adapt to new ways of working may overcome this hesitation. The ongoing use of telehealth in psychosocial interventions and supports will likely see continuing creative adaptations that will benefit those using these services; however, our findings highlight the crucial role of choice and caution against wholly telehealth service provision; most service users did not want this.

Strengths and Limitations

Our study strengths were the naturalistic examination of telehealth implementation, with a research team combining lived experience and multidisciplinary expertise. However, while the researchers tried to source people who had less positive experiences of telehealth, workers informed us that many of these service users had ceased contact with Neami. Furthermore, the use of telephone or videoconferencing to conduct interviews were barriers to including a sample representing all views, and impressions of these experiences relied upon reports from workers. It should also be noted that experiences of telehealth implementation during the unique circumstances of the COVID-19 pandemic may not be generalizable to other situations. Indeed, the government restrictions on in-person contact were especially strict and prolonged for many participants in this study.

Conclusions

The findings of this study suggest that the rapid transition to telehealth created the opportunity for a whole new way of thinking about and providing psychosocial support. While access difficulties of a subgroup of service users should not be ignored, most service users and workers were able to adapt to telehealth adoption by focusing on maintaining the relationship and using choice and flexibility to maintain service delivery. Together with opportunities for increasing access, this suggests that within community mental health services, telehealth has value as a new domain of practice.

Acknowledgments

The authors would like to thank all of the people who participated in the study and also Ms Greta Baumgartel for her contribution to study design. Neami National provided funding assistance for data collection, transcription, and analysis.

Conflicts of Interest

None declared.

References

1. Dunstan DA, Tooth SM. Treatment via videoconferencing: a pilot study of delivery by clinical psychology trainees. Aust J Rural Health 2012 Apr;20(2):88-94. [doi: 10.1111/j.1440-1584.2012.01260.x] [Medline: 22435769]
2. Simpson S, Bell L, Knox J, Mitchell D. Therapy via videoconferencing: a route to client empowerment? Clin Psychol Psychother 2005 Mar;12(2):156-165. [doi: 10.1002/cpp.436]
3. Simpson S, Guerrini L, Rochford S. Telepsychology in a university psychology clinic setting: a pilot project. Aust Psychol 2020 Nov 12;50(4):285-291. [doi: 10.1111/ap.12131]
4. Yuen EK, Herbert JD, Forman EM, Goetter EM, Juarascio AS, Rabin S, et al. Acceptance based behavior therapy for social anxiety disorder through videoconferencing. J Anxiety Disord 2013 May;27(4):389-397. [doi: 10.1016/j.janxdis.2013.03.002] [Medline: 23764124]
5. Thomas N, McDonald C, de Boer K, Brand RM, Nedeljkovic M, Seabrook L. Review of the current empirical literature on using videoconferencing to deliver individual psychotherapies to adults with mental health problems. Psychol Psychother 2021 Feb 23:1-33 [FREE Full text] [doi: 10.1111/papt.12332] [Medline: 33620133]
6. Baker AL, Turner A, Beck A, Berry K, Haddock G, Kelly PJ, et al. Telephone-delivered psychosocial interventions targeting key health priorities in adults with a psychotic disorder: systematic review. Psychol Med 2018 Dec;48(16):2637-2657 [FREE Full text] [doi: 10.1017/S0033291718001125] [Medline: 29798730]
7. Wiederhold B. Connecting through technology during the coronavirus disease 2019 pandemic: avoiding "zoom fatigue". Cyberpsychol Behav Soc Netw 2020 Jul;23(7):437-438. [doi: 10.1089/cyber.2020.29188.bkw] [Medline: 32551981]
8. Chadi N, Weisbaum E, Vo DX, Ahola Kohut S. Mindfulness-based interventions for adolescents: time to consider telehealth. J Altern Complement Med 2020 Mar 01;26(3):172-175. [doi: 10.1089/acm.2019.0302] [Medline: 31765222]
9. Chang JE, Sequeira A, McCord CE, Garney WR. Videoconference grief group counseling in rural texas: outcomes, challenges, and lessons learned. J Spec Group Work 2016 Mar 03;41(2):140-160. [doi: 10.1080/01933922.2016.1146376]
10. Banbury A, Nancarrow S, Dart J, Gray L, Parkinson L. Telehealth interventions delivering home-based support group videoconferencing: systematic review. J Med Internet Res 2018 Feb 02;20(2):e25 [FREE Full text] [doi: 10.2196/jmir.8090] [Medline: 29396387]
11. Interian A, King AR, St Hill LM, Robinson CH, Damschroder LJ. Evaluating the implementation of home-based videoconferencing for providing mental health services. Psychiatr Serv 2018 Jan 01;69(1):69-75. [doi: 10.1176/appi.ps.201700004] [Medline: 28859586]
12. Johnson S, Dalton-Locke C, Vera San Juan N, Foye U, Oram S, Papamichail A, COVID-19 Mental Health Policy Research Unit Group. Impact on mental health care and on mental health service users of the COVID-19 pandemic: a mixed methods survey of UK mental health care staff. Soc Psychiatry Psychiatr Epidemiol 2021 Jan;56(1):25-37 [FREE Full text] [doi: 10.1007/s00127-020-01927-4] [Medline: 32857218]
13. Sheridan Rains L, Johnson S, Barnett P, Steare T, Needle JJ, Carr S, COVID-19 Mental Health Policy Research Unit Group. Early impacts of the COVID-19 pandemic on mental health care and on people with mental health conditions: framework synthesis of international experiences and responses. Soc Psychiatry Psychiatr Epidemiol 2021 Jan;56(1):13-24 [FREE Full text] [doi: 10.1007/s00127-020-01924-7] [Medline: 32804258]
14. Green J, Thorogood N. Qualitative Methods for Health Research. London: Sage; 2018.
15. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
16. Isobel S, Wilson A, Gill K, Howe D. 'What would a trauma-informed mental health service look like?' perspectives of people who access services. Int J Ment Health Nurs 2021 Apr;30(2):495-505. [doi: 10.1111/inm.12813] [Medline: 33219725]
17. Simpson S, Richardson L, Pietrabissa G, Castelnuovo G, Reid C. Videotherapy and therapeutic alliance in the age of COVID-19. Clin Psychol Psychother 2021 Mar 24;28(2):409-421 [FREE Full text] [doi: 10.1002/cpp.2521] [Medline: 33037682]
18. Choi NG, Hegel MT, Marti N, Marinucci ML, Sirrianni L, Bruce ML. Telehealth problem-solving therapy for depressed low-income homebound older adults. Am J Geriatr Psychiatry 2014 Mar;22(3):263-271 [FREE Full text] [doi: 10.1097/JGP.0b013e318266b356] [Medline: 23567376]
19. Franklin CL, Cuccurullo L, Walton JL, Arseneau JR, Petersen NJ. Face to face but not in the same place: a pilot study of prolonged exposure therapy. J Trauma Dissociation 2017;18(1):116-130. [doi: 10.1080/15299732.2016.1205704] [Medline: 27348462]
20. Muir SD, de Boer K, Thomas N, Seabrook E, Nedeljkovic M, Meyer D. Videoconferencing psychotherapy in the public sector: synthesis and model for implementation. JMIR Ment Health 2020 Jan 21;7(1):e14996 [FREE Full text] [doi: 10.2196/14996] [Medline: 31961334]
21. Irvine A, Drew P, Bower P, Brooks H, Gellatly J, Armitage CJ, et al. Are there interactional differences between telephone and face-to-face psychological therapy? a systematic review of comparative studies. J Affect Disord 2020 Mar 15;265:120-131 [FREE Full text] [doi: 10.1016/j.jad.2020.01.057] [Medline: 32090733]
22. Moreno C, Wykes T, Galderisi S, Nordentoft M, Crossley N, Jones N, et al. How mental health care should change as a consequence of the COVID-19 pandemic. Lancet Psychiatry 2020 Sep;7(9):813-824 [FREE Full text] [doi: 10.1016/S2215-0366(20)30307-2] [Medline: 32682460]

23. Torous J, Wykes T. Opportunities from the coronavirus disease 2019 pandemic for transforming psychiatric care with telehealth. JAMA Psychiatry 2020 Dec 01;77(12):1205-1206. [doi: 10.1001/jamapsychiatry.2020.1640] [Medline: 32391857]
24. Wind TR, Rijkeboer M, Andersson G, Riper H. The COVID-19 pandemic: the 'black swan' for mental health care and a turning point for e-health. Internet Interv 2020 Apr;20:100317 [FREE Full text] [doi: 10.1016/j.invent.2020.100317] [Medline: 32289019]

Edited by G Eysenbach; submitted 25.04.21; peer-reviewed by J Nicholas; comments to author 18.05.21; revised version received 03.06.21; accepted 17.06.21; published 12.08.21.

Please cite as:
Venville A, O'Connor S, Roeschlein H, Ennals P, McLoughlan G, Thomas N
Mental Health Service User and Worker Experiences of Psychosocial Support Via Telehealth Through the COVID-19 Pandemic: Qualitative Study
JMIR Ment Health 2021;8(8):e29671
URL: https://mental.jmir.org/2021/8/e29671
doi:10.2196/29671
PMID:34182461

Letter to the Editor

Acceptability of Using Social Media Content in Mental Health Research: A Reflection. Comment on “Twitter Users’ Views on Mental Health Crisis Resolution Team Care Compared With Stakeholder Interviews and Focus Groups: Qualitative Analysis”

Nicola Morant[1], PhD; Natasha Chilman[1,2], MSc; Brynmor Lloyd-Evans[1], PhD; Jane Wackett[1], BSc, PGCert; Sonia Johnson[1,3], DM

[1]Division of Psychiatry, University College London, London, United Kingdom

[2]Department of Psychological Medicine, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, United Kingdom

[3]Camden and Islington NHS Foundation Trust, London, United Kingdom

Corresponding Author:
Nicola Morant, PhD
Division of Psychiatry
University College London
Maple House, 149 Tottenham Court Rd, Bloomsbury
London, W1T 7NF
United Kingdom
Phone: 44 7969688554
Email: n.morant@ucl.ac.uk

Related Article:

Comment on: https://mental.jmir.org/2021/6/e25742/

(JMIR Ment Health 2021;8(8):e32475) doi:10.2196/32475

KEYWORDS

Twitter; social media; qualitative; crisis resolution team; home treatment team; mental health; acute care; severe mental illness

Our recently published paper [1] that analyzed tweets about mental health crisis teams in the United Kingdom has sparked debate and some objections on social media, and we would like to clarify our position.

Our intention in conducting this research was to amplify voices and perspectives that may not be captured in more traditional qualitative research. Our findings highlight how, compared to views obtained using interviews and focus groups [2], Twitter users reported more negative experiences of mental health crisis services and described difficulties not identified by more standard qualitative methods. Social media research enables access to a broader range of voices, as it does not rely on recruitment via services. This has potential to usefully inform service developments, particularly in highlighting problems of access, engagement, or acceptability, and can help providers become more aware of how their services could be improved.

In planning this study, we obtained university ethics approval and followed guidance for social media research [3,4]. Only tweets from public accounts were included, and we did not record Twitter usernames or profiles. To avoid tweets being traceable to specific individuals, we paraphrased tweets rather than quoting them directly, using forms of words that were common across a body of similar material, and avoided using material about specific personal difficulties or circumstances.

Although we took care with our approach and felt we were able to reflect important concerns about a major service model, the publication of our paper has raised concerns that we feel other researchers should be aware of. The paper has divided opinion, with some social media users from a range of backgrounds strongly feeling that this form of social media data use is intrusive and that, even though material was in the public domain and was paraphrased, obtaining more specific consent was warranted. Conversations on Twitter may blur public and private, especially where it is used to share experiences among a community that includes people in vulnerable states.

Acceptability of research approaches, especially to people with relevant lived experience is clearly important. We would therefore not ourselves use a similar approach again without clear guidance being available that reflects what is acceptable to a fuller range of stakeholders. We recommend to other researchers being aware that use of material from public tweets may not be found acceptable even where ethical approval has been obtained and guidance followed. We hope this letter can contribute to the development of consensus and clear guidance on how researchers should navigate the complexity and variety of views that exist about social media content.

Conflicts of Interest

None declared.

References

1. Chilman N, Morant N, Lloyd-Evans B, Wackett J, Johnson S. Twitter users' views on mental health crisis resolution team care compared with stakeholder interviews and focus groups: qualitative analysis. JMIR Ment Health 2021 Jun 29;8(6):e25742 [FREE Full text] [doi: 10.2196/25742] [Medline: 34185017]
2. Morant N, Lloyd-Evans B, Lamb D, Fullarton K, Brown E, Paterson B, CORE Service User and Carer Working groups. Crisis resolution and home treatment: stakeholders' views on critical ingredients and implementation in England. BMC Psychiatry 2017 Jul 17;17(1):254 [FREE Full text] [doi: 10.1186/s12888-017-1421-0] [Medline: 28716022]
3. Ethics guidelines for internet-mediated research. The British Psychological Society. 2013. URL: https://www.bps.org.uk/sites/www.bps.org.uk/files/Policy%20-%20Files/Ethics%20Guidelines%20for%20Internet-Mediated%20Research%20%282013%29.pdf [accessed 2021-05-28]
4. Ethics guidelines for internet-mediated research. The British Psychological Society. 2017. URL: https://beta.bps.org.uk/sites/beta.bps.org.uk/files/Policy%20-%20Files/Ethics%20Guidelines%20for%20Internet-mediated%20Research%20%282017%29.pdf [accessed 2021-07-29]

Edited by T Derrick, J Torous; submitted 29.07.21; this is a non–peer-reviewed article;accepted 29.07.21; published 17.08.21.

Please cite as:
Morant N, Chilman N, Lloyd-Evans B, Wackett J, Johnson S
Acceptability of Using Social Media Content in Mental Health Research: A Reflection. Comment on "Twitter Users' Views on Mental Health Crisis Resolution Team Care Compared With Stakeholder Interviews and Focus Groups: Qualitative Analysis"
JMIR Ment Health 2021;8(8):e32475
URL: https://mental.jmir.org/2021/8/e32475
doi:10.2196/32475
PMID:34402799

Original Paper

Factors Associated With Psychological Disturbances During the COVID-19 Pandemic: Multicountry Online Study

Martyna Plomecka[1*], MSc; Susanna Gobbi[2*], MSc; Rachael Neckels[3], MSc; Piotr Radzinski[4], MSc; Beata Skorko[5]; Samuel Lazzeri[6], MSc; Kristina Almazidou[7]; Alisa Dedic[8]; Asja Bakalovic[8], MD; Lejla Hrustic[8], MD; Zainab Ashraf[9], BA; Sarvin Es Haghi[10]; Luis Rodriguez-Pino[11], MD; Verena Waller[12], MSc; Hafsa Jabeen[13], MD; A Beyza Alp[14]; Mehdi Behnam[1], MSc; Dana Shibli[15]; Zofia Baranczuk-Turska[16], PhD; Zeeshan Haq[17], MBA; Salah Qureshi[17], MD; Adriana M Strutt[18], PhD; Ali Jawaid[19], MD, PhD

[1]Neuroscience Center Zurich, University of Zurich, Zurich, Switzerland
[2]Zurich Center for Neuroeconomics, University of Zurich, Zurich, Switzerland
[3]Department of Biomolecular Sciences, Boise State University, Boise, ID, United States
[4]Faculty of Mathematics, Informatics and Mechanics, University of Warsaw, Warsaw, Poland
[5]Medical University of Warsaw, Warsaw, Poland
[6]Faculty of Science and Engineering, University of Groningen, Groningen, Netherlands
[7]Faculty of Veterinary Medicine, Aristotle University of Thessaloniki, Thessaloniki, Greece
[8]Faculty of Medicine, University of Tuzla, Tuzla, Bosnia and Herzegovina
[9]Faculty of Arts, University of Waterloo, Waterloo, ON, Canada
[10]Faculty of Medicine, Shahid Beheshti University of Medical Sciences, Tehran, Iran
[11]Faculty of Medicine, University of Valencia, Valencia, Spain
[12]Laboratory for Applied Radiobiology, Department of Radiation Oncology, University Hospital Zurich, University of Zurich, Zurich, Switzerland
[13]Dow University of Health Sciences, Karachi, Pakistan
[14]Faculty of Medicine, Maltepe University, Istanbul, Turkey
[15]Faculty of Medicine, University of Jordan, Amman, Jordan
[16]Institute of Mathematics, University of Zurich, Zurich, Switzerland
[17]Texas Behavioral Health, Houston, TX, United States
[18]Baylor Neurology, Baylor College of Medicine, Houston, TX, United States
[19]Center of Excellence for Neural Plasticity and Brain Disorders: BRAINCITY, Nencki Institute of Experimental Biology, Warsaw, Poland
[*]these authors contributed equally

Corresponding Author:
Ali Jawaid, MD, PhD
Center of Excellence for Neural Plasticity and Brain Disorders: BRAINCITY
Nencki Institute of Experimental Biology
Ludwika Pasteura 3
Warsaw, 02-093
Poland
Phone: 48 518892048
Email: a.jawaid@nencki.edu.pl

Abstract

Background: Accumulating evidence suggests that the COVID-19 pandemic has negatively impacted the mental health of individuals. However, the susceptibility of individuals to be impacted by the pandemic is variable, suggesting potential influences of specific factors related to participants' demographics, attitudes, and practices.

Objective: We aimed to identify the factors associated with psychological symptoms related to the effects of the first wave of the pandemic in a multicountry cohort of internet users.

Methods: This study anonymously screened 13,332 internet users worldwide for acute psychological symptoms related to the COVID-19 pandemic from March 29 to April 14, 2020, during the first wave of the pandemic amidst strict lockdown conditions.

A total of 12,817 responses were considered valid. Moreover, 1077 participants from Europe were screened a second time from May 15 to May 30, 2020, to ascertain the presence of psychological effects after the ease down of restrictions.

Results: Female gender, pre-existing psychiatric conditions, and prior exposure to trauma were identified as notable factors associated with increased psychological symptoms during the first wave of COVID-19 (*P*<.001). The same factors, in addition to being related to someone who died due to COVID-19 and using social media more than usual, were associated with persistence of psychological disturbances in the limited second assessment of European participants after the restrictions had relatively eased (*P*<.001). Optimism, ability to share concerns with family and friends like usual, positive prediction about COVID-19, and daily exercise were related to fewer psychological symptoms in both assessments (*P*<.001).

Conclusions: This study highlights the significant impact of the COVID-19 pandemic at the worldwide level on the mental health of internet users and elucidates prominent associations with their demographics, history of psychiatric disease risk factors, household conditions, certain personality traits, and attitudes toward COVID-19.

(JMIR Ment Health 2021;8(8):e28736) doi:10.2196/28736

KEYWORDS

COVID-19; pandemic; mental health; depression; posttraumatic stress disorder; general psychological disturbance; global

Introduction

The emergence of novel SARS-CoV-2 in December 2019 and the global spread of COVID-19 have become the most severe and publicized human crises in recent history. As of June 29, 2021, the global burden of COVID-19 has exceeded 180 million cases worldwide [1].

The impact of COVID-19 on mental health has recently emerged as a matter of enormous concern [2]. A number of factors related to the pandemic can adversely affect the mental health of individuals, with an even higher risk in those predisposed to psychological conditions [3]. Being in quarantine or isolation for extended periods of time has been associated with depression, anger, anxiety, and suicide as reported in several studies. Similarly, the uncertainty of economic recovery and loss of job security are important factors previously associated with psychological conditions [4-6]. Concerns have also been raised about an increase in the incidents of domestic violence and "screen time" of individuals during the COVID-19 pandemic [7-9], both of which are known risk factors for the development or worsening of psychological conditions [10]. Furthermore, the fear and paranoia of being infected with SARS-CoV-2 and social discrimination could negatively impact mental well-being [11]. The fear of losing a loved one and the grief following loss are other potential disturbances to mental health accompanying serious disease outbreaks [12,13].

Therefore, an assessment of the mental health impact of the COVID-19 pandemic on a global scale is paramount for optimization of mental health services to reduce the long-term morbidity and mortality related to the COVID-19 crisis. Furthermore, this information could aid policymakers in improving the compliance of the general public to lockdown measures [3]. Importantly, COVID-19 and the resulting physical distancing measures have established an unprecedented need to implement and optimize digital mental health services. The experiences and opinions of computer-literate individuals could help in tailoring the services according to their needs, as they are the most likely beneficiaries of digital mental health services [14-16]. The identification of specific individual or community-based vulnerability patterns could also assist in developing strategies to more efficiently deliver mental health services to vulnerable groups. Similarly, by elucidating potential resilience factors that are negatively associated with psychological symptoms, digitally-based strategies could be developed to guide susceptible individuals toward activities that could lessen their distress.

To address this, we assembled a team of health professionals (neuroscientists, psychiatrists, psychologists, data scientists, and medical students) across multiple countries to develop a global online study on the mental health impact of the COVID-19 pandemic. Our first assessment employed a fully anonymous online survey screening individuals in multiple countries for indicators and/or risks of general psychological disturbance, posttraumatic stress disorder (PTSD), and depression. The prevalence of these conditions was then cross-analyzed with participants' demographics, opinions/outlooks, certain personality traits, current household conditions, previous psychiatric disease history, and factors associated with COVID-19 to identify specific risk and resilience factors. The analysis revealed alarming trends for general psychological disturbances, and risks for PTSD and depression that were specifically associated with participant demographics, personality traits, household conditions, previous psychiatric disease and/or risk factor history, and prediction about COVID-19 resolution. One month later, a limited second assessment was performed targeting European participants when lockdown restrictions had been slightly eased.

Methods

Study Design

The study included two assessments separated by 1 month. The first assessment involved a cross-sectional electronic survey–based assessment of individuals above the age of 18 years willing to participate in the study. The anonymous survey was conducted among participants from diverse demographic groups across several continents using standardized self-report scales to screen for general psychological disturbance, risk for PTSD, and symptoms of depression. The survey was available online for a period of 15 consecutive days starting at 6 pm Central European time (CET) on March 29, 2020, and concluding at 6 pm CET on April 14, 2020. The second assessment was performed 1 month after completion of the first assessment for a period of 15 consecutive days starting at 6 pm CET on May 15, 2020, and concluding at 6 pm CET on May 30, 2020. The second assessment was limited to European participants, and the participants were asked to fill the survey only if they had completed the first assessment.

Questionnaire Development

The questionnaire was developed via close consultation among a neuroscientist, a neuropsychologist, a psychiatrist, a data scientist, and a psychiatry clinic manager. The questionnaire included closed-ended questions that assessed participant characteristics and opinions, and screened for psychological conditions through standardized and validated self-report scales. The questionnaire prototype was prepared in English (Multimedia Appendix 1) and translated into 10 additional languages (Arabic, Bosnian, French, German, Greek, Italian, Persian, Polish, Spanish, and Turkish) by bilingual native speakers and vetted by volunteers native to those countries. The feasibility of each questionnaire was confirmed using pilot studies of 10 participants each. These responses were excluded from the final analysis.

The questionnaires (Multimedia Appendix 1) included a section on participant demographics (age, gender, country, residential setting, educational status, and current employment status), household conditions (working/studying from home, home isolation conditions, pet ownership, level of social contact, social media usage, and time spent exercising), COVID-19–related factors (knowing a co-worker, friend, or family member who tested positive for COVID-19 or was thought to have died due to COVID-19, and prediction about pandemic resolution), certain personality traits (level of optimism and level of extroversion), history of psychiatric disease and/or trauma, previous exposure to human crisis, and levels of satisfaction with the actions of the state and employer during the current crisis. All questionnaires were rated on binary (yes/no) responses or Likert-type scales.

The other sections contained assessments based on the World Health Organization (WHO) Self-Reporting Questionnaire-20 (SRQ), Impact of Event Scale (IES), and Beck Depression Inventory II (BDI) [17-19]. These scales were chosen based on their common usage and efficacy in previously employed work studying the psychological impact of human crises including the SARS epidemic [20-29]. The IES was purposefully adjusted to assess the impact of an ongoing event rather than a past event. For this purpose, the past tense was converted to the present tense in each question without changing the subject matter. This adjustment was performed in consultation with an independent neuropsychologist not involved in the study. For all scales, participants were prompted to think of and report their physical and psychological states during the preceding week. The second assessment was only limited to the SRQ.

Ethical Considerations

Informed consent was obtained from each participant to allow for anonymous recording, analysis, and publication of their answers. The data were collected in a completely anonymous fashion without recording any personal identifiers, ensuring that the confidentiality of the participants was maintained in all phases of the study. The study procedures were reviewed and approved by the University of Zurich Research Office for Scientific Integrity and Cantonal Ethics Commission for the canton of Zurich (Switzerland; Multimedia Appendix 2), BRAINCITY Centre of Excellence for Neural Plasticity and Brain Disorders, Nencki Institute of Experimental Biology, Warsaw (Poland; Multimedia Appendix 3), and Faculty of Medicine, University of Tuzla, Tuzla (Bosnia and Herzegovina; Multimedia Appendix 4).

Data Collection

First Assessment

Using a nonrandomized referral sampling (snowball sampling) method, participants were contacted by a team of 70 researchers of diverse nationalities (study authors and volunteers who have been acknowledged in the Acknowledgment section) via electronic communication channels that included posts on social media platforms, direct digital messaging, and personal and professional email lists. For this assessment, the data collection procedures were repeated at least thrice during the data collection period (March 29 to April 14, 2020).

The data collection strategy resulted in a total of 13,332 responses during the first assessment. Surveys in which participants were younger than 18 years (n=34), responses were missing for any dependent variables (n=112), individuals had participated a second time (n=325), and geographic location was missing (n=20), as well as those that originated from the WHO African region (n=24) were excluded from the final analysis. When the responses were missing for individual items, the missing data were considered null and excluded from the analysis for that particular variable. The number of participants for 12 featured countries and the regions encompassing the other countries is represented in Multimedia Appendix 5.

Second Assessment

For the second assessment, data collection was limited to European participants only. The data collection team from Europe called upon potential participants using the same electronic communication channels that were used for data collection during the first assessment. The participants were prompted to fill the survey only if they had previously completed the first assessment. Data collection procedures were repeated three times during the data collection period, resulting in a total of 1077 responses during the second assessment. Against the 6207 responses collected from Europe during the first assessment, this established a response rate of 17.35%.

Statistical Analysis

All statistical analyses were performed using R version 3.6.3 and *Rstudio* [30]. All figures were produced using the packages *ggplot2* [31] and *CGPfunctions* [32].

Nonadjusted Analysis for SRQ, IES, and BDI scores

Mean scores with standard deviations were calculated for the SRQ, IES, and BDI from all valid responses (n=12,817) and compared across all of the below categorical factors via Kruskal-Wallis tests with the chi-square function. The categorical factors included gender, residential status, education level, employment status, being a medical professional, working remotely from home, satisfaction with the response of the employer to the pandemic, satisfaction with the response of the state (country government) to the pandemic, home isolation status, level of interaction with family and friends, social media usage, ability to share concerns with a mental health professional, ability to share concerns with family and friends, prior exposure to a human crisis situation, previous exposure to trauma, level of extroversion, optimism about COVID-19 resolution, and one's self-determined role in the pandemic.

Multiple Regression Models for the SRQ, IES, and BDI

Multiple linear and logistic regression models were built for the SRQ, IES, and BDI, using mean scores and cutoffs for respective categorical classification.

For linear regression, generalized linear models with the *glm* function were devised using the *lme4* package [33]. The three univariate linear regression models, one each for the SRQ, IES, and BDI, were fitted and corrected for multiple comparisons followed by *glm* function analyses. Following Bonferroni correction for multiple comparisons, the *P*-value threshold was set to *P*=.017. For each linear regression model, "age" was entered as a continuous independent predictor, whereas all aforementioned predictors were entered as categorical fixed effects. Poisson family and log link function were used to model BDI and SRQ factors. In order to choose the best model (based on Akaike information criterion [AIC] or Bayesian information criterion [BIC]) from the set of predictors, stepwise model selection was performed from the *MASS* package [34].

Logistic regression was performed to generate odds ratios (ORs) for the SRQ, IES, and BDI using the following categorization scheme: SRQ: 0=normal (0-7 points), 1=concern for general psychological disturbance (8-20 points); IES: 0=normal (0-23 points), 1=PTSD is a clinical concern (24-32 points), 2=threshold for a probable PTSD diagnosis (33-36 points), 3=severe condition (high enough to induce immunosuppression) (≥37 points), and for generating ORs, the variables were regrouped as 0=no concern versus any type of concern (levels 1/2/3); BDI: 0=these ups and downs are considered normal (1-10 points), 1=mild mood disturbance (11-16 points), 2=borderline clinical depression (17-20 points), 3=moderate depression (21-30 points), 4=severe depression (31-40 points), 5=extreme depression (>40 points), and for generating ORs, the variables were regrouped as 0=no concern versus any type of concern (levels 1/2/3/4/5). Cutoffs for the SRQ, IES, and BDI were defined using least stringent thresholds for each of these measures from previous literature to ensure high sensitivity of the screening [17-20,35]. Furthermore, separate OR analysis was performed with the reference level set to 0=absence of symptom compared to presence of symptom (varying severity levels of the symptom regrouped into one category). Correlations among the SRQ, IES, and BDI were assessed through the Pearson correlation test and illustrated as x-y plots.

For the second assessment, a generalized linear model with the *glm* function was fitted using the *lme4* package [33]. All predictors were entered as categorical fixed effects. Poisson family and log link function were used to model the SRQ factor. An interaction effect was introduced to inspect whether the second assessment and working from home, satisfaction with the employer, having a pre-existing psychiatric condition, closely knowing someone who died of COVID-19, and residence (urban or rural) had a significant effect on SRQ score progression during the first and second assessments.

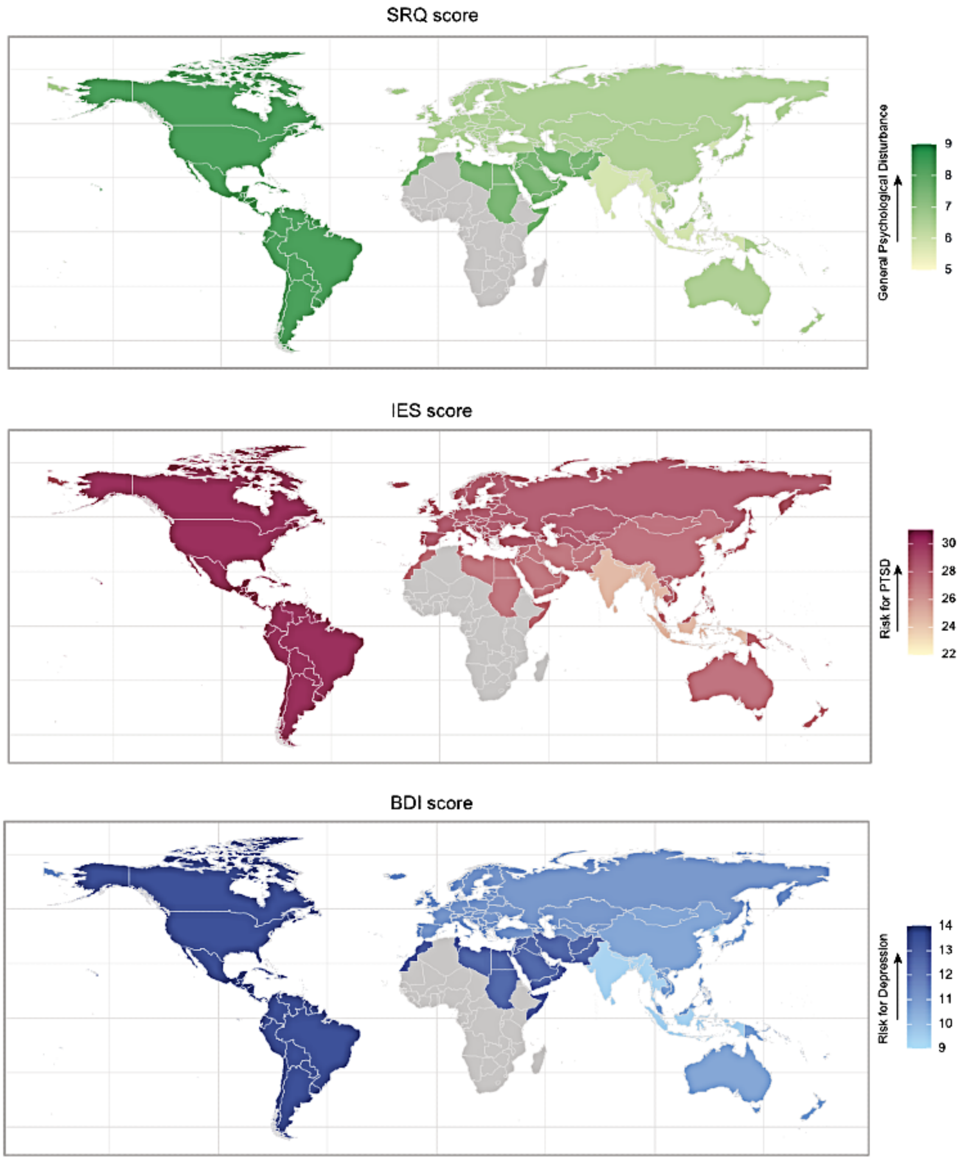
Results

First Assessment

A total of 12,817 valid responses were divided across the United States (n=1864), Iran (n=1198), Pakistan (n=1173), Poland (n=1110), Italy (n=1096), Spain (n=972), Bosnia and Herzegovina (n=885), Turkey (n=539), Canada (n=538), Germany (n=534), Switzerland (n=489), and France (n=337). The remaining countries were grouped according to WHO regions, that is, the European region (EURO; n=784), East Mediterranean region (EMRO; n=459), Western Pacific region (WPRO; n=326), South East Asian region (SEARO; n=259), and region of the Americas (PAHO; n=254). Overall, a prominent psychological impact of COVID-19 was evident worldwide with the highest SRQ scores (indicating general psychological disturbance) in Bosnia and Herzegovina, Canada, Pakistan, and the United States, and highest IES (indicating risk of PTSD) and BDI (indicating risk of depression) scores in Canada, Pakistan, and the United States (Figure 1).

There was an evident disproportion in valid responses overall, with higher numbers from those participants who reported being female (n=9314, 72.4%), residing in urban areas (n=10,666, 82.9%), having an advanced educational qualification (ie, bachelor's degree or higher) (n=9653, 75.0%), working/studying remotely from home (n=8289, 64.4%), and being under home isolation with a partner/family (n=10,691, 83.1%). Moreover, of notable prevalence were factors such as expressing satisfaction with the COVID-19–related employer response (n=4364, 33.9%), being somewhat satisfied with the COVID-19–related state response (n=4772, 37.1%), and spending less than 15 minutes on daily physical exercise (n=6306, 49.0%). A majority of participants also reported increased social media usage (n=8385, 65.1%), less than usual or minimal interaction with family and friends (n=7723, 60.0%), and feeling a sense of control in protecting themselves and others during the COVID-19 pandemic (n=10,408, 80.9%). Details of participant demographics, household conditions, history of psychiatric conditions, previous exposure to trauma/crisis, personality traits, and COVID-19–related factors and opinions are presented in Multimedia Appendix 6.

Figure 1. Geodemographic representation of global mental health burden. The three maps present mean scores from the World Health Organization (WHO) Self-Reporting Questionnaire-20 (SRQ), Impact of Event Scale (IES), and Beck Depression Inventory II (BDI). The means were calculated separately for each of the countries and for each WHO region. The total number of responders is 12,817. First panel: mean scores for the SRQ, indicating general psychological disturbance; Second panel: mean scores for the IES, indicating risk for posttraumatic stress disorder (PTSD); Third panel: mean scores for the BDI, indicating risk for depression. All mean scores were calculated separately for the featured countries and WHO regions.



Unadjusted Analysis of Risk and Resilience Factors for General Psychological Disturbance (SRQ), PTSD Risk (IES), and Depression (BDI)

Unadjusted analyses of SRQ, IES, and BDI scores between different participant demographics/characteristics showed a significantly greater prevalence (*P*<.017) of psychological symptoms in participants who were female, unemployed, working remotely from home, dissatisfied with the response of their employer/state to COVID-19, home isolated alone or with a pet, interacting with friends/family less than usual, and using social media more than usual, as well as those with a less than usual ability to share concerns with friends/family. Significantly higher scores (*P*<.017) on the SRQ, IES, and BDI were also seen in participants who self-reported being a pessimist or introvert, not feeling in control during COVID-19, and having an overall negative prediction about COVID-19 resolution. The means and standard deviations for all comparisons are presented in Table 1.

Table 1. Comparison of psychological symptoms between different participant demographics/characteristics.

Factor	Score[a], mean (SD)		
	SRQ[b]	IES[c]	BDI[d]
Gender			
Male	5.29 (4.64)[e]	23.57 (14.06)[e]	9.17 (9.07)[e]
Female	7.62 (5.05)[e]	30.22 (14.16)[e]	12.88 (10.05)[e]
Nonbinary	9.98 (5.87)[e]	34.18 (16.81)[e]	18.58 (11.78)[e]
Not disclosed	7.09 (5.32)[e]	27.78 (15.80)[e]	13.11 (10.61)[e]
Residence			
Rural	6.88 (5.08)	28.07 (14.58)	11.74 (9.60)
Urban	7.08 (5.06)	28.63 (14.43)	12.04 (10.04)
Education			
Compulsory	7.05 (5.09)[e]	27.64 (14.58)[e]	12.56 (10.51)[e]
Advanced	7.05 (5.07)[e]	28.87 (14.42)[e]	11.84 (9.81)[e]
Work status			
Private employed	6.35 (4.84)[e]	26.54 (14.05)[e]	10.30 (9.02)[e]
Public employed	6.63 (5.17)[e]	28.22 (14.71)[e]	11.02 (9.56)[e]
Freelancer	6.30 (4.81)[e]	27.19 (14.42)[e]	10.67 (9.32)[e]
Unemployed	8.14 (5.26)[e]	29.90 (15.07)[e]	13.96 (11.12)[e]
Medical or health care professional			
No	7.09 (5.09)[e]	28.61 (14.44)	12.12 (10.04)[e]
Yes	6.50 (4.87)[e]	28.01 (14.89)	10.76 (9.19)[e]
Remotely working from home			
No	6.63 (5.01)[e]	27.60 (14.88)[e]	11.70 (10.10)[e]
Yes	7.25 (5.08)[e]	29.04 (14.22)[e]	12.15 (9.91)[e]
Opinion about employer response to COVID-19			
Not satisfied	8.70 (5.22)[e]	32.39 (15.24)[e]	15.18 (11.31)[e]
Somewhat satisfied	7.64 (5.01)[e]	29.80 (14.18)[e]	12.71 (9.76)[e]
Satisfied	5.92 (4.83)[e]	26.42 (14.15)[e]	9.83 (8.99)[e]
Opinion about state response to COVID-19			
Not satisfied	7.78 (5.14)[e]	30.83 (14.76)[e]	13.74 (10.66)[e]
Somewhat satisfied	7.08 (4.96)[e]	28.55 (13.88)[e]	11.89 (9.42)[e]
Satisfied	6.25 (5.00)[e]	26.31 (14.48)[e]	10.37 (9.61)[e]
Home isolation			
Not isolated	5.29 (4.58)[e]	25.20 (14.68)[e]	9.44 (9.01)[e]
Individual home isolation	7.68 (5.37)[e]	30.04 (15.15)[e]	13.25 (10.58)[e]
Home isolation with family or partner	7.14 (5.05)[e]	28.70 (14.34)[e]	12.10 (9.97)[e]
Presence of a pet at home			
No pet at home	6.81 (5.00)[e]	27.92 (14.37)[e]	11.55 (9.85)[e]

Factor	Score[a], mean (SD)		
	SRQ[b]	IES[c]	BDI[d]
Pet at home	7.48 (5.16)[e]	29.74 (14.57)[e]	12.85 (10.16)[e]
Interaction with family or friends			
Less than usual	7.57 (5.02)[e]	29.77 (14.18)[e]	12.62 (9.87)[e]
Minimal interaction	7.34 (5.26)[e]	28.69 (14.69)[e]	12.74 (10.64)[e]
Like usual	6.41 (4.89)[e]	27.45 (14.38)[e]	10.89 (9.42)[e]
Use of social media			
Less than usual	7.61 (5.37)[e]	29.89 (16.06)[e]	13.47 (11.42)[e]
Like usual	5.56 (4.70)[e]	25.28 (14.20)[e]	10.17 (9.33)[e]
More than usual	7.64 (5.07)[e]	29.89 (14.22)[e]	12.69 (10.03)[e]
Time dedicated to physical exercise			
Less than 15 minutes	7.70 (5.17)[e]	29.33 (14.82)[e]	13.22 (10.61)[e]
More than 15 minutes	6.65 (4.90)[e]	28.26 (13.91)[e]	11.06 (9.04)[e]
More than 1 hour	5.72 (4.75)[e]	26.56 (14.30)[e]	10.06 (9.27)[e]
Close person positive for COVID-19			
No	6.97 (5.09)[e]	28.25 (14.55)[e]	12.00 (10.07)
Yes	7.26 (5.02)[e]	29.43 (14.16)[e]	12.01 (9.71)
Close person died due to COVID-19			
No	7.04 (5.08)	28.53 (14.52)	12.00 (9.99)
Yes	7.07 (4.95)	28.71 (13.67)	11.76 (9.81)
Psychiatric condition			
No psychiatric condition	6.21 (4.70)[e]	26.80 (13.88)[e]	10.34 (8.83)[e]
No change in pre-existing psychiatric condition	6.16 (4.31)[e]	25.74 (13.14)[e]	10.63 (8.53)[e]
Worsening of pre-existing psychiatric condition	12.5 (4.12)[e]	40.57 (12.84)[e]	22.53 (10.75)[e]
Ability to share concerns with a health professional			
No	8.44 (5.16)[e]	31.79 (14.46)[e]	14.50 (10.74)[e]
Yes	7.52 (5.11)[e]	30.09 (14.87)[e]	12.88 (10.35)[e]
Ability to share concerns with family or friends			
No	9.32 (5.69)[e]	31.59 (16.29)[e]	17.87 (13.25)[e]
Less than usual	9.78 (4.99)[e]	34.68 (14.23)[e]	17.06 (10.60)[e]
Like usual	5.95 (4.59)[e]	26.37 (13.67)[e]	9.78 (8.35)[e]
Previous exposure to a crisis			
No	7.05 (5.02)	28.52 (14.21)	11.99 (9.92)
Yes	7.03 (5.20)	28.79 (15.12)	12.11 (10.15)
Previous exposure to traumatic experiences			
No	6.21 (4.75)[e]	26.40 (14.05)[e]	10.46 (9.07)[e]
Yes	8.03 (5.30)[e]	31.48 (14.80)[e]	13.99 (10.87)[e]
Yes (before the age of 17 years)	7.81 (5.10)[e]	29.57 (13.87)[e]	12.92 (10.04)[e]

Factor	Score[a], mean (SD)		
	SRQ[b]	IES[c]	BDI[d]
Personality type			
Extrovert	6.36 (4.89)[e]	27.49 (14.36)[e]	10.42 (9.09)[e]
Introvert	7.65 (5.16)[e]	29.05 (14.42)[e]	13.16 (10.45)[e]
Personality			
Pessimist	9.99 (4.98)[e]	34.89 (14.46)[e]	18.41 (11.23)[e]
Optimist	5.57 (4.62)[e]	25.81 (13.81)[e]	8.86 (7.92)[e]
Realist	7.33 (4.98)[e]	28.86 (14.24)[e]	12.61 (9.95)[e]
Prediction about COVID-19 outcome/resolution			
It might be the end of the human race	10.00 (5.42)[e]	38.41 (16.48)[e]	21.88 (13.85)[e]
It will resolve after many months or years	7.81 (5.20)[e]	30.62 (14.92)[e]	13.64 (10.68)[e]
It will resolve in the summer but not within a month	6.76 (4.93)[e]	27.94 (13.93)[e]	11.23 (9.41)[e]
It will resolve within a month	6.36 (5.21)[e]	26.63 (14.80)[e]	10.62 (9.70)[e]
Self-opinion in the COVID-19 pandemic			
It is not in my control at all	10.11 (5.39)[e]	34.77 (16.50)[e]	18.65 (13.70)[e]
It is not in my control, but I can take precautions to protect myself	7.83 (5.30)[e]	30.39 (15.23)[e]	13.45 (10.69)[e]
It is not in my control, but I can take precautions to protect myself and others	6.77 (4.96)[e]	28.03 (14.10)[e]	11.48 (9.51)[e]

[a]The scores are divided according to different participant demographics/characteristics and compared through unadjusted Kruskal-Wallis tests.

[b]SRQ: Self-Reporting Questionnaire-20.

[c]IES: Impact of Event Scale.

[d]BDI: Beck Depression Inventory II.

[e]Significant differences (*P* value threshold set to *P*<.017 after multiple comparison correction) in mean scores are indicated. Each association indicates a difference in categories reported in the column vertically.

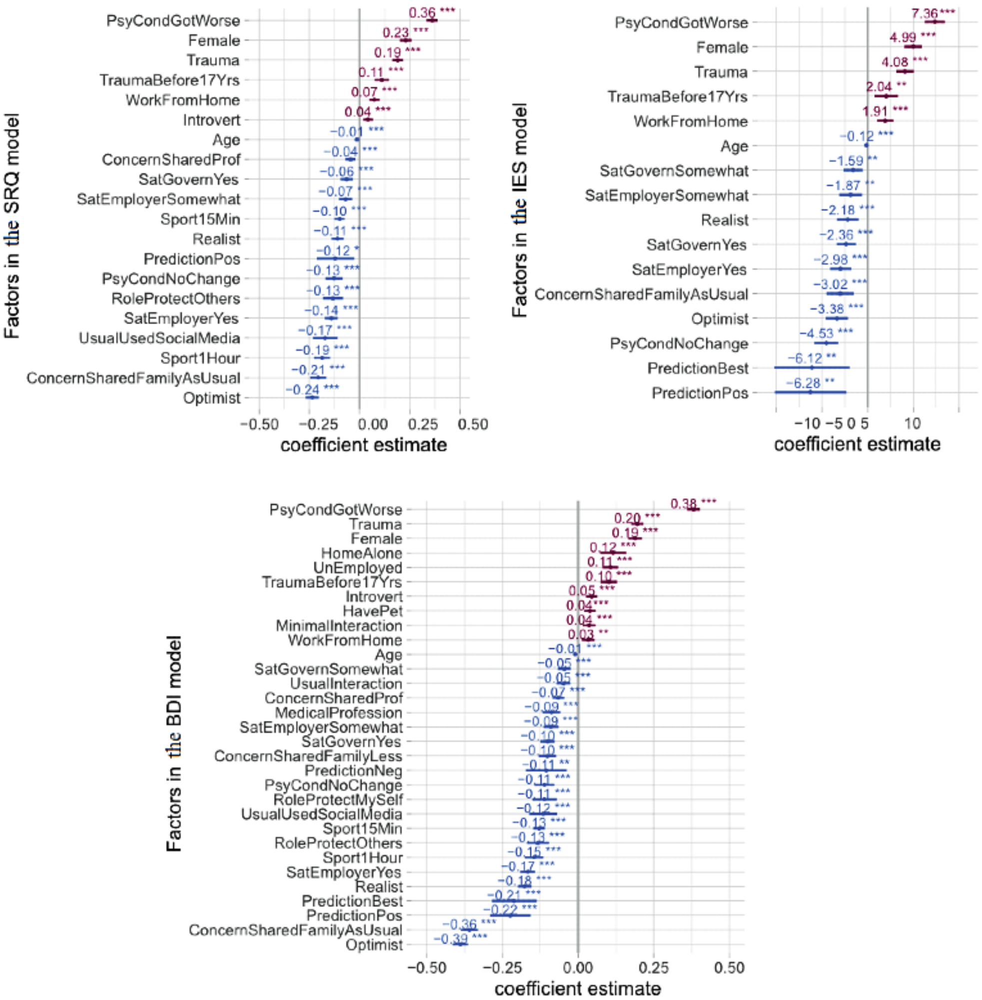
Adjusted Analysis of Factors Associated With General Psychological Disturbance (SRQ), PTSD Risk (IES), and Depression (BDI)

Adjusted analysis using different general linear models for each of the questionnaires is reported in Figure 2. Across all three questionnaires, we found the following factors increasing general psychological disturbance, PTSD, and depression: a psychiatric condition that worsened during the COVID-19 pandemic (SRQ mean coefficient: 0.36, 95% CI 0.33-0.39; IES mean coefficient: 7.36, 95% CI 6.26-8.46; BDI mean coefficient: 0.38, 95% CI 0.36-0.40), previous exposure to trauma (SRQ mean coefficient: 0.19, 95% CI 0.16-0.22; IES mean coefficient: 4.08, 95% CI 3.14-5.03; BDI mean coefficient: 0.20, 95% CI 0.17-0.22), and working remotely from home (SRQ mean coefficient: 0.07, 95% CI 0.05-0.10; IES mean coefficient: 1.91, 95% CI 1.01-2.82; BDI mean coefficient: 0.03, 95% CI 0.01-0.05).

Moreover, significant gender differences were observed, with higher risk in women versus men for general psychological disturbances (SRQ mean coefficient: 0.23, 95% CI 0.20-0.26), PTSD (IES mean coefficient: 4.99, 95% CI 4.03-5.95), and depression (BDI mean coefficient: 0.19, 95% CI 0.17-0.21).

Having an optimistic attitude, having a positive prediction about COVID-19, and being able to share concerns with family/friends decreased SRQ, IES, and BDI scores, indicating the protective effect of these factors against general psychological disturbance, PTSD, and depression (as shown in Figure 2 and Figure 3). Furthermore, daily physical activity/sports decreased both SRQ (mean coefficient: −0.19, 95% CI −0.23 to −0.15) and BDI (mean coefficient: −0.15, 95% CI −0.18 to −0.12) scores, with greater reductions resulting from the duration of physical activity/sports (exercise for ≥1 hour was more effective in decreasing SRQ and BDI scores compared to exercise for >15 minutes but <1 hour). In addition, health care professionals reported significantly lower BDI scores compared to nonhealth care professionals, suggesting this status to have a protective effect against depression (mean coefficient: −0.09, 95% CI −0.12 to −0.06).

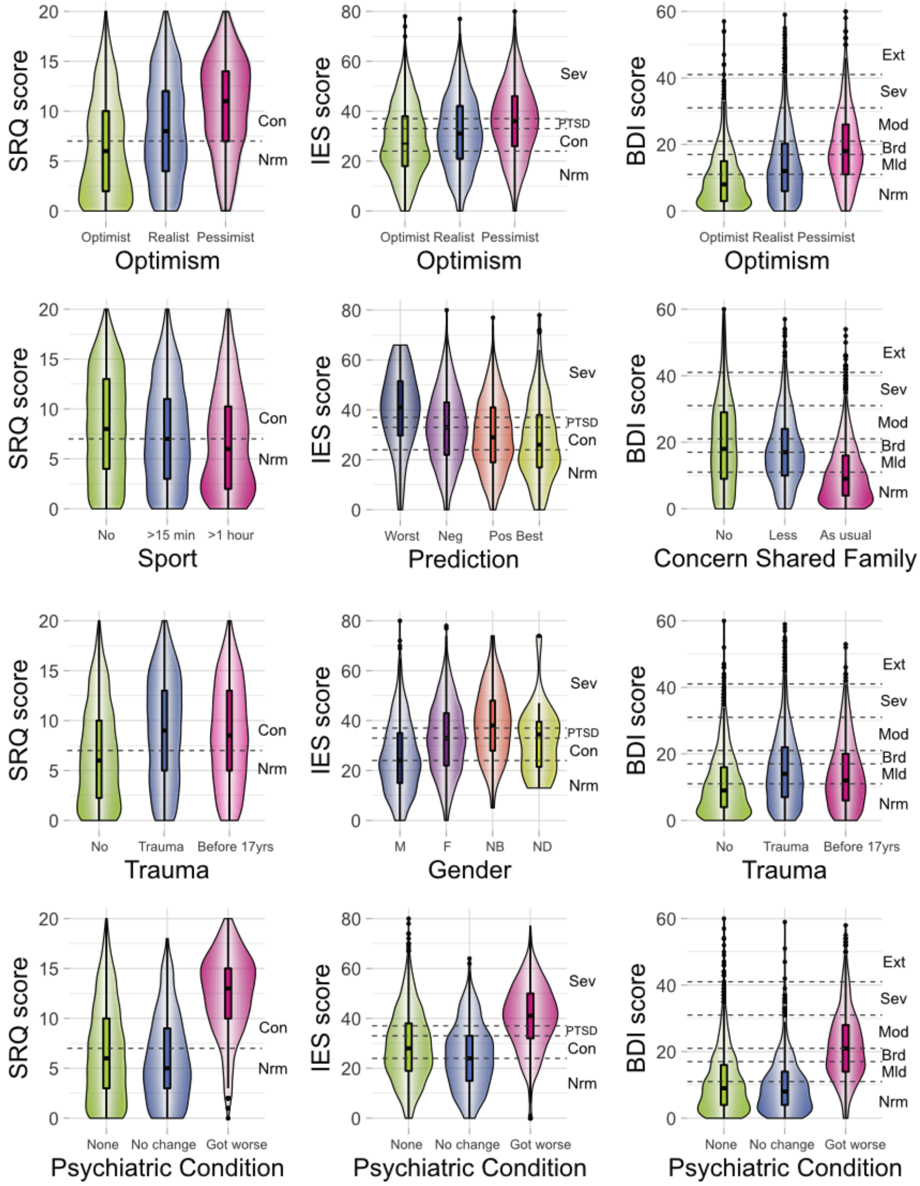
Figure 2. Risk and resilience factors for general psychological disturbance (Self-Reporting Questionnaire-20 [SRQ]), risk for posttraumatic stress disorder (PTSD) (Impact of Event Scale [IES]), and depression (Beck Depression Inventory II [BDI]). These foster plots show the mean estimates and the 95% CIs for adjusted coefficients significantly affecting SRQ, IES, and BDI scores generated through multiple regression models. Only factors that survived Bonferroni correction for multiple comparisons (*P*<.017) are listed. Factors associated with an increase in scores are shown in red, while those associated with a decrease in scores are in blue.



The logistic regression analyses performed after classifying SRQ, IES, and BDI scores into categorical cutoffs confirmed the primary results from the linear regression models (Multimedia Appendix 7). An individual with pre-existing psychiatric conditions that worsened during COVID-19 showed seven times higher odds of being depressed (OR 7.10, 95% CI 6.03-8.35), 1.6 times higher odds of having PTSD (OR 1.60, 95% CI 1.38-1.84), and two times higher odds of having general psychological disturbance (OR 2.64, 95% CI 1.99-3.48). As expected, individuals with previous trauma exposure exhibited greater ORs than their counterparts for these conditions according to the BDI (OR 1.61, 95% CI 1.46-1.76) and SRQ (OR 2.62, 95% CI 2.08-3.30). Still, an optimistic attitude and the opportunity to share concerns with family/friends like usual served as protective factors for general psychological disturbance according to the SRQ (OR 0.51, 95% CI 0.43-0.62 and OR 0.19, 95% CI 0.15-0.23, respectively) and depression according to the BDI (OR 0.23, 95% CI 0.20-0.26 and OR 0.39, 95% CI 0.33-0.45, respectively).

For visual aid, the association of participant-related factors with categorical classifications for general psychological disturbance (SRQ), PTSD (IES), and depression (BDI) are indicated through box plots in Multimedia Appendix 8. Owning a pet, having a pre-existing psychiatric condition, having previous exposure to trauma, considering oneself an introvert, and working remotely from home were associated with decreased percentages of responses in the unaffected ("normal") category based on the SRQ, IES, and BDI, suggesting these as risk factors. In contrast, a majority of responses from health care professionals landed in the unaffected ("normal") category for the BDI.

Figure 3. Violin plots indicating the effects of selected factors on general psychological disturbance (Self-Reporting Questionnaire-20 [SRQ]), risk for posttraumatic stress disorder (PTSD) (Impact of Event Scale [IES]), and depression (Beck Depression Inventory II [BDI]). These plots provide a relation between the participant scores on the SRQ, IES, and BDI and participant characteristics (previous history of a psychiatric condition, past exposure to trauma, prediction about COVID-19 resolution, level of optimism, gender, and daily physical activity/sports) adjusted for confounding variables through multiple regression models. Boxplots display the distribution of the selected factors with the visualization of five summary statistics (minimum, maximum, median, first quartile, and third quartile), and all outliers individually. Violin plots added behind the boxplots visualize the probability density of selected factors. Parallel to the x-axis, dashed lines present cutoffs for the scales used. For the BDI, Ext is "extreme," 40+ points, extreme depression; Sev is "severe," 31-40 points, severe depression; Mod is "moderate," 21-30 points, moderate depression; Brd is "borderline," 17-20 points, borderline clinical depression; Mld is "mild," 11-16 points, mild mood disturbance; and Nrm is "normal," 1-10 points, considered normal. For the SRQ, Con is "concern," 8-20 points, clinical concern for general psychological disturbance and Nrm is "normal," 0-7 points. For the IES, Sev is "severe," 37+ points, symptoms high enough to suppress the immune system; PTSD is "posttraumatic stress disorder," 34-36 points; Con is "clinical concern for possible PTSD," 24-33 points; and Nrm is "Normal," 0-23 points.



Correlation Among Scales

The continuous scores of all responses on the SRQ, BDI, and IES were also analyzed by Pearson correlations using all possible combinations on x-y plotting (SRQ vs IES, IES vs BDI, and BDI vs SRQ). All combinations yielded significant correlations, with the strongest correlation (R=0.79) between the BDI and SRQ (Multimedia Appendix 9).

Second Assessment

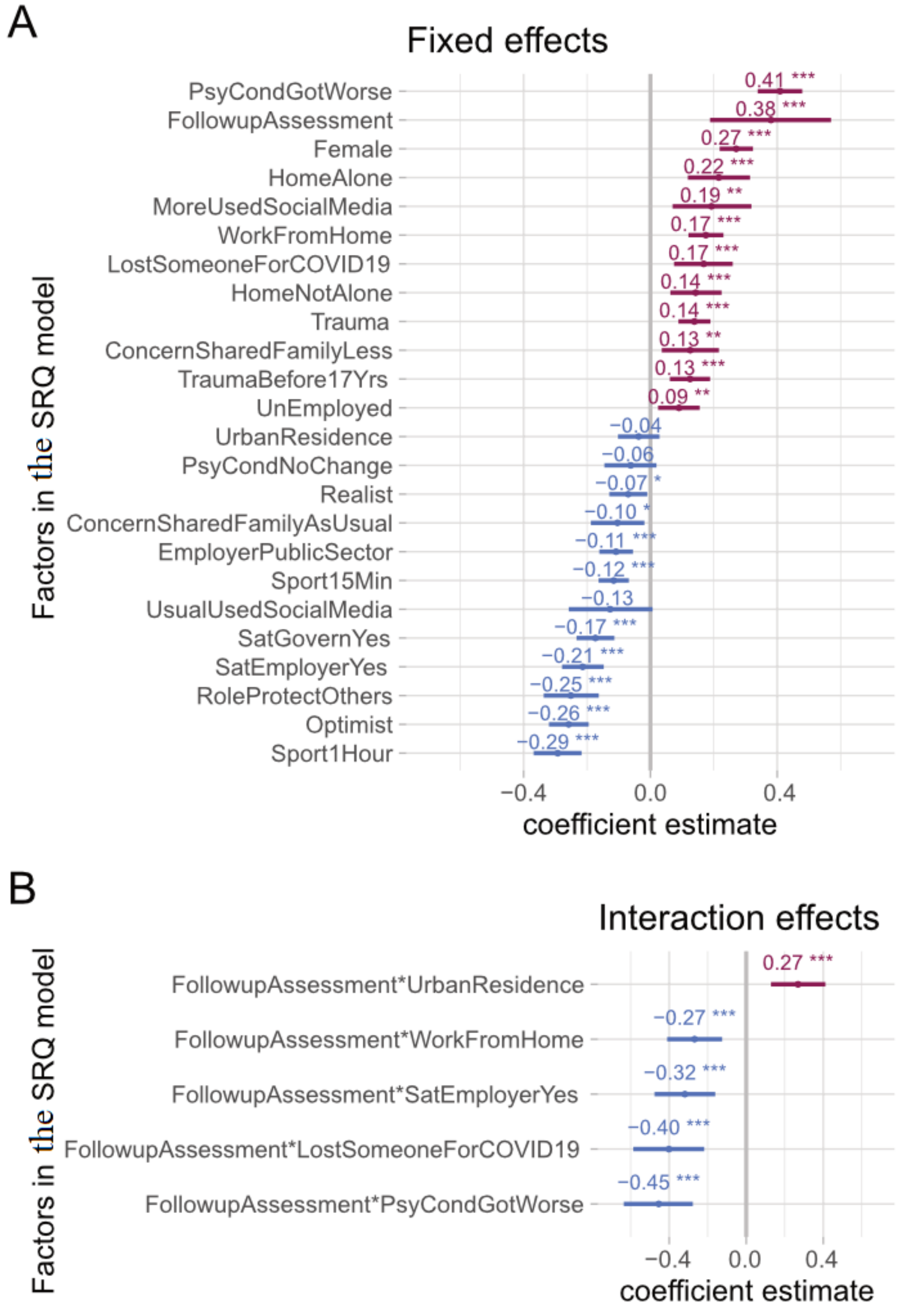
The demographic distribution of European participants included in the second assessment was similar to that in the first assessment, with higher number of responses from those participants who were female (n=803, 74.6%), working/studying remotely from home (n=613, 56.9%), and currently under home isolation with a partner/family (n=703, 65.3%). A majority of participants also reported increased social media usage (n=667, 61.9%), less than usual or minimal interaction with family and friends (n=703, 65.3%), and feeling a sense of control in protecting themselves and others during the COVID-19 pandemic (n=666, 61.9%).

Unadjusted analyses of SRQ scores between different participant demographics/characteristics showed a significantly higher prevalence of psychological symptoms (*P*<.05) in participants who were female, medical or health care professionals, dissatisfied with the response of their employer/state to COVID-19, interacting with friends/family less than usual, and using social media more than usual, as well as those with a less than usual ability to share concerns with friends/family. Significantly higher scores on the SRQ (*P*<.05) were also seen in participants with pre-existing psychiatric conditions and previous exposure to traumatic experiences, and those who self-reported being a pessimist or introvert. Means and standard deviations for all comparisons are presented in Multimedia Appendix 10.

Adjusted analysis utilizing a generalized linear model for the SRQ is reported in Figure 4. The following factors were independently associated with increased SRQ scores on the second assessment: a psychiatric condition that worsened during the COVID-19 pandemic (SRQ mean coefficient: 0.41, 95% CI 0.33-0.48), previous exposure to trauma before and after age 17 years (SRQ mean coefficient: 0.13, 95% CI 0.06-0.19 and SRQ mean coefficient: 0.14, 95% CI 0.08-0.19, respectively), and isolating at home alone (SRQ mean coefficient: 0.22, 95% CI 0.12-0.31). In addition, increased social media usage, working from home, and death of a family member due to COVID-19 significantly increased SRQ scores (SRQ mean coefficient: 0.19, 95% CI 0.07-0.32; SRQ mean coefficient: 0.17, 95% CI 0.12-0.23; and SRQ mean coefficient: 0.17, 95% CI 0.07-0.26). Moreover, significant gender differences were observed, with higher scores in women versus men (SRQ mean coefficient: 0.27, 95% CI 0.22-0.32). Having an optimistic attitude and feeling a sense of control in protecting oneself and others during the COVID-19 pandemic were associated with decreased SRQ scores in the second assessment, indicating a potentially protective effect of these factors against persistent general psychological disturbance (SRQ mean coefficient: −0.26, 95% CI −0.32 to −0.20 and SRQ mean coefficient: −0.25, 95% CI 0.12 to 0.23, respectively). Furthermore, participants who were satisfied with the employer/state response to COVID-19 and were able to share concerns with family/friends had lower SRQ scores overall (SRQ mean coefficient: −0.21, 95% CI −0.27 to −0.15; SRQ mean coefficient: 0.17, 95% CI −0.23 to −0.11; and SRQ mean coefficient: −0.10, 95% CI −0.19 to −0.02, respectively). Furthermore, daily physical activity/sports significantly decreased the SRQ score (mean coefficient: −0.29, 95% CI −0.37 to −0.22), with a greater protective effect associated with a higher duration of the physical activity/sport (exercise for ≥1 hour was more effective in decreasing the SRQ score compared to exercise for >15 minutes but <1 hour).

Finally, by including interaction terms in our regression model, we found that there was a relationship between residence type and SRQ score changes between the first and second assessments. Notably, SRQ scores increased in people living in urban areas compared to those living in rural areas (mean coefficient of interaction between acute/persistent and residence type: 0.27, 95% CI 0.13-0.41). Additionally, both people working from home and not working from home demonstrated a difference in responses between the two surveys (mean coefficient of interaction between acute/persistent and working from home: −0.27, 95% CI −0.41 to −0.12). Moreover, people who reported worsening of pre-existing psychiatric conditions during the first assessment reported lower SRQ scores in the second assessment, whereas those with no pre-existing psychiatric condition or a psychiatric condition that did not worsen showed an increase in SRQ scores in the second assessment (mean coefficient of interaction between acute/persistent and psychiatric condition: −0.45, 95% CI −0.64 to −0.28).

Figure 4. Factors associated with general psychological disturbance in the second assessment. These foster plots show the mean estimates and the 95% CIs for adjusted coefficients affecting the Self-Reporting Questionnaire-20 (SRQ) generated through multiple regression. Panel A shows fixed factors for SRQ scores during the second assessment. Panel B indicates interaction terms included in our regression model, indicating a significant difference between the fixed effects and SRQ scores during the first phase of the data collection and the second assessment. Factors increasing the SRQ score are shown in red, and factors decreasing the SRQ score are shown in blue.



Data and Material Availability

All data presented in the main text and supplementary items are deposited in a repository [36].

Discussion

This study, performed on a global scale, highlights the significant impact of the COVID-19 pandemic on the mental health of internet users during the first wave of the pandemic when the strictest lockdown restrictions were in place. It also provides evidence for the presence of these effects in a population subset of European participants 1 month later when restrictions were comparably less strict.

A major aim of this study was to identify specific factors that were positively or negatively associated with psychological perturbations in the immediate aftermath of the COVID-19 pandemic. Notably, the study was conducted when the strictest lockdown measures were in place, and the internet became the default mode of personal and professional communication. Worsening of a pre-existing psychiatric condition, female gender identification, previous exposure to trauma, and working remotely were associated with higher risks for general psychological disturbance, PTSD, and depression. Additionally, considering oneself an introvert was associated with the heightened risk of general psychological disturbance and depression, as was being unemployed, living alone, and having limited interaction with family and friends. An overall protective effect against all major psychological conditions was observed for the following factors: increasing age, considering oneself an optimist, optimism about the COVID-19 pandemic outcome, ability to share concerns with family and friends like usual, daily physical exercise/sports for 15 minutes or more, and being satisfied with the actions of the employer and state in the response to COVID-19.

To ensure that the psychological symptoms assessed in this study were related to the COVID-19 pandemic, the participants were repeatedly prompted to consider COVID-19 and their feelings during the preceding week while filling in the survey. Furthermore, the phrase "this crisis" was present in all the screening questions, for example, "I am unable to sleep well during this crisis." Further considerations about the attributability of psychological symptoms to the COVID-19 pandemic include the difference in the proportion (22%) of participants who reported pre-existing psychiatric conditions versus those who reported general psychological disturbance (43%) assessed through the SRQ during the first assessment. Furthermore, we compared the prevalence of depression in all of our featured countries based on different BDI cutoffs for depression versus the most recent available statistics from the WHO (2017) and found a remarkable difference.

To the best of our knowledge, this study is one of the few worldwide assessments of the mental health effects of COVID-19 performed during the first global wave of COVID-19. Earlier studies on the psychological impact of the COVID-19 outbreak were mostly from China [9,17,35,37-42]. However, a large number of studies on pandemic-related psychopathology have since been published, mostly focusing on populations from specific cities or countries [43-45]. Nevertheless, assessments performed on a global scale have been accumulating [46-51]. A study of almost 30,000 individuals across four South Asian countries showed that anxiety and depression were more common in those with chronic diseases and lower socioeconomic status [47]. Another study (n=4612) across eight countries showed that excessive and contradictory health information related to COVID-19 contributed to the psychological effects reported during the pandemic [46]. Another study (n=7091) across 13 countries found that, in alignment with our results, female individuals were more likely to report distress during the pandemic [48]. Similarly, a study of 9565 participants from 78 different countries showed that social support, finances, and psychological flexibility were the strongest predictors of being psychologically impacted by the pandemic [50]. Finally, our results are in agreement with findings from a recent study performed across Europe (n=15,790), which showed that lack of social contact has been a major stressor for individuals during the pandemic [51].

Identification of specific factors that are associated with an increased or decreased susceptibility to being psychologically affected by the pandemic could be crucial to mitigate the negative mental health impact of the COVID-19 pandemic at regional and global levels. For example, the vulnerability of females indicated in this study warrants further investigation for both the contributing factors and the resulting implications of such an increased risk. These include social factors, such as increased reporting of domestic violence in relation to COVID-19 [52], possible caregiver stress, and the impact of changes in familial roles and responsibilities secondary to the current health emergency. Furthermore, an increased risk of psychological symptoms in individuals with pre-existing psychiatric conditions and/or previous trauma exposure necessitates the initiation and/or expansion of mental health support systems available remotely [53]. Emerging evidence now supports the efficacy of web and social media–based interventions in promoting mental health via paradigms based on mindfulness, positive psychology, and exercise [54-56]. Such interventions could be developed at the governmental and institutional levels and delivered to the general public via mainstream and social media. Indeed, media outlets could play a major role in promoting optimism and a positive attitude toward COVID-19 resolution, both of which were identified in our study as important resilience factors. Furthermore, the association between remote working and increased psychological symptoms calls for optimization of the work-from-home model and a greater emphasis on the general well-being of employees. This is further corroborated by the observation in this study that participant satisfaction with the employer response to the COVID-19 pandemic is associated with reduced psychological symptoms.

Specifically, the results of this global online survey could be beneficial in the optimization of digital mental health services tailored to the needs of the target populations. Importantly, the study included lower- and middle-income countries such as Pakistan, Iran, and Bosnia and Herzegovina, where telemedicine/telepsychiatry services are likely less represented. The results of our study could help facilitate the organization and implementation of such services and their delivery to vulnerable populations. Furthermore, the administered measures in our study allowed for simultaneous screening of some psychiatric comorbidities that were found to be correlative to one another. These findings can provide invaluable insights for improving digital mental health services, whereby the presence of one psychopathology could prompt screening for the other. Regarding the optimization of digital mental health services, it is also important to note that the availability of the questionnaire in 11 different languages in this study provides insights to help extrapolate the results to individuals using web tools in different languages. Furthermore, the timing of this study is an important strength. The initial assessment was performed from March 29 to April 14, 2020. This timing coincides with the peak of the COVID-19 pandemic in North America and Europe, a time when almost one-half of the world remained in complete lockdown [57]. The second assessment, targeted at participants in Europe, was performed after a 1-month period when the situation had improved considerably in Europe and lockdown measures had been relatively eased. Finally, the identification of resilience factors identified in this study could have implications for digital mental health services. For example, the protective effect of exercise calls for efforts to promote exercise and physical activity through web-based outlets such as mobile health apps. Similarly, the protective effects of having a positive prediction and an optimistic attitude about the resolution of the COVID-19 pandemic should be taken into consideration when designing applications for information related to the pandemic or general promotion of positive psychology. The protective effect of maintaining contact with friends and family could also be incorporated into such resources. As an example, health apps that allow individuals to coordinate physical exercise with their friends and/or family members could be designed. It is noteworthy that a volunteer-based telehealth program for supporting the mental health of the elderly during the COVID-19 pandemic referred to many of our findings, including those related to pre-existing psychiatric conditions and impact on the elderly [58].

The study has potential limitations that warrant consideration when interpreting the results. First, the study employed a nonrandomized sampling strategy, and we advise caution in generalizing the findings of the study. The disproportionate demographic representation combined with the online nature of the study raises the potential for some level of participation bias. The association between self-reports of increased social media use and increased psychological symptoms must be interpreted with caution considering the study itself was conducted online. Furthermore, the unexpected result regarding medical or health care professionals reporting lower SRQ scores than the rest of the participants could have resulted from the disproportionate demographic representation; only 102 medical professionals participated in our survey.

The second considerable limitation is the use of self-report scales rather than clinical verification. However, the anonymous nature of the survey and widespread social distancing measures precluded such verification. Additionally, it is not possible to adjust for the confounding effect of non-COVID-19-related individual crisis situations on participant responses. We tried to reduce this effect by formatting survey questions in such a way that would prompt participants to consider their mental state specific to the COVID pandemic. Finally, any interpretation of the results from the second assessment warrants even more caution, as (1) the anonymous nature of the survey prevents verification of whether these are the same participants who filled the first survey; (2) it is unclear if the symptoms have persisted or are newly developed in the interval between the two assessments; and (3) only a fraction of individuals took part in the second assessment despite repeated reminders, leaving a viable possibility of participation bias.

In conclusion, this effort highlights the impact of the COVID-19 pandemic at both the regional and global levels on the mental health of internet users. Further, our study elucidates prominent associations with participant demographics, history of psychiatric disease risk factors, household conditions, personality traits, and attitudes toward COVID-19. These results could serve to inform health professionals and policymakers across the globe, aiding in dynamic optimization of digital mental health services during and following the COVID-19 pandemic, and potentially reducing its long-term morbidity and mortality.

Acknowledgments

We gratefully acknowledge the contributions of Luciana Armengol (Argentina); Professor Anthony Hannan, Maxine Mason, and Qi Hui Poh (Australia); Taria Brkić (Bosnia and Herzegovina); Barbara Levinsky (Brazil); Alexandra Schimmel and Léa Caya-Bissonnette (Canada); Claudia Valenzuela Rios (Chile); Marc Scherlinger, Alice Tondre, Lola Kouroma, and Morgane Roth (France); Katharina Schlerka, Lisa Garrelts, and Romy Seifert (Germany); Lena Heck (Germany/Switzerland); Varsha Hooda, Deepak Tanwar, and Chakradhar Yakkala (India); Professor Mohammad Es haghi and Sepehr Namirad (Iran); Darren Kelly (Ireland); Nour Mosawy (Jordan); Dayra Lorenzo (Mexico); Chirine Katrib (Lebanon/France); Usman Mukhtar, Uzair Jaswal, and Mubaris Bashir (Pakistan); Professor Kornelia Kedziora-Kornatowska, Milena Czarnocka, and Juli Davis (Poland); Ana Alexandra Moraru (Romania); Shoaib Jawaid (Saudi Arabia/United Arab Emirates); Myriam Merarchi (Singapore/France); Michelle McLuckie, Doman Obrist, Niharika Gaur, and Graciela Huber (Switzerland); Aurelia Muller (Taiwan/Germany/Switzerland); Burak Ozan (Turkey); Carmen Neagoe and Aleena Malik (United Kingdom); Anastasiia Timmer (Netherlands/United States); and Colette Rausch, Professor Paul Schulz, Professor Mo Salman, Saleha Tahir, Laura Luebbert, Sarish Khan, Rebecca Sager, Lupita Lozano, and American Physician Scientist Association (United States) for their dedicated help in data collection. We are also thankful to Lena Heck and Giuseppe Parente (University of Zurich) for technical support. Finally, we would like to express our gratitude to Professor Selmira Brkić (Faculty of Medicine, University of Tuzla, Bosnia and Herzegovina), Professor Leszek Kaczmarek (Director, Nencki-EMBL BRAINCITY, Warsaw, Poland), University of Zurich Research Office, Zurich Cantonal Ethics Commission, and Texas Behavioral Health for their expedited review of the study procedures under extraordinary circumstances and for their organizational support.

This work has been supported by the Nencki-EMBL Center of Excellence for Neural Plasticity and Brain Disorders: BRAINCITY project carried out within the International Research Agendas program of the Foundation for Polish Science co-financed by the European Union under the European Regional Development Fund.

Authors' Contributions

MP and SG contributed in conceptualization, questionnaire development, data collection, data mining, data analysis, visualization, review, and editing. RN contributed in data collection, manuscript writing, review, and editing. BS, SL, KA, AD, AB, LH, SE, HJ, LRP, VW, BA, MB, and DS contributed in questionnaire translation, data collection, data mining, review, editing, and project co-ordination. PR contributed in data analysis and visualization. ZA contributed in data collection, manuscript writing, review, and editing. ZB contributed in data analysis. ZH and SUQ contributed in data collection and project co-ordination. AMS contributed in data collection, project administration, and editing. AJ contributed in conceptualization, questionnaire development, study approval, data collection, data analysis, data visualization, manuscript writing, review, editing, project administration, and supervision. All authors have reviewed and approved the final draft.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Study questionnaire.
[PDF File (Adobe PDF File), 484 KB - mental_v8i8e28736_app1.pdf]

Multimedia Appendix 2
Switzerland ethical approval.
[PDF File (Adobe PDF File), 35 KB - mental_v8i8e28736_app2.pdf]

Multimedia Appendix 3
Poland ethical approval.
[PDF File (Adobe PDF File), 207 KB - mental_v8i8e28736_app3.pdf]

Multimedia Appendix 4
Bosnia and Herzegovina ethical approval.
[PDF File (Adobe PDF File), 122 KB - mental_v8i8e28736_app4.pdf]

Multimedia Appendix 5
Number of participants per country and World Health Organization region included in the primary assessment.
[DOCX File , 17 KB - mental_v8i8e28736_app5.docx]

Multimedia Appendix 6
Demographics and characteristics of the participants included in the primary assessment.
[DOCX File , 24 KB - mental_v8i8e28736_app6.docx]

Multimedia Appendix 7
Predictors for acute general psychological disturbance, posttraumatic stress disorder risk, and depression.
[DOCX File , 20 KB - mental_v8i8e28736_app7.docx]

Multimedia Appendix 8
Association of participant demographics/characteristics and categorical classifications for general psychological disturbance (Self-Reporting Questionnaire-20), posttraumatic stress disorder risk (Impact of Event Scale), and depression (Beck Depression Inventory II) in the primary assessment.
[DOCX File , 513 KB - mental_v8i8e28736_app8.docx]

Multimedia Appendix 9
Correlations between the Self-Reporting Questionnaire-20, Impact of Event Scale, and Beck Depression Inventory II scores for the primary assessment.
[DOCX File , 123 KB - mental_v8i8e28736_app9.docx]

Multimedia Appendix 10
Comparison of psychological symptoms between different participant demographics/characteristics included in the follow-up assessment.
[DOCX File , 24 KB - mental_v8i8e28736_app10.docx]

References

1. Coronavirus disease 2019 (COVID-19) Situation Report – 90. World Health Organization. URL: https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200419-sitrep-90-covid-19.pdf [accessed 2021-06-29]

2. Pfefferbaum B, North CS. Mental Health and the Covid-19 Pandemic. N Engl J Med 2020 Aug 06;383(6):510-512. [doi: 10.1056/NEJMp2008017] [Medline: 32283003]
3. Brooks SK, Webster RK, Smith LE, Woodland L, Wessely S, Greenberg N, et al. The psychological impact of quarantine and how to reduce it: rapid review of the evidence. The Lancet 2020 Mar 14;395(10227):912-920. [doi: 10.1016/S0140-6736(20)30460-8] [Medline: 32112714]
4. Mucci N, Giorgi G, Roncaioli M, Fiz Perez J, Arcangeli G. The correlation between stress and economic crisis: a systematic review. NDT 2016 Apr;12:983. [doi: 10.2147/ndt.s98525]
5. Bünnings C, Kleibrink J, Weßling J. Fear of Unemployment and its Effect on the Mental Health of Spouses. Health Econ 2017 Jan 06;26(1):104-117. [doi: 10.1002/hec.3279] [Medline: 26542072]
6. Nordt C, Warnke I, Seifritz E, Kawohl W. Modelling suicide and unemployment: a longitudinal analysis covering 63 countries, 2000–11. The Lancet Psychiatry 2015 Mar;2(3):239-245. [doi: 10.1016/s2215-0366(14)00118-7]
7. Dorn S. Domestic violence victims facing higher risks amid coronavirus quarantine. New York Post. 2020. URL: https://nypost.com/2020/03/28/domestic-violence-victims-facing-higher-risks-amid-coronavirus-quarantine/ [accessed 2020-03-28]
8. Collie M. Quarantine that stress: Limit screen time during coronavirus outbreak, experts say. Global News. 2020. URL: https://globalnews.ca/news/6700880/coronavirus-social-media-mental-health/ [accessed 2020-03-23]
9. Zhang W, Wang K, Yin L, Zhao W, Xue Q, Peng M, et al. Mental Health and Psychosocial Problems of Medical Health Workers during the COVID-19 Epidemic in China. Psychother Psychosom 2020 Apr 9;89(4):242-250 [FREE Full text] [doi: 10.1159/000507639] [Medline: 32272480]
10. Dhir A, Yossatorn Y, Kaur P, Chen S. Online social media fatigue and psychological wellbeing—A study of compulsive use, fear of missing out, fatigue, anxiety and depression. International Journal of Information Management 2018 Jun;40:141-152 [FREE Full text] [doi: 10.1016/j.ijinfomgt.2018.01.012]
11. Person B, Sy F, Holton K, Govert B, Liang A, National Center for Inectious Diseases/SARS Community Outreach Team. Fear and stigma: the epidemic within the SARS outbreak. Emerg Infect Dis 2004 Feb;10(2):358-363 [FREE Full text] [doi: 10.3201/eid1002.030750] [Medline: 15030713]
12. Elizarrarás-Rivas J, Vargas-Mendoza JE, Mayoral-García M, Matadamas-Zarate C, Elizarrarás-Cruz A, Taylor M, et al. Psychological response of family members of patients hospitalised for influenza A/H1N1 in Oaxaca, Mexico. BMC Psychiatry 2010 Dec 3;10(1):104. [doi: 10.1186/1471-244x-10-104]
13. Morris KL, Goldenberg JL, Arndt J, McCabe S. The enduring influence of death on health: insights from the terror management health model. Self and Identity 2018 Apr 30;18(4):378-404. [doi: 10.1080/15298868.2018.1458644]
14. Zhou P, Yang X, Wang X, Hu B, Zhang L, Zhang W, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. Nature 2020 Feb 03;579(7798):270-273. [doi: 10.1038/s41586-020-2012-7]
15. Morgenstern-Kaplan D, Rocha-Haro A, Canales-Albarrán SJ, Núñez-García E, León-Mayorga Y. An App-Based Telemedicine Program for Primary Care and Specialist Video Consultations During the COVID-19 Pandemic in Mexico. Telemed J E Health 2021 Mar 17. [doi: 10.1089/tmj.2021.0055] [Medline: 33733868]
16. Torous J, Keshavan M. COVID-19, mobile health and serious mental illness. Schizophr Res 2020 Apr;218:36-37 [FREE Full text] [doi: 10.1016/j.schres.2020.04.013] [Medline: 32327314]
17. Sartorius N, Janca A. Psychiatric assessment instruments developed by the World Health Organization. Soc Psychiatry Psychiatr Epidemiol 1996;31(2):55-69. [doi: 10.1007/bf00801901]
18. Weiss DS, Marmar CR. The impact of event scale – revised. In: Wilson JP, Keane TM, editors. Assessing psychological trauma and PTSD. New York, NY: The Guilford Press; 1997:399-411.
19. Upton J. Beck Depression Inventory (BDI). In: Gellman MD, editor. Encyclopedia of Behavioral Medicine. Cham: Springer; 2020.
20. Chen C, Yang P, Yen C, Wu H. Validation of Impact of Events Scale in nurses under threat of contagion by severe acute respiratory syndrome. Psychiatry Clin Neurosci 2005 Apr;59(2):135-139 [FREE Full text] [doi: 10.1111/j.1440-1819.2005.01347.x] [Medline: 15823157]
21. Cheng SKW, Wong CW, Tsang J, Wong KC. Psychological distress and negative appraisals in survivors of severe acute respiratory syndrome (SARS). Psychol Med 2004 Oct;34(7):1187-1195. [doi: 10.1017/s0033291704002272] [Medline: 15697045]
22. Koh D, Lim MK, Chia SE, Ko SM, Qian F, Ng V, et al. Risk perception and impact of Severe Acute Respiratory Syndrome (SARS) on work and personal lives of healthcare workers in Singapore: what can we learn? Med Care 2005 Jul;43(7):676-682. [doi: 10.1097/01.mlr.0000167181.36730.cc] [Medline: 15970782]
23. Hong X, Currier G, Zhao X, Jiang Y, Zhou W, Wei J. Posttraumatic stress disorder in convalescent severe acute respiratory syndrome patients: a 4-year follow-up study. Gen Hosp Psychiatry 2009;31(6):546-554 [FREE Full text] [doi: 10.1016/j.genhosppsych.2009.06.008] [Medline: 19892213]
24. Mak IWC, Chu CM, Pan PC, Yiu MGC, Chan VL. Long-term psychiatric morbidities among SARS survivors. Gen Hosp Psychiatry 2009 Jul;31(4):318-326 [FREE Full text] [doi: 10.1016/j.genhosppsych.2009.03.001] [Medline: 19555791]
25. Nicastri E, Balestra P, Ricottini M, Petrosillo N, Di Caro A, Capobianchi MR, INMI's Ebola Team. Temporary neurocognitive impairment with Ebola virus. J Neurol Neurosurg Psychiatry 2016 Dec;87(12):1386-1387. [doi: 10.1136/jnnp-2016-313695] [Medline: 27435087]

26. Styra R, Hawryluck L, Robinson S, Kasapinovic S, Fones C, Gold WL. Impact on health care workers employed in high-risk areas during the Toronto SARS outbreak. J Psychosom Res 2008 Feb;64(2):177-183 [FREE Full text] [doi: 10.1016/j.jpsychores.2007.07.015] [Medline: 18222131]
27. Wang Y, Xu B, Zhao G, Cao R, He X, Fu S. Is quarantine related to immediate negative psychological consequences during the 2009 H1N1 epidemic? Gen Hosp Psychiatry 2011;33(1):75-77 [FREE Full text] [doi: 10.1016/j.genhosppsych.2010.11.001] [Medline: 21353131]
28. Wu KK, Chan SK, Ma TM. Posttraumatic stress after SARS. Emerg Infect Dis 2005 Aug;11(8):1297-1300 [FREE Full text] [doi: 10.3201/eid1108.041083] [Medline: 16102324]
29. Yip PS, Cheung Y, Chau P, Law Y. The impact of epidemic outbreak: the case of severe acute respiratory syndrome (SARS) and suicide among older adults in Hong Kong. Crisis 2010 Mar;31(2):86-92. [doi: 10.1027/0227-5910/a000015] [Medline: 20418214]
30. RStudio. URL: http://www.rstudio.com/ [accessed 2021-08-10]
31. Wickham H. Data Analysis. In: ggplot2. Use R!. Cham: Springer; 2016.
32. Powell C. Powell Miscellaneous Functions for Teaching and Learning Statistics. Search R-Project. URL: https://search.r-project.org/CRAN/refmans/CGPfunctions/html/00Index.html [accessed 2021-08-09]
33. Bates D, Mächler M, Bolker B, Walker S. Fitting Linear Mixed-Effects Models Using. J. Stat. Soft 2015;67(1). [doi: 10.18637/jss.v067.i01]
34. Venables W, Ripley BD. Modern Applied Statistics with S. New York, NY: Springer; 2002.
35. Cao W, Fang Z, Hou G, Han M, Xu X, Dong J, et al. The psychological impact of the COVID-19 epidemic on college students in China. Psychiatry Res 2020 May;287:112934 [FREE Full text] [doi: 10.1016/j.psychres.2020.112934] [Medline: 32229390]
36. Mental Health Impact of COVID-19: A global study of risk and resilience factors. OSF. URL: https://osf.io/3vupe/?view_only=80f71b6f0c8d49b08573ea12eab10d33 [accessed 2021-08-07]
37. Lai J, Ma S, Wang Y, Cai Z, Hu J, Wei N, et al. Factors Associated With Mental Health Outcomes Among Health Care Workers Exposed to Coronavirus Disease 2019. JAMA Netw Open 2020 Mar 02;3(3):e203976 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.3976] [Medline: 32202646]
38. Dong L, Bouey J. Public Mental Health Crisis during COVID-19 Pandemic, China. Emerg Infect Dis 2020 Jul;26(7):1616-1618 [FREE Full text] [doi: 10.3201/eid2607.200407] [Medline: 32202993]
39. Zhang Y, Ma ZF. Impact of the COVID-19 Pandemic on Mental Health and Quality of Life among Local Residents in Liaoning Province, China: A Cross-Sectional Study. Int J Environ Res Public Health 2020 Mar 31;17(7):2381 [FREE Full text] [doi: 10.3390/ijerph17072381] [Medline: 32244498]
40. Li W, Yang Y, Liu Z, Zhao Y, Zhang Q, Zhang L, et al. Progression of Mental Health Services during the COVID-19 Outbreak in China. Int J Biol Sci 2020;16(10):1732-1738 [FREE Full text] [doi: 10.7150/ijbs.45120] [Medline: 32226291]
41. Wang C, Pan R, Wan X, Tan Y, Xu L, Ho CS, et al. Immediate Psychological Responses and Associated Factors during the Initial Stage of the 2019 Coronavirus Disease (COVID-19) Epidemic among the General Population in China. Int J Environ Res Public Health 2020 Mar 06;17(5) [FREE Full text] [doi: 10.3390/ijerph17051729] [Medline: 32155789]
42. Modi PD, Nair G, Uppe A, Modi J, Tuppekar B, Gharpure AS, et al. COVID-19 Awareness Among Healthcare Students and Professionals in Mumbai Metropolitan Region: A Questionnaire-Based Survey. Cureus 2020 Apr 02;12(4):e7514 [FREE Full text] [doi: 10.7759/cureus.7514] [Medline: 32377462]
43. Phiri P, Ramakrishnan R, Rathod S, Elliot K, Thayanandan T, Sandle N, et al. An evaluation of the mental health impact of SARS-CoV-2 on patients, general public and healthcare professionals: A systematic review and meta-analysis. EClinicalMedicine 2021 Apr;34:100806 [FREE Full text] [doi: 10.1016/j.eclinm.2021.100806] [Medline: 33842872]
44. Kola L, Kohrt BA, Hanlon C, Naslund JA, Sikander S, Balaji M, et al. COVID-19 mental health impact and responses in low-income and middle-income countries: reimagining global mental health. The Lancet Psychiatry 2021 Jun;8(6):535-550. [doi: 10.1016/s2215-0366(21)00025-0]
45. Hall PA, Sheeran P, Fong GT, Cheah CSL, Oremus M, Liu-Ambrose T, et al. Biobehavioral Aspects of the COVID-19 Pandemic: A Review. Psychosom Med 2021;83(4):309-321. [doi: 10.1097/psy.0000000000000932]
46. Wang C, Chudzicka-Czupała A, Tee ML, Núñez MIL, Tripp C, Fardin MA, et al. A chain mediation model on COVID-19 symptoms and mental health outcomes in Americans, Asians and Europeans. Sci Rep 2021 Mar 19;11(1):6481 [FREE Full text] [doi: 10.1038/s41598-021-85943-7] [Medline: 33742072]
47. Kusuma D, Pradeepa R, Khawaja KI, Hasan M, Siddiqui S, Mahmood S, NIHR Global Health Research Unit for DiabetesCardiovascular Disease in South Asia. Low uptake of COVID-19 prevention behaviours and high socioeconomic impact of lockdown measures in South Asia: Evidence from a large-scale multi-country surveillance programme. SSM Popul Health 2021 Mar;13:100751 [FREE Full text] [doi: 10.1016/j.ssmph.2021.100751] [Medline: 33665333]
48. Marzo RR, Ismail Z, Nu Htay MN, Bahari R, Ismail R, Villanueva EQ, et al. Psychological distress during pandemic Covid-19 among adult general population: Result across 13 countries. Clin Epidemiol Glob Health 2021;10:100708 [FREE Full text] [doi: 10.1016/j.cegh.2021.100708] [Medline: 33619459]

49. Nurunnabi M, Almusharraf N, Aldeghaither D. Mental health and well-being during the COVID-19 pandemic in higher education: Evidence from G20 countries. J Public Health Res 2020 Nov 17;9(Suppl 1):2010 [FREE Full text] [doi: 10.4081/jphr.2020.2010] [Medline: 33575229]
50. Gloster AT, Lamnisos D, Lubenko J, Presti G, Squatrito V, Constantinou M, et al. Impact of COVID-19 pandemic on mental health: An international study. PLoS One 2020;15(12):e0244809 [FREE Full text] [doi: 10.1371/journal.pone.0244809] [Medline: 33382859]
51. Veer IM, Riepenhausen A, Zerban M, Wackerhagen C, Puhlmann LMC, Engen H, et al. Psycho-social factors associated with mental resilience in the Corona lockdown. Transl Psychiatry 2021 Jan 21;11(1):67 [FREE Full text] [doi: 10.1038/s41398-020-01150-4] [Medline: 33479211]
52. John N, Casey S, Carino G, McGovern T. Lessons Never Learned: Crisis and gender-based violence. Dev World Bioeth 2020 Jun;20(2):65-68 [FREE Full text] [doi: 10.1111/dewb.12261] [Medline: 32267607]
53. Qiu J, Shen B, Zhao M, Wang Z, Xie B, Xu Y. A nationwide survey of psychological distress among Chinese people in the COVID-19 epidemic: implications and policy recommendations. Gen Psychiatr 2020;33(2):e100213 [FREE Full text] [doi: 10.1136/gpsych-2020-100213] [Medline: 32215365]
54. Mak WW, Chio FH, Chan AT, Lui WW, Wu EK. The Efficacy of Internet-Based Mindfulness Training and Cognitive-Behavioral Training With Telephone Support in the Enhancement of Mental Health Among College Students and Young Working Adults: Randomized Controlled Trial. J Med Internet Res 2017 Mar 22;19(3):e84 [FREE Full text] [doi: 10.2196/jmir.6737] [Medline: 28330831]
55. Ploeg J, Markle-Reid M, Valaitis R, McAiney C, Duggleby W, Bartholomew A, et al. Web-Based Interventions to Improve Mental Health, General Caregiving Outcomes, and General Health for Informal Caregivers of Adults With Chronic Conditions Living in the Community: Rapid Evidence Review. J Med Internet Res 2017 Jul 28;19(7):e263 [FREE Full text] [doi: 10.2196/jmir.7564] [Medline: 28754652]
56. Chang S, Lin Y, Lin C, Chang H, Chong P. Promoting positive psychology using social networking sites: a study of new college entrants on Facebook. Int J Environ Res Public Health 2014 Apr 29;11(5):4652-4663 [FREE Full text] [doi: 10.3390/ijerph110504652] [Medline: 24785540]
57. A third of the global population is on coronavirus lockdown — here's our constantly updated list of countries and restrictions. Business Insider. URL: https://tinyurl.com/5f4k2a79 [accessed 2021-08-10]
58. Dikaios E, Sekhon H, Allard A, Vacaflor B, Goodman A, Dwyer E, et al. Connecting During COVID-19: A Protocol of a Volunteer-Based Telehealth Program for Supporting Older Adults' Health. Front Psychiatry 2020 Dec 2;11:598356 [FREE Full text] [doi: 10.3389/fpsyt.2020.598356] [Medline: 33343425]

Abbreviations

BDI: Beck Depression Inventory II
CET: Central European time
IES: Impact of Event Scale
OR: odds ratio
PTSD: posttraumatic stress disorder
SRQ: Self-Reporting Questionnaire-20
WHO: World Health Organization

Edited by J Torous; submitted 12.03.21; peer-reviewed by S Gordon, M Marciniak; comments to author 27.06.21; revised version received 06.07.21; accepted 06.07.21; published 19.08.21.

Please cite as:
Plomecka M, Gobbi S, Neckels R, Radzinski P, Skorko B, Lazzeri S, Almazidou K, Dedic A, Bakalovic A, Hrustic L, Ashraf Z, Es Haghi S, Rodriguez-Pino L, Waller V, Jabeen H, Alp AB, Behnam M, Shibli D, Baranczuk-Turska Z, Haq Z, Qureshi S, Strutt AM, Jawaid A
Factors Associated With Psychological Disturbances During the COVID-19 Pandemic: Multicountry Online Study
JMIR Ment Health 2021;8(8):e28736
URL: https://mental.jmir.org/2021/8/e28736
doi:10.2196/28736
PMID:34254939

Original Paper

Psychosomatic Rehabilitation Patients and the General Population During COVID-19: Online Cross-sectional and Longitudinal Study of Digital Trainings and Rehabilitation Effects

Franziska Maria Keller[1], MSc; Alina Dahmen[2], MD; Christina Derksen[1], MSc; Lukas Kötting[1], MSc; Sonia Lippke[1], PhD

[1]Department of Psychology & Methods, Jacobs University Bremen gGmbH, Bremen, Germany
[2]Dr. Becker Klinikgruppe, Cologne, Germany

Corresponding Author:
Sonia Lippke, PhD
Department of Psychology & Methods, Jacobs University Bremen gGmbH
Campus Ring 1
Bremen, 28759
Germany
Phone: 49 421200 ext 4730
Email: s.lippke@jacobs-university.de

Abstract

Background: The COVID-19 pandemic has largely affected people's mental health and psychological well-being. Specifically, individuals with a pre-existing mental health disorder seem more impaired by lockdown measures posing as major stress factors. Medical rehabilitation treatment can help people cope with these stressors. The internet and digital apps provide a platform to contribute to regular treatment and to conduct research on this topic.

Objective: Making use of internet-based assessments, this study investigated individuals from the general population and patients from medical, psychosomatic rehabilitation clinics. Levels of depression, anxiety, loneliness, and perceived stress during the COVID-19 pandemic, common COVID-19–related worries, and the intention to use digital apps were compared. Furthermore, we investigated whether participating in internet-delivered digital trainings prior to and during patients' rehabilitation stay, as well as the perceived usefulness of digital trainings, were associated with improved mental health after rehabilitation.

Methods: A large-scale, online, cross-sectional study was conducted among a study sample taken from the general population (N=1812) in Germany from May 2020 to April 2021. Further, a longitudinal study was conducted making use of the internet among a second study sample of psychosomatic rehabilitation patients at two measurement time points—before (N=1719) and after (n=738) rehabilitation—between July 2020 and April 2021. Validated questionnaires and adapted items were used to assess mental health and COVID-19–related worries. Digital trainings were evaluated. Propensity score matching, multivariate analyses of covariance, an exploratory factor analysis, and hierarchical regression analyses were performed.

Results: Patients from the psychosomatic rehabilitation clinics reported increased symptoms with regard to depression, anxiety, loneliness, and stress ($F_{4,2028}$=183.74, P<.001, η^2_p=0.27) compared to the general population. Patients perceived greater satisfaction in communication with health care professionals ($F_{1,837}$=31.67, P<.001, η^2_p=0.04), had lower financial worries ($F_{1,837}$=38.96, P<.001, η^2_p=0.04), but had higher household-related worries ($F_{1,837}$=5.34, P=.02, η^2_p=0.01) compared to the general population. Symptoms of depression, anxiety, loneliness, and perceived stress were lower postrehabilitation ($F_{1,712}$=23.21, P<.001, η^2_p=0.04) than prior to rehabilitation. Psychosomatic patients reported a higher intention to use common apps and digital trainings ($F_{3,2021}$=51.41, P<.001, η^2_p=0.07) than the general population. With regard to digital trainings offered prior to and during the rehabilitation stay, the perceived usefulness of digital trainings on rehabilitation goals was associated with decreased symptoms of depression (β=–.14, P<.001), anxiety (β=–.12, P<.001), loneliness (β=–.18, P<.001), and stress postrehabilitation (β=–.19, P<.001). Participation in digital group therapy for depression was associated with an overall change in depression ($F_{1,725}$=4.82, P=.03, η^2_p=0.01) and anxiety ($F_{1,725}$=6.22, P=.01, η^2_p=0.01) from pre- to postrehabilitation.

Conclusions: This study validated the increased mental health constraints of psychosomatic rehabilitation patients in comparison to the general population and the effects of rehabilitation treatment. Digital rehabilitation components are promising tools that could prepare patients for their rehabilitation stay, could integrate well with face-to-face therapy during rehabilitation treatment, and could support aftercare.

Trial Registration: ClinicalTrials.gov NCT04453475; https://clinicaltrials.gov/ct2/show/NCT04453475 and ClinicalTrials.gov NCT03855735; https://clinicaltrials.gov/ct2/show/NCT03855735

(JMIR Ment Health 2021;8(8):e30610) doi:10.2196/30610

KEYWORDS

mental health; COVID-19; medical rehabilitation; psychosomatic rehabilitation; internet-delivered digital trainings

Introduction

Mental Health and the COVID-19 Pandemic

The COVID-19 pandemic has led to rapid changes in the lives of people all over the world, thus affecting both physical health as well as mental health and well-being [1]. Worries about one's own health, the health of family and friends, as well as worries associated with the future are indicative of decreased mental health and psychological well-being. Hence, for many individuals, the COVID-19 pandemic evoked feelings of uncertainty, social isolation due to contact regulations, stress reactions, symptoms of depression and anxiety, and general fear of the virus [2,3]. In case of prolonged concerns or worries, individuals are at risk of developing serious mental health disorders [4].

A study by Wang et al asked respondents to assess the psychological impact of the COVID-19 pandemic on their mental health. Results highlighted that 54% of the respondents rated the psychological impact of the COVID-19 pandemic as moderate to severe. Further, 29% estimated their own anxiety symptoms to be between moderate and severe, and 17% estimated symptoms of depression as moderate to severe [5]. Another study by Sønderskov et al revealed lower psychological well-being in the general public compared to before the COVID-19 pandemic [6]. Recent studies from the United States highlighted the worldwide increase in depressive symptoms as well as in symptoms of anxiety, which occurred about three times more frequently during the COVID-19 pandemic than before. Research has indicated that pre-existing mental health conditions may worsen due to COVID-19 [7,8].

The conjectured decrease in mental health worldwide may be explained by two developments associated with the ongoing course of the COVID-19 pandemic. On the one hand, the ramifications associated with the COVID-19 pandemic, such as uncertainties, unemployment, short-term employment, or social isolation, may pose a mental health threat. A cross-national comparison of Norway, the United Kingdom, the United States, and Australia found that secure employment status was associated with lower levels of loneliness and mental health distress as well as higher levels of well-being and quality of life during the early social distancing requirements of the pandemic [9]. Correspondingly, returning to work during the pandemic was associated with low levels of psychiatric problems [10].

On the other hand, the way most people live, work, study, socialize, or travel has been abruptly disrupted or shifted online. The associated containment measures, such as quarantining and physical distancing, restrict people in their freedom, but are necessary to control the disease's spread. Literature has shown that quarantining or physically distancing oneself from others may lead to problems associated with a decreased mental health status [11]. It can precipitate feelings related to fear, anger, anxiety, or even panic about possible negative outcomes and is associated with increased perceived loneliness and boredom.

The World Health Organization (WHO) has also expressed concerns with regard to the mental health and psychological well-being of individuals due to containment measures. According to the WHO, restrictions may interfere with people's daily activities and routines and may consequently lead to an increased perception of loneliness, depression and anxiety, insomnia, substance misuse, self-harm, or even suicidal behavior [12]. It has been shown that increased loneliness and reduced interactions due to social distancing are risk factors for several mental health disorders, such as depression, anxiety, and schizophrenia. Especially for women, young people, and those living with young children, mental health problems have increased over time [13].

COVID-19 and Patients With Pre-existing Mental Health Disorders

Literature on the impact of the 2003 severe acute respiratory syndrome (SARS) outbreak and the COVID-19 pandemic underlined more negative feelings associated with worry [14,15]. Worry, which can be defined as an attempt to engage in mental problem solving or to deal with outcome uncertainty under some circumstances [16], is a central feature of anxiety disorders [17] and is associated with depressive rumination [18]. Several studies have identified worries associated with the COVID-19 pandemic, such as health-, future-, or employment-related worries and their associated consequences, such as sleep hygiene, drinking behavior, changes in social interactions, or changes in physical exercise [19,20].

For patients with a pre-existing mental health disorder or a decreased perception of well-being, lockdown measures are major stress factors affecting daily routine and social rhythms. A study by van Rheenen et al examined the mental health status of individuals with a mood disorder during the COVID-19 pandemic in an Australian sample as compared to individuals without a prior mood disorder. Their results underlined that distress in response to the COVID-19 pandemic is highlighted in individuals with a mood disorder [20]. Patients with a pre-existing mental health disorder increasingly reported worries related to infecting themselves or infecting others [21].

COVID-19 and Psychosomatic Medical Rehabilitation

A population that has been especially concerned by the COVID-19 pandemic because of pre-existing mental health problems comprises psychosomatic patients in medical rehabilitation. They may be afraid of visiting a doctor and receiving inpatient treatment in a hospital; on the other hand, they report worsening physical and mental well-being [21]. This development aggravates the already worrisome situation of psychosocial and psychosomatic rehabilitation programs, causing patients to remain untreated. If they decide to use medical services, they are confronted with many changes in therapy programs: due to contact regulations and hygiene measures, as well as the general lack of therapists in health care systems, it is necessary to develop and establish internet-based programs and trainings as one component of therapy as well as digital support systems and platforms.

Due to the pandemic and, accordingly, its restrictions, resources had to be re-allocated and therapies had to be paused, which caused a decrease in the availability of on-site services [22]. Particularly for older people, the fear of infection can prevent hospital or rehabilitation stays [23]. Possible solutions can be home-based or telehealth rehabilitation programs [22,24], or to shift parts of the rehabilitation treatment to online preparation in the form of a home-based telehealth intervention. This is especially innovative because during the past several years, patients have frequently been prepared for rehabilitation as well as treated during the rehabilitation stay with written material. While the focus has increasingly shifted toward integrative online trainings and interventions as the basis of psychotherapy, which are considered emerging technologies in health care and therapy, so far this is rather rare in the German rehabilitation system with its focus on inpatient treatment of “chronic psychosomatic conditions at risk of resulting in long-term sickness leave and disability” (page 79 in Scheidt [25]). Such interventions and trainings are independent of time and location. They can, therefore, be used in preparation for a rehabilitation stay [26], during a rehabilitation stay to supplement and support in-person therapy [27], as well as for aftercare and stabilization processes [28,29].

Digital interventions and trainings allow for adherence to hygiene measures as well as allow for therapeutic services to be offered on a large-scale basis. In addition, patients may be offered digital treatment options if they refrain from entering a rehabilitation stay due to worries and fears associated with the current COVID-19 pandemic, such as their own health and well-being or worries associated with losing their work placement [20]. Several studies have examined the usefulness of electronically delivered cognitive behavioral therapy (eCBT), which has proven to be effective with regard to the treatment of anxiety and depression compared to regular in-person therapy [30,31]. However, it remains to be evaluated whether digital trainings and therapies are useful measures to reduce symptoms of anxiety, depression, loneliness, and perceived stress in psychosomatic rehabilitation patients and how they can be implemented in practice. A crucial aspect of digital or mobile health (mHealth) interventions is the users’ acceptance, often operationalized as perceived usefulness and ease of use. Both constructs determine the current or future usage and, thus, pose an important prerequisite for possible intervention effects [32]. Both perceived usefulness and ease of use should, therefore, be considered in mHealth interventions.

The Health Action Process Approach

Drawing on the Health Action Process Approach (HAPA) [33,34], which is separated into motivational and volitional phases, higher intentions, planning, as well as self-regulatory strategies are needed to perform a health behavior change. During the motivational phase, an intention is formed, and after the formation, self-regulatory strategies ensure that the target behavior is realized and maintained as part of the volitional phase. Therefore, planning bridges the gap between intentions and the respective behavior [34]. Literature has shown that the HAPA as a theoretical basis for digital trainings and interventions in the sector of care after psychosomatic rehabilitation has proven to be an effective model in explaining behavior change with consequent improvements in mental health [35]. Especially for psychosomatic rehabilitation patients diagnosed with a pre-existing affective disorder, it is necessary to specifically promote competencies, such as formulating intentions, action plans, as well as coping plans, and to foster the development of outcome expectancies. Hereby, patients can increase their own control over their actions and can be supported by means of digital trainings and interventions to change from a situation-focused orientation, which is considered typical for depression, to an action-focused orientation [36]. Future research is necessary with regard to the HAPA being implemented in digital trainings and interventions for psychosomatic rehabilitation patients.

Goal of This Study

Based on the described background, we posed a number of research questions in order to understand worries and associated consequences regarding the COVID-19 pandemic in different populations, especially medical, psychosomatic rehabilitation patients, by means of internet technology:

1. What differences are there in reported psychological variables such as depression, anxiety, loneliness, and perceived stress between the general population and patients assigned to medical, psychosomatic rehabilitation clinics? The patients were diagnosed with a mental illness and were, thus, hypothesized to be at a higher risk for an exaggeration of their illness due to the pandemic, as shown above.
2. Which worries are associated with the current COVID-19 pandemic and are there differences in the perception between the two groups? We hypothesized that individuals from the psychosomatic rehabilitation group experienced more worries with regard to the pandemic.
3. Is medical, psychosomatic rehabilitation treatment effective in terms of a decrease in symptomatology for depressive symptoms, symptoms of anxiety, loneliness, and perceived stress?
4. Do the general population and patient groups intend to make use of internet-delivered treatment components?
5. Is there a relationship between the usage, as well as perceived usefulness, of digital trainings that are offered before as well as during the rehabilitation stay in association

with the intensity of mental health symptoms (eg, depression) of patients after their medical rehabilitation?

By testing these research questions, we aimed to close the research gap of evaluating mental health and COVID-19–related worries between the general population and psychosomatic rehabilitation. Furthermore, this study assessed the usefulness of internet-delivered trainings and their association with the mental health status of psychosomatic rehabilitation patients in Germany. To our knowledge, this has not been done before systematically. It is warranted to implement not only innovative but also effective internet-delivered interventions into the provision of medical services.

Methods

Overview

The study was conducted as part of the project “Anhand-COVID19 – Offer to achieve treatment and rehabilitation goals in compliance with hygiene and social-distancing rules” (ClinicalTrials.gov Identifier: NCT04453475), which is supported by the Dr Becker clinic group. In addition, data collection and analyses on the general population was part of the research project “TeamBaby – Safe, digitally supported communication in obstetrics and gynecology” (ClinicalTrials.gov Identifier: NCT03855735), which is funded by the German Innovation Fund (Project No. 01VSF18023) of The Federal Joint Committee (G-BA).

First Sample: Recruitment and Procedure of the General Population

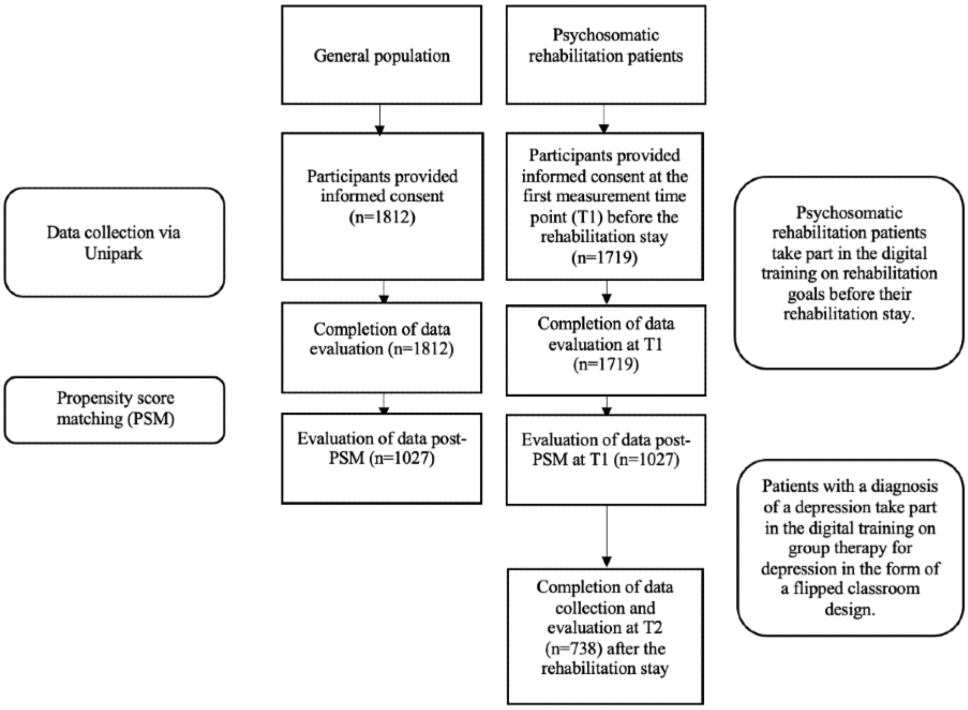
Data were collected anonymously through a nationwide recruitment campaign, press releases, social media posts, and the study home page of the TeamBaby project. No market research company or public sample was involved; however, the sample might have been selective. For data collection purposes, the software tool Unipark was used. The nationwide cross-sectional survey aimed to examine worries and coping mechanisms during the COVID-19 pandemic. All participants were informed about the purpose of the survey beforehand and provided online informed consent. Participants from the general population were not offered any form of compensation for participation. Data collection from the general population took place between May 2020 and April 2021. Time to complete the survey took, on average, 15.18 minutes (SD 11.50). Ethical approval for the online survey for the general population was given by the Ethics Committee at Jacobs University Bremen on September 17, 2019.

Second Sample: Recruitment and Procedure of Psychosomatic Rehabilitation Patients

The second group of participants were recruited through four psychosomatic clinics from the Dr Becker clinic group and attended regular treatment at the recruiting clinics, consisting of psychological and physical interventions (ie, individual and group psychotherapy, physiotherapy, or occupational therapy) as part of the incoming process for their rehabilitation stay. The German rehabilitation system focuses not on curation but on reintegration and social participation. “Interventions in rehabilitation include psychoeducation, physical training, psychotherapy, and the training of skills particularly with regard to working ability” (page 81 in Scheidt [25]). Participants from the four psychosomatic clinics were informed about the study in writing on the hospital group's original online portal. Therefore, only patients who had access to this digital portal via smartphone, tablet, or computer before the start of rehabilitation were included. Participation was only possible after the patients had read the participation information and had given their informed consent in writing; data were pseudonymized. Rehabilitation patients were not offered any form of compensation for their participation in the online study.

The online survey at the psychosomatic clinics was administered between July 2020 and April 2021. Data collection at the rehabilitation clinics was longitudinal and took place at two time points: 6 weeks before the start until the first day of rehabilitation (T1) and after their rehabilitation stay (T2). Four medical, psychosomatic rehabilitation clinics took part in this study and supported the recruitment of participants as well as provided psychosomatic rehabilitation between measurement points T1 and T2. For the recruitment process and data collection process, see Figure 1. Time to complete the survey at measurement point T1 took, on average, 29.28 minutes (SD 33.10) and at measurement point T2 took, on average, 30.16 minutes (SD 52.37). Ethical approval for the online survey concerning psychosomatic rehabilitation patients was given by the Ethics Committee at Jacobs University Bremen on June 25, 2020.

Figure 1. Study design of the cross-sectional and longitudiunal study.



Digital Intervention Only for the Second Sample of Psychosomatic Rehabilitation Patients

Digital trainings were provided via the internet prior to patients' rehabilitation stay to optimally prepare them for their medical rehabilitation treatment and to make good use of the treatment components, including psychoeducation, physical training, psychotherapy, and the training of skills, particularly with regard to working ability in the clinic. Such trainings could address rehabilitation goals.

The digital training on rehabilitation goals was offered to the patients in a digital PowerPoint (Microsoft) presentation training without face-to-face elements. Participants were able to participate in the digital training before their rehabilitation stay with a computer, laptop, tablet, or smartphone. The training included exercises on formulating precise plans for the rehabilitation stay. After the training, the patients were encouraged to make use of a digital exercise booklet containing exercises on formulating plans as well as for writing journal entries. Participation was on a voluntary basis. In addition, participants from two out of four psychosomatic rehabilitation clinics with a diagnosis of depression took part in group therapy for depression in the form of a flipped classroom as part of the rehabilitation treatment program (ie, digital group training for depression).

The digital group training for depression was a combination of digital and face-to-face components. The training was divided into six sessions, each consisting of a 5-minute digital training followed by a 45-minute analog group session. The digital training, including input from a therapist with flip chart accompaniment, was either viewed independently or was watched as a group at the beginning of the analog group session. Participation for patients with a diagnosed depressive disorder was mandatory. The digital group training for depression was based on cognitive behavioral therapy (CBT) and contained evidence-based components of eCBT and internet-delivered CBT interventions, based on the current state of the art [37-39]. Contents of the group sessions included, for example, an explanation of depression symptoms and how to cope with them in the form of psychoeducation, underlying models, and different available treatments (ie, pharmacotherapy and psychotherapy).

Instruments

Table 1 provides an overview of all questionnaires and scales used for the two subsamples as part of this study.

Table 1. Overview of questionnaires and scales used for the general population and the psychosomatic rehabilitation patients.

Variable	General population (N=1812)	Psychosomatic rehabilitation patients (N=1719)
Questionnaire or scale, mean (SD)[a]		
Overall worries related to the COVID-19 pandemic[b]	48.10 (9.07)	51.47 (7.51)
Depression (PHQ-2[c])	2.21 (1.89)	3.47 (1.65)
Anxiety (GAD-2[d])	1.95 (1.85)	3.61 (1.69)
Perceived stress (PSS-4[e])	7.40 (3.56)	9.44 (2.56)
Loneliness[f] (CES-D[g]; UCLA[h] Loneliness Scale)	4.14 (2.12)	4.49 (1.77)
Intention to use apps or digital trainings (HAPA[i])	5.43 (3.24)	6.51 (3.08)
Perceived usefulness of digital trainings (TAM[j])	N/A[k]	5.22 (1.90)
Sociodemographic characteristics, n (%)[a]		
Age (years)		
≤29	407 (22.5)	70 (4.1)
30-39	416 (23.0)	216 (12.6)
40-49	352 (19.4)	390 (22.7)
50-59	385 (21.2)	803 (46.7)
≥60	252 (13.9)	236 (13.7)
Sex		
Male	529 (29.2)	602 (35.0)
Female	1267 (70.0)	1104 (64.2)
Education		
10 or 11 years of schooling	193 (10.7)	398 (23.2)
12 or more years of schooling	421 (23.2)	241 (14.0)
Vocational training	507 (28.0)	791 (46.0)
University degree	690 (38.1)	266 (15.5)

[a]Mean (SD) and frequency values before propensity score matching.

[b]Overall worries were measured by 17 items on a self-constructed questionnaire.

[c]PHQ-2: 2-item Patient Health Questionnaire; items were rated on a 4-point Likert scale from 0 (not at all) to 3 (nearly every day), with summed scores from 0 to 6.

[d]GAD-2: 2-item Generalized Anxiety Disorder scale; items were rated on a 4-point Likert scale from 0 (not at all) to 3 (nearly every day), with summed scores from 0 to 6.

[e]PSS-4: 4-item Perceived Stress Scale; items were rated on a 5-point Likert scale from 0 (never) to 4 (very often), with summed scores from 0 to 16.

[f]Loneliness items were rated on a 4-point Likert scale from 1 (not at all) to 4 (almost every day), with summed scores from 0 to 8.

[g]CES-D: Center for Epidemiologic Studies–Depression scale.

[h]UCLA: University of California, Los Angeles.

[i]HAPA: Health Action Process Approach; items were rated on a 5-point Likert scale from 1 (no, I do not intend to) to 5 (yes, and it is very easy for me), with summed scores from 3 to 15.

[j]TAM: Technology Acceptance Model; items were rated on a 5-point Likert scale from 1 (not at all useful) to 5 (completely useful), with summed scores from 2 to 10.

[k]N/A: not applicable; this item was not relevant to the general population.

Instruments Used for the General Population and the Psychosomatic Rehabilitation Patients

Worries Related to the COVID-19 Pandemic

Items assessing worries related to the COVID-19 pandemic were derived from a study that measured frequently reported burdens and worries due to the COVID-19 pandemic [40]. Consequently, an item pool of 77 elements was developed, of which 17 items were of interest for further analysis, as they described common worries related to the COVID-19 pandemic. All items were refined by psychologists and a medical professional with expertise in the field of health psychology and psychosomatic rehabilitation.

Depressive Symptoms and Symptoms of Anxiety

For both subsamples, symptoms of depression and anxiety were measured with the 4-item Patient Health Questionnaire (PHQ-4), which is the composite measure of the 2-item Patient Health Questionnaire (PHQ-2) [41] and the 2-item Generalized Anxiety Disorder scale (GAD-2) [42], which measure symptoms of depression and anxiety, respectively [43]. The PHQ-4 consists of four items rated on a 4-point Likert scale from 0 (not at all) to 3 (nearly every day). Summed scores of 3 or higher for both the PHQ-2 (Spearman ρ=0.75) and the GAD-2 (Spearman ρ=0.74) indicated a probable case of depression and anxiety [42,44]. The PHQ-2 and the GAD-2 were not used as diagnostic tools in this study but, rather, were used to highlight symptoms associated with depression and anxiety.

Perceived Stress

The Perceived Stress Scale (PSS) [45] is a globally used self-report scale that measures perceived stress; the PSS was presented to the general population and to the psychosomatic rehabilitation patients. The scale assesses “the degree to which situations in one's life are appraised as stressful” (page 387 in Cohen et al [45]), situations that are, therefore, perceived as unpredictable, uncontrollable, and overloaded during the past month. For the purpose of this study, perceived stress was assessed using the short version of this scale, the 4-item PSS (PSS-4) [46]. It assesses perceived stress by rating four items on a 5-point Likert scale from 0 (never) to 4 (very often), with a Cronbach α of .79.

Loneliness

Loneliness was assessed with two items: “How often do you feel lonely?” stemming from the Center for Epidemiologic Studies–Depression scale [47], and “How often do you feel unhappy to be alone?” from the UCLA (University of California, Los Angeles) Loneliness Scale [48] (Spearman ρ=0.85). The items were rated on a 4-point Likert scale from 1 (not at all) to 4 (almost every day). Both items were presented to the general population and to the psychosomatic rehabilitation patients.

Intention to Use Apps or Digital Trainings During the COVID-19 Pandemic

Intention to use apps or digital trainings as supportive means during the COVID-19 pandemic was assessed by rating three items on a 5-point Likert scale from 1 (no, I do not intend to) to 5 (yes, and it is very easy for me). These items were adapted based on the stages of change as part of the HAPA, which suggests that individuals typically progress through stages of behavior change independently of any time frame [49,50].

Instruments Used Only for Psychosomatic Rehabilitation Patients: Perceived Usefulness of Digital Trainings

Based on the different digital trainings that psychosomatic rehabilitation patients took part in, the perceived usefulness of the offered digital trainings was measured by rating two items on a 5-point Likert scale from 1 (not at all useful) to 5 (completely useful): one item for the digital training on rehabilitation goals and one item for the digital, flipped classroom, group therapy for depression. Both items were adopted and modified based on the Technology Acceptance Model, which was originally designed to evaluate patients' responses to health information technology [32].

Sociodemographic and Additional Information

Additional data on sociodemographic information included participants' age, sex, and educational status. Age was categorized into five groups: ≤29 years, 30-39 years, 40-49 years, 50-59 years, and ≥60 years. Sex was categorized into three groups: male, female, and diverse. The highest obtained educational status was categorized into four groups: 10 or 11 years of schooling, 12 or more years of schooling, vocational training, and university degree. All variables were measured as categorical variables.

Data Analysis for Both Subsamples

Literature has shown that propensity score matching (PSM) has been able to effectively reduce biases of treatment selection in nonrandomized studies [51]. Through PSM, covariates can be balanced between groups [52]. Hence, in this study, a PSM analysis was used to minimize the effect of confounding variables as well as the uneven distribution of covariates in the two groups before comparing them. The matching algorithm was based on logistic regression. Participants were matched based on sex, age, and educational status; the match tolerance was 0.01 without any failures to match.

After PSM, 2054 participants were included for further analyses, and data from the general population and the psychosomatic clinics were examined for differences. To assess whether individuals recruited from the general population and individuals from the psychosomatic clinics differed in their expression concerning psychological symptoms of depression, anxiety, loneliness, and perceived stress, a multivariate analysis of covariance was performed, controlling for gender, age, and educational level. Afterward, an exploratory factor analysis (EFA) was carried out to determine factors within the worries related to the COVID-19 pandemic based on items' factor loadings. Regarding the EFA, meaningful factors to retain for further analysis were based on the scree plot as well as the percentage of common variance explained by a given factor with an eigenvalue above 1. Meaningful factors were retained for varimax rotation. Items with a factor loading above 0.40 were used for interpretation purposes. Hence, out of 17 items primarily used to analyze worries related to the COVID-19 pandemic, one item was eliminated due to a low item loading. After EFA, significant differences between the data from the general population and the psychosomatic clinics, regarding the defined factors measuring worries related to the COVID-19 pandemic, were examined by a multivariate analysis of covariance controlling for gender, age, educational status, perceived stress, loneliness, depressive symptoms, and symptoms of anxiety.

In addition, a repeated-measures analysis of covariance was performed, controlling for gender and age on 738 psychosomatic rehabilitation patients to examine whether individuals from the psychosomatic clinics showed a change in psychological symptoms on the variables of depression, anxiety, loneliness, and perceived stress prior to and after their rehabilitation stay. To evaluate whether taking part in digital trainings (ie, rehabilitation goals and group therapy for depression) was associated with a significant change in symptom intensity with regard to depression, anxiety, loneliness, and perceived stress, a repeated-measures analysis of covariance was performed, controlling for age and gender.

To examine the intention to use common digital apps and trainings with a focus on health that were not offered during the rehabilitation stay with regard to the general population and patients from the psychosomatic rehabilitation clinics, an analysis of covariance was performed controlling for age, gender, and educational status. Finally, to evaluate the perceived usefulness of internet trainings offered during the rehabilitation stay and the association with patients' mental health status after their rehabilitation stay, a hierarchical regression analysis was performed. All data analyses were carried out using SPSS, version 27 (IBM Corp).

Missing Data

The amount of missing data was below 5% for all items and 1.3% on average. Participants with missing data on the social-cognitive variables were included for further analysis if they had at least one nonmissing data point under the assumption of missing completely at random. However, no missing data points were imputed due to the overall low percentage of missing data points.

Results

Participants Before Propensity Score Matching: General Population

Overall, 3531 participants completed the online questionnaire. With regard to the general population, 1812 participants participated in the data collection. Out of these participants, 1267 (69.9%) were female and 16 (0.9%) did not respond. Age ranged from 18 to over 60 years. Out of 1812 participants, 193 (10.7%) had 10 or 11 years of schooling, 421 (23.2%) had 12 or more years of schooling, 507 (28.0%) had completed vocational training, and 690 (38.1%) had a university degree; there was 1 (0.1%) missing data point.

Participants Before Propensity Score Matching: Psychosomatic Rehabilitation Patients

Concerning participants from the psychosomatic rehabilitation clinics, 1719 participants participated in the survey before their rehabilitation stay. Of these participants, 1104 (64.2%) were female and there were 13 (0.8%) missing data points. Age ranged from 18 to over 60 years. Out of 1719 participants, 398 (23.2%) had 10 or 11 years of schooling, 241 (14.0%) had 12 or more years of schooling, 791 (46.0%) had completed vocational training, and 266 (15.5%) had a university degree; there were 23 (1.3%) missing data points. After the rehabilitation stay, 738 participants participated in the survey.

Participants After Propensity Score Matching: General Population

With regard to the general population of 1027 participants, 684 (66.6%) were female, their age ranged from 18 to over 60 years, 163 (15.9%) had 10 or 11 years of schooling, 173 (16.8%) had 12 or more years of schooling, 409 (39.8%) had completed vocational training, and 282 (27.5%) had a university degree.

Participants After Propensity Score Matching: Psychosomatic Rehabilitation Patients

With regard to the 1027 participants from the psychosomatic rehabilitation clinics, 659 (64.2%) were female, their age ranged from 18 to over 60 years, 167 (16.3%) had 10 or 11 years of schooling, 194 (18.9%) had 12 or more years of schooling, 404 (39.3%) had completed vocational training, and 262 (25.5%) had a university degree.

Difference in Psychological Symptoms

The multivariate analysis of covariance revealed significant differences in mental health between the general population and individuals from the psychosomatic clinics ($F_{4,2028}$=183.74, P<.001, η^2_p=0.27), with age, gender, and education being significant covariates. Individuals from the psychosomatic clinics displayed significantly higher scores on all four psychological variables compared to individuals recruited from the general population: depression ($F_{1,2036}$=460.51, P<.001, η^2_p=0.19), anxiety ($F_{1,2036}$=682.11, P<.001, η^2_p=0.25), loneliness ($F_{1,2036}$=90.31, P<.001, η^2_p=0.05), and perceived stress ($F_{1,2036}$=424.65, P<.001, η^2_p=0.17) (Table 2).

Table 2. Descriptive statistics and mean differences[a] between the general population and the sample from the psychosomatic clinics (ie, medical sample) across the test variables of depression, anxiety, loneliness, and perceived stress.

Test variable	General population, mean (SD)	Medical sample, mean (SD)	Mean difference	95% CI of the grand mean	*P* value	Cohen *d*[b]
Depression (PHQ-2[c])	1.85 (1.70)	3.43 (1.64)	–1.58	2.57-2.71	<.001	0.57
Anxiety (GAD-2[d])	1.65 (1.67)	3.58 (1.68)	–1.93	2.54-2.68	<.001	0.69
Loneliness[e] CES-D[f]; UCLA[g] Loneliness Scale	3.70 (1.92)	4.47 (1.75)	–0.77	4.01-4.17	<.001	0.23
Perceived stress (PSS-4[h])	6.75 (3.42)	9.46 (2.56)	–2.71	7.98-8.23	<.001	0.30

[a]Descriptive statistics and mean differences after propensity score matching.

[b]Cohen *d*: 0.20=small effect, 0.50=medium effect, and 0.80=large effect.

[c]PHQ-2: 2-item Patient Health Questionnaire; items were rated on a 4-point Likert scale from 0 (not at all) to 3 (nearly every day), with summed scores from 0 to 6.

[d]GAD-2: 2-item Generalized Anxiety Disorder scale; items were rated on a 4-point Likert scale from 0 (not at all) to 3 (nearly every day), with summed scores from 0 to 6.

[e]Loneliness items were rated on a 4-point Likert scale from 1 (not at all) to 4 (almost every day), with summed scores from 0 to 8.

[f]CES-D: Center for Epidemiologic Studies–Depression scale.

[g]UCLA: University of California, Los Angeles.

[h]PSS-4: 4-item Perceived Stress Scale; items were rated on a 5-point Likert scale from 0 (never) to 4 (very often), with summed scores from 0 to 16.

Analysis of Worries Related to the COVID-19 Pandemic

After varimax rotation, 16 items were retained in the analysis with factor loadings of ≥0.40 with the respective factor (Table 3). Four factors were able to explain 67.14% of the total variance. Factors identified included satisfaction with communication (six items measured on a 6-point Likert scale from 1 [completely disagree] to 6 [completely agree]; reliability indicator Cronbach α=.90), health-related worries (six items measured on a 5-point Likert scale from 1 [never] to 5 [always]; reliability indicator Cronbach α=.82), financial worries (two items measured on a 5-point Likert scale from 1 [no, completely disagree] to 5 [yes, completely agree]; reliability indicator Spearman ρ=0.65), and household-related worries (two items; reliability indicator Spearman ρ=0.28). With regard to household measures, one item was assessed on a 4-point Likert scale from 1 (not at all) to 4 (completely). The second item was assessed on a 6-point Likert scale from 1 (never or less than once per month) to 6 (daily or several times per day). Hence, the Likert scale of the second item was transformed to a 4-point Likert scale. For all factors, composite mean scores were computed.

Table 3. Exploratory factor analysis: factor loadings with all study participants from the general population and the medical sample (N=2054).

Scale and item	Label	Factor loadings[a]			
		1	2	3	4
Scale 1: Satisfaction with communication	Clear explanation	0.92	—[b]	—	—
	Early communication	0.91	—	—	—
	Sufficient information	0.89	—	—	—
	Taken seriously during communication	0.87	—	—	—
	Made sure that everything was understood	0.85	—	—	—
	Including accompanying persons and respecting situation	0.61	—	—	—
Scale 2: Health-related worries	Concerned about getting infected	—	0.86	—	—
	Concerned about becoming ill	—	0.84	—	—
	Concerned about visiting a doctor	—	0.71	—	—
	Concerned about infecting others	—	0.70	—	—
	Concerned about visiting the hospital	—	0.66	—	—
	Anxious when hearing news	—	0.62	—	—
Scale 3: Financial worries	Worries about one's job	—	—	0.87	—
	Afraid of financial difficulties	—	—	0.87	—
Scale 4: Household-related worries	Conflicts in household	—	—	—	0.80
	Supported each other as a household	—	—	—	0.75

[a]Exploratory factor analysis and factor loadings after propensity score matching.

[b]Factor loadings were reported for their corresponding scales.

The same factor structure was found in both samples. A total of 70.32% of the variance could be explained in the general population and 64.33% of the variance could be explained in the sample of psychosomatic rehabilitation patients.

Summarizing the results from the factor analysis, the factor structure of the evaluated worries associated with the current COVID-19 pandemic was equal across samples. Hence, the overall EFA across samples revealed four factors associated with the COVID-19 pandemic: satisfaction with communication, health-related worries, financial worries, and household-related worries.

Differences in Worries Related to the COVID-19 Pandemic Between Groups

Results from the multivariate analysis of covariance indicated significant differences between the two groups ($F_{4,835}$=17.17, P<.001, η^2_p=0.08) concerning worries related to the COVID-19 pandemic (Table 4): satisfaction with communication ($F_{1,838}$=31.66, P<.001, η^2_p=0.04), household-related worries ($F_{1,838}$=5.34, P=.02, η^2_p=0.01), and financial worries ($F_{1,837}$=38.87, P<.001, η^2_p=0.04). Age, gender, perceived stress, loneliness, depressive symptoms, and symptoms of anxiety were significant covariates. Hence, patients recruited from the psychosomatic clinics perceived a significantly greater satisfaction with communication, increased household-related worries, but significantly lower financial worries before their rehabilitation stay (Table 4).

Patients reported being unemployed more frequently prior to their rehabilitation stay (253/1027, 24.6%) compared to the general population (123/1027, 12.0%; Table 3). Furthermore, the patient sample reported more health-related worries (Table 3). However, the difference between the groups was revealed to be nonsignificant ($F_{1,837}$=0.13, P=.72, η^2_p=0.01).

Table 4. Descriptive statistics and mean differences[a] between the general population and the sample from the psychosomatic clinics (ie, medical sample) across COVID-19–related worries.

Test variable[b]	General population, mean (SD)	Medical sample, mean (SD)	Mean difference	95% CI of the grand mean	*P* value	Cohen *d*[c]
Satisfaction with communication[d]	24.45 (7.45)	26.53 (5.69)	–2.08	25.06-25.94	<.001	0.05
Health-related worries[e]	14.56 (5.60)	15.69 (4.88)	–1.13	14.77-15.46	.02	0.04
Financial worries[f]	4.48 (2.52)	4.31 (2.21)	0.17	4.24-4.55	<.001	0.03
Household-related worries[g]	6.06 (1.32)	5.93 (1.28)	0.13	5.90-6.08	.72	0.08

[a]Descriptive statistics and mean differences after propensity score matching.

[b]Scales were aggregated from items reported in Table 2.

[c]Cohen *d*: 0.20=small effect, 0.50=medium effect, and 0.80=large effect.

[d]Satisfaction with communication: 6 items were rated on a 6-point Likert scale from 1 (completely disagree) to 6 (completely agree), with summed scores from 6 to 36.

[e]Health-related worries: 6 items were rated on a 5-point Likert scale from 1 (never) to 5 (always), with summed scores from 5 to 25.

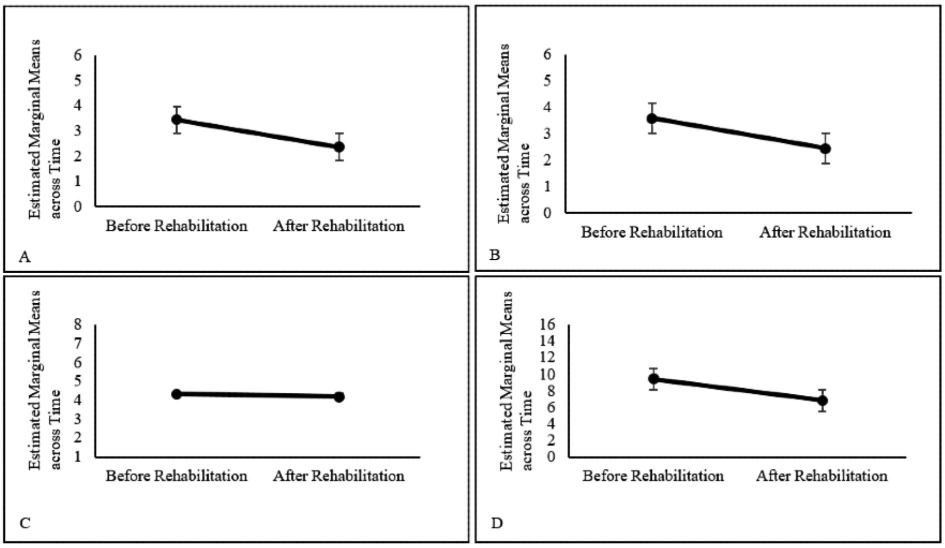
[f]Financial worries: 2 items were rated on a 5-point Likert scale from 1 (no, completely disagree) to 5 (yes, completely agree), with summed scores from 2 to 10.

[g]Household-related worries: the first item was rated on a 4-point Likert scale from 1 (not al all) to 4 (completely) and the second item was measured on a 6-point Likert scale from 1 (never of less than once per month) to 6 (daily or several times per day). After transformation to a 4-point Likert scale, summed scores ranged from 2 to 8.

Changes in Psychological Symptoms Before and After Rehabilitation

Results of the repeated-measures analysis revealed overall significant differences across time ($F_{1,712}$=23.21, P<.001, η^2_p=0.04). Taking a closer look at individual test variables, results revealed a significant reduction in depression ($F_{1,723}$=0.98, P<.001, η^2_p=0.02), anxiety ($F_{1,723}$=0.99, P<.001, η^2_p=0.01), perceived stress ($F_{1,720}$=19.69, P<.001, η^2_p>=0.03), and loneliness ($F_{1,722}$=0.99, P<.05, η^2_p=0.005) in psychosomatic patients after their rehabilitation (Figure 2).

Figure 2. Estimated marginal means after propensity score matching for symptoms of depression (A), symptoms of anxiety (B), perceived loneliness (C), and perceived stress (D). Error bars represent standard errors of the mean.



Intention to Use Common Digital Trainings and Apps

Results of the multivariate analysis of variance revealed an overall significant difference between individuals from the general population and individuals from the medical, psychosomatic rehabilitation clinics ($F_{3,2021}$=51.41, P<.001, η^2_p=0.07). Patients appeared more inclined to use common apps and digital trainings offered outside of their rehabilitation stay supporting them in their communication with health care professionals ($F_{1,2027}$=6.66, P=.01, η^2_p=0.01) as well as COVID-19–related health care apps ($F_{1,2027}$=144.51, P<.001, η^2_p=0.07).

Association Between Taking Part in Digital Trainings and Changes in Psychological Symptom Intensity

We examined whether taking part in the digital training on rehabilitation goals (ie, only digital training) and digital group therapy for depression (ie, combination of digital and face-to-face components) were associated with a decrease in symptom intensity after the rehabilitation stay compared to before. The results highlighted the following significant differences: taking part in the digital group therapy for depression was associated with a significant decrease in symptom intensity after the rehabilitation stay with regard to depression ($F_{1,725}$=4.82, P=.03, η^2_p=0.01) and anxiety ($F_{1,725}$=6.22, P=.01, η^2_p=0.01).

Perceived Usefulness of Digital Trainings and Association With Mental Health Status

Table 5 shows the association between perceived usefulness of the digital trainings (ie, digital training on rehabilitation goals and digital group therapy for depression) evaluated by participants after their rehabilitation stay and their mental health status after their rehabilitation stay. Overall, increased perceived usefulness of digital training on rehabilitation goals was significantly associated with a higher reduction in perceived depression, anxiety, loneliness, and stress postrehabilitation.

Table 5. Association between perceived usefulness of digital trainings and mental health status of psychosomatic rehabilitation patients after their rehabilitation stay[a].

Predictor	Dependent variable							
	Depression, β (95% CI)	*P* value	Anxiety, β (95% CI)	*P* value	Loneliness, β (95% CI)	*P* value	Perceived stress, β (95% CI)	*P* value
Participation in digital depression group therapy	.08 (–.01 to .25)	.08	.08 (–.02 to .24)	.11	.09 (–.02 to .25)	.08	.05 (–.12 to .39)	.30
Participation in digital training on rehabilitation goals	–.14 (–.37 to –.07)	<.001	–.13 (–.36 to –.05)	<.001	–.19 (–.46 to –.16)	<.001	–.19 (–.90 to –.31)	<.001

[a]Each column represents a separate analysis after propensity score matching. Analyses controlled for age, gender, and education, with gender being significant at P<.05 for anxiety and perceived stress, age being significant for loneliness and stress, and education being significant for education.

Data Availability

The data that support the findings of this study are available from the corresponding author (SL) upon reasonable request.

Discussion

Principal Findings

In this study, differences between a sample of 1027 individuals from the general population and 1027 patients from medical, psychosomatic rehabilitation clinics in depression, anxiety, loneliness, and perceived stress during the COVID-19 pandemic were examined by means of the internet after PSM. The expression of symptoms and worries related to the COVID-19 pandemic were psychometrically assessed and tested for differences between the two samples (research questions 1 and 2). As a third research question, a potential decrease in symptom intensity on the test variables was examined for psychosomatic rehabilitation patients before starting and after their rehabilitation stay. Moreover, this paper evaluated the differences in intention to use digital apps and trainings during the COVID-19 pandemic between individuals from the general population and individuals from the psychosomatic rehabilitation clinics (research question 4). With regard to the potential decrease in symptoms, research question 5 evaluated the association between participation in digital trainings addressing rehabilitation goals and digital depression group therapy. Furthermore, the perceived usefulness of digital trainings before (ie, digital training on rehabilitation goals) and during (ie, digital group therapy for depression) the rehabilitation stay was evaluated with regard to the symptom intensity of depression, anxiety, loneliness, and perceived stress after the rehabilitation stay.

The findings from this study confirm that individuals felt affected by the COVID-19 pandemic in terms of their mental health and well-being. For individuals from the psychosomatic medical rehabilitation, symptoms of depression, anxiety, loneliness, and perceived stress were elevated compared to the general population. Thus, we can answer the research question 1 by showing that individuals assigned to medical, psychosomatic rehabilitation clinics perceive and express more mental health symptoms, which is in line with our hypothesis. In prior research, individuals with a pre-existing mental health disorder reported poorer access to support services since the beginning of the pandemic, had earlier discharges from psychiatric units, or had discontinuation of psychotherapy treatments [53-55]. The loss of such support systems due to the COVID-19 pandemic may have led to negative consequences, such as an increase in symptom intensity, increased social isolation, and perhaps even suicidal behavior [53]. Therefore, digital interventions and trainings that target positive thinking, active stress coping, and social support to reduce depression, anxiety, loneliness, and perceived stress need to be implemented for individuals with a pre-existing mental health disorder, irrespective of taking part in rehabilitation; these may also work as primary preventative measures for the general population [56]. Accordingly, research questions arise in the context of digital prevention as well as digital support interventions, which should be investigated further.

Surprisingly, psychosomatic patients perceived significantly greater satisfaction in communication with health care professionals and had significantly lower financial worries but higher household-related worries, even after statistically controlling for confounding variables. However, no significant difference between the groups was found with regard to health-related worries, which is contrary to the hypothesis that they would experience more worries (research question 2). The fact that psychosomatic rehabilitation patients perceived greater satisfaction with communication before their rehabilitation stay may be due to previous information obtained digitally (ie, through the digital training on rehabilitation goals) from the clinic as well as participating in surveys and tasks before their stay. Furthermore, contact with the rehabilitation clinics might have been perceived as an emerging support system by the rehabilitation patients, offering the hope that their situation would soon improve and that they would receive help during the pandemic. Literature has shown that effective communication with patients may prove empowering for patients [57].

The results concerning financial worries are partly in line with van Rheenen et al [20]. Their study indicated that individuals with a mood disorder expressed lower concerns with personal finances, as they were more commonly unemployed or unable to work [58]. As with the results of this study, psychosomatic rehabilitation patients increasingly indicated that they were either unemployed or unable to work before the rehabilitation stay. Due to pre-existing unemployment or lack of participation in the workforce, there was already a lower financial status and greater job insecurity as well as financial uncertainty [59]. Additionally, the inability to work due to disabilities is, in part, financially subsidized by the German social system [60]. Besides, pre-existing mental health disorders are associated with greater incapacity to work and may lead to an earlier disability pension [61]. Hence, these patients may not be aware of, nor concerned with, job uncertainty as a result of the COVID-19 pandemic due to their medical treatment, which was partially digitally supported.

Psychosomatic rehabilitation patients indicated greater worries associated with their household, which includes conflicts within the family or dissatisfaction with household dynamics before their rehabilitation stay, as compared to individuals from the general population. As the COVID-19 pandemic is characterized by strict travel restrictions, an increase in working from home and homeschooling, short-term employment, or unemployment, people tend to either spend more time with immediate family at home, leading to an increase in family conflicts, or experience isolation while quarantining. The results by Guo et al support the results of this study. They highlight that a risk factor associated with reduced mental health status during the COVID-19 pandemic is living alone [62]. Nevertheless, family conflicts as a correlate of the COVID-19 pandemic may, conversely, also be a stressor contributing to diminishing mental health. Digital solutions offer the option to bridge the gap to mobile rehabilitation, especially if family constraints prevent patients from attending rehabilitation treatment on-site. According to the research, questions arise and should be investigated further.

Moreover, individuals with and without mood disorders reported a similar frequency of worries related to health, such as worries about loved ones falling sick with COVID-19 as well as implications for one's own health and well-being. This is in line with results by van Rheenen et al highlighting equal concerns about the health and well-being of the social environment for individuals with mood disorders (ie, depression and anxiety) and those without a mental health disorder. Furthermore, individuals with and without a mental health disorder indicated almost equal concerns regarding their own health and well-being during the COVID-19 pandemic [20]. This shows that health concerns about others and oneself during the coronavirus pandemic are estimated as equally important, irrespective of the mental health status of individuals.

In addition, results indicated that for psychosomatic patients, symptoms of depression and anxiety as well as of perceived stress and loneliness decreased significantly between pre- and postrehabilitation, thus answering in favor of the third research question. This underlines the importance and necessity of medical rehabilitation treatment for patients with chronic mental disorders [63]; in particular, the group therapy for depression in the form of a flipped classroom design shows support for the decrease in symptom intensity with regard to depression and anxiety postrehabilitation compared to before the rehabilitation stay. Past evidence has shown that the combination of digital therapeutic elements and regular face-to-face therapy was able to improve the mental health outcomes of patients significantly [64-66], which is in line with the results of this study.

Interestingly, compared to individuals from the general population, psychosomatic rehabilitation patients reported a greater intention to use common apps and digital trainings focusing on health and that are not offered during rehabilitation. This offers important insights into research question 4. First of all, patients who have applied for a rehabilitation stay may already be more open to medical and lifestyle interventions. Hence, pre-existing motivation to change may foster intentions to pursue a change. Furthermore, patients with an affective disorder may be more prone to excessive reassurance-seeking, which may be defined as the repeated need for safety-related information [67,68]. Therefore, one may postulate that by reassurance-seeking through the use of health care–related apps, patients with an affective disorder may fulfill their desires for safety behaviors. However, upon the increased intention to use those apps, patients need to learn effective coping strategies and skills to perform and maintain the actual behavior without using excessive reassurance-seeking and relying on safety behaviors. Shafran et al highlighted the importance of daily self-monitoring through, for example, digital trainings in supporting patients with reduced mental health status in translating intentions into actual behavior [69].

Next to the increased intentions of psychosomatic rehabilitation patients to use digital trainings, those who evaluated the perceived usefulness of the digital training on rehabilitation goals as increasingly useful and helpful postrehabilitation stay also displayed lower symptoms regarding depression, anxiety, loneliness, and perceived stress compared to the general population. This finding clearly supports a positive association that was examined in the final research question. One might postulate that useful and helpful preparation for the rehabilitation stay provides the basis for effective digital training, such as digital group therapy for depression. Feeling well prepared, informed, and being offered additional material prior to the rehabilitation stay may motivate the patients to be an active agent and engage effectively in the subsequent digital treatment programs. Moreover, the result can be explained by the assumptions of the HAPA [33]. During the voluntary participation in the digital rehabilitation goals training, an intention to achieve a better mental health status postrehabilitation may be created. Based on this intention and in combination with the supplementary support provided during the digital training on rehabilitation goals, participants may develop adequate planning strategies to reach the desired health outcome.

With the support of the digital group therapy addressing depression, the desired behavior of achieving a better mental health status may be facilitated by means of developing coping strategies, learning new skills, and activating resources. Initial approaches in offering psychosomatic rehabilitation patients digital therapy tools for aftercare have been made by Schmädeke et al and have proven to be effective [35]. Hence, for future studies, a digital training focusing on the preparation for a medical rehabilitation stay, the support of face-to-face therapy, and empowering patients for the time after rehabilitation should be developed and evaluated based on the HAPA model, as our results are promising. In addition, such digital training should be assessed further with regard to its effectiveness in the form of a randomized controlled trial with a waiting control group.

Overall, psychosomatic rehabilitation is an effective treatment, especially during the pandemic, and should be offered to all people who either suffer from a pre-existing chronic mental disorder or who developed mental disorders due to the pandemic and its restrictions. In addition, digital trainings should be integrated with the rehabilitation process for patients with an affective mood disorder. As the COVID-19 pandemic poses several barriers toward the uptake of a psychosomatic treatment [53], it must, therefore, be ensured that people with pre-existing or newly developed mental disorders have simple, straightforward access to psychosomatic rehabilitation and additional internet-delivered supplements. Hence, possible access opportunities for psychosomatic patients may also be provided in the form of low-threshold digital trainings to offer support before a rehabilitation stay.

This study highlights the need to offer individuals support to maintain sufficient mental health, especially in times of a pandemic and its aftermath. This can be achieved in multiple, low-threshold ways that meet different needs and preferences. It may include offering individuals—not only limited to individuals with a prior mental health diagnosis to ensure prevention—facilitated access to video and telephone consultation hours, digital preventive programs, or psychosomatic rehabilitation stays. Facing a substantial lack of medical doctors, therapists, and other health care workers and the need to reduce physical contact, it is necessary to develop and establish internet-based programs and trainings as one component of therapy as well as digital support systems and platforms.

Internet-delivered treatment components offer different advantages that need to be planned more systematically. For instance, physicians or general practitioners should briefly screen all patients perceived to be at risk for stress, anxiety, or depression due to the pandemic (ie, using the GAD-2 and the PHQ-2) and then recommend further online services [70-72] or hybrid options to those with elevated symptoms. Moreover, individualized recommendations on how to deal with mental health difficulties for the general population, as well as for individuals with a pre-existing mental health disorder, should be created. Suggestions on how to deal with barriers, such as finding specialist care and waiting times during the pandemic, should be a key component of these recommendations [73], especially if individuals are confronted with hygiene regulations that might conflict with the need to connect socially with others or to seek professional help.

Limitations and Suggestions for Future Research

Several limitations need to be considered while interpreting the findings of this study. First of all, the data evaluated from the general population are cross-sectional. Also, any changes due to the COVID-19 pandemic (ie, situation, perception, behavior, well-being, or mental health) could not have been controlled for significant events such as lockdown measures in Germany.

Furthermore, the items assessing worries related to the COVID-19 pandemic were nonvalidated items based on research found in the literature so far. Hence, for future studies, the analyzed items should be examined with respect to validity, while this study provides the first results regarding reliability. Another critical point to highlight is that mental health, in the form of symptoms of depression, anxiety, loneliness, and stress, was self-reported. In addition, the questionnaires that were used (ie, the PHQ-2 and the GAD-2) only gave an indication about the symptom intensity but were not used as diagnostic tools.

Furthermore, we had no indication of symptoms or clinical mental health diagnoses before the start of the COVID-19 pandemic. Hence, we cannot be certain that the levels of psychological symptoms reported by participants were subject to the COVID-19 pandemic. However, it has been suggested that symptoms of anxiety and depression increased as a result of the COVID-19 pandemic compared to historical normative data [74].

In addition, the digital training on rehabilitation goals and group therapy for depression were not tested with regard to their effectiveness beforehand. Furthermore, with regard to the data collected with the psychosomatic clinics, it remains to be evaluated in future research whether somatic aspects, including a potential COVID-19 infection, are a confounding factor for the expression of symptoms of depression, anxiety, loneliness, and perceived stress.

Conclusions

This study provides insights into the mental health status and perceived well-being of psychosomatic rehabilitation patients compared to the general population during the COVID-19 pandemic. In addition, this study was able to evaluate COVID-19–related worries between the general population and psychosomatic rehabilitation patients. Further, the usefulness of internet-delivered trainings and their association with the mental health status of psychosomatic rehabilitation patients in Germany was assessed. The results suggest that psychosomatic rehabilitation patients perceived greater symptoms of depression, anxiety, loneliness, and stress compared to the general population before their rehabilitation stay, which validates their status as assigned to rehabilitation. Future studies should replicate these findings in other countries and with individuals from a different cultural background. In particular, the question remains as to whether different health care systems and rehabilitation treatments (eg, delivered at a higher proportion in a mobile, internet-delivered mode) would result in the same outcomes. It is also imperative to disentangle which components of the internet-delivered interventions were especially effective in which patients.

The general population perceived greater financial worries, whereas patients before their rehabilitation stay perceived greater worries associated with their health and household. In addition, our results underline that in comparison to before their medical rehabilitation stay, patients' symptoms of depression, anxiety, and perceived stress were significantly lower after their rehabilitation stay. This stresses the value and necessity of psychosomatic rehabilitation treatments, concerning the psychotherapy of chronic mood disorder, and their relevance during the COVID-19 pandemic, especially among individuals with elevated symptoms and needs.

Internet-delivered medical rehabilitation components integrated into face-to-face therapy have the option to accelerate mental health improvements due to rehabilitation, which is especially important in times of limited treatment capacities and the need to reduce the transmission of viruses (ie, physical contact between treatment providers and patients). Internet-delivered medical treatment can bridge the gap and can also help patients to cope with a potential aftermath of the COVID-19 pandemic in terms of more patients in need of care than available resources at the patients' residence.

Acknowledgments

We would like to thank the Dr Becker clinics Möhnesee, Norddeich, Juliana, and Burg for their assistance in data collection. Furthermore, we would like to thank Jing Liu and Nerma Pasic for proofreading the final version of this manuscript. This work was supported by the German Innovation Fund (Project No. 01VSF18023) of The Federal Joint Committee (G-BA) as part of the research project "TeamBaby – Safe, digitally supported communication in obstetrics and gynecology" (grant 01VSF18023). The open access fee was provided by the Konfuzius Institute Bremen/Hanban. The Konfuzius Institute Bremen/Hanban had no role in the study design, study implementation, data collection, data analysis, manuscript preparation, or publication decision. The work is the responsibility of the authors.

Authors' Contributions

FMK conceived the research questions, designed the study, and set up the data collection survey tool. FMK, AD, and SL conducted the statistical analysis and drafted the manuscript. FMK, AD, CD, LK, and SL finalized the manuscript. All authors assisted in the questionnaire design, supported the data collection and interpretation, and gave comments with regard to the interpretation of the findings of the manuscript. FMK, AD, CD, LK, and SL had full access to all data from the study and were responsible for the integrity of the data. FMK was responsible for the accuracy of the data analysis. SL acquired funding for TeamBaby and AD acquired funding for Anhand-COVID19.

Conflicts of Interest

None declared.

References

1. Shigemura J, Ursano RJ, Morganstein JC, Kurosawa M, Benedek DM. Public responses to the novel 2019 coronavirus (2019-nCoV) in Japan: Mental health consequences and target populations. Psychiatry Clin Neurosci 2020 Apr;74(4):281-282 [FREE Full text] [doi: 10.1111/pcn.12988] [Medline: 32034840]
2. Balkhi F, Nasir A, Zehra A, Riaz R. Psychological and behavioral response to the coronavirus (COVID-19) pandemic. Cureus 2020 May 02;12(5):e7923 [FREE Full text] [doi: 10.7759/cureus.7923] [Medline: 32499970]
3. Liu X, Luo W, Li Y, Li C, Hong Z, Chen H, et al. Psychological status and behavior changes of the public during the COVID-19 epidemic in China. Infect Dis Poverty 2020 May 29;9(1):58 [FREE Full text] [doi: 10.1186/s40249-020-00678-3] [Medline: 32471513]
4. Fiorillo A, Gorwood P. The consequences of the COVID-19 pandemic on mental health and implications for clinical practice. Eur Psychiatry 2020 Apr 01;63(1):e32 [FREE Full text] [doi: 10.1192/j.eurpsy.2020.35] [Medline: 32234102]
5. Wang C, Pan R, Wan X, Tan Y, Xu L, Ho CS, et al. Immediate psychological responses and associated factors during the initial stage of the 2019 coronavirus disease (COVID-19) epidemic among the general population in China. Int J Environ Res Public Health 2020 Mar 06;17(5):1729 [FREE Full text] [doi: 10.3390/ijerph17051729] [Medline: 32155789]
6. Sønderskov KM, Dinesen PT, Santini ZI, Østergaard SD. The depressive state of Denmark during the COVID-19 pandemic. Acta Neuropsychiatr 2020 Aug;32(4):226-228 [FREE Full text] [doi: 10.1017/neu.2020.15] [Medline: 32319879]
7. Ettman CK, Abdalla SM, Cohen GH, Sampson L, Vivier PM, Galea S. Prevalence of depression symptoms in US adults before and during the COVID-19 pandemic. JAMA Netw Open 2020 Sep 01;3(9):e2019686 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.19686] [Medline: 32876685]

8. Twenge J, Joiner TE. US Census Bureau-assessed prevalence of anxiety and depressive symptoms in 2019 and during the 2020 COVID-19 pandemic. Depress Anxiety 2020 Oct;37(10):954-956 [FREE Full text] [doi: 10.1002/da.23077] [Medline: 32667081]
9. Ruffolo M, Price D, Schoultz M, Leung J, Bonsaksen T, Thygesen H, et al. Employment uncertainty and mental health during the COVID-19 pandemic initial social distancing implementation: A cross-national study. Glob Soc Welf 2021 Jan 07:1-10 [FREE Full text] [doi: 10.1007/s40609-020-00201-4] [Medline: 33432284]
10. Tan W, Hao F, McIntyre RS, Jiang L, Jiang X, Zhang L, et al. Is returning to work during the COVID-19 pandemic stressful? A study on immediate mental health status and psychoneuroimmunity prevention measures of Chinese workforce. Brain Behav Immun 2020 Jul;87:84-92 [FREE Full text] [doi: 10.1016/j.bbi.2020.04.055] [Medline: 32335200]
11. Chatterjee SS, Barikar C, Mukherjee A. Impact of COVID-19 pandemic on pre-existing mental health problems. Asian J Psychiatr 2020 Jun;51:102071 [FREE Full text] [doi: 10.1016/j.ajp.2020.102071] [Medline: 32334407]
12. Mental health and psychosocial considerations during the COVID-19 outbreak. World Health Organization. 2020 Mar 18. URL: https://apps.who.int/iris/bitstream/handle/10665/331490/WHO-2019-nCoV-MentalHealth-2020.1-eng.pdf [accessed 2021-06-18]
13. Pierce M, Hope H, Ford T, Hatch S, Hotopf M, John A, et al. Mental health before and during the COVID-19 pandemic: A longitudinal probability sample survey of the UK population. Lancet Psychiatry 2020 Oct;7(10):883-892 [FREE Full text] [doi: 10.1016/S2215-0366(20)30308-4] [Medline: 32707037]
14. Reynolds DL, Garay JR, Deamond SL, Moran MK, Gold W, Styra R. Understanding, compliance and psychological impact of the SARS quarantine experience. Epidemiol Infect 2008 Jul;136(7):997-1007. [doi: 10.1017/S0950268807009156] [Medline: 17662167]
15. Liu PL. COVID-19 information seeking on digital media and preventive behaviors: The mediation role of worry. Cyberpsychol Behav Soc Netw 2020 Oct;23(10):677-682. [doi: 10.1089/cyber.2020.0250] [Medline: 32498549]
16. Borkovec TD, Costello E. Efficacy of applied relaxation and cognitive-behavioral therapy in the treatment of generalized anxiety disorder. J Consult Clin Psychol 1993 Aug;61(4):611-619. [doi: 10.1037//0022-006x.61.4.611] [Medline: 8370856]
17. Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition. Arlington, VA: American Psychiatric Association; 2013.
18. Watkins E, Moulds M, Mackintosh B. Comparisons between rumination and worry in a non-clinical population. Behav Res Ther 2005 Dec;43(12):1577-1585. [doi: 10.1016/j.brat.2004.11.008] [Medline: 16239152]
19. Barber SJ, Kim H. COVID-19 worries and behavior changes in older and younger men and women. J Gerontol B Psychol Sci Soc Sci 2021 Jan 18;76(2):e17-e23 [FREE Full text] [doi: 10.1093/geronb/gbaa068] [Medline: 32427341]
20. Van Rheenen TE, Meyer D, Neill E, Phillipou A, Tan EJ, Toh WL, et al. Mental health status of individuals with a mood-disorder during the COVID-19 pandemic in Australia: Initial results from the COLLATE project. J Affect Disord 2020 Oct 01;275:69-77 [FREE Full text] [doi: 10.1016/j.jad.2020.06.037] [Medline: 32658826]
21. Dahmen A, Keller FM, Kötting L, Derksen C, Lippke S. Fear of coronavirus, intention to follow the AHA rules and risk perception regarding visits to the doctor: Cross-sectional study with patients suffering from pre-existing mental illness [Article in German]. Gesundheitswesen 2021 Apr;83(4):274-281 [FREE Full text] [doi: 10.1055/a-1397-7214] [Medline: 33831971]
22. Drwal KR, Forman DE, Wakefield BJ, El Accaoui RN. Cardiac rehabilitation during COVID-19 pandemic: Highlighting the value of home-based programs. Telemed J E Health 2020 Nov;26(11):1322-1324. [doi: 10.1089/tmj.2020.0213] [Medline: 32552412]
23. Hau YS, Kim JK, Hur J, Chang MC. How about actively using telemedicine during the COVID-19 pandemic? J Med Syst 2020 Apr 30;44(6):108 [FREE Full text] [doi: 10.1007/s10916-020-01580-z] [Medline: 32350626]
24. Bryant MS, Fedson SE, Sharafkhaneh A. Using telehealth cardiopulmonary rehabilitation during the COVID-19 pandemic. J Med Syst 2020 May 28;44(7):125 [FREE Full text] [doi: 10.1007/s10916-020-01593-8] [Medline: 32462352]
25. Scheidt C. Psychosomatic medicine in Germany. Int J Body Mind Cult 2017;4(2):78-86. [doi: 10.22122/ijbmc.v4i2.95]
26. Becker J, Beutel ME, Gerzymisch K, Schulz D, Siepmann M, Knickenberg RJ, et al. Evaluation of a video-based internet intervention as preparation for inpatient psychosomatic rehabilitation: Study protocol for a randomized controlled trial. Trials 2016 Jun 13;17(1):287 [FREE Full text] [doi: 10.1186/s13063-016-1417-y] [Medline: 27296249]
27. Zwerenz R, Becker J, Knickenberg RJ, Hagen K, Dreier M, Wölfling K, et al. Enhancing inpatient psychotherapeutic treatment with online self-help: Study protocol for a randomized controlled trial. Trials 2015 Mar 17;16:98 [FREE Full text] [doi: 10.1186/s13063-015-0620-6] [Medline: 25872505]
28. Hennemann S, Farnsteiner S, Sander L. Internet- and mobile-based aftercare and relapse prevention in mental disorders: A systematic review and recommendations for future research. Internet Interv 2018 Dec;14:1-17 [FREE Full text] [doi: 10.1016/j.invent.2018.09.001] [Medline: 30510909]
29. Zwerenz R, Gerzymisch K, Edinger J, Holme M, Knickenberg RJ, Spörl-Dönch S, et al. Evaluation of an internet-based aftercare program to improve vocational reintegration after inpatient medical rehabilitation: Study protocol for a cluster-randomized controlled trial. Trials 2013 Jan 25;14:26 [FREE Full text] [doi: 10.1186/1745-6215-14-26] [Medline: 23351836]

30. Sztein DM, Koransky CE, Fegan L, Himelhoch S. Efficacy of cognitive behavioural therapy delivered over the internet for depressive symptoms: A systematic review and meta-analysis. J Telemed Telecare 2018 Sep;24(8):527-539. [doi: 10.1177/1357633X17717402] [Medline: 28696153]
31. Alavi N, Yang M, Stephenson C, Nikjoo N, Malakouti N, Layzell G, et al. Using the Online Psychotherapy Tool to address mental health problems in the context of the COVID-19 pandemic: Protocol for an electronically delivered cognitive behavioral therapy program. JMIR Res Protoc 2020 Dec 18;9(12):e24913 [FREE Full text] [doi: 10.2196/24913] [Medline: 33290245]
32. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep;13(3):319. [doi: 10.2307/249008]
33. Schwarzer R. Modeling health behavior change: How to predict and modify the adoption and maintenance of health behaviors. Appl Psychol 2008 Jan;57(1):1-29. [doi: 10.1111/j.1464-0597.2007.00325.x]
34. Schwarzer R, Lippke S, Luszczynska A. Mechanisms of health behavior change in persons with chronic illness or disability: The Health Action Process Approach (HAPA). Rehabil Psychol 2011 Aug;56(3):161-170. [doi: 10.1037/a0024509] [Medline: 21767036]
35. Schmädeke S, Kockler T, Olbich D. DE-RENA rehabilitation follow-up with a smartphone app for depressed patient outcomes on acceptability and effectiveness [Document in German]. In: Proceedings of the 28th Rehabilitation Colloquium. 2019 Presented at: 28th Rehabilitation Colloquium; April 15-17, 2019; Berlin, Germany p. 383.
36. Kuhl J. Motivation and Personality: Interaction of Mental Systems [Book in German]. Göttingen, Germany: Hogrefe Verl für Psychologie; 2001.
37. Luo C, Sanger N, Singhal N, Pattrick K, Shams I, Shahid H, et al. A comparison of electronically-delivered and face to face cognitive behavioural therapies in depressive disorders: A systematic review and meta-analysis. EClinicalMedicine 2020 Jul;24:100442 [FREE Full text] [doi: 10.1016/j.eclinm.2020.100442] [Medline: 32775969]
38. Karyotaki E, Kemmeren L, Riper H, Twisk J, Hoogendoorn A, Kleiboer A, et al. Is self-guided internet-based cognitive behavioural therapy (iCBT) harmful? An individual participant data meta-analysis. Psychol Med 2018 Nov;48(15):2456-2466 [FREE Full text] [doi: 10.1017/S0033291718000648] [Medline: 29540243]
39. Sasseville M, LeBlanc A, Boucher M, Dugas M, Mbemba G, Tchuente J, et al. Digital health interventions for the management of mental health in people with chronic diseases: A rapid review. BMJ Open 2021 Apr 05;11(4):e044437 [FREE Full text] [doi: 10.1136/bmjopen-2020-044437] [Medline: 33820786]
40. Varga TV, Bu F, Dissing AS, Elsenburg LK, Bustamante JJH, Matta J, et al. Loneliness, worries, anxiety, and precautionary behaviours in response to the COVID-19 pandemic: A longitudinal analysis of 200,000 Western and Northern Europeans. Lancet Reg Health Eur 2021 Mar;2:100020 [FREE Full text] [doi: 10.1016/j.lanepe.2020.100020] [Medline: 33870246]
41. Kroenke K, Spitzer RL, Williams JBW. The Patient Health Questionnaire-2: Validity of a two-item depression screener. Med Care 2003 Nov;41(11):1284-1292. [doi: 10.1097/01.MLR.0000093487.78664.3C] [Medline: 14583691]
42. Kroenke K, Spitzer RL, Williams JBW, Monahan PO, Löwe B. Anxiety disorders in primary care: Prevalence, impairment, comorbidity, and detection. Ann Intern Med 2007 Mar 06;146(5):317-325. [doi: 10.7326/0003-4819-146-5-200703060-00004] [Medline: 17339617]
43. Löwe B, Wahl I, Rose M, Spitzer C, Glaesmer H, Wingenfeld K, et al. A 4-item measure of depression and anxiety: Validation and standardization of the Patient Health Questionnaire-4 (PHQ-4) in the general population. J Affect Disord 2010 Apr;122(1-2):86-95. [doi: 10.1016/j.jad.2009.06.019] [Medline: 19616305]
44. Löwe B, Kroenke K, Gräfe K. Detecting and monitoring depression with a two-item questionnaire (PHQ-2). J Psychosom Res 2005 Feb;58(2):163-171. [doi: 10.1016/j.jpsychores.2004.09.006] [Medline: 15820844]
45. Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. J Health Soc Behav 1983 Dec;24(4):385-396. [Medline: 6668417]
46. Cohen S, Williamson GM. Perceived stress in a probability sample of the United States. In: Spacapan S, Oskamp S, editors. The Social Psychology of Health. Newbury Park, CA: SAGE Publications, Inc; 1988:31-67.
47. Radloff LS. The CES-D Scale. Appl Psychol Meas 2016 Jul 26;1(3):385-401. [doi: 10.1177/014662167700100306]
48. Russell DW. UCLA Loneliness Scale (Version 3): Reliability, validity, and factor structure. J Pers Assess 1996 Feb;66(1):20-40. [doi: 10.1207/s15327752jpa6601_2] [Medline: 8576833]
49. Lippke S, Ziegelmann JP, Schwarzer R, Velicer WF. Validity of stage assessment in the adoption and maintenance of physical activity and fruit and vegetable consumption. Health Psychol 2009 Mar;28(2):183-193 [FREE Full text] [doi: 10.1037/a0012983] [Medline: 19290710]
50. Lippke S, Fleig L, Pomp S, Schwarzer R. Validity of a stage algorithm for physical activity in participants recruited from orthopedic and cardiac rehabilitation clinics. Rehabil Psychol 2010 Nov;55(4):398-408. [doi: 10.1037/a0021563] [Medline: 21171799]
51. D'Agostino RB. Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. Stat Med 1998 Oct 15;17(19):2265-2281. [doi: 10.1002/(sici)1097-0258(19981015)17:19<2265::aid-sim918>3.0.co;2-b] [Medline: 9802183]
52. Loux TM. Randomization, matching, and propensity scores in the design and analysis of experimental studies with measured baseline covariates. Stat Med 2015 Feb 20;34(4):558-570. [doi: 10.1002/sim.6361] [Medline: 25384851]

53. Chevance A, Gourion D, Hoertel N, Llorca P, Thomas P, Bocher R, et al. Ensuring mental health care during the SARS-CoV-2 epidemic in France: A narrative review. Encephale 2020 Jun;46(3):193-201 [FREE Full text] [doi: 10.1016/j.encep.2020.04.005] [Medline: 32370982]
54. Hao F, Tan W, Jiang L, Zhang L, Zhao X, Zou Y, et al. Do psychiatric patients experience more psychiatric symptoms during COVID-19 pandemic and lockdown? A case-control study with service and research implications for immunopsychiatry. Brain Behav Immun 2020 Jul;87:100-106 [FREE Full text] [doi: 10.1016/j.bbi.2020.04.069] [Medline: 32353518]
55. Yao H, Chen J, Xu Y. Rethinking online mental health services in China during the COVID-19 epidemic. Asian J Psychiatr 2020 Apr;50:102015 [FREE Full text] [doi: 10.1016/j.ajp.2020.102015] [Medline: 32247261]
56. Budimir S, Probst T, Pieh C. Coping strategies and mental health during COVID-19 lockdown. J Ment Health 2021 Apr;30(2):156-163. [doi: 10.1080/09638237.2021.1875412] [Medline: 33502917]
57. Tsamakis K, Triantafyllis AS, Tsiptsios D, Spartalis E, Mueller C, Tsamakis C, et al. COVID-19 related stress exacerbates common physical and mental pathologies and affects treatment (Review). Exp Ther Med 2020 Jul;20(1):159-162 [FREE Full text] [doi: 10.3892/etm.2020.8671] [Medline: 32509006]
58. Topper M, Emmelkamp PM, Ehring T. Improving prevention of depression and anxiety disorders: Repetitive negative thinking as a promising target. Appl Prev Psychol 2010 Jun;14(1-4):57-71. [doi: 10.1016/j.appsy.2012.03.001]
59. Lippke S, Ricken L, Zschucke E, Hessel A, Schüz N. Subjective financial resources of temporary disability pensioners: Interrelation with loneliness and life satisfaction [Article in German]. Rehabilitation (Stuttg) 2020 Dec;59(6):341-347. [doi: 10.1055/a-1242-9408] [Medline: 33285613]
60. Brenke K, Ziemendorff J. In need of assistance despite work: Not a mass phenomenon in Germany [Article in German]. DIW Wochenbericht 2008;75(4):33-40.
61. Zielke M. Erwerbsminderungsrenten infolge psychischer Erkrankungen. Psychotherapeut 2014 Jun 26;59(4):323-328. [doi: 10.1007/s00278-014-1060-6]
62. Guo Y, Cheng C, Zeng Y, Li Y, Zhu M, Yang W, et al. Mental health disorders and associated risk factors in quarantined adults during the COVID-19 outbreak in China: Cross-sectional study. J Med Internet Res 2020 Aug 06;22(8):e20328 [FREE Full text] [doi: 10.2196/20328] [Medline: 32716899]
63. Haaf HG. Findings on the effectiveness of rehabilitation [Article in German]. Rehabilitation (Stuttg) 2005 Oct;44(5):259-276. [doi: 10.1055/s-2005-867015] [Medline: 16208590]
64. Davies F, Shepherd HL, Beatty L, Clark B, Butow P, Shaw J. Implementing web-based therapy in routine mental health care: Systematic review of health professionals' perspectives. J Med Internet Res 2020 Jul 23;22(7):e17362 [FREE Full text] [doi: 10.2196/17362] [Medline: 32706713]
65. Davies EB, Morriss R, Glazebrook C. Computer-delivered and web-based interventions to improve depression, anxiety, and psychological well-being of university students: A systematic review and meta-analysis. J Med Internet Res 2014 May 16;16(5):e130 [FREE Full text] [doi: 10.2196/jmir.3142] [Medline: 24836465]
66. Huckvale K, Nicholas J, Torous J, Larsen ME. Smartphone apps for the treatment of mental health conditions: Status and considerations. Curr Opin Psychol 2020 Dec;36:65-70 [FREE Full text] [doi: 10.1016/j.copsyc.2020.04.008] [Medline: 32553848]
67. Joiner TE, Metalsky GI, Katz J, Beach SRH. Depression and excessive reassurance-seeking. Psychol Inq 2009 Nov 19;10(3):269-278. [doi: 10.1207/s15327965pli1004_1]
68. Parrish CL, Radomsky AS. Why do people seek reassurance and check repeatedly? An investigation of factors involved in compulsive behavior in OCD and depression. J Anxiety Disord 2010 Mar;24(2):211-222. [doi: 10.1016/j.janxdis.2009.10.010] [Medline: 19939622]
69. Shafran R, Gyani A, Rostron J, Allen S, Myles-Hooton P, Allcott-Watson H, et al. Translating the intention to seek treatment into action: Does symptom monitoring make a difference? Results from a randomized controlled trial. Behav Cogn Psychother 2018 Aug 23;47(1):114-128. [doi: 10.1017/s1352465818000498]
70. Rogers MA, Lemmen K, Kramer R, Mann J, Chopra V. Internet-delivered health interventions that work: Systematic review of meta-analyses and evaluation of website availability. J Med Internet Res 2017 Mar 24;19(3):e90 [FREE Full text] [doi: 10.2196/jmir.7111] [Medline: 28341617]
71. Sin J, Galeazzi G, McGregor E, Collom J, Taylor A, Barrett B, et al. Digital interventions for screening and treating common mental disorders or symptoms of common mental illness in adults: Systematic review and meta-analysis. J Med Internet Res 2020 Sep 02;22(9):e20581 [FREE Full text] [doi: 10.2196/20581] [Medline: 32876577]
72. Torous J, Jän Myrick K, Rauseo-Ricupero N, Firth J. Digital mental health and COVID-19: Using technology today to accelerate the curve on access and quality tomorrow. JMIR Ment Health 2020 Mar 26;7(3):e18848 [FREE Full text] [doi: 10.2196/18848] [Medline: 32213476]
73. Dong L, Bouey J. Public mental health crisis during COVID-19 pandemic, China. Emerg Infect Dis 2020 Jul;26(7):1616-1618 [FREE Full text] [doi: 10.3201/eid2607.200407] [Medline: 32202993]
74. Nelson BW, Pettitt A, Flannery JE, Allen NB. Rapid assessment of psychological and epidemiological correlates of COVID-19 concern, financial strain, and health-related behavior change in a large online sample. PLoS One 2020;15(11):e0241990 [FREE Full text] [doi: 10.1371/journal.pone.0241990] [Medline: 33175882]

Abbreviations

CBT: cognitive behavioral therapy
eCBT: electronically delivered cognitive behavioral therapy
EFA: exploratory factor analysis
GAD-2: 2-item Generalized Anxiety Disorder scale
HAPA: Health Action Process Approach
mHealth: mobile health
PHQ-2: 2-item Patient Health Questionnaire
PHQ-4: 4-item Patient Health Questionnaire
PSM: propensity score matching
PSS: Perceived Stress Scale
PSS-4: 4-item Perceived Stress Scale
SARS: severe acute respiratory syndrome
T1: first time point (6 weeks before the start until the first day of rehabilitation)
T2: second time point (after rehabilitation)
UCLA: University of California, Los Angeles
WHO: World Health Organization

Edited by G Eysenbach; submitted 21.05.21; peer-reviewed by J Apolinário-Hagen; comments to author 14.06.21; revised version received 18.06.21; accepted 02.07.21; published 26.08.21.

Please cite as:
Keller FM, Dahmen A, Derksen C, Kötting L, Lippke S
Psychosomatic Rehabilitation Patients and the General Population During COVID-19: Online Cross-sectional and Longitudinal Study of Digital Trainings and Rehabilitation Effects
JMIR Ment Health 2021;8(8):e30610
URL: https://mental.jmir.org/2021/8/e30610
doi:10.2196/30610
PMID:34270444

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org