

JMIR Mental Health

Internet interventions, technologies and digital innovations for mental health and behaviour change
Volume 8 (2021), Issue 5 ISSN: 2368-7959 Editor in Chief: John Torous, MD

Contents

Original Papers

The Use of Closed-Circuit Television and Video in Suicide Prevention: Narrative Review and Future Directions (e27663) Sandersan Onie, Xun Li, Morgan Liang, Arcot Sowmya, Mark Larsen.	3
Assessing Suicide Reporting in Top Newspaper Social Media Accounts in China: Content Analysis Study (e26654) Kaisheng Lai, Dan Li, Huijuan Peng, Jingyuan Zhao, Lingnan He.	20
Intention to Use Behavioral Health Data From a Health Information Exchange: Mixed Methods Study (e26746) Randyl Cochran, Sue Feldman, Nataliya Ivankova, Allyson Hall, William Opoku-Agyeman.	33
Learning From Clinical Consensus Diagnosis in India to Facilitate Automatic Classification of Dementia: Machine Learning Study (e27113) Haomiao Jin, Sandy Chien, Erik Meijer, Pranali Khobragade, Jinkook Lee.	63
A Web-Based Group Cognitive Behavioral Therapy Intervention for Symptoms of Anxiety and Depression Among University Students: Open-Label, Pragmatic Trial (e27400) Jason Bantjes, Alan Kazdin, Pim Cuijpers, Elsie Breet, Munita Dunn-Coetzee, Charl Davids, Dan Stein, Ronald Kessler.	75
Social Representations of e-Mental Health Among the Actors of the Health Care System: Free-Association Study (e25708) Margot Morgiève, Pierre Mesdjian, Olivier Las Vergnas, Patrick Bury, Vincent Demassiet, Jean-Luc Roelandt, Déborah Sebbane.	91
Problematic Social Media Use in Sexual and Gender Minority Young Adults: Observational Study (e23688) Erin Vogel, Danielle Ramo, Judith Prochaska, Meredith Meacham, John Layton, Gary Humfleet.	103
Knowledge-Infused Abstractive Summarization of Clinical Diagnostic Interviews: Framework Development Study (e20865) Gaur Manas, Vamsi Aribandi, Ugur Kursuncu, Amanuel Alambo, Valerie Shalin, Krishnaprasad Thirunarayan, Jonathan Beich, Meera Narasimhan, Amit Sheth.	115

Review

Initial Training for Mental Health Peer Support Workers: Systematized Review and International Delphi Consultation ([e25528](#))

Ashleigh Charles, Rebecca Nixdorf, Nashwa Ibrahim, Lion Meir, Richard Mpango, Fileuka Ngakongwa, Hannah Nudds, Soumitra Pathare, Grace Ryan, Julie Repper, Heather Wharrad, Philip Wolf, Mike Slade, Candelaria Mahlke. 47

Original Paper

The Use of Closed-Circuit Television and Video in Suicide Prevention: Narrative Review and Future Directions

Sandersan Onie¹, PhD; Xun Li², PhD; Morgan Liang²; Arcot Sowmya², PhD; Mark Erik Larsen¹, PhD

¹Black Dog Institute, University of New South Wales, Sydney, Sydney, Australia

²School of Computer Science and Engineering, University of New South Wales, Sydney, Sydney, Australia

Corresponding Author:

Sandersan Onie, PhD

Black Dog Institute

University of New South Wales, Sydney

Hospital Road

Sydney, 2031

Australia

Phone: 61 432359134

Email: sandy.onie@gmail.com

Abstract

Background: Suicide is a recognized public health issue, with approximately 800,000 people dying by suicide each year. Among the different technologies used in suicide research, closed-circuit television (CCTV) and video have been used for a wide array of applications, including assessing crisis behaviors at metro stations, and using computer vision to identify a suicide attempt in progress. However, there has been no review of suicide research and interventions using CCTV and video.

Objective: The objective of this study was to review the literature to understand how CCTV and video data have been used in understanding and preventing suicide. Furthermore, to more fully capture progress in the field, we report on an ongoing study to respond to an identified gap in the narrative review, by using a computer vision-based system to identify behaviors prior to a suicide attempt.

Methods: We conducted a search using the keywords “suicide,” “cctv,” and “video” on PubMed, Inspec, and Web of Science. We included any studies which used CCTV or video footage to understand or prevent suicide. If a study fell into our area of interest, we included it regardless of the quality as our goal was to understand the scope of how CCTV and video had been used rather than quantify any specific effect size, but we noted the shortcomings in their design and analyses when discussing the studies.

Results: The review found that CCTV and video have primarily been used in 3 ways: (1) to identify risk factors for suicide (eg, inferring depression from facial expressions), (2) understanding suicide after an attempt (eg, forensic applications), and (3) as part of an intervention (eg, using computer vision and automated systems to identify if a suicide attempt is in progress). Furthermore, work in progress demonstrates how we can identify behaviors prior to an attempt at a hotspot, an important gap identified by papers in the literature.

Conclusions: Thus far, CCTV and video have been used in a wide array of applications, most notably in designing automated detection systems, with the field heading toward an automated detection system for early intervention. Despite many challenges, we show promising progress in developing an automated detection system for preattempt behaviors, which may allow for early intervention.

(*JMIR Ment Health* 2021;8(5):e27663) doi:[10.2196/27663](https://doi.org/10.2196/27663)

KEYWORDS

suicide; suicide prevention; CCTV; video; computer vision; machine learning

Introduction

Suicide is a recognized public health priority, with approximately 800,000 people dying by suicide each year [1].

The use of different means of suicide can vary by geographic region, with public means such as jumping from a height being more common in certain areas [2]. Data from reviews of coronial records in the UK suggest approximately 30% of all suicide

deaths occur in public places [3]. Suicides in public places can attract media attention, potentially introducing a degree of “notoriety” about a specific location, and thus increase its frequency of use [4]. Furthermore, these incidents can adversely affect bystanders [5]. Thus, efforts to prevent suicide in public places are of particular importance.

A range of suicide prevention initiatives in public places have been evaluated, with the use of CCTV cameras being proposed as a means to increase the likelihood of a third-party intervention [6], for example, by initiating a police call-out after climbing a safety fence [7]. Studies from the metro/underground railway settings have retrospectively analyzed behaviors which precede a suicide attempt, with observable behaviors being identified [8,9]. These studies have further suggested that it may be possible to automatically detect behaviors prior to a suicide attempt, potentially allowing for earlier intervention and the interruption of a suicide attempt.

This paper builds on these foundations by addressing 2 core aims. First, it provides a scoping review on the existing use of CCTV and other video data to understand and prevent suicide. Second, it reports progress on an ongoing study to respond to an identified gap in this literature, by using a computer vision-based system to identify behaviors prior to a suicide attempt.

Methods

Searches of the literature were performed using the PubMed, Inspec, and Web of Science databases, as pilot searches indicated that these databases provided unique literature from medical and computer science publications. Searches were performed using the terms “(video OR cctv) AND suicid*” on all fields. Searches were performed on August 26, 2020, and included all entries from conception to that date.

We sought to understand the broad use of video, and therefore included studies which (1) related to understanding suicide or suicide prevention; (2) described the use, analysis, or consideration of video footage or CCTV; and (3) related to the understanding of risk factors for suicide, used video monitoring to detect suicide or suicide risk, or used video data pertaining

to individuals. We excluded studies of video gaming (with no use of video data), video-based interventions (ie, showing a video to a person or group, or portrayal of suicide in video), or video-based telehealth interventions, suicide bombing, music videos, video vignettes, nonhuman subjects, or EEG-CCTV. Furthermore, we excluded any commentaries and opinion pieces without data or studies described, or studies which were not reported in English.

Abstracts were independently screened for eligibility by 2 reviewers (SO and ML), with disagreements resolved by discussion until consensus was achieved. Eligible papers were then downloaded for full-text review, and 1 reviewer (SO) extracted a narrative summary of the reported use of video data.

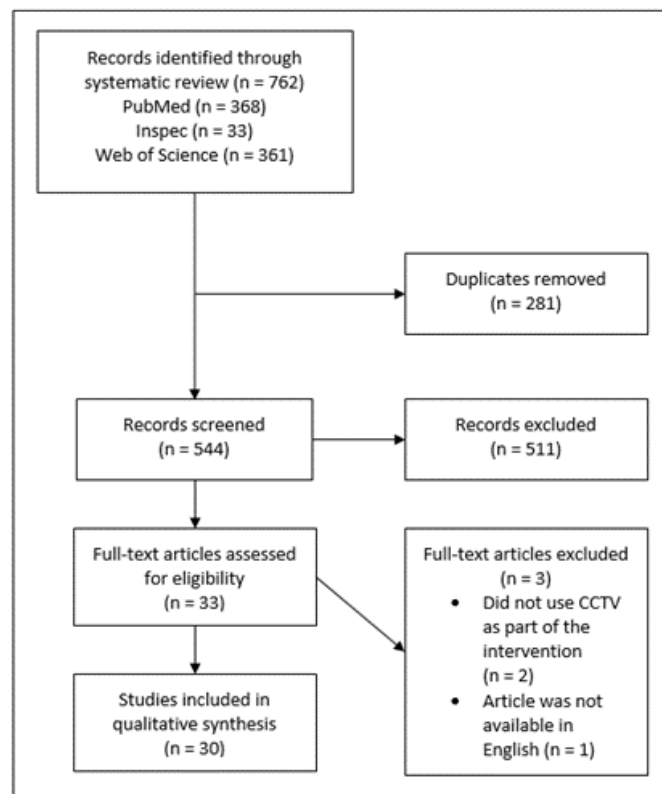
Results

Database Search and Review of Articles Retrieved

Database searches identified 544 unique articles to screen, as indicated in Figure 1. Following review of the title and abstracts, 33 articles were downloaded for full-text review. Two studies were excluded as they did not report relevant use of video or CCTV data, and 1 was excluded as it was not available in English. The remaining 30 articles were retained for the narrative synthesis.

In the review, we identified 3 main categories of papers corresponding to their temporal relationship with suicidal behavior. These are characterized as (1) studies that use CCTV and video to understand risk factors for suicide (6 papers); (2) studies that describe and evaluate interventions using CCTV or video footage (16 papers); and (3) studies that use CCTV or video footage to understand suicide after an attempt (8 papers). In this section, we discuss the findings of the review in these categories; however, we discuss the studies within the second category, that is, using CCTV or video as an intervention, last due to its length.

As many of the studies described below use machine learning and computer vision techniques to obtain desired results, for example, predicting depression from facial cues, a short introduction to relevant computer vision techniques is provided below.

Figure 1. Flowchart of Study Selection.

Computer Vision Overview

In this section we describe some basic concepts in computer vision to aid understanding of the described studies. Computer vision is a subfield of artificial intelligence that trains computers to interpret and develop a high-level understanding of the visual world. It seeks to replicate the workings of human perception by extracting semantic information from image data. It aims to automate the time-consuming human processes, analyzing a much larger volume of data with real-time efficiency. Two examples of its application discussed in this paper include pedestrian tracking, in which the algorithm is able to identify people within a scene and interpret their behavior, as well as facial analysis, in which the algorithm detects key points on a face and attempts to infer emotional states from the configuration of these points.

Applications of computer vision techniques in analyzing video footage to understand human behaviors in a natural environment (such as outdoor scenes captured by CCTV cameras) have typically followed a processing pipeline involving object detection, object tracking, and action recognition. Object detection refers to the ability to recognize and locate objects of interest in each frame of a video [10]. Object tracking is the process of locating the object of interest in multiple sequential frames over time [11]. Action recognition aims to identify the actions and goals of the objects of interest from a series of observations [12]. There are numerous algorithms and methods available to achieve a single goal, introducing the need for experimentation, which involves applying a variety of different methods to the same data and measuring the outputs against relevant performance metrics.

Paramount to the success of the aforementioned pipeline is the ability to gather a considerable amount of relevant data for a given task. This relates to the fact that these algorithms can be “taught” the required task by training them with annotated data, which is denoted as supervised learning in the field of machine learning. The data are annotated with labels that denote the location of the object of interest within the image and other semantic information about the object in the video frame. These annotations also help in adapting the models to the characteristics of the CCTV data such as camera environment, camera angle, and key object features. Overall, this increases the accuracy of the algorithm on the task at hand, for example, identification, tracking as well as action recognition.

In the following sections, we note specific names of algorithms for thoroughness. However, we do not go into their technical details, and the names are mentioned alongside their functionality.

Understanding Suicide Risk Factors

As described above, the first category of studies identified in this review related to the use of video and CCTV data to understand suicide risk factors. Within this category, 2 main types of studies were identified. The first analyzed videos of individuals who died by suicide, in order to understand risk factors. In one study, researchers searched video footage of suicides posted on Facebook in Bangladesh to understand the circumstances that may have contributed to the suicides [13]. The authors found 19 videos, predominantly of male students. Hanging was the most frequently used means, with relationship problems, academic stress, and mental health disorders being commonly identified causes.

The second type of study attempted to infer depression from facial cues using both human and computer vision. For example, in one study [14], the authors attempted to assess whether humans are able to perceive depression in another person at both a conscious and a physiological level. Participants viewed videos of individuals with depression while various physiological responses of the viewer were recorded (eg, galvanic skin response, temperature, and heart rate). The study found that despite not being able to consciously determine whether the person they were viewing had depression, the viewer's physiological response differed as a function of that person's depressive state.

Beyond manually assessing footage, automatic video analysis from the field of computer vision and machine learning has been used. Many studies in this category attempt to identify depression from facial cues [15-18]. For example, in one study [17] participants were shown a sad video, a neutral video, and a text to read and were interviewed while their facial expressions were analyzed using a Microsoft Kinect camera. The study found that when watching a neutral video, and using information primarily from the eyebrows and mouth, it was possible to predict the presence of depression above chance level (86.8% for females and 79.4% for males).

While different studies adopted different methodological approaches, the critical difference between the reported studies is in the experimentation protocols adopted when applying various computer vision algorithms and approaches to best achieve the desired outcome. For example, to infer depression from facial expressions, Girard et al [19] used video to locate 66 facial points and extract facial features surrounding those points, converting the visual stimuli into features. After extracting this information, the authors used a dimension-reduction approach to reduce the amount of data processed. Following that, a support vector machine with a radial basis function kernel classifier (a machine learning algorithm that classifies data into categories) was trained using the data, which was then able to detect depression comparable to manual coding of facial features present in individuals with depression. In another approach, Scherer et al [20] used MultiSense, a system built upon a combination of a variety of head and facial tracking algorithms, to analyze the following behaviors: head gaze (orientation), eye gaze, smile intensity, and smile duration. Features such as gaze direction, duration, and intensity of smiles were measured and their correlation with distress, depression, anxiety, and posttraumatic stress disorder analyzed. Therefore, different algorithms and approaches can be used for the same goal, with differing rates of success.

Understanding Suicide After an Attempt

In the third category of using videos and CCTV to understand suicide after an attempt, all the studies related to forensic examinations. As a result of this, many of these studies are quite graphic in nature, and while we attempt to describe these studies as simply as possible, there may be text that readers may find disturbing. If this is the case, please skip over this to the next section.

Most videos sought to understand the sequence of bodily reactions from the suicide attempt to eventual death, called the

agonal sequence. Understanding these sequences can help forensic scientists to retrospectively determine details surrounding an individual's death. By far, the most commonly reported means of suicide was hanging, in which researchers viewed recorded hangings and analyzed the process by which the individual dies [21-24]. All the studies noted that despite hanging being the most common method of suicide, there is little known about the sequence of events leading to eventual death. In all studies, the researchers viewed the hangings and noted how long it took for certain events to occur (eg, loss of consciousness, convulsions, rigidities).

In addition to death by hanging, some studies examined other cases of suicide to understand details of interest. For example, in one study, forensic experts studied a case of self-immolation [25]. In another, researchers studied the death of an individual at a Swiss right-to-die organization by oxygen deprivation [26]. The organization typically used barbiturates, and the researchers were assessing the feasibility of using oxygen deprivation, noting ways to improve the process.

Two forensic studies highlighted the importance of CCTV and video footage in determining whether a death was due to homicide or suicide. In one study, an individual presented with stab wounds to the neck, which was initially assessed as a homicide. However, the authors noted that due to the presence of CCTV, they were able to conclude that the individual had died by suicide [25]. In another study, an individual was found to have died due to impact by a motor vehicle [26]. While the individual presented with self-harm scars and prior substance use, the CCTV of a neighboring business was able to definitively conclude that this was in fact a suicide. This is incredibly important as certain deaths are difficult to rule as suicide, and thus CCTV can enable us to determine them as such.

Describing and Evaluating Interventions Using CCTV or Video Footage

In the second category of studies, researchers sought to develop new ways to utilize CCTV and monitoring to identify persons in crisis or to allow an intervention during an attempt. This has been identified as 1 of 3 key interventions outlined in a meta-analysis of suicide prevention interventions at hotspots (ie, increasing the possibility of an intervention by a third party) [6]. This often takes the form of CCTV monitoring with or without an accompanying automated alert system or analyzing past CCTV footage to understand crisis behaviors.

Studies have shown that having human-monitored CCTV cameras at a site may reduce the incidence of suicide. For example, in prisons where hangings occur, a study compared the incidence of various behaviors in areas covered and not covered by CCTV and found that there were fewer incidents of suicides in monitored areas [27]. However, this conclusion is undermined by the fact that hangings may most likely occur in areas that do not have CCTV cameras (eg, cells). In another study, the authors investigated the effect of CCTV at metro stations in the state of Victoria in Australia [28]. The authors found that CCTV was able to reduce the incidence of suicides. It is not known whether this is due to individuals being more wary in the presence of cameras or the station staff being able to intervene earlier. Finally, another study investigated the

efficacy of a virtual monitoring system for patients with psychiatric disorders in a general hospital setting, in which 1 person was able to monitor up to 10 rooms with patients. While the study results were not conclusive, it appeared that the system allowed staff to monitor the safety of more patients at the same time [29].

One key factor associated with the systems described above is that they rely heavily on manual monitoring. Therefore, automated or semiautomated systems have been developed and tested. For example, a study examined a bridge in Seoul which used manual CCTV monitoring along with infrared motion sensors [30]. The sensors would detect when a person climbs over the fencing or jumps into the water below, alerting the guards stationed close by.

Another similar approach to automated monitoring is to draw regions of interest in a static CCTV scene. For example, Mukherjee and Ghosh [31] described an automated detection system in a metro station. The system detects movement within a predefined region of the video footage. Any movement in that region beyond the arrival of trains would send an alert to station staff. Similarly, 2 other studies described a similar approach at a cliff located in Sydney, Australia [5,7]. Unlike the metro station, this outdoor site was much larger and with extremely uneven geography, leading to less complete CCTV coverage. The region beyond the fence is specified as a region of interest and if movement is detected on the wrong side of the fence, an alarm is triggered at a monitoring station. While this approach of identifying static regions of interest within scenes has allowed the successful delivery of suicide prevention interventions, this method is only effective in locations where the means of suicide relied on individuals crossing over a boundary or a fence (eg, a cliff or metro station).

Another form of automated detection uses behavioral identification using automated vision methods to detect a hanging. Rather than using motion sensors or an outlined region of interest in the CCTV scene, these methods use computer vision analysis to detect a person in the scene, their limbs, as well as when the position and movement of those limbs indicate a suicide attempt by hanging [32–36]. In one study, the authors developed a video surveillance system to detect a suicide attempt by hanging using color cameras that can also detect depth (red, green, blue—depth cameras; RGB-D) to understand the three-dimensional positions of body joints [33]. Hand-crafted features are extracted based on the joint positions. More specifically, pair-wise joint distances within each frame and between 2 consecutive frames are used as feature vectors for classification. A linear discriminant classifier (LDC) was trained to classify 2 groups of actions: nonsuspicious actions (such as move, sit, wear clothes) and suicide by hanging. However, the training data set was only based on simulated video recordings taken in a room environment rather than genuine, unplanned events. Similar methods have also been adopted in other studies [20,36].

With the increased emergence of multipurpose intelligent computer vision systems, a paper proposed hierarchical evaluation metrics to evaluate such a system at metro stations, using three dimensions of performance [37]. The first dimension

is object detection, which is the system's ability to accurately detect an object of interest (eg, a person or a bicycle). This is important to prevent triggering false alarms if a nonhuman object is identified within a defined region of interest. The second dimension is object tracking, which is whether the object of interest is able to be tracked within the scene across time and different image frames. Within this object tracking, there are 3 factors: trajectory similarity, which identifies the path taken within the frame; ID changing, which is whether an object is consistently identified as the same object during tracking; and latency, which is how quickly the system can track the object. Finally, the third performance dimension is context awareness, which is how accurately the system is able to identify the situation (eg, a person currently in an unsafe location, such as walking on railway tracks).

However, even when using automated systems, the ability to intervene is limited to when an attempt is made or a very short period prior to an event, such as entering an unsafe location. As a previous study [33] notes, the next step forward is to find ways to detect presuicidal behaviors to reach out to people in crisis prior to an attempt, and a necessary step is to first and foremost identify these behaviors. By investigating predictive behaviors, a third party may be able to intervene earlier or more quickly, which could greatly reduce the incidence of suicide. Thus far, there have been only 3 studies investigating such predictive behaviors, including 1 study using videos from social media, and 2 studies at metro stations.

In the first study, the authors used videos of livestreamed suicides posted on social media and observed behaviors preceding the suicide [38]. In particular, the authors investigated 3 behavioral markers: verbal markers to investigate whether the individual talks more about self-harm and uses more negative language; acoustic markers to investigate whether pitch and pauses in speech are indicative of suicide; and visual behavior markers to investigate whether there are any visually identifiable markers (eg, frequently shifting eye gaze or pose changes). The analysis revealed that frequent silences, slouched shoulders, and use of profanity are indicative of elevated suicide risk.

In one study, over the course of 2 experiments, the authors attempted to identify behaviors indicative of a subsequent suicide attempt at metro stations, and tested how well these behaviors predict an attempt [8]. In the first experiment, the authors identified key behaviors by having 2–3 observers code both easily observable behaviors (eg, standing near the edge of the platform) and those behaviors requiring interpretation (eg, looking depressed). In the second experiment, the authors identified several behaviors by analyzing 5-minute segments of a video footage, with or without an attempt at the end, to check whether these behaviors were indeed present in individuals who were in crisis. The results suggested that 2 behaviors (pacing back and forth from the edge of the platform, and leaving belongings on the platform) could identify 24% of the attempters with no false positives. This suggests that there are indeed crisis behaviors that can be observed, and within 5 minutes prior to an attempt.

Another study [9] expanded on the findings from the aforementioned study [8] and used semistructured interviews

and questionnaires administered to railway staff in addition to CCTV footage to identify these potential indicator behaviors. The process identified 5 main behaviors: station hopping and platform switching, limited contact with people, allowing trains to pass by, where they stood on the platform, and repetitive behaviors such as walking up and down the length of the platform.

CCTV and videos have been used in a variety of ways to understand and to prevent suicide. Studies provide evidence that manual CCTV monitoring may in fact reduce incidence of suicide in various settings, such as in a hospital or in metro stations. Newer approaches utilize automated systems, such as using thermal motion sensors or defining a region of interest in CCTV scenes, while others utilize computer vision to detect behaviors indicative of a suicide attempt by hanging [20,33,36]. Finally, others have examined crisis behaviors as a starting point to develop interventions that are able to reach out even earlier [8,9].

Discussion

Interim Findings

This review shows that CCTV and video data are an important addition in understanding and preventing suicide. This review focused on the peer-reviewed literature, and did not include gray literature such as reports, working papers, government documents, white papers, and evaluation. Future reviews may seek to additionally include such gray literature. Nevertheless, we found that computer vision was able to objectively identify depression, an important risk factor for suicide, above chance level. This may potentially provide new avenues for screening or triage, given its objectivity and ability to process large amounts of data at once.

Other uses help forensic scientists understand how death occurs during certain means of suicide, and critically was able to provide evidence that helped rule distinct incidents as suicides. One of the challenges in understanding suicide is ruling when it has occurred. By incorporating CCTV and video into this process, we are able to better understand when and how suicides occur. Without fully understanding the means by which people die by suicide, we are unable to prevent it. CCTV and video are able to help with this process.

Finally, CCTV has been used in preventing suicide, with novel approaches being developed or investigated. These approaches include the use of thermal sensors, defining a region of interest in the scene, and more recently using computer vision. Thus far, while studies have been able to identify behaviors of a hanging attempt using computer vision, and some other studies have been able to identify signs of a subsequent attempt [8,9], there have been no studies published that identify crisis behaviors and attempt to identify them using computer vision.

Current Work: Integration of CCTV and Machine Learning at Hotspots

Aims

As identified in the preceding review of the literature, the use of CCTV has the potential to play a crucial role in suicide

prevention at certain locations. While computer vision has been used to identify hanging, it has not yet been used to identify presuicidal behaviors in public places such as suicide hotspots. Critically, while available systems are able to send an alert once an attempt is in progress (eg, after climbing a fence or entering the railway tracks), our work aims to detect early behavioral indicators that may precede a suicide attempt, which will may therefore allow more time for intervention and prevention. In this section, we briefly discuss our work in progress, where we use CCTV data and machine learning to identify crisis behaviors at hotspots. We present our approach to fill this gap for automated presuicide action detection.

Setting

Data are collected from a coastal tourist destination in Sydney, Australia. In 2010, the local council, along with several partners, implemented a multilevel self-harm and suicide prevention strategy, which included installing various types of fencing, improving amenities, installing signage and hotline booths, as well as installing CCTV cameras. At this location, there are 3 key types of cameras: static RGB color cameras; pan, tilt, and zoom (PTZ) RGB color cameras; and thermal cameras. The thermal cameras primarily cover the fences where people often cross over and thus have a much wider range. Further, given that many incidents happen at night, thermal cameras are used to trigger an alert if an individual crosses the border, by defining a region of interest. Once an alert has been sent, the police are notified. During daytime, PTZ RGB cameras can be used to locate the person of interest. This study has been approved by the UNSW Human Research Ethics Committee.

Methodology

To identify signs of a person in crisis, we propose conducting a human coding study similar to previous studies [8,9] and adapting the findings to a computer vision-based framework. The identification of these behaviors is critical to the design and structure of the information pipeline. In earlier works from metro settings [8,9], several key behaviors were identified, including pacing back and forth from the edge, neatly placing shoes and bags on the platform, and standing at the end of the platform where the train would approach. One challenge this poses is that while as humans we can identify these patterns as noted by Mishara and colleagues [8], in computer vision identifying different types of behaviors may require different approaches. Within each behavior type, the underlying system has different levels of complexity. For instance, placing items on the platform would require the algorithm to identify that the person has interacted with another, separate object. Methods to detect pacing back and forth from the edge, by contrast, would require pedestrian tracking along with formulating a trajectory—a path illustrating the object's movement within the scene. To recognize multiple humans' actions in untrimmed videos from CCTV cameras, 3 levels of vision processing are required: human detection (low-level vision), human tracking (intermediate-level vision), and behavior understanding methods (high-level vision) [39]. To identify certain high risk and potential indicator behaviors, a human coding study is required to provide annotated data for training the subsequent machine learning processes.

Behavioral Annotation

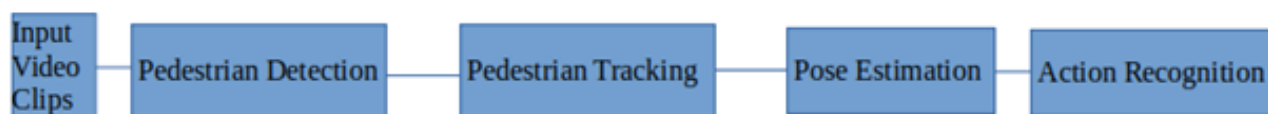
Incidents where an individual has climbed the safety fence and triggered an alarm will be retrospectively identified from site incident logs. Incidents which are noted as a noncrisis incident (eg, on the wrong side of the fence to take photos) will be excluded. Recordings of incidents will be extracted, and individuals traced back, as far as possible, to their entry to the site. Based on the 2 human coding studies [8,9] that previously reported analyses of 60 and 16 clips, respectively, we aim to analyze 16 incident clips. These will be matched by an equal number and duration of control clips of routine behaviors.

Following again the 2 studies identifying crisis behaviors at metro stations [8,9], relevant behaviors which have been identified from the metro setting will be generalized to form a codebook of behaviors which are likely to be relevant to this setting. Two researchers will review each clip and note any observable behaviors, their time of onset, and duration. Whether a behavior is thought to be positively or negatively associated with distress will also be recorded. These behavioral annotations will then be used as training data for the computer vision algorithms.

Behavior Identification Through Computer Vision

In this stage, we aim to develop an automated system that can detect specific human behaviors in real time for suicide prevention. The video clips will be fed through a data pipeline consisting of 4 layers of function modules, each with their own goal: the first is a pedestrian detection module to detect objects and persons in the clip; the second is a pedestrian tracking module to track the individual in the scene; the third is a pose estimation module to outline the location and configuration of the human joints and limbs (similar to that used by [32-36] to identify hanging); and the fourth is action recognition which interprets the configuration and motion of the joints and limbs to infer behavior (Figure 2). Using this approach, we are able to cover a wide range of behaviors. For example, if the individual is taking a photograph (a hypothetical routine behavior), the action recognition module should be able to identify it and eliminate that person from interest, or if the person walks back and forth from the fence, a hypothetical crisis behavior, then tracking should be able to capture this back-and-forth trajectory. We briefly discuss the 4 stages, noting the selected algorithms. These algorithms were not selected on the basis of their performance alone, as each camera, location, and context is different. Rather it is the result of extensive experimentation, the full details of which are out of the scope of this review.

Figure 2. Proposed Pipeline for Footage Processing.



In the first level, the algorithm should detect pedestrians or objects of interest. In order to do this, we use a deep learning-based object detector named YOLOv5 (You Only Look Once, version 5; [40]), as it provides a good balance between accuracy and processing speed in our context and reflects the most recent progress in the field. As noted above, these algorithms require training, and the algorithm is pretrained on a large public data set named MSCOCO [41]. However, the camera parameters such as angle and depth in the public data set differ largely from those of the cameras used at the site. This results in large differences in the visual appearance of pedestrians in the scene (such as shape and scale); therefore, we further trained (fine-tuned) the model using annotated footage from the location for customized pedestrian detection.

The algorithm output before retraining on our data set is illustrated in Figure 3, and Figure 4 illustrates the output after retraining on our data set. In both figures, pedestrians are indicated using bounding boxes, which are the smallest possible enclosing box around the individual. A confidence score (0-1) is shown above each pedestrian, indicating the level of confidence that the algorithm has detected a person within the box. As can be observed, the algorithm without fine-tuning (Figure 3) identified 3 individuals with a relatively low confidence (0.50-0.58), whereas after fine-tuning by training the algorithm with an annotated footage from this location (Figure 4) 4 individuals were correctly identified with higher confidence (0.86-0.88).

Figure 3. Pedestrian Detection Output Prior To Finetuning. Only three of the four individuals were detected, and all had low confidence values.



Figure 4. Pedestrian Detection Output using the Fine-tuned Algorithm. All four pedestrians were identified, and with high confidence values.



The output of the pedestrian detection module is then used as the input of the pedestrian tracking module that can track the same individual through multiple frames of footage. To that end, we chose a classical detection-based tracking method, named DeepSORT [42], which is a tracking algorithm with good speed and accuracy. The inputs to this module are the detected pedestrian regions (in the form of bounding boxes). The tracking algorithm assigns a unique identifier to each pedestrian, and tracks each target throughout the duration when it is visible in the scene. A challenge in tracking is ID switching,

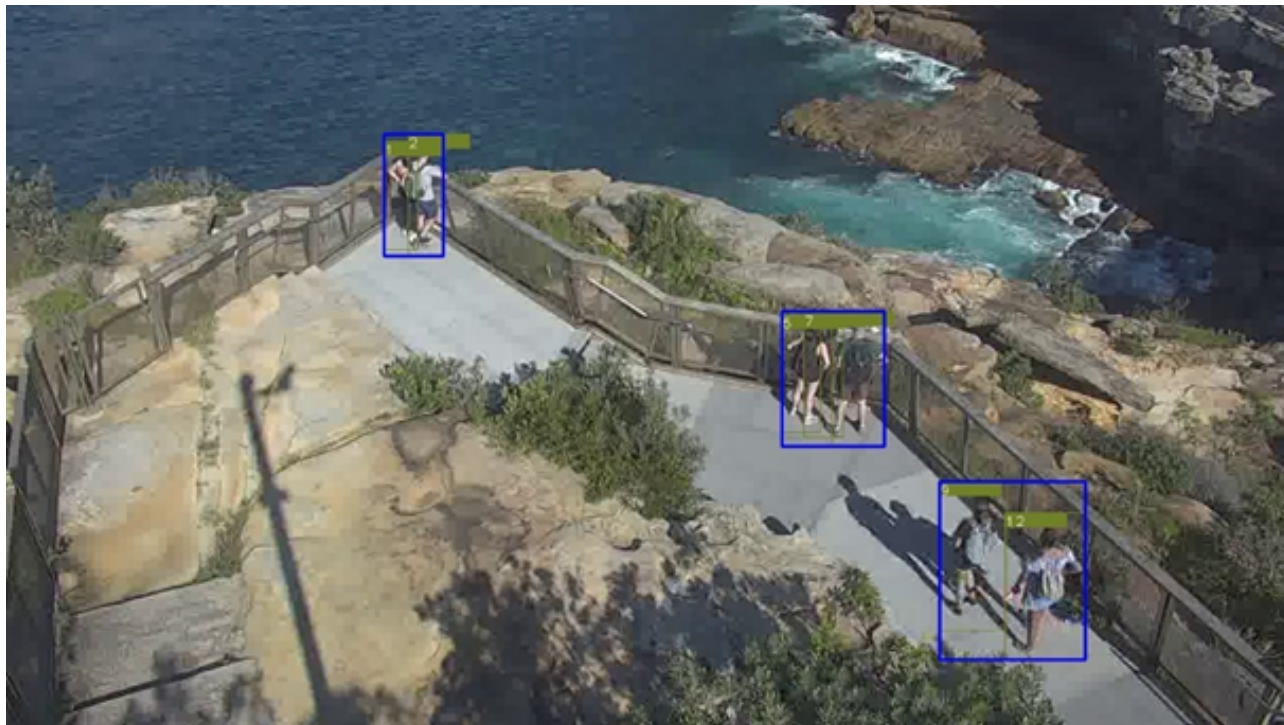
which is when 2 people overlap and thus 1 person is occluded in the camera's view, and the algorithm switches their initially allocated IDs. The current DeepSORT algorithm is able to alleviate such problems by including association metrics that use similarities between appearance features extracted by deep neural networks.

At this stage, we also include 2 other modules, namely, grouping and trajectory analysis. We include grouping analysis, as individuals who travel in a group are hypothesized to have a

substantially lower risk of being in crisis compared with individuals. These groupings are calculated by measuring how far apart individuals are from each other and whether they travel at similar speeds, as well as their scale differences. If the

algorithm perceives that they are in the group, a bounding box is used to group them together—see Figure 5 for an example output.

Figure 5. Pedestrian Tracking with Group Functionality. Three groups of people are detected in the scene: (1,2), (5,7), and (9,12). Each pedestrian is represented by a unique ID, and groups enclosed by blue rectangles.



A second functionality for this module is trajectory analysis. The trajectory for each pedestrian is stored for a predefined duration. This functionality is included as past studies investigating crisis behaviors have found that walking or pacing

back and forth may be indicative of a future attempt. Figure 6 shows example trajectories for groups of stationary and walking individuals.

Figure 6. Pedestrian Tracking with Trajectory Functionality. Pedestrians 1 and 2 have been stationary, so the trajectories are collapsed to green dots. Pedestrians 5 and 6 have walked along the fence line, as shown by the green path.



The third component of our pipeline is pose estimation, which is the estimation of the joint positions in the human body and connecting them. Using information from the previous modules, the pose estimation module analyses the areas within the bounding boxes and estimates the position of various joints.

This is shown as stick-figure outlines overlaid on top of the individual (Figure 7). The accuracy of the pose estimation algorithm depends on the image quality, with a closer view, of higher resolution, resulting in more accurate estimates.

Figure 7. Pose Estimation Module Output. The joint locations and limbs are identified by a pose estimation model. For better visibility, the other visualizations for grouping and trajectory analysis have been omitted.



In previous modules, each pedestrian has been detected, tracked, and their joints located. In the final module, pose-based features are extracted from each individual pedestrian over a time window and used to detect behavior. Once features have been extracted from video footage containing different actions, they are used to train machine learning-based classifiers to classify the actions, such as support vector machines [43], random forests [44], and multilayer perceptrons [45].

We have thus far integrated 4 preliminary actions or behaviors: standing, walking, taking a photo, and waving. These are actions

which are readily observed in routine footage and will be augmented with actions identified through the behavioral annotation of incident footage upon completion of data collection.

In summary, our current pipeline is able to detect an individual, track, and illustrate his/her path through the scene, identify groups, estimate the configuration of his/her limbs, and infer behaviors. Figures 8 and 9 show visual summaries of the overall analysis pipeline.

Figure 8. Example Output of the Analysis Pipeline. Individual pedestrians are shown with their unique IDs, recent trajectories, and action class (stand/walk). No groups were identified in this frame. Pose estimation has been omitted to improve visibility of the other outputs.

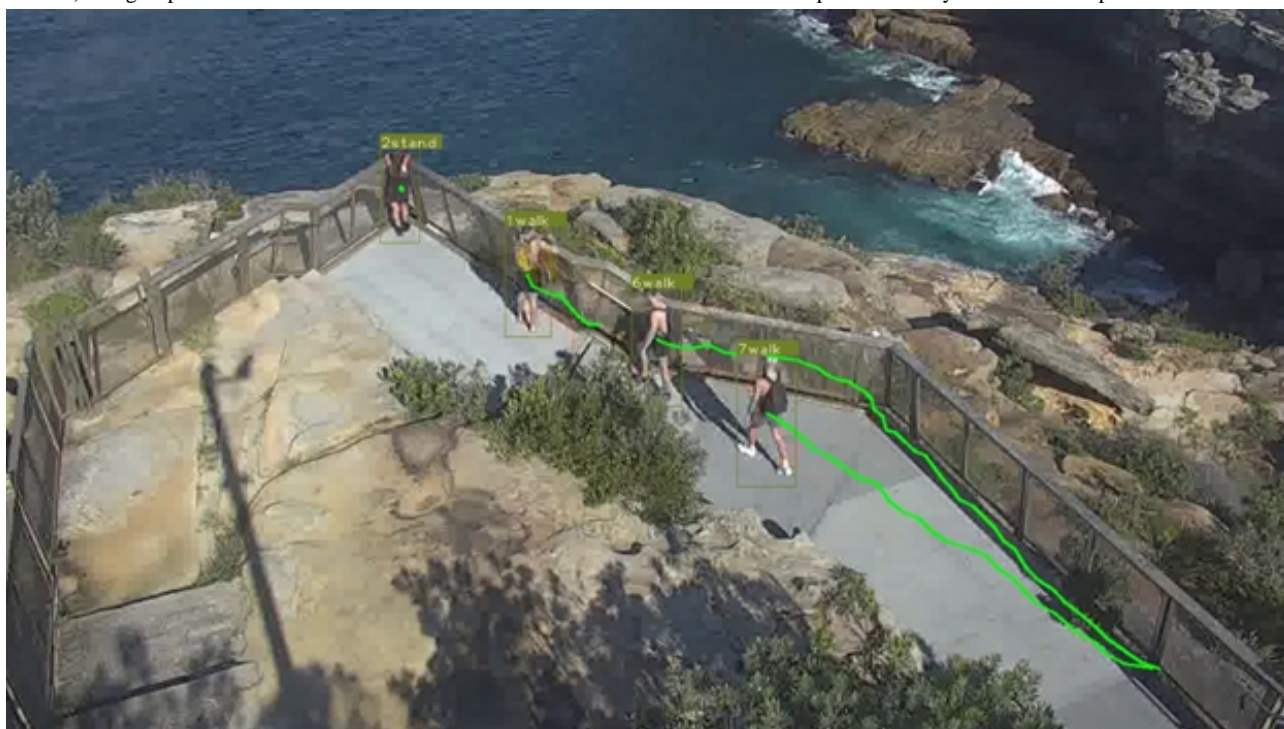
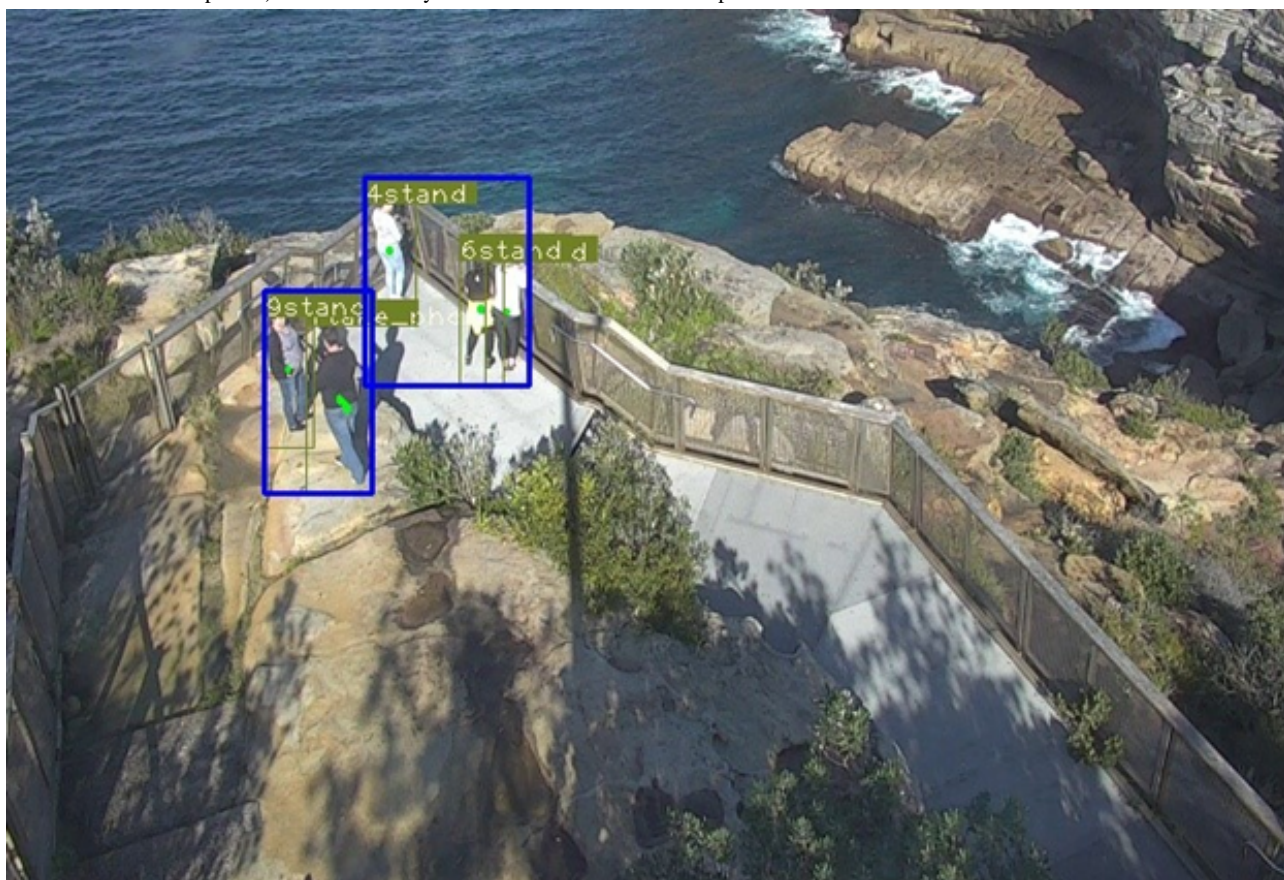


Figure 9. Example Output of the Analysis Pipeline 2. In this frame, five pedestrians within two groups have been identified. Their actions have been identified, with four individuals standing (“stand”) and one individual taking a photo (“take photo”). Their trajectories (in the form of dots and short lines at the centre of each person) indicate that they have remained in their current positions for the recent duration.



Thermal Cameras

The pipeline described in the previous section was implemented using footage from the RGB color cameras. However, a number of cameras at the site are thermal based, and the algorithms cannot be directly applied to these cameras. Therefore, a similar pipeline is also currently being developed for the thermal cameras. To date, pedestrian detection has been implemented using the EfficientDet algorithm [46]. This algorithm was selected based on its comparatively high accuracy performance compared with other available object detection algorithms on thermal data. Similar to YOLOv5 [40], this algorithm is

pretrained on the large MSCOCO [41] public data set, and Figure 10 illustrates the output. Further data collection and annotation are required to allow fine-tuning to increase the accuracy of the algorithm (similar to the improvement shown between Figures 3 and 4).

As with the RGB color camera recordings, following pedestrian detections, the DeepSORT [42] algorithm is applied to track multiple individuals in the thermal footage. An ID is assigned to each individual who is then tracked through subsequent video frames, as illustrated in Figure 11.

Figure 10. Pedestrian Identification using Thermal Cameras. The output of the EfficientDet pedestrian detection algorithm using data from a thermal camera. Individuals in the foreground are shown within bounding boxes with confidence levels between 61-80%.



Figure 11. Pedestrians Tracking using Thermal Cameras. Individuals are assigned a unique ID which persists until they leave the camera view.



Challenges and Future Directions

This above section described the computer vision pipeline which will be used to identify behaviors associated with crisis incidents. However, several challenges remain which require further consideration.

While previous studies have examined footage from metro settings, our study location is a very large outdoor space which experiences large variations in lighting and weather conditions. This can have a large impact on the performance of the detection algorithms, to the extent that RGB color cameras provide no footage for unlit areas at night. Furthermore, as a significant proportion of incidents occur after dark, the use of footage from thermal cameras is critical. However, thermal cameras provide less visual information that can be used in vision-based analysis. They also typically have a much lower resolution, resulting in additional challenges for pedestrian detection and tracking, and also suffer from blooming (blurred borders and outlines), and thus may not be suitable for pose estimation approaches that require identification of human limbs.

Therefore, a different strategy for identifying individual in crisis may need to be used using thermal cameras. For example, if the human coding study confirms that some individuals in crisis crouch down to leave possessions on the ground (a behavior which can be detected using pose estimation), or pace back and

forth from the fence (a behavior that could be detected using trajectory analysis, and not requiring pose estimation), a thermal camera may only be able to detect the latter crisis behavior. Given that behaviors which rely on location—and not limb—analysis have been identified as key crisis behaviors in previous papers [8,9], this approach alone is likely to provide good accuracy in identifying individuals in crisis.

While thermal cameras provide less information for computer vision algorithms, thermal cameras of sufficient quality and minimal distance may still be able to identify behaviors through pose estimation. However, this requires experimental testing of the computer vision algorithm and possible postprocessing to improve the clarity of the image.

It is also important to consider the diversity of camera angles and distances to pedestrians within the scenes. These angles and distances are different for each camera at the site, and these may differ considerably from those which have been used to pretrain the selected algorithms, thus negatively affecting performance. Further annotations and fine-tuning may therefore be required for each scene included in the analysis.

Our current implementation uses a supervised learning approach to train models to recognize certain unit actions, such as walking, standing, waving, and taking a photo. However, a substantial human time effort is required to annotate these behaviors across video frames. It may be possible to incorporate

unsupervised as well as active learning methods in conjunction with transfer learning to optimally leverage the existing annotated data without requiring additional manual annotations. To detect and differentiate between longer and more complex activities, higher-level reasoning may be required to aid the decision-making process. Not only will unit actions be detected, but also our system will learn movement patterns to allow the identification of individuals who appear to be in crisis. These results will be reported elsewhere upon the conclusion of data collection.

In addition to the technological aspects identified in this review, it is also important to assess the acceptability of this approach with decision makers, people with lived experience, and other stakeholders. For example, the Joint Commission in the United States released a statement prohibiting the use of video monitoring of patients with high suicide risk unless in-person monitoring is also present [47]. However, the statement notes that this decision is due to inability to immediately intervene with video monitoring alone. Therefore, future studies will also need to understand the concerns surrounding implementing such technologies, as well as how to address them.

Conclusions

This paper has reviewed the existing literature pertaining to the use of CCTV and video data in understanding and preventing suicide. The narrative review found that studies have reported on 3 broad approaches: assessing depression risk by facial cues as a suicide risk factor; using CCTV and video to understand suicide after the fact through agonal sequences and determining suicidal intent; and finally using CCTV as a suicide prevention initiative. Studies have reported systems to generate alerts when a suicide attempt is detected, for example detecting hanging or jumping from a bridge; however these focus on the period after an attempt is made. Others have reported human coding studies to describe behaviors which can be observed in individuals during crisis prior to a suicide attempt. Building on these previously reported research themes, we reported progress on a current project to identify behaviors associated with crisis incidents at a specific site. The proposed methodology for a human annotation study is described, along with the development of a computer vision processing pipeline. We describe several challenges and considerations which are required to extend computer vision techniques to different settings. Ultimately, this research aims to automatically detect behaviors which may be indicative of a suicide attempt, allowing earlier intervention to save lives.

Acknowledgments

We thank Kathy Woodcock, Simon Baker, Gihan Samarasinghe, and Yang Song for their support in the early stages of establishing this project. This work was supported by a Suicide Prevention Research Fund Innovation Grant, managed by Suicide Prevention Australia, and the NHMRC Centre of Research Excellence in Suicide Prevention (APP1152952). We also thank Woollahra Municipal Council for sharing information on the use of CCTV as part of their commitment to self-harm minimization within their local area and the work they are doing with police and emergency response personnel and mental health support agencies. The contents of this manuscript are the responsibility of the authors, and have not been approved or endorsed by the funders.

Authors' Contributions

According to the CRediT Taxonomy, SO, MEL, and AS were responsible for conceptualization; all authors were responsible for data curation; MEL and AS were responsible for funding acquisition; XL, AS, and ML were responsible for software and formal analysis; MEL and AS were responsible for supervision; XL and ML were responsible for visualization; and all authors took part in writing the manuscript.

Conflicts of Interest

None declared.

References

1. World Health Organization. Suicide. 2019. URL: <https://www.who.int/news-room/fact-sheets/detail/suicide> [accessed 2021-04-15]
2. Torok M, Shand F, Phillips M, Meteoro N, Martin D, Larsen M. Data-informed targets for suicide prevention: a small-area analysis of high-risk suicide regions in Australia. *Soc Psychiatry Psychiatr Epidemiol* 2019 Oct;54(10):1209-1218 [FREE Full text] [doi: [10.1007/s00127-019-01716-8](https://doi.org/10.1007/s00127-019-01716-8)] [Medline: [31041467](https://pubmed.ncbi.nlm.nih.gov/31041467/)]
3. Owens C, Lloyd-Tomlins S, Emmens T, Aitken P. Suicides in public places: findings from one English county. *Eur J Public Health* 2009 Dec;19(6):580-582 [FREE Full text] [doi: [10.1093/eurpub/ckp052](https://doi.org/10.1093/eurpub/ckp052)] [Medline: [19380332](https://pubmed.ncbi.nlm.nih.gov/19380332/)]
4. Pirkis J, Spittal MJ, Cox G, Robinson J, Cheung YTD, Studdert D. The effectiveness of structural interventions at suicide hotspots: a meta-analysis. *Int J Epidemiol* 2013 Apr;42(2):541-548. [doi: [10.1093/ije/dyt021](https://doi.org/10.1093/ije/dyt021)] [Medline: [23505253](https://pubmed.ncbi.nlm.nih.gov/23505253/)]
5. Ross V, Koo YW, Kölves K. A suicide prevention initiative at a jumping site: A mixed-methods evaluation. *EClinicalMedicine* 2020 Feb;19:100265 [FREE Full text] [doi: [10.1016/j.eclinm.2020.100265](https://doi.org/10.1016/j.eclinm.2020.100265)] [Medline: [32140675](https://pubmed.ncbi.nlm.nih.gov/32140675/)]

6. Pirkis J, Too LS, Spittal MJ, Krysinska K, Robinson J, Cheung YTD. Interventions to reduce suicides at suicide hotspots: a systematic review and meta-analysis. *Lancet Psychiatry* 2015 Nov;2(11):994-1001. [doi: [10.1016/S2215-0366\(15\)00266-7](https://doi.org/10.1016/S2215-0366(15)00266-7)] [Medline: [26409438](https://pubmed.ncbi.nlm.nih.gov/26409438/)]
7. Lockley A, Cheung YTD, Cox G, Robinson J, Williamson M, Harris M, et al. Preventing suicide at suicide hotspots: a case study from Australia. *Suicide Life Threat Behav* 2014 Aug 15;44(4):392-407. [doi: [10.1111/sltb.12080](https://doi.org/10.1111/sltb.12080)] [Medline: [25250406](https://pubmed.ncbi.nlm.nih.gov/25250406/)]
8. Mishara B, Dupont S. Can CCTV identify people in public transit stations who are at risk of attempting suicide? An analysis of CCTV video recordings of attempters and a comparative investigation. *BMC Public Health* 2016 Dec 15;16(1):1245 [FREE Full text] [doi: [10.1186/s12889-016-3888-x](https://doi.org/10.1186/s12889-016-3888-x)] [Medline: [27974046](https://pubmed.ncbi.nlm.nih.gov/27974046/)]
9. Mackenzie J, Borrill J, Hawkins E, Fields B, Kruger I, Noonan I, et al. Behaviours preceding suicides at railway and underground locations: a multimethodological qualitative approach. *BMJ Open* 2018 Apr 10;8(4):e021076 [FREE Full text] [doi: [10.1136/bmjopen-2017-021076](https://doi.org/10.1136/bmjopen-2017-021076)] [Medline: [29643167](https://pubmed.ncbi.nlm.nih.gov/29643167/)]
10. Jiao L, Zhang F, Liu F, Yang S, Li L, Feng Z, et al. A Survey of Deep Learning-Based Object Detection. *IEEE Access* 2019;7:128837-128868. [doi: [10.1109/access.2019.2939201](https://doi.org/10.1109/access.2019.2939201)]
11. Ciaparrone G, Luque Sánchez F, Tabik S, Troiano L, Tagliaferri R, Herrera F. Deep learning in video multi-object tracking: A survey. *Neurocomputing* 2020 Mar;381:61-88. [doi: [10.1016/j.neucom.2019.11.023](https://doi.org/10.1016/j.neucom.2019.11.023)]
12. Poppe R. A survey on vision-based human action recognition. *Image and Vision Computing* 2010 Jun;28(6):976-990 [FREE Full text] [doi: [10.1016/j.imavis.2009.11.014](https://doi.org/10.1016/j.imavis.2009.11.014)]
13. Soron T, Shariful Islam SM. Suicide on Facebook-the tales of unnoticed departure in Bangladesh. *Glob Ment Health (Camb)* 2020;7:e12 [FREE Full text] [doi: [10.1017/gmh.2020.5](https://doi.org/10.1017/gmh.2020.5)] [Medline: [32742670](https://pubmed.ncbi.nlm.nih.gov/32742670/)]
14. Zhu X, Gedeon T, Caldwell S, Jones R. Visceral versus Verbal: Can We See Depression? *Acta Polytechnica Hungarica* 2019;16(9):113-133 [FREE Full text]
15. de Melo WC, Granger A, Hadid A. Combining Global and Local Convolutional 3D Networks for Detecting Depression from Facial Expressions. 2019 Presented at: 14th IEEE International Conference on Automatic Face & Gesture Recognition (FG); 2019; Lille, France p. 1-8. [doi: [10.1109/fg.2019.8756568](https://doi.org/10.1109/fg.2019.8756568)]
16. Dadiz B, Marcos N. Analysis of Depression Based on Facial Cues on A Captured Motion Picture. 2018 Presented at: IEEE 3rd International Conference on Signal and Image Processing (ICSIP); 2018; Shenzhen, China p. 49-54. [doi: [10.1109/siprocess.2018.8600523](https://doi.org/10.1109/siprocess.2018.8600523)]
17. Li J, Liu Z, Ding Z, Wang G. A novel study for MDD detection through task-elicited facial cues. 2019 Presented at: IEEE International Conference on Bioinformatics and Biomedicine (BIBM); 2018; Madrid, Spain p. 1003-1008. [doi: [10.1109/bibm.2018.8621236](https://doi.org/10.1109/bibm.2018.8621236)]
18. Graham S, Depp C, Lee EE, Nebeker C, Tu X, Kim HC, et al. Artificial Intelligence for Mental Health and Mental Illnesses: an Overview. *Curr Psychiatry Rep* 2019 Nov 07;21(11):116 [FREE Full text] [doi: [10.1007/s11920-019-1094-0](https://doi.org/10.1007/s11920-019-1094-0)] [Medline: [31701320](https://pubmed.ncbi.nlm.nih.gov/31701320/)]
19. Girard JM, Cohn JF, Mahoor MH, Mavadati S, Rosenwald DP. Social Risk and Depressionvidence from Manual and Automatic Facial Expression Analysis. 2013 Presented at: Proceedings of the International Conference on Automatic Face and Gesture Recognition; 2013; Shanghai, China p. 1-8. [doi: [10.1109/fg.2013.6553748](https://doi.org/10.1109/fg.2013.6553748)]
20. Scherer S, Stratou G, Mahmoud M, Boberg J, Gratch J, Rizzo A, et al. Automatic behavior descriptors for psychological disorder analysis. 2013 Presented at: 10th IEEE Conference on Automatic Face and Gesture Recognition; 2013; Shanghai, China p. 1-8. [doi: [10.1109/fg.2013.6553789](https://doi.org/10.1109/fg.2013.6553789)]
21. Yamasaki S, Kobayashi AK, Nishi K. Evaluation of suicide by hanging : From the video recording. *Forensic Sci Med Pathol* 2007 Mar;3(1):45-51. [doi: [10.1385/FSMP.3.1.45](https://doi.org/10.1385/FSMP.3.1.45)] [Medline: [25868889](https://pubmed.ncbi.nlm.nih.gov/25868889/)]
22. Sauvageau A. Agonal sequences in four filmed hangings: analysis of respiratory and movement responses to asphyxia by hanging. *J Forensic Sci* 2009 Jan;54(1):192-194. [doi: [10.1111/j.1556-4029.2008.00910.x](https://doi.org/10.1111/j.1556-4029.2008.00910.x)] [Medline: [19040672](https://pubmed.ncbi.nlm.nih.gov/19040672/)]
23. Sauvageau A, Laharpe R, King D, Dowling G, Andrews S, Kelly S, Working Group on Human Asphyxia. Agonal sequences in 14 filmed hangings with comments on the role of the type of suspension, ischemic habituation, and ethanol intoxication on the timing of agonal responses. *Am J Forensic Med Pathol* 2011 Jun;32(2):104-107. [doi: [10.1097/PAF.0b013e3181efba3a](https://doi.org/10.1097/PAF.0b013e3181efba3a)] [Medline: [20683242](https://pubmed.ncbi.nlm.nih.gov/20683242/)]
24. Sauvageau A, Racette S. Agonal sequences in a filmed suicidal hanging: analysis of respiratory and movement responses to asphyxia by hanging. *J Forensic Sci* 2007 Jul;52(4):957-959. [doi: [10.1111/j.1556-4029.2007.00459.x](https://doi.org/10.1111/j.1556-4029.2007.00459.x)] [Medline: [17524058](https://pubmed.ncbi.nlm.nih.gov/17524058/)]
25. Alunni V, Grevin G, Buchet L, Gaillard Y, Quatrehomme G. An amazing case of fatal self-immolation. *Forensic Science International* 2014 Nov;244:e30-e33. [doi: [10.1016/j.forsciint.2014.08.030](https://doi.org/10.1016/j.forsciint.2014.08.030)]
26. Ogden RD, Hamilton WK, Whitcher C. Assisted suicide by oxygen deprivation with helium at a Swiss right-to-die organisation. *Journal of Medical Ethics* 2010 Mar 08;36(3):174-179. [doi: [10.1136/jme.2009.032490](https://doi.org/10.1136/jme.2009.032490)]
27. Allard TJ, Wortley RK, Stewart AL. The Effect of CCTV on Prisoner Misbehavior. *The Prison Journal* 2008 Sep 01;88(3):404-422. [doi: [10.1177/0032885508322492](https://doi.org/10.1177/0032885508322492)]
28. Too LS, Spittal MJ, Bugeja L, Milner A, Stevenson M, McClure R. An investigation of neighborhood-level social, economic and physical factors for railway suicide in Victoria, Australia. *J Affect Disord* 2015 Sep 01;183:142-148. [doi: [10.1016/j.jad.2015.05.006](https://doi.org/10.1016/j.jad.2015.05.006)] [Medline: [26005775](https://pubmed.ncbi.nlm.nih.gov/26005775/)]

29. Kroll DS, Stanghellini E, DesRoches SL, Lydon C, Webster A, O'Reilly M, et al. Virtual monitoring of suicide risk in the general hospital and emergency department. *Gen Hosp Psychiatry* 2020;63:33-38. [doi: [10.1016/j.genhosppsych.2019.01.002](https://doi.org/10.1016/j.genhosppsych.2019.01.002)] [Medline: [30665667](https://pubmed.ncbi.nlm.nih.gov/30665667/)]
30. Lee J, Lee C, Park N. Application of sensor network system to prevent suicide from the bridge. *Multimed Tools Appl* 2015 Dec 16;75(22):14557-14568 [FREE Full text] [doi: [10.1007/s11042-015-3134-z](https://doi.org/10.1007/s11042-015-3134-z)]
31. Mukherjee A, Ghosh B. An Automated Approach to Prevent Suicide in Metro Stations. 2017 Mar 17 Presented at: Proceedings of the 5th International Conference on Frontiers in Intelligent Computing: Theory and Applications. *Advances in Intelligent Systems and Computing*; 2017; Singapore URL: https://doi.org/10.1007/978-981-10-3153-3_75 [doi: [10.1007/978-981-10-3153-3_75](https://doi.org/10.1007/978-981-10-3153-3_75)]
32. Chiranjeevi V, Elangovan D. A Review on Human Action Recognition Machine Learning Techniques for Suicide Detection System. In: *Innovations in Bio-Inspired Computing and Applications*. 2019 Presented at: IBICA 2018: Innovations in Bio-Inspired Computing and Applications; 17 - 19 December 2018; Kochi, India p. 46-55 URL: https://doi.org/10.1007/978-3-030-16681-6_5 [doi: [10.1007/978-3-030-16681-6_5](https://doi.org/10.1007/978-3-030-16681-6_5)]
33. Bouachir W, Noumeir R. Automated video surveillance for preventing suicide attempts. 2016 Presented at: 7th International Conference on Imaging for Crime Detection and Prevention (ICDP), Madrid; 2016; Madrid, Spain p. 1-6. [doi: [10.1049/ic.2016.0081](https://doi.org/10.1049/ic.2016.0081)]
34. Li B, Bouachir W, Gouiaa R, Noumeir R. Real-time recognition of suicidal behavior using an RGB-D camera. 2017 Presented at: Seventh International Conference on Image Processing Theory, Tools Applications (IPTA); 28 November - 1 December 2017; Montreal, Canada p. 1-6. [doi: [10.1109/ipta.2017.8310087](https://doi.org/10.1109/ipta.2017.8310087)]
35. Bouachir W, Gouiaa R, Li B, Noumeir R. Intelligent video surveillance for real-time detection of suicide attempts. *Pattern Recognition Letters* 2018 Jul;110:1-7. [doi: [10.1016/j.patrec.2018.03.018](https://doi.org/10.1016/j.patrec.2018.03.018)]
36. Lee S, Kim H, Lee S, Kim Y, Lee D, Ju J, et al. Detection of a Suicide by Hanging Based on a 3-D Image Analysis. *IEEE Sensors J* 2014 Sep;14(9):2934-2935. [doi: [10.1109/jsen.2014.2332070](https://doi.org/10.1109/jsen.2014.2332070)]
37. Eom K, Ahn T, Kim G, Jang G, Jung J, Kim M. Hierarchically Categorized Performance Evaluation Criteria for Intelligent Surveillance System. 2009 Presented at: Proceedings of the International Symposium on Web Information Systems and Applications (WISA'09); 2009; Nanchang, China p. 218-221 URL: <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.499.6411&rep=rep1&type=pdf>
38. Shah AP, Vaibhav, Sharma V, Ismail MA, Girard J, Morency L. Multimodal Behavioral Markers Exploring Suicidal Intent in Social Media Videos. 2019 Presented at: International Conference on Multimodal Interaction (ICMI '19); 2019; Suzhou, China p. 409-413. [doi: [10.1145/3340555.3353718](https://doi.org/10.1145/3340555.3353718)]
39. Zhang H, Zhang Y, Zhong B, Lei Q, Yang L, Du J, et al. A Comprehensive Survey of Vision-Based Human Action Recognition Methods. *Sensors (Basel)* 2019 Feb 27;19(5):1005 [FREE Full text] [doi: [10.3390/s19051005](https://doi.org/10.3390/s19051005)] [Medline: [30818796](https://pubmed.ncbi.nlm.nih.gov/30818796/)]
40. ultralytics/yolov5. GitHub. URL: <https://github.com/ultralytics/yolov5> [accessed 2010-01-01]
41. Lin T, Maire M, Belongie S, Hays J, Perona P, Ramanan D, et al. Microsoft COCO: Common objects in Context. In: *Computer Vision – ECCV 2014*. 2014 Presented at: ECCV 2014; 2014; Zurich, Switzerland p. 740-755. [doi: [10.1007/978-3-319-10602-1_48](https://doi.org/10.1007/978-3-319-10602-1_48)]
42. Wojke N, Bewley A, Paulus D. Simple online and realtime tracking with a deep association metric. 2018 Presented at: IEEE International Conference on Image Processing (ICIP); 2017; Beijing, China. [doi: [10.1109/icip.2017.8296962](https://doi.org/10.1109/icip.2017.8296962)]
43. Cortes C, Vapnik V. Support-vector networks. *Mach Learn* 1995 Sep;20(3):273-297. [doi: [10.1007/BF00994018](https://doi.org/10.1007/BF00994018)]
44. Flaxman AD, Vahdatpour A, Green S, James SL, Murray CJ, Population Health Metrics Research Consortium (PHMRC). Random forests for verbal autopsy analysis: multisite validation study using clinical diagnostic gold standards. *Popul Health Metr* 2011 Aug 04;9:29 [FREE Full text] [doi: [10.1186/1478-7954-9-29](https://doi.org/10.1186/1478-7954-9-29)] [Medline: [21816105](https://pubmed.ncbi.nlm.nih.gov/21816105/)]
45. Collobert R, Bengio S. Links between Perceptrons, MLPs and SVMs. In: *Proc. Int'l Conf. on Machine Learning (ICML)*. 2004 Presented at: 21st International Conference on Machine Learning; 4 - 8th July 2004; Banff, Alberta p. 123-128. [doi: [10.1145/1015330.1015415](https://doi.org/10.1145/1015330.1015415)]
46. Tan M, Pang R, Le QV. EfficientDet: Scalable and Efficient Object Detection. 2020 Jun 13 Presented at: 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2020; 13-19 June 2020; Seattle, WA, USA URL: <https://doi.org/10.1109/cvpr42600.2020.01079>
47. The Joint Commission. Can video monitoring/electronic-sitters be used to monitor patients at high risk for suicide? 2021. URL: <https://www.jointcommission.org/standards/standard-faqs/hospital-and-hospital-clinics/national-patient-safety-goals-npsg/000002227> [accessed 2021-03-01]

Abbreviations

CCTV: closed-circuit television
LDC: linear discriminant classifier
PTZ: pan, tilt, and zoom
RGB-D: red, green, blue—depth

Edited by J Torous, G Eysenbach; submitted 01.02.21; peer-reviewed by JM Mackenzie, D Kroll; comments to author 16.02.21; revised version received 17.03.21; accepted 17.03.21; published 07.05.21.

Please cite as:

Onie S, Li X, Liang M, Sowmya A, Larsen ME

The Use of Closed-Circuit Television and Video in Suicide Prevention: Narrative Review and Future Directions

JMIR Ment Health 2021;8(5):e27663

URL: <https://mental.jmir.org/2021/5/e27663>

doi: [10.2196/27663](https://doi.org/10.2196/27663)

PMID: [33960952](https://pubmed.ncbi.nlm.nih.gov/33960952/)

©Sandersan Onie, Xun Li, Morgan Liang, Arcot Sowmya, Mark Erik Larsen. Originally published in JMIR Mental Health (<https://mental.jmir.org>), 07.05.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Mental Health, is properly cited. The complete bibliographic information, a link to the original publication on <https://mental.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Assessing Suicide Reporting in Top Newspaper Social Media Accounts in China: Content Analysis Study

Kaisheng Lai¹, PhD; Dan Li¹, BA; Huijuan Peng¹, BA; Jingyuan Zhao¹, BA; Lingnan He^{2,3,4}, PhD

¹School of Journalism and Communication, Jinan University, Guangzhou, China

²School of Communication and Design, Sun Yat-Sen University, Guangzhou, China

³Department of Psychology, Sun Yat-Sen University, Guangzhou, China

⁴Guangdong Key Laboratory for Big Data Analysis and Simulation of Public Opinion, Guangzhou, China

Corresponding Author:

Lingnan He, PhD

School of Communication and Design

Sun Yat-Sen University

No. 132 Waihuan East Road

Higher Education Mega Center

Guangzhou,

China

Phone: 86 020 3933 1935

Email: hel3@mail.sysu.edu.cn

Abstract

Background: Previous studies have shown that suicide reporting in mainstream media has a significant impact on suicidal behaviors (eg, irresponsible suicide reporting can trigger imitative suicide). Traditional mainstream media are increasingly using social media platforms to disseminate information on public-related topics, including health. However, there is little empirical research on how mainstream media portrays suicide on social media platforms and the quality of their coverage.

Objective: This study aims to explore the characteristics and quality of suicide reporting by mainstream publishers via social media in China.

Methods: Via the application programming interface of the social media accounts of the top 10 Chinese mainstream publishers (eg, People's Daily and Beijing News), we obtained 2366 social media posts reporting suicide. This study conducted content analysis to demonstrate the characteristics and quality of the suicide reporting. According to the World Health Organization (WHO) guidelines, we assessed the quality of suicide reporting by indicators of harmful information and helpful information.

Results: Chinese mainstream publishers most frequently reported on suicides stated to be associated with conflict on their social media (eg, 24.47% [446/1823] of family conflicts and 16.18% [295/1823] of emotional frustration). Compared with the suicides of youth (730/1446, 50.48%) and urban populations (1454/1588, 91.56%), social media underreported suicides in older adults (118/1446, 8.16%) and rural residents (134/1588, 8.44%). Harmful reporting practices were common (eg, 54.61% [1292/2366] of the reports contained suicide-related words in the headline and 49.54% [1172/2366] disclosed images of people who died by suicide). Helpful reporting practices were very limited (eg, 0.08% [2/2366] of reports provided direct information about support programs).

Conclusions: The suicide reporting of mainstream publishers on social media in China broadly had low adherence to the WHO guidelines. Considering the tremendous information dissemination power of social media platforms, we suggest developing national suicide reporting guidelines that apply to social media. By effectively playing their separate roles, we believe that social media practitioners, health institutions, social organizations, and the general public can endeavor to promote responsible suicide reporting in the Chinese social media environment.

(*JMIR Ment Health* 2021;8(5):e26654) doi:[10.2196/26654](https://doi.org/10.2196/26654)

KEYWORDS

suicide; suicide reporting; mainstream publishers; social media; WHO guidelines

Introduction

Suicide is a public health problem of global concern. More than 800,000 suicide deaths occur every year, meaning that more than 2000 people die by suicide every day [1]. To deal with this problem, the World Health Organization (WHO) has proposed the need to highly prioritize suicide prevention on global public health and public policy issues [2]. Suicide is a multidimensional issue related to various factors that can be facilitated by psychological, biological, social, and environmental factors [3,4]. Mass media coverage is a significant agent in the social construction of reality in today's world, which may affect people's exposure to suicide behaviors, especially for vulnerable groups [5]. Accordingly, researchers have determined that the role of mass media in suicidal behavior warrants serious and focused attention [6].

Research on the negative effect of suicide reporting (commonly known as the Werther effect) in the academic field can be traced back to 1974. Philips [7] found that publishing suicide stories in British and American newspapers led to an immediate increase in the number of suicides in the region. Since then, studies have confirmed the imitation effect caused by media reporting [8,9] and further illustrated that this imitation effect can be aggravated when suicide is repeatedly reported [10] or depicted sensationally and graphically [11] or when the subject is a celebrity [12,13]. Subsequently, scholars have noted the potential effect of the media on suicide prevention. Empirical research conducted in Switzerland [14], Australia [15], and Hong Kong [16] has indicated that promoting education by introducing reporting guidelines to media practitioners could effectively improve the quality of suicide reporting. By conducting long-term experiments in Austria, researchers found that the application of media guidelines and media campaigns in the city of Vienna led to a reduction of more than 80% in the subway suicide rate in 6 months [17-19].

Given the effectiveness of media guidelines in suicide prevention, in 2008, WHO and the International Association for Suicide Prevention released professional criteria for suicide reporting for media practitioners [20]. Consequently, some scholars have begun to evaluate the extent to which suicide reporting followed these guidelines; these evaluation studies were concentrated in Western countries and showed that the implementation of the guidelines varied [21,22]. Recently, research and evaluation on suicide reporting in Asia has increased; generally, studies have found a low level of compliance with WHO guidelines [23-26]. When assessing the quality of suicide reporting in Indian newspapers, Armstrong et al [24] found a considerable amount of harmful information and minimal educational information. If we consider that Asia has the largest suicide numbers on the globe [27], however, we find that the amount of empirical evidence we have on the quality of suicide reporting in the continent is quite limited.

Overall, the literature on the evaluation of suicide reporting has focused on traditional media, especially newspapers; with the emergence of Web 2.0 technology, however, social media is becoming an indispensable part of public information dissemination [28]. In the health field, people increasingly rely

on social media for health communication, and social media plays a crucial role in the construction of general cognition, attitude, and behavior in the field [29,30]. Thus, many media outlets have allocated their resources toward social media platforms [31], and these accounts have shown tremendous influence. The Twitter account of the New York Times has nearly 500 million followers; similarly, People's Daily has more than 100 million followers on Sina Weibo. Language style, staff structure, and frequency of posting of mainstream media on social media platforms differs from traditional media platforms [32-34], and research has shown that mainstream media's language style on social media is more vivid, concise, and easy to understand [34]. Do the differences between traditional and social media platforms affect mainstream media's reporting preferences or content quality? Based on this question, investigating how the mainstream media reports on public issues via social media has recently become a challenging research topic.

An empirical study found that American mainstream outlets showed preferential interest in disseminating information related to psychiatric disorders on Twitter, and media attention to different mental health disorders determines followers' retweet responses [35]. Regarding suicide, a recent study evaluated news publishers' reporting of suicide on Facebook and found that news articles often provided harmful elements to readers, whereas positive elements were relatively rare [36]. To the best of our knowledge, the study has been the only one to investigate news publishers' suicide reporting on a social media platform, and it focused primarily on English-speaking countries. It remains unknown how mainstream media reports suicide via social media in different cultural contexts, especially in China.

Suicide remains a significant public health issue in China today. In 2019, nearly 140,000 people died by suicide in China, second only to India in the world [37]. Although a few scholars have assessed suicide reporting in China, these studies were limited to specific regions or focused on traditional media [23,25,38]. Using WHO guidelines, Fu et al [23] compared suicide reporting of newspapers in Hong Kong, Taiwan, and Guangzhou; the findings indicated that reports in the three regions were mostly inconsistent with WHO recommendations. Chu et al [25] evaluated suicide reporting in China's most influential newspapers and internet-based media, and the results showed that compliance with the guidelines was very low for 4 of the 12 recommendations (eg, less than 5% of stories provided information on where a person could go for help). To the best of our knowledge, no nationwide research has investigated the compliance of suicide reporting published by Chinese mainstream publishers on social media platforms. Thus, the aim of this study is 2-fold and the research questions are as follows: What are the characteristics of suicide reporting of mainstream publishers via social media? What is the extent to which their suicide reporting follows WHO guidelines?

Methods

Newspapers Selected

We obtained the top 100 media list of comprehensive communication power from a 2018 national online survey [39].

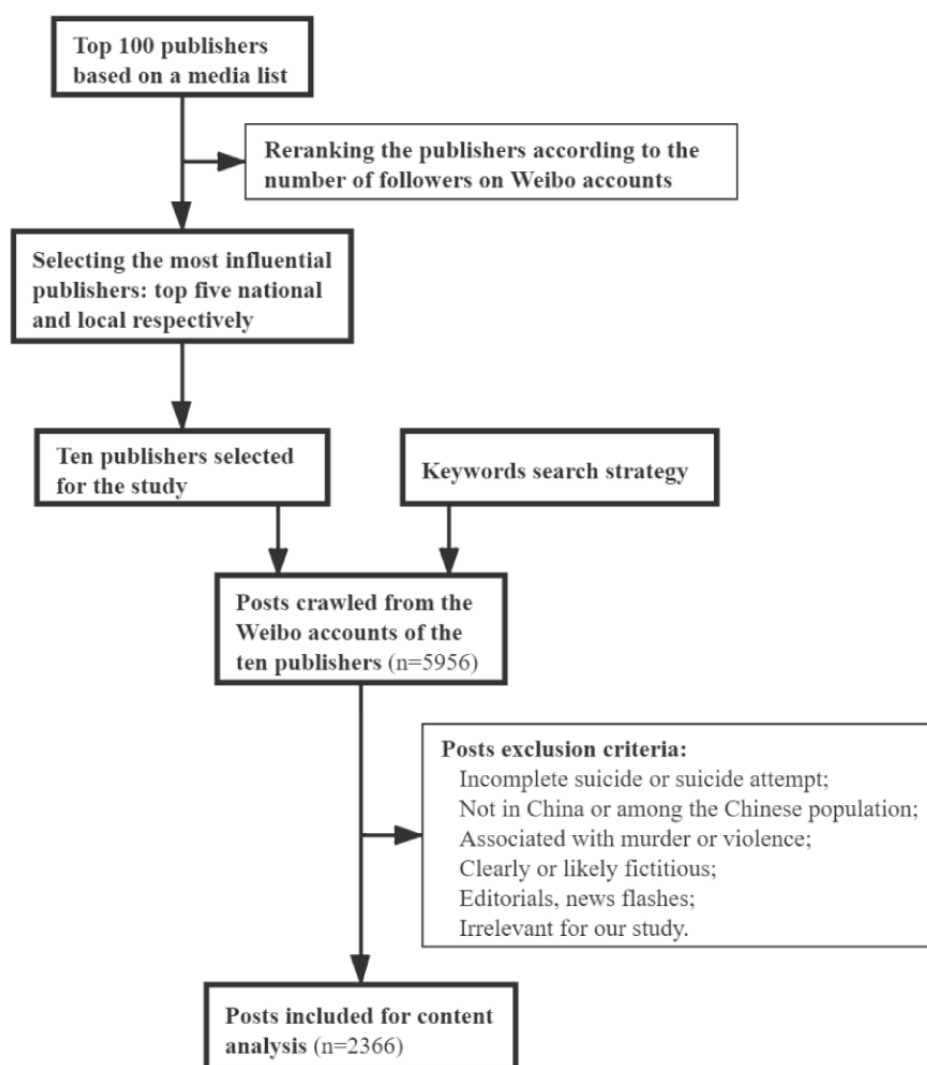
Since the number of followers directly affects the visibility of reports, the number of followers on Weibo, one of the most popular platforms in China [40,41], can represent the publisher's influence. We ranked publisher Weibo accounts according to the number of followers on November 1, 2019, and categorized them as national or local. We then selected the top 5 national (People's Daily, China Daily, Global Times, Guangming Daily, and China Youth News) and local (Beijing News, Shanghai Morning Post, New Express, Yangtze Evening Post, and Southern Metropolis Daily) newspapers for further study. Due to access restrictions, we were not able to collect data from New Express and replaced it with Chutian Metropolis Daily, which ranked sixth in local newspapers. Each publisher had more than 10 million followers on their Weibo accounts.

Data Extraction

We used the open-source web data crawling tool Octopus to crawl the data by keyword search via the application programming interface of Weibo. Specifically, we designed a keyword search strategy that contained a core word (ie,

"suicide"), a synonym (ie, "Zijin," which means suicide in Chinese), and 8 hyponyms of common suicide methods in China (ie, "jumping off a building," "jumping off a bridge," "jumping into a river," "drinking pesticides," "taking drugs," "wrist cutting," "hanging," and "charcoal burning") [42]. Initially, we obtained 5956 posts published between August 11, 2011, and November 16, 2019, with at least one of the above keywords. These posts were then filtered based on the following criteria taken from prior research: event depicted in the post must have been a completed suicide or suicide attempt [24,25]; event must have occurred in China and among the Chinese population (ie, posts on foreigners who died by suicide in China and on Chinese people who died by suicide overseas were omitted) [23]; suicide events associated with murder or violence were omitted [25]; post must have been a narrative (ie, fictional, editorials, and single-sentence posts, called flashes, were omitted) [23]; and posts irrelevant for our study were omitted (eg, a post that stated "Staying up late equals to chronic suicide"). In total, we obtained 2366 posts (2363 original posts and 3 reposts) about suicide reporting (Figure 1).

Figure 1. Flowchart of mainstream publishers selected and data extraction.



Content Analysis

We conducted content analysis based on prior evaluation research [21-26]. First, we extracted descriptive data to address the first research question that included the following 3 categories of suicide reporting characteristics: (1) text characteristics (ie, publisher name, publication time, and whether the post contained pictures or videos); (2) suicide event characteristics, including the type (ie, completed suicide or attempt), site (ie, urban or rural area), method (ie, jumping off a building, jumping off the bridge, jumping into the river, drinking pesticides, taking drugs, wrist cutting, hanging, charcoal burning, other ways, or multiple ways), and cause (ie, family conflict, emotional frustration, financial difficulty, interpersonal conflict, mental disorder, study or work pressure, to avoid responsibility, physical disease, loneliness or solitude, other causes, or multiple factors) [42]; and (3) demographic characteristics of people who die by suicide, including gender (ie, female, male, or group suicidal event), age (ie, juvenile, youth, middle-aged, or older adults group), marital status (ie, unmarried, married, divorced, or widowed), and economic activity status (ie, student, economically active–employed, economically active–unemployed, or economically inactive) [43].

Second, we designed a social media evaluation framework based on the WHO guidelines to assess the quality of suicide reporting by mainstream publishers. We divided the content of the suicide reports into the categories harmful and helpful information [20,24,36]. The harmful information category included level of detail of the suicide description, disclosure of private information, and vividness of the reporting.

Despite their usefulness, the WHO guidelines were derived from press practices of traditional media, and the recommendations were not aligned with the features of social media. Hence, we adjusted several of the original items to improve the flexibility and applicability of the evaluation framework; for example, because social media posts do not have a layout order, avoid the prominent placement of suicide reporting was modified to avoid the inclusion of suggestive signs or emojis in suicide reporting. To improve the evaluation framework's scientificity and operability, we removed the item about disclosure of private information about the person who died by suicide from harmful information after consulting with relevant experts because the meaning of *information* is too broad to operationalize. Finally, we achieved an evaluation framework containing 11 dichotomous items (ie, present=1, absent=0). The structure of the evaluation framework and item definitions and examples are shown in Table 1.

Table 1. Evaluation framework for the quality of social media suicide reporting.

Dimension, subdimension, and item	Example
Harmful information	
Level of detail of the suicide description	
Describes in detail the tools used for suicide or specific suicide actions, such as drug dosage.	"She took nearly 500 tablets of chlorpheniramine in her car."
Provides specific information about the site of suicide event, such as the exact name of the bridge, street, community, and so on.	"Weikang Hotel, Nujiang Town, Tongjiang County, Sichuan Province."
Disclosure of private information	
Includes photographs or videos showing images of the person who died.	— ^a
Describes the content of the suicide note in detail.	"He expressed his apology to parents and left his bank card password in the suicide note."
Includes interviews of immediate family members of the person who died by suicide.	"His father said in the interview that his daughter was often depressed after the accident."
Vividness of the reporting	
Attaches hashtags at the beginning or end of the post.	#The 13-year-old boy left home after being blamed by his father#
Uses suggestive characters or emoji to highlight the content.	"!!!," "~~"
Includes the word <i>suicide</i> directly in the headline or discloses the cause of suicide.	"Confession failed, a man climbed to a high building to suicide."
Helpful information	
Relates suicidal behaviors with mental disorders such as depression.	"This pregnant woman suffered from postpartum depression."
Provides information on where to seek help such as a psychological consultation hotline or other public health service.	"Beijing Huilongguan Hospital psychological crisis intervention hotline: 800-810-1117."
Provides professional knowledge about suicide prevention from psychologist experts or scholars.	"The psychiatrists said that [positive psychological intervention] can completely change the situation of mental illness."

^aNot applicable.

We selected 3 well-trained graduate students who had majored in journalism and communication for coding; they used an explicit coding framework. We conducted a reliability test based on 100 sample data items, finding the following: the interrater reliability of suicide reporting characteristics ranged from 0.78 to 0.98 with an average of .91 and that of suicide reporting quality ranged from 0.73 to 1.0, with an average of 0.88. Based on prior research, both parts of the framework indicated strong interrater reliability [44]. To bridge understanding deviations regarding the coding and reach consensus throughout the research process, we held regular meetings with the coders.

Results

Analysis of Suicide Reporting

The number of suicide reports varied considerably among the analyzed media accounts (see Table 2). Overall, the Global Times showed the largest number of suicide reports (21.26%, 503/2366), whereas Guangming Daily showed the smallest number (0.89%, 21/2366), both being national publishers. In local publishers, the Shanghai Morning Posts showed the largest number (19.65%, 465/2366), second only to the national publisher Global Times. From the perspective of time, the number of suicide reports varied considerably between 2011 to 2019; the year 2018 showed the largest number of suicide reports (19.57%, 463/2366) and 2011, the lowest (1.61%, 38/2366).

Table 2. Suicide reporting by national and local Chinese publishers (n=2366).

Newspaper name	Suicide reports, n (%)
Global Times	503 (21.26)
Shanghai Morning Post	465 (19.65)
Chutian Metropolis Daily	420 (17.75)
Beijing News	330 (13.95)
Yangtze Evening Post	246 (10.40)
People’s Daily	121 (5.11)
Southern Metropolis Daily	92 (3.89)
China Daily	87 (3.68)
China Youth News	81 (3.42)
Guangming Daily	21 (0.89)

Characteristics of Suicide Reporting on Social Media

Most reported suicide events occurred in urban areas (1454/1588, 91.56%; excluding 778 posts lacking information on site; see Table 3). Almost all suicide reporting disclosed the suicide method (2195/2366, 92.77%), with jumping off buildings (884/2195, 40.27%) and jumping into rivers (391/2195, 17.81%) being the most commonly reported. Despite being one of the most popular suicide methods in South and East Asia [45], reported events of drinking pesticide ranked fourth (172/2195, 7.84%).

Nearly 80% (1823/2366, 77.05%) of the reporting described the suicide attribution as stated in media reports, among these family conflict (446/1823, 24.47%), followed by emotional

frustration (295/1823, 16.18%), financial difficulty (199/1823, 10.92%), and interpersonal conflicts (196/1823, 10.75%).

Regarding the demographic characteristics of people who died by suicide (Table 3), the number of females (1096/2326, 47.12%) was very close to males (1117/2326, 48.02%), and 4.86% (113/2326) reports were group suicidal events. More than 60% (1446/2366, 61.12%) of suicide reporting disclosed the person’s age, of which the youth group accounted for more than half (730/1446, 50.48%), followed by the juvenile (424/1446, 29.32%), middle-aged (174/1446, 12.03%), and older adults groups (118/1446, 8.16%). In addition, suicide reports on social media more frequently reported on suicidal behaviors of unmarried groups and students.

Table 3. Descriptive characteristics of suicide reporting on social media (n=2366).

Characteristics	Values, n (%)
Suicide event characteristics	
Event (n=2366)	
Suicide attempt	1308 (55.28)
Suicide death	1058 (44.72)
Site (n=1588)	
Urban area	1454 (91.56)
Rural area	134 (8.44)
Method (n=2195)	
Jumping off building	884 (40.27)
Jumping into river	391 (17.81)
Other way	212 (9.66)
Drink pesticide	172 (7.84)
Jumping off bridge	165 (7.52)
Hanging	119 (5.42)
Taking medicine	94 (4.28)
Wrist cutting	60 (2.73)
Charcoal burning	58 (2.64)
Multiples ways	40 (1.82)
Attribution (factor) as stated in media reports (n=1823)	
Family conflict	446 (24.47)
Emotional frustration	295 (16.18)
Financial difficulty	199 (10.92)
Interpersonal conflict	196 (10.75)
Mental disorder	172 (9.43)
Study or work pressure	168 (9.22)
Other cause	157 (8.61)
Avoid responsibility	60 (3.29)
Physical disease	55 (3.02)
Loneliness or solitude	42 (2.30)
Multiple factors	33 (1.81)
Demographic characteristics	
Gender (n=2326)	
Male	1117 (48.02)
Female	1096 (47.12)
Group suicidal event	113 (4.86)
Age group (years; n=1446)	
Youth (19 to 35)	730 (50.48)
Juvenile (below 18)	424 (29.32)
Middle-age (36 to 65)	174 (12.03)
Older adults (over 65)	118 (8.16)
Marital status (n=1292)	
Unmarried	927 (71.75)

Characteristics	Values, n (%)
Married	324 (25.08)
Divorced	31 (2.40)
Widowed	10 (0.77)
Economic activity status (n=1062)	
Student	593 (55.84)
Economically active–employed	364 (34.27)
Economically inactive	60 (5.65)
Economically active–unemployed	45 (4.24)

Assessing Suicide Reporting Quality on Social Media Against WHO Guidelines

Our study demonstrated that harmful reporting practices were widespread on Chinese social media (Table 4); 41.50% (982/2366) of suicide reporting contained 3 or more harmful instances of information, while only 5.92% (140/2366) of the suicide reporting did not have any harmful information. Further, the vividness of the reporting subdimension had the highest mean (mean 0.35), followed by disclosure of private information (mean 0.27), and level of detail of the suicide description (mean 0.22).

Regarding vividness of the reporting, more than half of the suicide reporting either directly used the word “suicide” or disclosed the reasons for suicide in the headlines, practices that deviate considerably from the WHO guidelines to word headlines carefully. It is worth noting that the use of internet elements has become common on social media platforms; on the topic, this study showed that 37.19% (880/2366) of the suicide reporting used suggestive symbols (eg, multiple exclamation points or emojis). Furthermore, 13.99% (331/2366) of the suicide reporting attached hashtags at the beginning of the articles. These hashtags were highly visible because they were discussed by the public to a certain extent (eg, the hashtag #26-year-old female teacher fell off the building#).

Regarding the disclosure of private information, half (1172/2366, 49.54%) of the suicide reporting exposed images of the people who died, and nearly 30% (313/1172, 26.71%) were not pixelated. Additionally, 11.33% (268/2366) of suicide reporting disclosed the suicide note, with 41.8% (112/268) of the notes coming from the person’s social media platform (eg, on Weibo or WeChat). According to the WHO guidelines, social

media posts and email information of people who die by suicide should not be disclosed, as these suicide notes may contain sensitive private information (eg, their debts or information on their interpersonal relationships).

Regarding level of detail of the suicide description, almost 24% (568/2366, 24.01%) of the suicide reporting described the suicide methods in detail (eg, “She closed the window, locked the door, and lit the charcoal fire”). Moreover, 20.08% (475/2366) of the suicide reporting provided specific information on the suicide site (eg, the exact names of streets, communities, or bridges).

Our results also indicated that Chinese mainstream publishers’ suicide reporting on social media has significant deficiencies regarding the provision of helpful information: only 1 of all 2366 suicide reports included all the helpful information outlined in evaluation framework for this study; in contrast, more than 85% (2017/2366, 85.25%) of the reporting did not provide any helpful information. WHO considers it important to emphasize the causal relationship between suicidal behavior and mental illness for suicide prevention efforts [20]. However, only 10.86% (257/2366) of the analyzed suicide reporting highlighted this relationship. Moreover, less than 6% (137/2366, 5.79%) of the suicide reporting provided suicide prevention knowledge for the public (eg, psychology experts’ advice on suicide prevention or official suicide statistics to highlight the importance of the suicide issue). Furthermore, only 2 reports provided direct information about support programs, which are essential for vulnerable groups seeking help; in contrast, the WHO guidelines suggest mass media publishers list available sources of support (eg, psychological intervention hotlines and community resources) for those in need [20].

Table 4. Assessing suicide reporting quality on social media against the WHO guidelines (n=2366).

Items	Value, n (%)
Harmful information	
Vividness of the reporting	
Headline contains suicide-related words	1292 (54.61)
Suggestive symbols or emojis used	880 (37.91)
Hashtag used	331 (13.99)
Disclosure of private information	
Exposure of images of people who died by suicide	1172(49.54)
Interviews with relatives of people who died by suicide	376 (15.89)
Disclosure of the suicide note	268 (11.33)
Level of detail of the suicide description	
Detailed description of the suicide method	568 (24.01)
Detailed description of the suicide site	475 (20.08)
Helpful information	
Emphasize relationship between suicide and mental disorders	257 (10.86)
Provide suicide prevention knowledge	137 (5.79)
Provide information about support programs	2 (0.08)

Discussion

Principal Findings

This study analyzed the characteristics of suicide reporting of the top 10 publisher accounts (in terms of number of followers) on Sina Weibo and assessed their suicide reporting quality against the WHO guidelines. Our analyses yielded the following findings: (1) Chinese mainstream publishers on social media most frequently reported on suicides that were stated to be associated with conflict and underreported suicides in older adults and rural residents, (2) these suicide reports provided widespread harmful information, especially concerning vividness of the reporting and disclosure of private information, and (3) these reports provided limited helpful information such as direct information about suicide support programs.

Analysis of suicide reporting characteristics revealed two features. First, the reporting of suicide events was found to be clickbait oriented, in which publishers tried their best to attract readers with a sensational and dramatic reporting style. Suicide is a complex phenomenon with multiple contributing factors. However, in this study, almost all suicide reports interpreted suicide as an unexpected event caused by a single factor, with interpersonal conflicts and conflicts leading to suicide (eg, by family conflict or emotional frustration) being the most commonly reported causes. Only 1.81% of posts published a multicausal explanation of suicides. This may be explained by the publisher intention to catch the reader's eye and gain clicks by describing suicide news in a conflicted and dramatic manner [38].

Second, suicides of older adults and rural residents were underreported. Our study demonstrated that the analyzed social media accounts put a disproportionate amount of focus on young

adult suicide while rarely reporting on suicides among older adults; in contrast, according to China's official health data reports, the suicide rate of people aged over 65 years is the highest among all age groups. For example, the average suicide rate of urban residents aged 65 to 85 years in 2018 was about 5.5 times higher than that of people aged 20 to 40 years (13.10 vs 2.39 per 100,000) [46]. In addition, the suicide rate of rural residents in China has always been higher than that of urban residents [46]. The underreporting of older adult suicide found in this study is consistent with the findings of Fu et al [23] for traditional newspapers. This reporting bias may pose a considerable risk for suicide prevention efforts; for example, policymakers may acknowledge the lack of reports and promote an unbalanced distribution of health resources. This type of bias may lead to the neglect of the suicides of vulnerable people by the general population, thus hindering people's ability to intervene and support those at risk for suicide. Given the specific context of the Chinese population, which has been experiencing rapid aging [47], such bias and the risk that it produces warrant great focus from researchers and practitioners.

In terms of suicide reporting quality, our results revealed that the analyzed social media accounts had low adherence to the WHO guidelines; this result echoed results found in previous Asian research [23-26]. Regarding harmful information, vividness of the reporting showed the worst performance; for traditional printed media, to reduce the risk of imitative suicide, publishers are advised to avoid placing suicide reporting in prominent positions of the printed news [20,38]. The research of Chu et al [25] found low compliance—23% for newspaper articles—with the guideline to avoid prominent placement like the front page, boxes, or similar, but 85% for internet-based media articles. Our findings are likely to explain these differences further. Namely, although there is no fixed layout

on social media, our results indicated that it uses alternative ways to capture readers' attention, such as using diverse internet elements (eg, hashtags and emojis). However, there are currently no standardized guidelines for their use. We are worried that this type of reporting style may be turning a serious public health issue into something that may be entertaining or even frivolous.

We also found that disclosure of private information showed a moderately problematic performance; almost half of the suicide reporting on the analyzed social media accounts exposed images of the person who died by suicide. This finding is concordant with the study of Fu et al [23] of Chinese newspapers (57.5%) [23]; one possible explanation is that, as prior studies showed, Chinese people may be generally insensitive to privacy issues [48,49]. This conclusion can also be supported by research conducted in India, where 21.5% of the reporting contained photos of the person who died by suicide [24]. Also, a study on Facebook found that only 24% of suicide news disclosed a photo or video with information about the suicide site [36].

Our findings demonstrated that level of detail of the suicide description showed relevant noncompliance ratios; specifically, 24.01% of the suicide reporting provided an explicit description of the suicide method. This result is roughly consistent with prior studies on print newspapers in China and India [23,24]; however, since the speed and breadth of social media is greater than that of traditional media, we believe that even if it is roughly an equal proportion, social media may have a more profound negative impact.

Regarding helpful information, we found that the analyzed Chinese social media accounts provided very limited information on suicide prevention. WHO has suggested that mass media endeavor eliminate the general public's misconception about suicide and convey the view that mental illnesses and suicide are inseparable [20,50]. Unfortunately, less than 11% of the suicide reports in this study emphasized the relationship between mental health and suicidal behaviors. This result is consistent with previous research by Chu et al [25] which showed that less than 20% of the reported suicides (including in newspapers and websites) acknowledged the relationship between suicidal behavior and mental illness. However, in New Zealand, this proportion was more than 30% [51] and in Australia, it was more than 50% of the analyzed samples [15]. To some extent, this shows the different attribution patterns of suicidal behavior in Chinese and Western media. Regarding information on support programs, our results showed that only 2 reports attached direct information on support programs like psychological intervention hotlines. This result echoes the findings on traditional media, suggesting that both traditional and social media do an inefficient job of providing information about support programs [23-25]. In addition, a US study on suicide reporting on Facebook revealed that 16% of suicide news contained a hotline phone number, a higher percentage than in China but still low overall [36]. In summary, while social media has tremendous convenience, such as link support resources, the potential of mainstream publishers to provide helpful information on suicide is currently underused.

Implications

For decades, academic and practical fields have been concerned with the quality of suicide reporting in the mainstream media. However, we have little information on how mainstream publishers report suicide via social media platforms. Although a recent study assessed suicide news on Facebook in English-speaking countries [36], empirical evidence on suicide reporting on Asian social media platforms, especially in China, is still very limited. Thus, we endeavored to examine adherence to WHO suicide guidelines on Chinese social media platforms; we intended to address the above mentioned gap by empirically investigating these reporting qualities.

Various studies have indicated that the degree of compliance with the WHO guidelines are affected by complex factors, including economic and media policy issues present in different countries [24,36,52]. Among them, culture is an important factor that cannot be ignored. This is evident in the Chinese and Western attributions regarding suicide. It is generally believed that China is an acquaintance society that attaches great importance to interpersonal relations. Therefore, Chinese society (including the media) focuses on suicide caused by interpersonal conflicts such as marriage and family conflicts. In contrast, the West tends to explain suicide in terms of pathology, emphasizing suicides that occur due to mental illness [53]. Our findings also confirm the cultural differences in suicide attribution between China and the West. Our study complements the current literature on suicide reporting in different cultural contexts, thereby indirectly supporting the cross-cultural study of suicide.

Our study has enormous practical implications for national—and potentially international—suicide prevention programs. First, for media practitioners, since we found that there was generally low compliance with WHO guidelines, we reiterate Chinese mainstream publishers' role as gatekeepers on social media platforms by encouraging the strengthening of editorial review policies to limit harmful information and disseminate helpful information. Since media practitioners' understanding and support of the guidelines will also affect the implementation of the guidelines [52], these publishers should focus on training and educating their social media professionals to enhance their knowledge of critical public health issues (eg, suicide).

Second, our study clarifies the role of the Ministry of Health in promoting responsible media practices; given that in China there are currently no national suicide reporting guidelines [54], one of the first tasks may be to develop national guidelines that align with the current condition of the media in the country. We suggest that an updated suicide reporting guideline document should fully consider the internet elements. Furthermore, considering that social media is constantly evolving and richer forms of media are likely to emerge in the future, the Ministry of Health should support the implementation of these guidelines and the development of supervisory tools to monitor and evaluate the quality of suicide reporting on social media on an ongoing basis.

Third, social organizations (eg, mental health institutions) can actively cooperate with mainstream publishers to increase dissemination of helpful information related to suicide on social media platforms. Prior studies have shown that stakeholders'

active participation in the development of guidelines effectively boosts compliance with the guidelines [52,55]. Therefore, while social organizations provide professional knowledge on suicide prevention, social media can ensure that this knowledge is highly disseminated.

Last, users can play an active role in suicide prevention. On an interactive information platform, every click made by social media users can be regarded as a second transmission. The promotion and discussion of suicide topics on social networks may increase negative thoughts of suicide by members of vulnerable groups. It is necessary to help users safely convey information about suicide through social media [56]. Our results suggest that users should refrain from retweeting suicide reports that contain harmful information, such as those containing the private information of the people who died by suicide, and instead spread suicide-related articles with rich educational materials (eg, on psychological interventions and suicide prevention-related knowledge).

Limitations

This study has some limitations. First, since our data on suicide reporting came exclusively from the 10 most influential Chinese mainstream media accounts on Weibo, we did not include data from small-size media accounts, which may still play a

meaningful role in particular and local issues. Thus, future evaluation research may include a greater variety of media types. In addition, our study only assessed the suicide reporting of the analyzed publishers on Weibo; thus, future research can extend the analysis to other Chinese social media platforms (eg, WeChat) and attempt to conduct cross-cultural research on the topic of suicide on English-language social media platforms.

Conclusions

This study analyzed the characteristics of suicide reporting published by Chinese mainstream publishers via a social media platform and assessed the suicide reporting quality against WHO guidelines. Our findings illustrated that suicide reporting on social media by mainstream publishers in China disseminated too much harmful information, especially concerning the vividness of the reporting and disclosure of private information. Conversely, they provided minimal helpful information. Considering the tremendous information dissemination power of social media platforms, we highlight the need for the development of national suicide reporting guidelines that apply to the new media environment and local cultural backgrounds. We also recommend that social media practitioners, health care institutions, social organizations, and the public work together to promote responsible suicide reporting in the social media environment.

Acknowledgments

This work was partially supported by grant 19ZDA332 from the Major Project of the National Social Science Foundation of China.

Conflicts of Interest

None declared.

References

1. World Health Organization. Preventing suicide: a global imperative. 2014. URL: http://apps.who.int/iris/bitstream/10665/131056/1/9789241564779_eng.pdf?ua=1&ua=1 [accessed 2020-11-19]
2. Public health action for the prevention of suicide: a framework. Geneva: World Health Organization; 2012. URL: http://apps.who.int/iris/bitstream/10665/75166/1/9789241503570_eng.pdf?ua=1 [accessed 2020-12-03]
3. Maris RW. Suicide. *Lancet* 2002 Jul 27;360(9329):319-326. [doi: [10.1016/S0140-6736\(02\)09556-9](https://doi.org/10.1016/S0140-6736(02)09556-9)] [Medline: [12147388](https://pubmed.ncbi.nlm.nih.gov/12147388/)]
4. Hawton K, van Heeringen K. *The International Handbook of Suicide and Attempted Suicide*. Chichester: Wiley; 2000.
5. Sisask M, Värnik A. Media roles in suicide prevention: a systematic review. *Int J Environ Res Public Health* 2012 Jan;9(1):123-138 [FREE Full text] [doi: [10.3390/ijerph9010123](https://doi.org/10.3390/ijerph9010123)] [Medline: [22470283](https://pubmed.ncbi.nlm.nih.gov/22470283/)]
6. Hawton K, Williams K. The connection between media and suicidal behavior warrants serious attention. *Crisis* 2001;22(4):137-140. [doi: [10.1027//0227-5910.22.4.137](https://doi.org/10.1027//0227-5910.22.4.137)] [Medline: [11848654](https://pubmed.ncbi.nlm.nih.gov/11848654/)]
7. Phillips DP. The influence of suggestion on suicide: substantive and theoretical implications of the Werther effect. *Am Sociol Rev* 1974;39(3):340-354. [Medline: [11630757](https://pubmed.ncbi.nlm.nih.gov/11630757/)]
8. Pirkis J, Blood RW. Suicide and the media. Part I: reportage in nonfictional media. *Crisis* 2001;22(4):146-154. [doi: [10.1027//0227-5910.22.4.146](https://doi.org/10.1027//0227-5910.22.4.146)] [Medline: [11848658](https://pubmed.ncbi.nlm.nih.gov/11848658/)]
9. Stack S. Suicide in the media: a quantitative review of studies based on non-fictional stories. *Suicide Life Threat Behav* 2005 Apr;35(2):121-133. [doi: [10.1521/suli.35.2.121.62877](https://doi.org/10.1521/suli.35.2.121.62877)] [Medline: [15843330](https://pubmed.ncbi.nlm.nih.gov/15843330/)]
10. Niederkrotenthaler T, Voracek M, Herberth A, Till B, Strauss M, Etzersdorfer E, et al. Role of media reports in completed and prevented suicide: Werther v. Papageno effects. *Br J Psychiatry* 2010 Sep;197(3):234-243. [doi: [10.1192/bjp.bp.109.074633](https://doi.org/10.1192/bjp.bp.109.074633)] [Medline: [20807970](https://pubmed.ncbi.nlm.nih.gov/20807970/)]
11. Gould M, Jamieson P, Romer D. Media contagion and suicide among the young. *Am Behav Sci* 2003 May 01;46(9):1269-1284. [doi: [10.1177/0002764202250670](https://doi.org/10.1177/0002764202250670)] [Medline: [22973420](https://pubmed.ncbi.nlm.nih.gov/22973420/)]

12. Niederkrotenthaler T, Fu K, Yip PSF, Fong DYT, Stack S, Cheng Q, et al. Changes in suicide rates following media reports on celebrity suicide: a meta-analysis. *J Epidemiol Community Health* 2012 Nov;66(11):1037-1042. [doi: [10.1136/jech-2011-200707](https://doi.org/10.1136/jech-2011-200707)] [Medline: [22523342](https://pubmed.ncbi.nlm.nih.gov/22523342/)]
13. Cheng ATA, Hawton K, Lee CTC, Chen THH. The influence of media reporting of the suicide of a celebrity on suicide rates: a population-based study. *Int J Epidemiol* 2007 Dec;36(6):1229-1234. [doi: [10.1093/ije/dym196](https://doi.org/10.1093/ije/dym196)] [Medline: [17905808](https://pubmed.ncbi.nlm.nih.gov/17905808/)]
14. Michel K, Frey C, Wyss K, Valach L. An exercise in improving suicide reporting in print media. *Crisis* 2000;21(2):71-79. [doi: [10.1027//0227-5910.21.2.71](https://doi.org/10.1027//0227-5910.21.2.71)] [Medline: [11019482](https://pubmed.ncbi.nlm.nih.gov/11019482/)]
15. Pirkis J, Dare A, Blood RW, Rankin B, Williamson M, Burgess P, et al. Changes in media reporting of suicide in Australia between 2000/01 and 2006/07. *Crisis* 2009;30(1):25-33. [doi: [10.1027/0227-5910.30.1.25](https://doi.org/10.1027/0227-5910.30.1.25)] [Medline: [19261565](https://pubmed.ncbi.nlm.nih.gov/19261565/)]
16. Fu KW, Yip PSF. Changes in reporting of suicide news after the promotion of the WHO media recommendations. *Suicide Life Threat Behav* 2008 Oct;38(5):631-636. [doi: [10.1521/suli.2008.38.5.631](https://doi.org/10.1521/suli.2008.38.5.631)] [Medline: [19014313](https://pubmed.ncbi.nlm.nih.gov/19014313/)]
17. Etzersdorfer E, Sonneck G. Preventing suicide by influencing mass-media reporting. The Viennese experience 1980–1996. *Arch Suicide Res* 1998 Jan;4(1):67-74. [doi: [10.1080/13811119808258290](https://doi.org/10.1080/13811119808258290)]
18. Niederkrotenthaler T, Herberth A, Sonneck G. [The "Werther-effect": legend or reality?]. *Neuropsychiatr* 2007;21(4):284-290. [Medline: [18082110](https://pubmed.ncbi.nlm.nih.gov/18082110/)]
19. Niederkrotenthaler T, Sonneck G. Assessing the impact of media guidelines for reporting on suicides in Austria: interrupted time series analysis. *Aust N Z J Psychiatry* 2007 May;41(5):419-428. [doi: [10.1080/00048670701266680](https://doi.org/10.1080/00048670701266680)] [Medline: [17464734](https://pubmed.ncbi.nlm.nih.gov/17464734/)]
20. Preventing suicide: a resource for media professionals. Geneva: World Health Organization; 2008. URL: https://apps.who.int/iris/bitstream/handle/10665/43954/9789241597074_eng.pdf?sequence=1&isAllowed=y [accessed 2020-11-19]
21. Tatum PT, Canetto SS, Slater MD. Suicide coverage in U.S. newspapers following the publication of the media guidelines. *Suicide Life Threat Behav* 2010 Oct;40(5):524-534 [FREE Full text] [doi: [10.1521/suli.2010.40.5.524](https://doi.org/10.1521/suli.2010.40.5.524)] [Medline: [21034215](https://pubmed.ncbi.nlm.nih.gov/21034215/)]
22. Creed M, Whitley R. Assessing fidelity to suicide reporting guidelines in Canadian news media: the death of Robin Williams. *Can J Psychiatry* 2017 May;62(5):313-317 [FREE Full text] [doi: [10.1177/0706743715621255](https://doi.org/10.1177/0706743715621255)] [Medline: [27600531](https://pubmed.ncbi.nlm.nih.gov/27600531/)]
23. Fu K, Chan Y, Yip PSF. Newspaper reporting of suicides in Hong Kong, Taiwan and Guangzhou: compliance with WHO media guidelines and epidemiological comparisons. *J Epidemiol Community Health* 2011 Oct;65(10):928-933. [doi: [10.1136/jech.2009.105650](https://doi.org/10.1136/jech.2009.105650)] [Medline: [20889589](https://pubmed.ncbi.nlm.nih.gov/20889589/)]
24. Armstrong G, Vijayakumar L, Niederkrotenthaler T, Jayaseelan M, Kannan R, Pirkis J, et al. Assessing the quality of media reporting of suicide news in India against World Health Organization guidelines: A content analysis study of nine major newspapers in Tamil Nadu. *Aust N Z J Psychiatry* 2018 Sep;52(9):856-863. [doi: [10.1177/0004867418772343](https://doi.org/10.1177/0004867418772343)] [Medline: [29726275](https://pubmed.ncbi.nlm.nih.gov/29726275/)]
25. Chu X, Zhang X, Cheng P, Schwebel DC, Hu G. Assessing the use of media reporting recommendations by the World Health Organization in suicide news published in the most influential media sources in China, 2003-2015. *Int J Environ Res Public Health* 2018 Mar 05;15(3):1 [FREE Full text] [doi: [10.3390/ijerph15030451](https://doi.org/10.3390/ijerph15030451)] [Medline: [29510591](https://pubmed.ncbi.nlm.nih.gov/29510591/)]
26. Arafat SMY, Khan MM, Niederkrotenthaler T, Ueda M, Armstrong G. Assessing the quality of media reporting of suicide deaths in Bangladesh against World Health Organization guidelines. *Crisis* 2020 Jan;41(1):47-53. [doi: [10.1027/0227-5910/a000603](https://doi.org/10.1027/0227-5910/a000603)] [Medline: [31140319](https://pubmed.ncbi.nlm.nih.gov/31140319/)]
27. Bertolote /, Fleischmann A. A global perspective in the epidemiology of suicide. *Suicidology* 2015 Jun 11;7(2):6-8. [doi: [10.5617/suicidologi.2330](https://doi.org/10.5617/suicidologi.2330)]
28. Whiting A, Williams D. Why people use social media: a uses and gratifications approach. *Qualitative Mrkt Res: An Int J* 2013 Aug 30;16(4):362-369. [doi: [10.1108/QMR-06-2013-0041](https://doi.org/10.1108/QMR-06-2013-0041)]
29. Moorhead SA, Hazlett DE, Harrison L, Carroll JK, Irwin A, Hoving C. A new dimension of health care: systematic review of the uses, benefits, and limitations of social media for health communication. *J Med Internet Res* 2013;15(4):e85 [FREE Full text] [doi: [10.2196/jmir.1933](https://doi.org/10.2196/jmir.1933)] [Medline: [23615206](https://pubmed.ncbi.nlm.nih.gov/23615206/)]
30. Xu S, Markson C, Costello KL, Xing CY, Demissie K, Llanos AA. Leveraging social media to promote public health knowledge: example of cancer awareness via twitter. *JMIR Public Health Surveill* 2016;2(1):e17 [FREE Full text] [doi: [10.2196/publichealth.5205](https://doi.org/10.2196/publichealth.5205)] [Medline: [27227152](https://pubmed.ncbi.nlm.nih.gov/27227152/)]
31. García-Avilés JA, Kaltenbrunner A, Meier K. Media convergence revisited. *J Pract* 2014 Feb 28;8(5):573-584. [doi: [10.1080/17512786.2014.885678](https://doi.org/10.1080/17512786.2014.885678)]
32. Cao MY. [The writing and characteristics of Weibo news reports—take the news reports in People's Daily Weibo as an example]. *J News Res* 2015;6(11):152-153.
33. Tian H, Chang J. [The cultural transformation of party newspaper in the social media era: a case study based on the emotional expression of People's Daily]. *Shanghai J Re* 2019;1:1. [doi: [10.16057/j.cnki.31-1171/g2.2019.01.008](https://doi.org/10.16057/j.cnki.31-1171/g2.2019.01.008)]
34. Xiong Y. [Research on the status quo of People's Daily official Weibo communication]. *Youth J* 2014;17:84-85. [doi: [10.15997/j.cnki.qnjz.2014.17.032](https://doi.org/10.15997/j.cnki.qnjz.2014.17.032)]
35. Alvarez-Mon MA, Asunsolo Del Barco A, Lahera G, Quintero J, Ferre F, Pereira-Sanchez V, et al. Increasing interest of mass communication media and the general public in the distribution of tweets about mental disorders: observational study. *J Med Internet Res* 2018 May 28;20(5):e205 [FREE Full text] [doi: [10.2196/jmir.9582](https://doi.org/10.2196/jmir.9582)] [Medline: [29807880](https://pubmed.ncbi.nlm.nih.gov/29807880/)]

36. Sumner SA, Burke M, Kooti F. Adherence to suicide reporting guidelines by news shared on a social networking platform. *Proc Natl Acad Sci U S A* 2020 Jul 14;117(28):16267-16272 [FREE Full text] [doi: [10.1073/pnas.2001230117](https://doi.org/10.1073/pnas.2001230117)] [Medline: [32631982](https://pubmed.ncbi.nlm.nih.gov/32631982/)]
37. World Health Organization. World health statistics overview 2019: monitoring health for the sustainable development goals. 2019. URL: <https://apps.who.int/iris/bitstream/handle/10665/311696/WHO-DAD-2019.1-eng.pdf?sequence=1&isAllowed=y> [accessed 2020-11-19]
38. Chiang Y, Chung F, Lee C, Shih H, Lin D, Lee M. Suicide reporting on front pages of major newspapers in Taiwan violating reporting recommendations between 2001 and 2012. *Health Commun* 2016 Nov;31(11):1395-1404. [doi: [10.1080/10410236.2015.1074024](https://doi.org/10.1080/10410236.2015.1074024)] [Medline: [27007575](https://pubmed.ncbi.nlm.nih.gov/27007575/)]
39. China media convergence communication index in 2018. people.cn. URL: <http://media.people.com.cn/n1/2019/0326/c120837-30994743.html> [accessed 2020-11-19]
40. Qi G, Abel F, Houben G, Yong Y. A comparative study of users' microblogging behavior on Sina Weibo and Twitter. 2012 Presented at: Proceedings of the International Conference on User Modeling; 2012; Montreal. [doi: [10.1007/978-3-642-31454-4_8](https://doi.org/10.1007/978-3-642-31454-4_8)]
41. Tian X, Yu G, He F. An analysis of sleep complaints on Sina Weibo. *Comput Human Behav* 2016 Sep;62:230-235. [doi: [10.1016/j.chb.2016.04.014](https://doi.org/10.1016/j.chb.2016.04.014)]
42. Lu P. [Content analysis of media suicide news: a perspective of spiritual healthy communication]. *J Commun* 2005;12:31-41.
43. Au JSK, Yip PSF, Chan CLW, Law YW. Newspaper reporting of suicide cases in Hong Kong. *Crisis* 2004;25(4):161-168. [doi: [10.1027/0227-5910.25.4.161](https://doi.org/10.1027/0227-5910.25.4.161)] [Medline: [15580851](https://pubmed.ncbi.nlm.nih.gov/15580851/)]
44. McHugh ML. Interrater reliability: the kappa statistic. *Biochem Med (Zagreb)* 2012;22(3):276-282 [FREE Full text] [Medline: [23092060](https://pubmed.ncbi.nlm.nih.gov/23092060/)]
45. Eddleston M, Karalliedde L, Buckley N, Fernando R, Hutchinson G, Isbister G, et al. Pesticide poisoning in the developing world—a minimum pesticides list. *Lancet* 2002 Oct 12;360(9340):1163-1167. [doi: [10.1016/s0140-6736\(02\)11204-9](https://doi.org/10.1016/s0140-6736(02)11204-9)] [Medline: [12387969](https://pubmed.ncbi.nlm.nih.gov/12387969/)]
46. China Health Statistics Yearbook. China Union Medical University Press. 2019. URL: <https://data.cnki.net/area/Yearbook/Single/N2020020200?z=D09> [accessed 2020-11-19]
47. Nie J. Erosion of eldercare in China: a socio-ethical inquiry in aging, elderly suicide and the government's responsibilities in the context of the one-child policy. *Ageing Int* 2016 Nov 7;41(4):350-365. [doi: [10.1007/s12126-016-9261-7](https://doi.org/10.1007/s12126-016-9261-7)]
48. McDougall BS. Privacy in modern China. *History Compass* 2004 Jan;2(1):1-8. [doi: [10.1111/j.1478-0542.2004.00097.x](https://doi.org/10.1111/j.1478-0542.2004.00097.x)]
49. Wang Y, Norice G, Cranor L. Who is concerned about what? A study of American, Chinese and Indian users' privacy concerns on social network sites. *Trust Trustworthy Comput* 2011;6740:146-153. [doi: [10.1007/978-3-642-21599-5_11](https://doi.org/10.1007/978-3-642-21599-5_11)]
50. Nock MK, Hwang I, Sampson N, Kessler RC, Angermeyer M, Beautrais A, et al. Cross-national analysis of the associations among mental disorders and suicidal behavior: findings from the WHO World Mental Health Surveys. *PLoS Med* 2009 Aug;6(8):e1000123 [FREE Full text] [doi: [10.1371/journal.pmed.1000123](https://doi.org/10.1371/journal.pmed.1000123)] [Medline: [19668361](https://pubmed.ncbi.nlm.nih.gov/19668361/)]
51. McKenna B, Thom K, Edwards G. Reporting of Suicide in New Zealand Media: Content and Case Study Analysis. Auckland: Centre for Metal Health Research; 2010.
52. Bohanna I, Wang X. Media guidelines for the responsible reporting of suicide: a review of effectiveness. *Crisis* 2012;33(4):190-198. [doi: [10.1027/0227-5910/a000137](https://doi.org/10.1027/0227-5910/a000137)] [Medline: [22713977](https://pubmed.ncbi.nlm.nih.gov/22713977/)]
53. Qin P, Mortensen PB. Specific characteristics of suicide in China. *Acta Psychiatr Scand* 2001 Feb;103(2):117-121. [doi: [10.1034/j.1600-0447.2001.00008.x](https://doi.org/10.1034/j.1600-0447.2001.00008.x)] [Medline: [11167314](https://pubmed.ncbi.nlm.nih.gov/11167314/)]
54. Beautrais A, Hendin H, Yip P. Improving portrayal of suicide in the media in Asia. In: Hendin H, editor. *Suicide and Suicide Prevention in Asia*. Geneva: World Health Organization; 2008:39-50.
55. Collings SC, Kemp CG. Death knocks, professional practice, and the public good: the media experience of suicide reporting in New Zealand. *Soc Sci Med* 2010 Jul;71(2):244-248. [doi: [10.1016/j.socscimed.2010.03.017](https://doi.org/10.1016/j.socscimed.2010.03.017)] [Medline: [20398990](https://pubmed.ncbi.nlm.nih.gov/20398990/)]
56. Robinson J, Hill NTM, Thorn P, Battersby R, Teh Z, Reavley NJ, et al. The #chatsafe project. Developing guidelines to help young people communicate safely about suicide on social media: a Delphi study. *PLoS One* 2018 Nov 15;13(11):e0206584 [FREE Full text] [doi: [10.1371/journal.pone.0206584](https://doi.org/10.1371/journal.pone.0206584)] [Medline: [30439958](https://pubmed.ncbi.nlm.nih.gov/30439958/)]

Abbreviations

WHO: World Health Organization

Edited by G Eysenbach; submitted 19.12.20; peer-reviewed by G Hu, A Teo, M Alvarez de Mon, S Byrne; comments to author 13.01.21; revised version received 07.03.21; accepted 15.04.21; published 13.05.21.

Please cite as:

Lai K, Li D, Peng H, Zhao J, He L

Assessing Suicide Reporting in Top Newspaper Social Media Accounts in China: Content Analysis Study

JMIR Ment Health 2021;8(5):e26654

URL: <https://mental.jmir.org/2021/5/e26654>

doi: [10.2196/26654](https://doi.org/10.2196/26654)

PMID: [33983127](https://pubmed.ncbi.nlm.nih.gov/33983127/)

©Kaisheng Lai, Dan Li, Huijuan Peng, Jingyuan Zhao, Lingnan He. Originally published in JMIR Mental Health (<https://mental.jmir.org>), 13.05.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Mental Health, is properly cited. The complete bibliographic information, a link to the original publication on <https://mental.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Intention to Use Behavioral Health Data From a Health Information Exchange: Mixed Methods Study

Randyl A Cochran¹, PhD; Sue S Feldman², RN, MEd, PhD; Nataliya V Ivankova², MPH, PhD; Allyson G Hall², PhD; William Opoku-Agyeman³, MSc, MPH, PhD

¹Department of Health Sciences, College of Health Professions, Towson University, Towson, MD, United States

²Department of Health Services Administration, School of Health Professions, University of Alabama at Birmingham, Birmingham, AL, United States

³School of Health and Applied Human Sciences, College of Health and Human Services, University of North Carolina Wilmington, Wilmington, NC, United States

Corresponding Author:

Randyl A Cochran, PhD

Department of Health Sciences

College of Health Professions

Towson University

8000 York Road

Linthicum Hall 121C

Towson, MD, 21252

United States

Phone: 1 410 704 2345

Email: rcochran@towson.edu

Abstract

Background: Patients with co-occurring behavioral health and chronic medical conditions frequently overuse inpatient hospital services. This pattern of overuse contributes to inefficient health care spending. These patients require coordinated care to achieve optimal health outcomes. However, the poor exchange of health-related information between various clinicians renders the delivery of coordinated care challenging. Health information exchanges (HIEs) facilitate health-related information sharing and have been shown to be effective in chronic disease management; however, their effectiveness in the delivery of integrated care is less clear. It is prudent to consider new approaches to sharing both general medical and behavioral health information.

Objective: This study aims to identify and describe factors influencing the intention to use behavioral health information that is shared through HIEs.

Methods: We used a mixed methods design consisting of two sequential phases. A validated survey instrument was emailed to clinical and nonclinical staff in Alabama and Oklahoma. The survey captured information about the impact of predictors on the intention to use behavioral health data in clinical decision making. Follow-up interviews were conducted with a subsample of participants to elaborate on the survey results. Partial least squares structural equation modeling was used to analyze survey data. Thematic analysis was used to identify themes from the interviews.

Results: A total of 62 participants completed the survey. In total, 63% (n=39) of the participants were clinicians. Performance expectancy ($\beta=.382$; $P=.01$) and trust ($\beta=.539$; $P<.001$) predicted intention to use behavioral health information shared via HIEs. The interviewees (n=5) expressed that behavioral health information could be useful in clinical decision making. However, privacy and confidentiality concerns discourage sharing this information, which is generally missing from patient records altogether. The interviewees also stated that training for HIE use was not mandatory; the training that was provided did not focus specifically on the exchange of behavioral health information.

Conclusions: Despite barriers, individuals are willing to use behavioral health information from HIEs if they believe that it will enhance job performance and if the information being transmitted is trustworthy. The findings contribute to our understanding of the role HIEs can play in delivering integrated care, particularly to vulnerable patients.

(JMIR Ment Health 2021;8(5):e26746) doi:[10.2196/26746](https://doi.org/10.2196/26746)

KEYWORDS

behavioral health; integrated care; health information exchange; behavioral intention; patient care; mixed methods research

Introduction

Background

Behavioral health conditions adversely affect an individual's well-being, quality of life, and life expectancy. Behavioral health conditions are defined as mental illnesses (eg, depression, anxiety, schizophrenia, and bipolar disorder) and substance use disorders that impair the functioning of an individual [1]. Patients with these conditions require the delivery of team-based care to achieve optimal health outcomes [2,3]. Nearly half of the adult population will meet the criteria for a mental illness during their lifetime [4]. Recent projections indicate that in the United States, approximately 19% of the adult population may have a mental illness at any point in time [5-7]. Behavioral health conditions are expected to surpass chronic medical conditions to become the leading cause of disability worldwide [8], with depression alone projected to become the second most prevalent cause of disability [7].

The relationship between physical health and mental health is complex [9-11]. Patients diagnosed with a behavioral health condition frequently have chronic medical comorbidities [4,5,12-15]. The association of mental illness with chronic disease contributes to the inefficient use of health care services (eg, increased inpatient hospital utilization) [5,16-18]. Other factors such as infrequent use of primary care and preventive services, poor chronic disease management, and the use of antipsychotic medications [19-23] also contribute to the overutilization of hospital services. Patients with behavioral health conditions, coupled with chronic medical comorbidities, contribute substantially to health care spending. A population-based cohort study in Alberta, Canada, revealed that patients with a chronic medical illness and a co-occurring mental health condition had a higher mean of 3-year adjusted health care costs (Can \$38,250 [US \$31,700]) than chronically ill patients without a mental health condition (Can \$22,250 [US \$18,440]) [24]. A similar pattern is expected in the United States. In fact, the costs are likely to be higher [25]. The costs alone indicate that it is both financially and clinically prudent to consider new ways to manage behavioral health conditions in chronically ill patients [25]. A novel approach to control costs is a coordinated care delivery system.

Behavioral health patients require coordinated care but often fail to receive it [17,26]. This level of care requires better communication and information sharing between general medical and behavioral health care providers. Therefore, it is necessary to develop new approaches that facilitate the effective treatment of both medical and behavioral health conditions [27]. Health information exchanges (HIEs), a component of health information technology (IT), can facilitate the integration of different health care services provided to behavioral health patients with chronic conditions.

HIEs are organizations that enable the digital exchange of health-related data [28], and they are one type of health IT thought to facilitate such integration. The exchange of health information through HIEs has many benefits, including safer, more efficient care that effectively manages both the behavioral and physical health needs of individual patients [13,29-31].

HIEs lead to efficient information sharing between providers in emergencies that require prompt, accurate diagnosis and treatment [29]. The effectiveness of HIEs has been studied in care coordination in general, especially in chronic disease management, but there are few studies that examine HIEs in the context of behavioral health [6,32-35]. The dearth of studies in behavioral health and HIE is a result of rising challenges with behavioral health data. Specifically, there is a disjoint between the clinical language, codes, and data reporting between behavioral health and general medicine among providers [6]. In addition, federal regulations complicate behavioral health information sharing. One of these regulations, that is, Section 42 Part 2, was adopted in the Code of Federal Regulations (42 CFR Part 2) in 1972 to ensure privacy protection for individuals receiving treatment for substance use disorders. It prohibited the disclosure of records related to substance use treatment without express authorization from the patient (with certain exceptions in emergency situations) [36]. Revisions of 42 CFR Part 2 in 2017 and 2018 have aimed to modernize the regulation and to facilitate the integration of behavioral health in general medical settings. Despite these efforts, challenges remain [6,36]. Taken together, these factors create barriers to exchanging behavioral health information and delivering integrated care for this patient population.

Therefore, the aim of this study is to identify and describe various factors that may influence health care providers' intention to use and actual use of behavioral health information obtained from an HIE. This is accomplished by using a mixed methods approach guided by the theoretical backdrop of the unified theory of acceptance and use of technology (UTAUT) and diffusion of innovation (DOI). This study addresses the following research questions:

1. What factors are associated with the intention to use behavioral health information obtained from HIEs?
2. From the perspective of health care providers, what are the facilitators of and barriers to behavioral health information use from HIEs?

Literature Review and Conceptual Framework

Integrated behavioral health care has demonstrated effectiveness in treating patients with behavioral health disorders and co-occurring medical conditions [37,38]. Such models of care delivery can improve the quality of care by delivering treatments that align with the clinical needs of patients [38]. A systematic review revealed that among 4 randomized controlled trials, 2 of the studies showed minor improvements in physical health, whereas the other 2 studies showed no significant changes in physical health [39]. Furthermore, improvements in seeking preventive care were also observed. However, none of these studies examined the effect of these interventions on functional or clinical outcomes [39]. Improvements in both access and quality of care have been attributed to communication and information sharing between clinicians, particularly when treating patients with complex medical and behavioral needs [40]. This finding suggests that establishing effective communication channels is crucial for the success of integrated care delivery models. However, there are barriers to

communication between medical and behavioral health care providers.

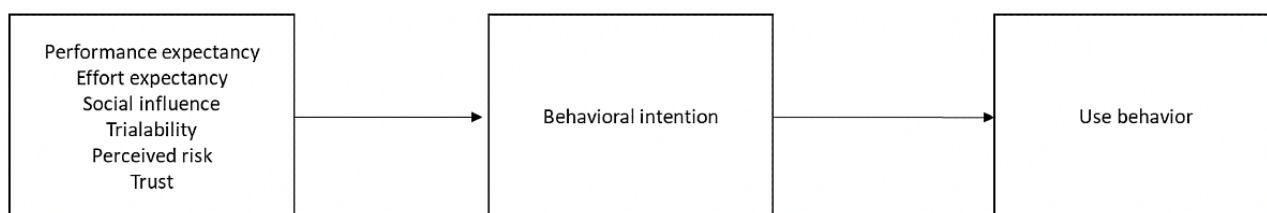
A 2011 study examined the perspectives of behavioral health care providers regarding the benefits and barriers of HIE utilization [41]. According to the study, behavioral health care providers stated that the use of HIEs would (1) improve the quality of care and communication between different types of providers, (2) increase cost and time burdens, (3) raise concerns related to access to and vulnerability of patient and client information, and (4) impact workflow and control. Similarly, a 2019 qualitative study explored the perspectives of patients with mental health conditions to gain an understanding of privacy in the context of exchanging personal health information [42]. All 14 participants acknowledged the importance of privacy in health care, particularly concerning information related to sensitive diagnoses (eg, mental health conditions). However, the degree of concern varied and was most frequently related to negative experiences and perceptions related to privacy. The interviewees expressed that, overall, they trusted that health care professionals and organizations would take appropriate measures to protect the privacy of information related to mental health diagnoses. The participants were largely uninformed about existing policies and regulations designed to protect the privacy of mental health information, and those who had negative experiences related to privacy expressed doubt about the effectiveness of these regulations. However, the participants acknowledged that the electronic exchange of health information was a practical approach to ensuring high-quality patient care. It is important to note that this study was conducted in Canada, where a single-payer system exists.

The 2 theories that are deemed appropriate to explore our overall research questions are the UTAUT and the DOI theory.

UTAUT integrates constructs from earlier IT adoption theories to examine their influence on individuals' intentions to use technology and the subsequent use of the technology [43]. UTAUT explains nearly 70% of the variance related to the intention to use technology [44]. The original UTAUT model attributes behavioral intention and subsequent technology use to 4 main constructs: (1) *performance expectancy* (perceptions of the ability of the technology to assist in meeting job-related goals), (2) *effort expectancy* (ease of use of technology), (3) *social influence* (perception that important figures support the use of technology), and (4) *facilitating conditions* (infrastructure that supports the use of technology) [43,45]. Extensions of the UTAUT model have considered the influence of additional factors on behavioral intention to use technology, such as trust and perceived risk [46].

DOI examines 5 factors that determine the rate of adoption of a new technology (innovation) within an organization: (1) *relative advantage* (the innovation is perceived as a better solution than its predecessor), (2) *compatibility* (the innovation fits with the organization's culture and norms), (3) *complexity* (the innovation is easy to understand and use), (4) *trialability* (individuals can experiment with innovation before adoption), and (5) *observability* (the results of using the innovation are visible to others) [47,48]. The trialability construct was extracted from this theory because it was not captured in the UTAUT framework. The exchange of behavioral health data via an HIE is a relatively novel concept, and it is possible that pilot testing this type of data exchange could influence behavioral intention. The conceptual model for this study, which is derived from these 2 theories, is shown in Figure 1. The research questions were identified using this model.

Figure 1. Conceptual model.



Methods

Overview

A sequential explanatory mixed methods design was used in this study. This mixed methods design enables researchers to provide more thorough explanations of the trends that emerge from the quantitative phase of the study by further exploring participants' perspectives [49,50]. The quantitative and qualitative study phases are described in the following sections.

Quantitative Phase

Sampling

Survey participants were selected through convenience sampling. Alabama and Oklahoma were selected for participation because both these states have established HIEs and reported making some efforts to incorporate behavioral

health data into patient records. General medical care providers (physicians, nurses, and physician assistants) and nonclinical staff (directors, network administrators, and patient experience representatives) were identified as potential participants.

Data Collection

A validated survey instrument was emailed to clinicians and staff in Alabama and Oklahoma [45,46]. The survey prompt and questionnaire items are provided in [Multimedia Appendix 1](#). The influence of several predictors was examined through the conceptual model developed for the study: (1) performance expectancy, (2) effort expectancy, (3) social influence, (4) perceived risk, (5) trust, and (6) trialability. The first 5 constructs were derived from the UTAUT model [45]. The sixth construct, trialability, was derived from the DOI model [47,48]. The 6 constructs from the UTAUT and DOI frameworks were the independent variables for this study.

The final survey was adapted from the UTAUT questionnaire items developed by Venkatesh et al [45]. Two of the researchers (RC and SF) developed, pilot tested with 4 respondents not involved in the study, and refined the survey (eg, question order and wording) before it was administered to the study sites. The survey used in this study contained statements related to each construct. Respondents expressed their level of agreement with each statement using a 7-point Likert scale, with 1 being *Strongly agree* and 7 being *Strongly disagree*. The constructs within the conceptual model have been found to have high reliability and validity in previous studies: standardized factor loadings and average variance extracted values exceeded the recommended threshold of 0.50 [46,51,52]. Furthermore, the composite reliability values were above 0.90 for each of the constructs, surpassing the minimum recommended value of 0.70 [46,53]. Similar findings have emerged for the behavioral intention construct [46]. Preliminary analysis of the survey data collected for this study also met the recommended thresholds for reliability and validity.

An initial prompt to participate in the survey was sent via email. The prompt briefly described the study (eg, purpose, procedures, risks, and benefits). In addition, the prompt included a link to the web-based survey and a consent form that provided more details about the study. Data collection began on June 4, 2018, and ended on August 14, 2018.

Variables

Behavioral intention was the outcome of the combination of the independent variables. It was also a predictor of use behavior. Similar to the independent variables, behavioral intention was operationalized as an ordinal-level variable. Survey items related to behavioral intention were answered using the same 7-point Likert scale adopted for the independent variables.

Use behavior was operationalized as a dichotomous variable. Participants answered *Yes* or *No* to the initial screening question: "Have you ever used behavioral health information obtained from the HIE your organization uses?" Participants who responded *Yes* answered questions related to performance expectancy, effort expectancy, social influence, perceived risk, and trust. Participants who responded *No* answered an additional set of questions pertaining to trialability. On the basis of the response to the initial screening question, the wording of the subsequent questions was modified to reflect actual experiences versus perceptions of obtaining behavioral health information from the HIE.

Finally, the survey instrument included questions designed to collect demographic information, which included data related to gender, education, job title, age, race or ethnicity, and level of comfort and experience with technology. This study was conducted with the approval of the University of Alabama at Birmingham IRB #300001017.

Data Analysis

For the survey data, preliminary descriptive statistics were generated. Furthermore, we tested the reliability and validity of the survey items. We used partial least squares structural equation modeling (PLS-SEM) to test the relationships between

the constructs in the conceptual model. PLS-SEM is a nonparametric alternative to covariance-based structural equation modeling, which makes it more applicable to this study, as we have a small sample and data that are not normally distributed [54-57]. PLS-SEM can be used in both exploratory and confirmatory studies [58]. All statistical analyses were performed using Stata version 15 at an α level of .05.

Qualitative Phase

Sampling

Semistructured interviews were conducted with a subsample of the survey participants to better understand the role of these factors from the subjects' perspectives and to discover any additional factors that were not captured in the survey. A total of 18% (11/62) survey participants provided their contact information for a follow-up interview at the end of the quantitative survey. Of these, 5 completed the interviews.

Interview Guide

Three of the authors (RC, SF, and NI) collaborated on the development of the interview guide for this phase. The interview questions were developed to elaborate on predictors of intention to use behavioral health information from an HIE that emerged from the initial quantitative survey. Therefore, the interview guide was designed to facilitate a closer examination of the constructs from the UTAUT and DOI frameworks. Likewise, the interview questions also captured information about findings that were contradictory to previous research to better understand the emergence of these inconsistencies [59]. Before deployment, the interview guide was pilot tested with 4 individuals with expertise in health informatics (n=3) and psychiatric nursing (n=1). Minor revisions were made to the wording and ordering of the interview questions. [Multimedia Appendix 2](#) provides the final interview guide.

Data Collection

The first author (RC) conducted the interviews with the survey respondents. An initial request for interview email was sent to the participants, and follow-up emails were sent 2 and 4 weeks after the initial request. Interviews were conducted either in person or via teleconference (Zoom), based on interviewee preference and timing. All interviews were recorded and transcribed using the Otter Voice Notes mobile app. The interviews were conducted from March 19 to May 2, 2019.

Data Analysis

Verbatim interview transcripts were analyzed using inductive and deductive thematic analysis with NVivo 12. The development of the interview questions was partially guided by the constructs in the conceptual model; additional themes were derived from the content within the interview transcripts [60]. The following 4-stage analytical process was adopted for the study: (1) becoming familiar with data, (2) generating themes by aggregating coded segments of data, (3) creating comparative categories, and (4) revising and refining the themes [61]. Two of the authors (RC and SF) engaged in peer debriefing to discuss the interview process and the resulting themes, thus generating additional insights into the interview findings [62]. To increase the trustworthiness of data, we used member checking:

participants received a 1- to 2-page summary of the interview transcript from the interviewee (RC) by email and were asked to provide clarification or ask additional questions. Of the 5 participants, 2 responded to the follow-up email. One participant provided additional clarification for one of the discussion points; the other participant found no errors in the summary document. Quantitative and qualitative findings were further integrated to create meta-inferences using a joint display [63].

Results

Due to the sequential nature of this mixed methods study design, we present the results in sequence: quantitative findings followed by qualitative findings and then followed by the integration of quantitative and qualitative findings.

Quantitative

Descriptive Statistics

Descriptive statistics of the survey participants are presented in [Table 1](#). A total of 90% (56/62) of the respondents reported that they did not exchange behavioral health data via the HIE. A majority of the respondents (49/62, 79%) were female, and 83% (52/62) of the respondents were aged between 30 and 59 years. A large proportion of the health care providers in this sample were identified as nurses (27/62, 44%). Nonclinical roles (23/62, 37%) consisted of directors, network administrators, and patient experience representatives. With regard to the reported level of computer experience, most participants (45/62, 72%) identified as *average* users.

Table 1. Descriptive statistics of behavioral health information exchange survey participants (N=62).

Variables	Values, n (%)
Retrieved behavioral health data from their organization's HIE^a	
Yes	6 (10)
No	56 (90)
Gender	
Female	49 (79)
Male	12 (19)
Prefer not to state	1 (2)
Age group (years)	
21-29	3 (5)
30-39	18 (29)
40-49	18 (29)
50-59	16 (25)
60-69	6 (10)
≥70	1 (2)
Job title	
Nurses	27 (44)
Physicians	10 (16)
Physician's assistants	2 (3)
Others	23 (37)
Education	
High school diploma or GED ^b	3 (5)
Some college	8 (13)
2-year degree	20 (32)
4-year college degree	10 (16)
Master's degree	10 (16)
Professional degree	10 (16)
Doctorate	1 (2)
Hispanic or Latino origin	
Yes	3 (5)
No	59 (95)
Race	
American Indian or Alaska Native	2 (3)
Asian	5 (8)
Black or African American	6 (10)
White	47 (76)
Other	2 (3)
Level of computer experience	
Novice	1 (2)
Average	45 (73)
Advanced	16 (26)

^aHIE: health information exchange.

^bGED: general education development.

Measurement Validity

The measurement model determined the reliability and validity of the survey items associated with each latent construct in the conceptual model. Cronbach α scores of .7 or higher indicate internal consistency [64,65]. The measurement model is illustrated in [Multimedia Appendix 3](#). With the exception of

one item on the performance expectancy construct, all items met the criteria for internal consistency.

PLS-SEM Results

The results of PLS-SEM are presented in [Table 2](#). The standardized path coefficients (β) denote the strength and direction of the relationship between the constructs in the conceptual model.

Table 2. Path analysis of the constructs within the conceptual model.

Predictors of intention to use behavioral health information from the HIE ^a	β ^b	P value
Performance expectancy	.382	.01
Effort expectancy	.055	.72
Social influence	-.043	.72
Perceived risk	.061	.47
Trust	.539	<.001
Triability	.093	.34
Use behavior	.127	.68

^aHIE: health information exchange.

^b β represents the standardized path coefficients from the partial least squares structural equation modeling results. The standardized path coefficients indicate the strength and direction of the relationship between the constructs.

The PLS-SEM results revealed that performance expectancy and trust are the significant predictors of behavioral intention to exchange behavioral health information via HIEs. On average, survey participants who reported higher scores on the performance expectancy measures also reported higher scores on the behavioral intention measures ($\beta=.383$; $P=.01$). Similarly, participants who scored higher on the trust measures also had higher scores on the behavioral intention measures ($\beta=.539$; $P<.001$). In this study, trust was found to be the strongest predictor of behavioral intention. Furthermore, behavioral intention was not significantly associated with the actual use of behavioral health information from HIEs.

Qualitative

Interview Participants and Overview

Of the 5 individuals who participated in the interviews, only 1 was employed in Oklahoma. The majority of the participants were nonclinical employees, with only 1 participant (a nurse practitioner) involved in direct patient care. The length of the interviews ranged from 19 minutes to 60 minutes, with an average interview length of 34 (SD 16) minutes.

Themes From the Interviews

A total of 6 themes emerged from the interviews: (1) usefulness of behavioral health information in care delivery, (2) regulations restricting behavioral HIE, (3) behavioral health and stigma, (4) missing or difficult-to-locate behavioral health data, (5) lack of mandatory training for behavioral HIE, and (6) future utilization of the HIEs. Each theme is explored in detail in the following sections. More details about each theme and the related participants' quotes are given in [Multimedia Appendix 4](#).

Usefulness of Behavioral Health Information in Care Delivery

Sharing behavioral health information can improve care delivery. In various settings, exchanging health data via an HIE increases efficiency. More specifically, exchanging behavioral health information in an emergency department could help the attending physician to become familiar with a patient's medical history and could be used in discharge planning. Despite these perceived benefits, there are several barriers that prohibit the exchange of behavioral health information.

Regulations Restricting Behavioral HIE

Existing regulations and policies make behavioral HIE challenging. One of the most significant perceived barriers to behavioral HIE, 42 CFR Part 2, makes health care workers hesitant to share sensitive information electronically. 42 CFR Part 2 leaves room for interpretation with regard to the exchange of behavioral health information, and it is not uncommon for health care administrators to err on the side of caution.

Overall, the participants expressed doubt that behavioral HIE will become a common practice in the near future, given the influence of 42 CFR Part 2. The restrictions surrounding the exchange of behavioral health information will most likely discourage the sharing of this information, unless it is deemed essential to the provision of care. Even with the recent relaxation of 42 CFR Part 2, one interviewee specifically stated that behavioral health data will never be effectively shared via HIE: "[42 CFR Part 2] simply discourages it." 42 CFR Part 2 was enacted as a response to the stigma surrounding sensitive diagnoses, including those related to behavioral health. This stigma persists even today.

Behavioral HIE and Stigma

Stigma continues to surround behavioral health conditions and contributes to the suppression of behavioral health information in medical records. As such, mental illness continues to be considered taboo, and because of this, information related to a mental health diagnosis is often separated from the rest of a patient's medical record. Furthermore, the existing stigma may discourage patients from revealing their diagnoses to a new health care provider. This hesitation to disclose this information hinders the flow of health-related information and complicates the delivery of appropriate care, especially if the provider is unfamiliar with the patient's medical history.

Missing or Difficult-to-Locate Behavioral Health Information

One of the concerns raised about incorporating behavioral health information into the electronic medical record was centered on locating the information within the HIE. One interviewee suggested that if behavioral health information is difficult to identify, it could reduce efficiency in care provision. In addition, removing behavioral health patients' records from the HIE further complicates the delivery of appropriate care.

General medical practitioners who access the HIE do not specifically look for behavioral health information. As such, participants do not know where this information would be located within the patient records. A participant (interviewee 4) suggested that it might be beneficial to provide training to assist health care providers in identifying behavioral health information in the patient record. Adequate training could not only reduce the potential workflow issues associated with behavioral HIE but could also facilitate the provision of appropriate care.

Lack of Mandatory Training for Behavioral HIE

The local network administrator for the Alabama site (interviewee 2) trained participants on the proper use of the HIE. There was no component of the training that focused exclusively on the exchange of behavioral health information via the HIE. The training encompassed basic functions (eg, how to log in, what information can be found within the system, and where to locate it). The training provided was not mandatory. In fact, some participants learned about HIEs via word of mouth. One participant (interviewee 4) believed that providing formal training could allow care providers to provide input regarding the usefulness of clinical data. Likewise, the training could help them to understand why there are restrictions on exchanging certain types of clinical data, such as behavioral health information.

The lack of mandatory training for HIE use made it challenging for providers to figure out how to integrate behavioral health information. However, the interviewees expressed hope that they will be able to increase the utilization of HIEs in the future.

Future Utilization of HIEs

The interviewees acknowledged the value of exchanging general health data across HIEs and hoped to increase their utilization in the near future. Some participants also saw value in using the systems to exchange behavioral health information between multiple providers, despite the numerous barriers that currently

exist. One participant (interviewee 4) suggested that these HIEs might be necessary to provide effective treatment in appropriate care settings to these difficult-to-treat patients.

Integration of the Quantitative and Qualitative Findings

We further integrated the quantitative and qualitative findings using a joint display. The joint display enabled us to compare the quantitative and qualitative results, explain similarities and differences between the findings, and develop meta-inferences regarding the exchange of behavioral health information via an HIE ([Multimedia Appendix 5](#)). In summary, the participants stated that having access to a patient's full medical record, including behavioral health diagnoses, would be useful in delivering appropriate treatment. Despite the potential benefits of having access to this information, there are several barriers that prohibit the exchange of behavioral health information. These barriers are primarily related to technical capabilities, policies and regulations, stigma, and training. Even with these challenges, the interviewees acknowledged the value of having complete information and expressed some hope that behavioral HIE will become standard practice in the future.

Discussion

Principal Findings

The purpose of this mixed methods study is to identify and describe factors that may influence the intention to exchange and use behavioral health information obtained from an HIE. The preliminary findings from the quantitative phase of the study indicate that performance expectancy and trust are significant predictors of behavioral intention to exchange behavioral health information via HIEs. In other words, participants believed that exchanging behavioral health information via HIEs would improve their job performance. However, 90 % (56/62) of survey participants reported that they did not use HIEs to exchange behavioral health information, which likely contributed to the insignificant association between behavioral intention and actual use. The misalignment between performance expectancy and behavioral intention is supported in the literature [66,67]. One explanation for this finding is that understanding the potential job-related benefits of using a technology might not motivate health care professionals to adopt it if there are other contextual factors to consider [67].

Within health care, trust is a predictor of the intention to use technology among both patients and physicians [68,69]. Our findings were consistent with the literature in that participants suggested that the lack of full disclosure of behavioral health conditions, possibly due to stigma, may produce an incomplete record and not provide practitioners with valuable clinical information at the point of care.

In summary, 2 constructs (performance expectancy and trust) emerged as significant predictors in the initial quantitative phase of the study. One theme (usefulness of behavioral health information in care delivery) emerged from qualitative interviews that provided further insights into the significance of the performance expectancy construct. However, there were several divergent findings. Perceived risk did not emerge as a

significant predictor in the quantitative phase; however, the interviewees stated that regulations and stigma are barriers to behavioral HIE. The interviewees suggested some connection between trust and stigma of behavioral health diagnoses. One interviewee stated that people were hesitant to disclose behavioral health information; therefore, the information that could truly be helpful at the point of care is often missing in terms of behavioral health (interviewee 1). Trialability did not emerge as a significant predictor of behavioral intention in the quantitative phase; however, the interviewees discussed a lack of in-depth training in the subsequent qualitative phase. Finally, behavioral intention was not a significant predictor of the actual use of HIEs for the purpose of behavioral HIE. This is likely due to the fact that participation in behavioral HIE efforts at these 2 study sites is low. However, interviewees expressed hope that in the future, HIEs will be used to exchange the full patient record, including any behavioral health diagnoses or treatment.

Implications and Future Research

As behavioral health becomes a growing public health concern, health care practitioners have largely accepted the need for integrated care [70,71]. Early integrated care models have shown improvements in patients' mental and physical health [71-74]. In addition, health care providers have expressed the belief that omitting behavioral health data from patient records hinders their ability to make appropriate treatment recommendations [71]. The findings of this study provide additional support to these notions.

The successful delivery of integrated care requires a reconsideration of existing policies and regulations as well as the establishment of clear guidelines and best practices for handling sensitive health-related information [75]. In the wake of the COVID-19 pandemic, restrictions on the sharing of personal health information have been relaxed. The requirement to obtain patient consent may be waived if the provider determines that there is a bona fide medical emergency [76]. It is crucial for stakeholders to consider whether this regulatory change should be retained following the pandemic.

Limitations

Although the findings are valid in the context of this study, the small sample size may limit its generalizability. This is a common limitation of studies that are conducted during early uses of IT, thus necessitating the use of nonparametric methods such as PLS-SEM [77]. With the exchange of behavioral health

data via HIE being relatively nascent (at the time of this study), this limitation is expected. Furthermore, the relationships between these constructs cannot be deemed causal but rather provide some insights into perceptions and reality. Similarly, because there were few participants who exchanged behavioral health information via HIEs, there was not enough variation in the responses to determine whether there were different sets of predictors for those who used the HIEs to exchange behavioral health information and those who did not. It is possible that the use of convenience sampling contributed to the homogeneity in the responses. However, the lack of active participants in behavioral HIE efforts could also be partially explained by the novelty of both the Alabama and Oklahoma HIEs.

Even in light of these limitations, the findings provide early positive suggestions for the use of behavioral health data shared by HIEs. This study is the first to adopt a mixed methods approach to examine the use of HIEs to share patients' behavioral health information. The use of mixed methods allowed the researchers to identify and explore factors that may influence the adoption and use of HIEs to facilitate behavioral HIE. This study was conducted across multiple states. Finally, the study is applicable to both the health services and health IT arenas. It is relevant to the health services literature because it acknowledges the importance of a whole-person approach to care. Likewise, it is relevant to health IT because it considers the role of technology in delivering integrated care to patients with complex needs.

Conclusions

Patients with behavioral health conditions are readmitted to the hospital 30 days after discharge at a higher rate than patients without behavioral health disorders. Several factors contribute to this pattern of overutilization. Existing regulations have traditionally restricted the exchange of behavioral health information unless the patient has given authorization at the person level, meaning that the specific person must be identified. The prevalence of suboptimal treatment outcomes for behavioral health patients with chronic illnesses indicates that it is necessary to commit more efforts to providing higher quality care to some of the most vulnerable patients.

Health care providers acknowledge that it is necessary to deliver holistic care to improve the quality of care and health-related outcomes for this subset of difficult-to-treat patients. This study contributes to our understanding of the potential role of HIEs in integrating the traditionally fragmented behavioral and general health service arenas.

Acknowledgments

The authors would like to thank Brian Yeaman, who shares the relevance and importance of behavioral health data exchange, for the support and guidance. The authors would also like to thank Darrell Burke for helping with the completion of this study. In addition, the authors would like to thank colleagues, friends, and subject matter experts who participated in the pilot testing of the survey instrument and interview protocol. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Authors' Contributions

RC and SF developed a conceptual framework for this study. RC collected and analyzed the quantitative and qualitative data and created a joint display for the quantitative and qualitative findings with meta-inferences. SF served as the subject matter expert on HIEs, helped to develop the quantitative survey, and provided general project oversights. NI guided the selection and implementation of the mixed methods approach and facilitated the creation of the joint display. NI and SF assisted with the development of the qualitative interview guide. AH made substantial contributions to the framing of the study and emphasized the relevance of the findings to existing policies and regulations. WOA assisted with quantitative data analysis and the creation of tables and figures for the manuscript. All authors (RC, SF, NI, AH, and WOA) contributed to the writing and editing of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Survey prompt and questionnaire items.

[[DOCX File , 17 KB - mental_v8i5e26746_app1.docx](#)]

Multimedia Appendix 2

Interview questions.

[[DOCX File , 15 KB - mental_v8i5e26746_app2.docx](#)]

Multimedia Appendix 3

Measurement model.

[[PNG File , 169 KB - mental_v8i5e26746_app3.png](#)]

Multimedia Appendix 4

Themes and quotes from qualitative interviews (phase 2).

[[DOCX File , 14 KB - mental_v8i5e26746_app4.docx](#)]

Multimedia Appendix 5

Joint display of quantitative and qualitative findings with meta-inferences.

[[DOCX File , 19 KB - mental_v8i5e26746_app5.docx](#)]

References

1. Davis M, Balasubramanian BA, Waller E, Miller BF, Green LA, Cohen DJ. Integrating behavioral and physical health care in the real world: early lessons from advancing care together. *J Am Board Fam Med* 2013;26(5):588-602 [[FREE Full text](#)] [doi: [10.3122/jabfm.2013.05.130028](#)] [Medline: [24004711](#)]
2. Mechanic D. Seizing opportunities under the Affordable Care Act for transforming the mental and behavioral health system. *Health Aff (Millwood)* 2012 Feb;31(2):376-382. [doi: [10.1377/hlthaff.2011.0623](#)] [Medline: [22323168](#)]
3. Farmanova E, Baker G, Cohen D. Combining integration of care and a population health approach: a scoping review of redesign strategies and interventions, and their impact. *Int J Integr Care* 2019 Apr 11;19(2):5 [[FREE Full text](#)] [doi: [10.5334/ijic.4197](#)] [Medline: [30992698](#)]
4. American Hospital Association. Bringing behavioral health into the care continuum: opportunities to improve quality, costs and outcomes. American Hospital Association. 2012. URL: <https://www.aha.org/guidesreports/2012-01-20-bringing-behavioral-health-care-continuum-opportunities-improve-quality> [accessed 2017-06-10]
5. Miller JE, Glover RW, Gordon SY. Crossing the behavioral health digital divide: the role of health information technology in improving care for people with behavioral health conditions in state behavioral health systems. National Association of State Mental Health Program Directors. 2014. URL: <https://docplayer.net/9632568-Crossing-the-behavioral-health-digital-divide.html> [accessed 2017-07-15]
6. Cifuentes M, Davis M, Fernald D, Gunn R, Dickinson P, Cohen DJ. Electronic health record challenges, workarounds, and solutions observed in practices integrating behavioral health and primary care. *J Am Board Fam Med* 2015 Oct;28 Suppl 1:63-72 [[FREE Full text](#)] [doi: [10.3122/jabfm.2015.S1.150133](#)] [Medline: [26359473](#)]
7. Steinberg J, Daniel Jr JW. Depression as a major mental health problem for the behavioral health care industry. *J Health Sci Manag Pub Health*. 2020. URL: http://www.medportal.ge/journal/2003/volume4_n1/deppresione.pdf [accessed 2021-05-19]
8. Crowley RA, Kirschner N, Health and Public Policy Committee of the American College of Physicians. The integration of care for mental health, substance abuse, and other behavioral health conditions into primary care: executive summary

- of an American College of Physicians position paper. *Ann Intern Med* 2015 Aug 18;163(4):298-299. [doi: [10.7326/M15-0510](https://doi.org/10.7326/M15-0510)] [Medline: [26121401](https://pubmed.ncbi.nlm.nih.gov/26121401/)]
9. Kutlubaev MA, Hackett ML. Part II: predictors of depression after stroke and impact of depression on stroke outcome: an updated systematic review of observational studies. *Int J Stroke* 2014 Dec 26;9(8):1026-1036. [doi: [10.1111/ijss.12356](https://doi.org/10.1111/ijss.12356)] [Medline: [25156411](https://pubmed.ncbi.nlm.nih.gov/25156411/)]
 10. Shi Y, Yang D, Zeng Y, Wu W. Risk factors for post-stroke depression: a meta-analysis. *Front Aging Neurosci* 2017 Jul 11;9:218 [FREE Full text] [doi: [10.3389/fnagi.2017.00218](https://doi.org/10.3389/fnagi.2017.00218)] [Medline: [28744213](https://pubmed.ncbi.nlm.nih.gov/28744213/)]
 11. Thayabaranathan T, Andrew NE, Kilkenny MF, Stolwyk R, Thrift AG, Grimley R, et al. Factors influencing self-reported anxiety or depression following stroke or TIA using linked registry and hospital data. *Qual Life Res* 2018 Dec 4;27(12):3145-3155. [doi: [10.1007/s11136-018-1960-y](https://doi.org/10.1007/s11136-018-1960-y)] [Medline: [30078162](https://pubmed.ncbi.nlm.nih.gov/30078162/)]
 12. Goodell S, Druss B, Walker E. Mental disorders and medical comorbidity. Robert Wood Johnson Foundation. 2011. URL: <https://www.rwjf.org/en/library/research/2011/02/mental-disorders-and-medical-comorbidity.html> [accessed 2018-02-10]
 13. Office of the National Coordinator for Health Information Technology. URL: <https://www.healthit.gov/> [accessed 2019-03-08]
 14. Wu L, Ghitza UE, Zhu H, Spratt S, Swartz M, Mannelli P. Substance use disorders and medical comorbidities among high-need, high-risk patients with diabetes. *Drug Alcohol Depend* 2018 May 01;186:86-93 [FREE Full text] [doi: [10.1016/j.drugalcdep.2018.01.008](https://doi.org/10.1016/j.drugalcdep.2018.01.008)] [Medline: [29554592](https://pubmed.ncbi.nlm.nih.gov/29554592/)]
 15. Cheung S, Spaeth-Rublee B, Shalev D, Li M, Docherty M, Levenson J, et al. A model to improve behavioral health integration into serious illness care. *J Pain Symptom Manage* 2019 Sep;14(3):503-514. [doi: [10.1016/j.jpainsymman.2019.05.017](https://doi.org/10.1016/j.jpainsymman.2019.05.017)] [Medline: [31175941](https://pubmed.ncbi.nlm.nih.gov/31175941/)]
 16. Schiefelbein EL, Olson JA, Moxham JD. Patterns of health care utilization among vulnerable populations in Central Texas using data from a regional health information exchange. *J Health Care Poor Underserved* 2014 Feb;25(1):37-51. [doi: [10.1353/hpu.2014.0020](https://doi.org/10.1353/hpu.2014.0020)] [Medline: [24509011](https://pubmed.ncbi.nlm.nih.gov/24509011/)]
 17. Perrin J, Reimann B, Capobianco J, Wahrenberger JT, Sheitman BB, Steiner BD. A model of enhanced primary care for patients with severe mental illness. *N C Med J* 2018 Jul 10;79(4):240-244 [FREE Full text] [doi: [10.18043/ncm.79.4.240](https://doi.org/10.18043/ncm.79.4.240)] [Medline: [29991617](https://pubmed.ncbi.nlm.nih.gov/29991617/)]
 18. Germack HD, Caron A, Solomon R, Hanrahan NP. Medical-surgical readmissions in patients with co-occurring serious mental illness: a systematic review and meta-analysis. *Gen Hosp Psychiatry* 2018 Nov;55:65-71. [doi: [10.1016/j.genhosppsych.2018.09.005](https://doi.org/10.1016/j.genhosppsych.2018.09.005)] [Medline: [30414592](https://pubmed.ncbi.nlm.nih.gov/30414592/)]
 19. Desai MM, Rosenheck RA, Druss BG, Perlin JB. Mental disorders and quality of diabetes care in the veterans health administration. *Am J Psychiatry* 2002 Sep;159(9):1584-1590. [doi: [10.1176/appi.ajp.159.9.1584](https://doi.org/10.1176/appi.ajp.159.9.1584)] [Medline: [12202281](https://pubmed.ncbi.nlm.nih.gov/12202281/)]
 20. Druss BG, Rosenheck RA, Desai MM, Perlin JB. Quality of preventive medical care for patients with mental disorders. *Med Care* 2002 Feb;40(2):129-136. [doi: [10.1097/00005650-200202000-00007](https://doi.org/10.1097/00005650-200202000-00007)] [Medline: [11802085](https://pubmed.ncbi.nlm.nih.gov/11802085/)]
 21. Alakeson V, Frank RG, Katz RE. Specialty care medical homes for people with severe, persistent mental disorders. *Health Aff (Millwood)* 2010 May;29(5):867-873. [doi: [10.1377/hlthaff.2010.0080](https://doi.org/10.1377/hlthaff.2010.0080)] [Medline: [20439873](https://pubmed.ncbi.nlm.nih.gov/20439873/)]
 22. Weinstein LC, Stefancic A, Cunningham AT, Hurley KE, Cabassa LJ, Wender RC. Cancer screening, prevention, and treatment in people with mental illness. *CA Cancer J Clin* 2016 Dec 10;66(2):134-151 [FREE Full text] [doi: [10.3322/caac.21334](https://doi.org/10.3322/caac.21334)] [Medline: [26663383](https://pubmed.ncbi.nlm.nih.gov/26663383/)]
 23. Compton MT, Manseau MW, Dacus H, Wallace B, Seserman M. Chronic disease screening and prevention activities in mental health clinics in New York state: current practices and future opportunities. *Community Ment Health J* 2020 May 4;56(4):717-726. [doi: [10.1007/s10597-019-00532-3](https://doi.org/10.1007/s10597-019-00532-3)] [Medline: [31902049](https://pubmed.ncbi.nlm.nih.gov/31902049/)]
 24. Sporinova B, Manns B, Tonelli M, Hemmelgarn B, MacMaster F, Mitchell N, et al. Association of mental health disorders with health care utilization and costs among adults with chronic disease. *JAMA Netw Open* 2019 Aug 02;2(8):e199910 [FREE Full text] [doi: [10.1001/jamanetworkopen.2019.9910](https://doi.org/10.1001/jamanetworkopen.2019.9910)] [Medline: [31441939](https://pubmed.ncbi.nlm.nih.gov/31441939/)]
 25. Schwenk TL. Comorbid mental illness is associated with doubled healthcare costs in patients with chronic disease. *NEJM Journal Watch*. 2019. URL: <https://www.jwatch.org/na49812/2019/08/29/comorbid-mental-illness-associated-with-doubled-healthcare> [accessed 2020-05-15]
 26. Benzer JK, Singer SJ, Mohr DC, McIntosh N, Meterko M, Vimalananda VG, et al. Survey of patient-centered coordination of care for diabetes with cardiovascular and mental health comorbidities in the Department of Veterans Affairs. *J Gen Intern Med* 2019 May 16;34(Suppl 1):43-49 [FREE Full text] [doi: [10.1007/s11606-019-04979-8](https://doi.org/10.1007/s11606-019-04979-8)] [Medline: [31098975](https://pubmed.ncbi.nlm.nih.gov/31098975/)]
 27. Wulsin L, Pinkhasov A, Cunningham C, Miller L, Smith A, Oros S. Innovations for integrated care: The Association of Medicine and Psychiatry recognizes new models. *Gen Hosp Psychiatry* 2019 Nov;61:90-95. [doi: [10.1016/j.genhosppsych.2019.04.007](https://doi.org/10.1016/j.genhosppsych.2019.04.007)] [Medline: [31104827](https://pubmed.ncbi.nlm.nih.gov/31104827/)]
 28. Callan K, Fuller J, Galterio L, Just B, Reich K, Steigerwald C, et al. Making health information exchange work. *J AHIMA* 2014;85(11):32-36. [Medline: [25682655](https://pubmed.ncbi.nlm.nih.gov/25682655/)]
 29. Hu LL, Sparenborg S, Tai B. Privacy protection for patients with substance use problems. *Subst Abuse Rehabil* 2011;2:227-233 [FREE Full text] [doi: [10.2147/SAR.S27237](https://doi.org/10.2147/SAR.S27237)] [Medline: [24474860](https://pubmed.ncbi.nlm.nih.gov/24474860/)]
 30. Druss BG, Mauer BJ. Health care reform and care at the behavioral health--primary care interface. *Psychiatr Serv* 2010 Nov;61(11):1087-1092. [doi: [10.1176/ps.2010.61.11.1087](https://doi.org/10.1176/ps.2010.61.11.1087)] [Medline: [21041346](https://pubmed.ncbi.nlm.nih.gov/21041346/)]

31. Vest JR, Gamm LD. Health information exchange: persistent challenges and new strategies. *J Am Med Inform Assoc* 2010;17(3):288-294 [FREE Full text] [doi: [10.1136/jamia.2010.003673](https://doi.org/10.1136/jamia.2010.003673)] [Medline: [20442146](https://pubmed.ncbi.nlm.nih.gov/20442146/)]
32. Bates DW, Bitton A. The future of health information technology in the patient-centered medical home. *Health Aff (Millwood)* 2010 Apr;29(4):614-621. [doi: [10.1377/hlthaff.2010.0007](https://doi.org/10.1377/hlthaff.2010.0007)] [Medline: [20368590](https://pubmed.ncbi.nlm.nih.gov/20368590/)]
33. Vest J, Kern L, Silver M, Kaushal R, HITEC investigators. The potential for community-based health information exchange systems to reduce hospital readmissions. *J Am Med Inform Assoc* 2015 Mar;22(2):435-442. [doi: [10.1136/amiajnl-2014-002760](https://doi.org/10.1136/amiajnl-2014-002760)] [Medline: [25100447](https://pubmed.ncbi.nlm.nih.gov/25100447/)]
34. Chen M, Guo S, Tan X. Does health information exchange improve patient outcomes? Empirical evidence from Florida hospitals. *Health Aff (Millwood)* 2019 Feb;38(2):197-204. [doi: [10.1377/hlthaff.2018.05447](https://doi.org/10.1377/hlthaff.2018.05447)] [Medline: [30715992](https://pubmed.ncbi.nlm.nih.gov/30715992/)]
35. Golden RL, Emery-Tiburcio EE, Post S, Ewald B, Newman M. Connecting social, clinical, and home care services for persons with serious illness in the community. *J Am Geriatr Soc* 2019 May 10;67(S2):412-418. [doi: [10.1111/jgs.15900](https://doi.org/10.1111/jgs.15900)] [Medline: [31074858](https://pubmed.ncbi.nlm.nih.gov/31074858/)]
36. Campbell AN, McCarty D, Rieckmann T, McNeely J, Rotrosen J, Wu L, et al. Interpretation and integration of the federal substance use privacy protection rule in integrated health systems: a qualitative analysis. *J Subst Abuse Treat* 2019 Feb;97:41-46 [FREE Full text] [doi: [10.1016/j.jsat.2018.11.005](https://doi.org/10.1016/j.jsat.2018.11.005)] [Medline: [30577898](https://pubmed.ncbi.nlm.nih.gov/30577898/)]
37. Falconer E, Kho D, Docherty JP. Use of technology for care coordination initiatives for patients with mental health issues: a systematic literature review. *Neuropsychiatr Dis Treat* 2018;14:2337-2349 [FREE Full text] [doi: [10.2147/NDT.S172810](https://doi.org/10.2147/NDT.S172810)] [Medline: [30254446](https://pubmed.ncbi.nlm.nih.gov/30254446/)]
38. Schmidt EM, Behar S, Barrera A, Cordova M, Beckum L. Potentially preventable medical hospitalizations and emergency department visits by the behavioral health population. *J Behav Health Serv Res* 2018 Jul 13;45(3):370-388. [doi: [10.1007/s11414-017-9570-y](https://doi.org/10.1007/s11414-017-9570-y)] [Medline: [28905296](https://pubmed.ncbi.nlm.nih.gov/28905296/)]
39. Bradford DW, Cunningham NT, Slubicki MN, McDuffie JR, Kilbourne AM, Nagi A, et al. An evidence synthesis of care models to improve general medical outcomes for individuals with serious mental illness. *J Clin Psychiatry* 2013 Aug 15;74(08):754-764. [doi: [10.4088/jcp.12r07666](https://doi.org/10.4088/jcp.12r07666)]
40. Annamalai A, Staeheli M, Cole RA, Steiner JL. Establishing an integrated health care clinic in a community mental health center: lessons learned. *Psychiatr Q* 2018 Mar 30;89(1):169-181. [doi: [10.1007/s11126-017-9523-x](https://doi.org/10.1007/s11126-017-9523-x)] [Medline: [28664447](https://pubmed.ncbi.nlm.nih.gov/28664447/)]
41. Shank N. Behavioral health providers' beliefs about health information exchange: a statewide survey. *J Am Med Inform Assoc* 2012 Jul;19(4):562-569 [FREE Full text] [doi: [10.1136/amiajnl-2011-000374](https://doi.org/10.1136/amiajnl-2011-000374)] [Medline: [22184253](https://pubmed.ncbi.nlm.nih.gov/22184253/)]
42. Shen N, Sequeira L, Silver MP, Carter-Langford A, Strauss J, Wiljer D. Patient privacy perspectives on health information exchange in a mental health context: qualitative study. *JMIR Ment Health* 2019 Nov 13;6(11):e13306 [FREE Full text] [doi: [10.2196/13306](https://doi.org/10.2196/13306)] [Medline: [31719029](https://pubmed.ncbi.nlm.nih.gov/31719029/)]
43. Dwivedi YK, Rana NP, Chen H, Williams MD. A meta-analysis of the Unified Theory of Acceptance and Use of Technology (UTAUT). In: *Governance and Sustainability in Information Systems. Managing the Transfer and Diffusion of IT*. Berlin: Springer; 2011:155-170.
44. Kijisanayotin B, Pannarunothai S, Speedie SM. Factors influencing health information technology adoption in Thailand's community health centers: applying the UTAUT model. *Int J Med Inform* 2009 Jun;78(6):404-416. [doi: [10.1016/j.ijmedinf.2008.12.005](https://doi.org/10.1016/j.ijmedinf.2008.12.005)] [Medline: [19196548](https://pubmed.ncbi.nlm.nih.gov/19196548/)]
45. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. *MIS Q* 2003;27(3):425. [doi: [10.2307/30036540](https://doi.org/10.2307/30036540)]
46. Slade EL, Dwivedi YK, Piercy NC, Williams MD. Modeling consumers' adoption intentions of remote mobile payments in the United Kingdom: extending UTAUT with innovativeness, risk, and trust. *Psychol Mark* 2015 Jul 07;32(8):860-873. [doi: [10.1002/mar.20823](https://doi.org/10.1002/mar.20823)]
47. Rogers EM. *Diffusion of Innovations*, 4th Edition. New York: Free Press; 1995:1-518.
48. Rogers EM. Diffusion of preventive innovations. *Addict Behav* 2002 Nov;27(6):989-993. [doi: [10.1016/s0306-4603\(02\)00300-3](https://doi.org/10.1016/s0306-4603(02)00300-3)]
49. Ivankova NV, Creswell J, Stick S. Using mixed-methods sequential explanatory design: from theory to practice. *Field Methods* 2016 Jul 21;18(1):3-20. [doi: [10.1177/1525822X05282260](https://doi.org/10.1177/1525822X05282260)]
50. Subedi D. Explanatory sequential mixed method design as the third research community of knowledge claim. *Am J Educ Res* 2016;4(7):570-577 [FREE Full text] [doi: [10.12691/education-4-7-10](https://doi.org/10.12691/education-4-7-10)]
51. Fornell C, Larcker DF. Evaluating structural equation models with unobservable variables and measurement error. *J Mark Res* 2018 Nov 28;18(1):39-50. [doi: [10.1177/002224378101800104](https://doi.org/10.1177/002224378101800104)]
52. Gefen D, Straub D, Boudreau M. Structural equation modeling and regression: guidelines for research practice. *Commun Assoc Inf Syst* 2000;4(7):A. [doi: [10.17705/1CAIS.00407](https://doi.org/10.17705/1CAIS.00407)]
53. Nunnally J, Bernstein I. *Psychometric Theory*. New York: McGraw-Hill Education; 1993:1-784.
54. Haenlein M, Kaplan AM. A beginner's guide to partial least squares analysis. *Understand Sta* 2004 Nov;3(4):283-297. [doi: [10.1207/s15328031us0304_4](https://doi.org/10.1207/s15328031us0304_4)]
55. Henseler J, Ringle CM, Sinkovics RR. The use of partial least squares path modeling in international marketing. *Adv Int Mark* 2009;20:277-319. [doi: [10.1108/S1474-7979\(2009\)0000020014](https://doi.org/10.1108/S1474-7979(2009)0000020014)]

56. Hair JF, Ringle CM, Sarstedt M. PLS-SEM: indeed a silver bullet. *J Mark Theory Prac* 2014 Dec 08;19(2):139-152. [doi: [10.2753/mtp1069-6679190202](https://doi.org/10.2753/mtp1069-6679190202)]
57. Kazár K. PLS path analysis and its application for the examination of the psychological sense of a brand community. *Procedia Econ Financ* 2014;17:183-191. [doi: [10.1016/s2212-5671\(14\)00893-4](https://doi.org/10.1016/s2212-5671(14)00893-4)]
58. Sarstedt M, Ringle CM, Henseler J, Hair JF. On the emancipation of PLS-SEM: a commentary on Rigdon (2012). *Long Range Plann* 2014 Jun;47(3):154-160. [doi: [10.1016/j.lrp.2014.02.007](https://doi.org/10.1016/j.lrp.2014.02.007)]
59. Creswell JW. *Qualitative Inquiry and Research Design: Choosing Among Five Approaches*. Thousand Oaks, California, United States: Sage Publications Inc; 2012:1-472.
60. Clarke V, Braun V. Using thematic analysis in counselling and psychotherapy research: a critical reflection. *Couns Psychother Res* 2018 Mar 31;18(2):107-110. [doi: [10.1002/capr.12165](https://doi.org/10.1002/capr.12165)]
61. Barnett J, Vasileiou K, Djemil F, Brooks L, Young T. Understanding innovators' experiences of barriers and facilitators in implementation and diffusion of healthcare service innovations: a qualitative study. *BMC Health Serv Res* 2011 Dec 16;11(1):342 [FREE Full text] [doi: [10.1186/1472-6963-11-342](https://doi.org/10.1186/1472-6963-11-342)] [Medline: [22176739](https://pubmed.ncbi.nlm.nih.gov/22176739/)]
62. Onwuegbuzie AJ, Leech NL. Validity and qualitative research: an oxymoron? *Qual Quant* 2006 May 25;41(2):233-249. [doi: [10.1007/s11135-006-9000-3](https://doi.org/10.1007/s11135-006-9000-3)]
63. Guetterman TC, Fetter MD, Creswell JW. Integrating quantitative and qualitative results in health science mixed methods research through joint displays. *Ann Fam Med* 2015 Nov 09;13(6):554-561 [FREE Full text] [doi: [10.1370/afm.1865](https://doi.org/10.1370/afm.1865)] [Medline: [26553895](https://pubmed.ncbi.nlm.nih.gov/26553895/)]
64. Rutherford GS, Hair JF, Anderson RE, Tatham RL. Multivariate data analysis with readings. *J Royal Stat Soc - Series D (The Statistician)* 1988;37(4/5):484. [doi: [10.2307/2348783](https://doi.org/10.2307/2348783)]
65. Elkaseh AM, Wong KW, Fung CC. Perceived ease of use and perceived usefulness of social media for e-learning in libyan higher education: a structural equation modeling analysis. *Int J Inf Educ Technol* 2016;6(3):192-199. [doi: [10.7763/ijiet.2016.v6.683](https://doi.org/10.7763/ijiet.2016.v6.683)]
66. Schaper LK, Pervan GP. ICT and OTs: a model of information and communication technology acceptance and utilisation by occupational therapists. *Int J Med Inform* 2007 Jun;76 Suppl 1:212-221. [doi: [10.1016/j.ijmedinf.2006.05.028](https://doi.org/10.1016/j.ijmedinf.2006.05.028)] [Medline: [16828335](https://pubmed.ncbi.nlm.nih.gov/16828335/)]
67. Ifinedo P. Understanding information systems security policy compliance: an integration of the theory of planned behavior and the protection motivation theory. *Comput Secur* 2012 Feb;31(1):83-95. [doi: [10.1016/j.cose.2011.10.007](https://doi.org/10.1016/j.cose.2011.10.007)]
68. Alaiad A, Zhou L. Patients' behavioral intention toward using healthcare robots. URL: <https://tinyurl.com/5xkbcja9> [accessed 2020-11-09]
69. Hsieh P. Physicians' acceptance of electronic medical records exchange: an extension of the decomposed TPB model with institutional trust and perceived risk. *Int J Med Inform* 2015 Jan;84(1):1-14. [doi: [10.1016/j.ijmedinf.2014.08.008](https://doi.org/10.1016/j.ijmedinf.2014.08.008)] [Medline: [25242228](https://pubmed.ncbi.nlm.nih.gov/25242228/)]
70. Ranallo PA, Kilbourne AM, Whatley AS, Pincus HA. Behavioral health information technology: from chaos to clarity. *Health Aff (Millwood)* 2016 Jun 01;35(6):1106-1113. [doi: [10.1377/hlthaff.2016.0013](https://doi.org/10.1377/hlthaff.2016.0013)] [Medline: [27269029](https://pubmed.ncbi.nlm.nih.gov/27269029/)]
71. Grando MA, Murcko A, Mahankali S, Saks M, Zent M, Chern D, et al. A study to elicit behavioral health patients' and providers' opinions on health records consent. *J Law Med Ethics* 2017 Jun 14;45(2):238-259 [FREE Full text] [doi: [10.1177/1073110517720653](https://doi.org/10.1177/1073110517720653)] [Medline: [30976154](https://pubmed.ncbi.nlm.nih.gov/30976154/)]
72. Druss BG, Rohrbaugh RM, Levinson CM, Rosenheck RA. Integrated medical care for patients with serious psychiatric illness: a randomized trial. *Arch Gen Psychiatry* 2001 Sep 01;58(9):861-868. [doi: [10.1001/archpsyc.58.9.861](https://doi.org/10.1001/archpsyc.58.9.861)] [Medline: [11545670](https://pubmed.ncbi.nlm.nih.gov/11545670/)]
73. Unützer J, Katon W, Callahan CM, Williams JW, Hunkeler E, Harpole L, IMPACT (Improving Mood-Promoting Access to Collaborative Treatment) Investigators. Collaborative care management of late-life depression in the primary care setting: a randomized controlled trial. *J Am Med Assoc* 2002 Dec 11;288(22):2836-2845. [doi: [10.1001/jama.288.22.2836](https://doi.org/10.1001/jama.288.22.2836)] [Medline: [12472325](https://pubmed.ncbi.nlm.nih.gov/12472325/)]
74. Alexopoulos GS, Reynolds III CF, Bruce ML, Katz IR, Raue PJ, Mulsant BH, PROSPECT Group. Reducing suicidal ideation and depression in older primary care patients: 24-month outcomes of the PROSPECT study. *Am J Psychiatry* 2009 Aug;166(8):882-890 [FREE Full text] [doi: [10.1176/appi.ajp.2009.08121779](https://doi.org/10.1176/appi.ajp.2009.08121779)] [Medline: [19528195](https://pubmed.ncbi.nlm.nih.gov/19528195/)]
75. What is integrated care? SAMHSA-HRSA Center for Integrated Health Solutions (CIHS). URL: <https://www.samhsa.gov/integrated-health-solutions> [accessed 2020-12-10]
76. Alexander GC, Stoller KB, Haffajee RL, Saloner B. An epidemic in the midst of a pandemic: opioid use disorder and COVID-19. *Ann Intern Med* 2020 Jul 07;173(1):57-58. [doi: [10.7326/m20-1141](https://doi.org/10.7326/m20-1141)]
77. Kock N, Hadaya P. Minimum sample size estimation in PLS-SEM: the inverse square root and gamma-exponential methods. *Info Systems J* 2016 Nov 29;28(1):227-261. [doi: [10.1111/isj.12131](https://doi.org/10.1111/isj.12131)]

Abbreviations

CFR: Code of Federal Regulations
DOI: diffusion of innovation

HIE: health information exchange

IT: information technology

PLS-SEM: partial least squares structural equation modeling

UTAUT: unified theory of acceptance and use of technology

Edited by J Torous; submitted 23.12.20; peer-reviewed by H Pratomo, S Liu; comments to author 06.02.21; revised version received 31.03.21; accepted 13.04.21; published 27.05.21.

Please cite as:

Cochran RA, Feldman SS, Ivankova NV, Hall AG, Opoku-Agyeman W

Intention to Use Behavioral Health Data From a Health Information Exchange: Mixed Methods Study

JMIR Ment Health 2021;8(5):e26746

URL: <https://mental.jmir.org/2021/5/e26746>

doi: [10.2196/26746](https://doi.org/10.2196/26746)

PMID: [34042606](https://pubmed.ncbi.nlm.nih.gov/34042606/)

©Randyl A Cochran, Sue S Feldman, Nataliya V Ivankova, Allyson G Hall, William Opoku-Agyeman. Originally published in JMIR Mental Health (<https://mental.jmir.org>), 27.05.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Mental Health, is properly cited. The complete bibliographic information, a link to the original publication on <https://mental.jmir.org/>, as well as this copyright and license information must be included.

Review

Initial Training for Mental Health Peer Support Workers: Systematized Review and International Delphi Consultation

Ashleigh Charles¹, MSc; Rebecca Nixdorf², MSc; Nashwa Ibrahim^{1,3}, PhD; Lion Gai Meir⁴; Richard S Mpango^{5,6,7}, PhD; Fileuka Ngakongwa^{8,9}, MD; Hannah Nudds¹, BSc; Soumitra Pathare¹⁰, MD; Grace Ryan¹¹, MSc; Julie Repper¹², PhD; Heather Wharrad¹³, PhD; Philip Wolf¹⁴, MSc; Mike Slade¹, PhD; Candelaria Mahlke², PhD

¹School of Health Sciences, Institute of Mental Health, University of Nottingham, Nottingham, United Kingdom

²Department of Psychiatry, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

³Psychiatric and Mental Health Nursing Department, Faculty of Nursing, Mansoura University, Masoura, Egypt

⁴Department of Social Work, Ben Gurion University of the Negev, Beer Sheva, Israel

⁵Butabika National Referral Hospital, Butabika, Uganda

⁶School of Health Sciences, Soroti University, Soroti, Uganda

⁷MRC/UVRI and LSHTM Uganda Research Unit, Entebbe, Uganda

⁸Ifakara Health Institute, Dar es Salaam, United Republic of Tanzania

⁹Department of Psychiatry and Mental Health, Muhimbili University of Health and Allied Sciences, Dar es Salaam, United Republic of Tanzania

¹⁰Centre for Mental Health Law and Policy, Indian Law Society, Pune, India

¹¹Centre for Global Mental Health, London School of Hygiene and Tropical Medicine, London, United Kingdom

¹²ImROC, Nottinghamshire Healthcare NHS Foundation Trust, Nottingham, United Kingdom

¹³Faculty of Medicine and Health Sciences, University of Nottingham, Nottingham, United Kingdom

¹⁴Department of Psychiatry II, Ulm University II, Ulm, Germany

Corresponding Author:

Ashleigh Charles, MSc

School of Health Sciences, Institute of Mental Health

University of Nottingham

Jubilee Campus, University of Nottingham Innovation Park, Triumph Road
Nottingham,

United Kingdom

Phone: 44 (0)115 7484303

Email: ashleigh.charles@nottingham.ac.uk

Abstract

Background: Initial training is essential for the mental health peer support worker (PSW) role. Training needs to incorporate recent advances in digital peer support and the increase of peer support work roles internationally. There is a lack of evidence on training topics that are important for initial peer support work training and on which training topics can be provided on the internet.

Objective: The objective of this study is to establish consensus levels about the content of initial training for mental health PSWs and the extent to which each identified topic can be delivered over the internet.

Methods: A systematized review was conducted to identify a preliminary list of training topics from existing training manuals. Three rounds of Delphi consultation were then conducted to establish the importance and web-based deliverability of each topic. In round 1, participants were asked to rate the training topics for importance, and the topic list was refined. In rounds 2 and 3, participants were asked to rate each topic for importance and the extent to which they could be delivered over the internet.

Results: The systematized review identified 32 training manuals from 14 countries: Argentina, Australia, Brazil, Canada, Chile, Germany, Ireland, the Netherlands, Norway, Scotland, Sweden, Uganda, the United Kingdom, and the United States. These were synthesized to develop a preliminary list of 18 topics. The Delphi consultation involved 110 participants (49 PSWs, 36 managers, and 25 researchers) from 21 countries (14 high-income, 5 middle-income, and 2 low-income countries). After the Delphi consultation (round 1: n=110; round 2: n=89; and round 3: n=82), 20 training topics (18 universal and 2 context-specific) were identified. There was a strong consensus about the importance of five topics: *lived experience as an asset, ethics, PSW well-being,*

and PSW role focus on recovery and communication, with a moderate consensus for all other topics apart from the *knowledge of mental health*. There was no clear pattern of differences among PSW, manager, and researcher ratings of importance or between responses from participants in countries with different resource levels. All training topics were identified with a strong consensus as being deliverable through blended web-based and face-to-face training (rating 1) or fully deliverable on the internet with moderation (rating 2), with none identified as only deliverable through face-to-face teaching (rating 0) or deliverable fully on the web as a stand-alone course without moderation (rating 3).

Conclusions: The 20 training topics identified can be recommended for inclusion in the curriculum of initial training programs for PSWs. Further research on web-based delivery of initial training is needed to understand the role of web-based moderation and whether web-based training better prepares recipients to deliver web-based peer support.

(*JMIR Ment Health* 2021;8(5):e25528) doi:[10.2196/25528](https://doi.org/10.2196/25528)

KEYWORDS

peer support work; peer support worker training; Delphi consultation; mental health; mobile phone

Introduction

Background

Peer support is rapidly developing into a central approach to support mental health recovery [1]. It has been defined as “a system of giving and receiving help founded on the key principles of respect, shared responsibility, and a mutual agreement of what is helpful” [2]. Formal peer support involves individuals with lived experience of mental health conditions and/or mental health services—variously called peer providers, peer specialists, peer support volunteers, or as peer support workers (PSWs)—who are engaged in a peer support capacity to support others’ recovery from mental health conditions [3].

There is strong empirical evidence for the positive benefits of implementing PSW roles. PSWs contribute to an increase in service user engagement, a sense of empowerment [4], improved social relationships [5], self-efficacy [6], hope [7], self-management [8], and positive clinical outcomes [9,10]. PSW roles are being implemented internationally, increasingly in lower middle-income countries such as India [11] and Uganda [12], where PSWs are engaged to address the mental health care gap [13].

Numerous PSW training programs exist [14,15]. These programs are designed to prepare individuals for the PSW role and sometimes also to meet the required competencies for accreditation. However, PSW roles and activities vary depending on their particular setting and context. For example, PSWs may work one-on-one or in a group setting with people who use mental health services, they may support individuals through service transitions, from hospital to the community, into employment, and/or to access local opportunities or resources. Training programs vary depending on the setting, local context, and resources.

Peer support has traditionally been provided as a face-to-face intervention. However, with the growth of digital mental health interventions changing the ways in which mental health care is delivered, peer support is increasingly being offered via digital technologies known as digital peer support. Digital peer support is defined as automated or live peer support services that can be delivered through multiple modalities [16], such as web-based peer support [17], smartphone-supported interventions [18,19], and web-based peer-to-peer networks [20]. A recent systematic

review concluded that digital peer support appears to be both acceptable and feasible, with a strong potential for clinical effectiveness [16]. The use of web-based PSW approaches such as digital peer support is likely to increase in response to the COVID-19 global pandemic [21]. This new development means that PSW training may also be offered on the web, and some initial training topics may need to be updated or reviewed.

Initial training is essential for a formal PSW role. Several studies have reported the need to identify core peer support competencies in mental health [22,23]. However, there is no consensus on the core training topics required for PSW roles and accreditation [24]. Given the advances in digital peer support and lack of evidence about what topics are important in initial PSW training, there is a need to establish the level of consensus about the content of initial PSW training. A specific knowledge gap relates to the extent to which initial PSW training topics can be delivered on the web. Web-based training may offer benefits, including opportunities to access more prospective PSWs on a larger scale and the reduction of costs as compared with face-to-face training but may not be feasible for training in the relational aspects of the PSW role.

Aims and Objectives

The aim of this study is to conduct a Delphi consultation technique to establish the level of consensus regarding initial PSW training topics. The objectives were (1) to identify training topics; (2) to identify the extent to which each identified training topic can be delivered on the web; and (3) to assess the degree of consensus that exists with respect to findings from objectives (1) and (2) overall, between PSWs and PSW managers versus PSW researchers and between countries with different resource levels.

Methods

Overview

This study was conducted as a part of UPSIDES (Using Peer Support in Developing Empowering Mental Health Services), a 5-year (2018-2022) European Union-funded multinational study that aims to replicate and scale up peer support interventions for people with severe mental illness. Ethics approval was obtained from the University of Nottingham, Faculty of Medicine and Health Sciences Research Ethics

Committee (FMHS 377-1908). All participants provided web-based informed consent after reading the participant information sheet available on the web.

Design

A systematized review was conducted to develop a preliminary list of training topics, followed by a three-round Delphi consultation. Round 1 refined the preliminary list, and rounds 2 and 3 established a consensus.

Participants

Inclusion criteria were age 18 years or above, able to provide informed consent electronically, and one of the following: publication record relating to peer support work (researcher group) or experience of working as a PSW group or clinical or practical experience in training and/or supervision of peer support (manager group). Participants were not screened for inclusion but self-identified as belonging to one or more of the above groups.

Procedures

Systematized Review

First, we conducted a systematized review of initial PSW training manuals to identify a preliminary list of topics for use in initial PSW training. A systematized review includes elements of a systematic review process while stopping short of being a full systematic review [25]. The inclusion criteria were as follows:

- Training manuals designed specifically for mental health PSWs, that is, designed for people with lived experience of mental health conditions, to prepare them to work in a PSW role with other people with mental health conditions.
- Involved face-to-face training, web-based training, or a combination of both.
- Training with and without accreditation or certification was included, as countries vary in their progress toward an accreditation process.
- Training for both a generic PSW role or for work with a specific mental health subpopulation, for example, dual diagnosis with substance misuse, dementia, and young people.
- Published in the English or Arabic language.

Exclusion criteria were as follows:

- Training manuals not specifically for mental health PSWs.
- Does not include initial PSW training, for example, focus is on stigma related to mental health, human rights, peer leadership, consumer-run organizations, gender identity, and childhood trauma.
- Train the trainer manuals.

Training programs were identified from seven sources:

1. The MEDLINE database was searched using the search strategy shown in [Multimedia Appendix 1](#).
2. Gray literature databases (OpenGrey, New York Academy of Medicine's Grey Literature Report, TRIP, and Health Quality Ontario) were searched using the phrases: "mental health peer support work training manuals," "mental health

consumer provider training manuals," "certification training for mental health peer support workers," and "mental health peer support workers' core competencies."

3. Google and Google Scholar were searched using the same phrases mentioned above. The first 100 hits were then screened.
4. Massive open online course platforms (Coursera, edX, FutureLearn, Canvas Network, and Independent) were searched using the massive open online course list website [26], using the same phrases mentioned above.
5. The preliminary findings from a related ongoing systematic review [27].
6. The websites of mental health organizations include the Substance Abuse and Mental Health Services Administration [28], National Mental Health Commission [29], Scottish Recovery Network [30], Mental Health America [31], Mental Health Innovation Network [32], Depression and Bipolar Support Alliance [33], and REdeAmericas [34].
7. The websites of PSW certification, accreditation, and professional bodies include the Missouri Peer Specialist [35], Nevada Certification Board [36], National Association of Peer Supporters [37], and Global Mental Health Peer Network [38].

Searches were conducted in April 2019, and no date restrictions were applied. Endnote software was used to collate the identified manuals. Screening and data extraction were equally divided between AC and NI. Both researchers independently analyzed their allocated manuals and discussed any discrepancies with MS (eg, inclusion and exclusion of manuals). The data abstraction table was populated by extracting data relating to country (specific country and country income level), service setting (eg, community or inpatient), target population (eg, general mental health or dementia), training modality (face-to-face, internet-based, or both), training topics (using terms from source documents), definitions or training goals or learning objectives (where specified), and examples of the training exercises.

Both AC and NI independently conducted thematic analysis [39]. Vote counting for each theme was conducted, and the 2 analysts discussed the discrepancies. AC, NI, and MS then conducted a process of data reduction, which involved comparing the training topics within and across themes and merging and integrating subthemes to generate one coherent coding framework or codebook comprising the training topic name and definition.

Delphi Consultation

A three-round Delphi consultation was then conducted. Delphi consultation is a systematic method of determining how much agreement exists on a particular topic based on experts' opinions [40]. The method involves an iterative and multistage process, comprising multiple rounds of questions designed to combine opinions and assess group consensus.

Compared with the traditional Delphi method, which is delivered via questionnaires through face-to-face meetings, a web-based Delphi offers participants the time to deliberate their responses, thus increasing the validity of results, and requires less

resources, for example, time and costs [41]. In addition, given the growth of PSW roles internationally and the wide range of peer support programs that exist, the web-based Delphi was chosen as an appropriate method to answer the research question, as it has the potential to access a diverse and large group of experts engaged in peer support from around the world [42]. Compared with other research designs, the Delphi method can produce rigorous and rich data because of multiple rounds and refinement based on response feedback [43]. Limitations include a low level of reliability of judgment among experts, lack of clear methodological guidelines, and difficulty in assessing the degree of expertise included [44].

Participants for the Delphi consultation were recruited through advertising (1) across networks including the Recovery Research Network [45], UPSIDES consortium, Recovery College Network, and Being Network (Australia); (2) by partner organizations including Nottinghamshire Healthcare National Health Service Foundation Trust, Institute of Mental Health; (3) social media, including Twitter; and (4) snowball sampling. Purposive sampling was used to obtain a balance in role (researcher, PSW, and manager) and income level (high, middle, and low).

In each round, prospective participants were invited via email to complete an anonymous web-based survey using the Jisc survey platform (round 1 and 2) or Google Forms (round 3). Participants were sent 2 reminder emails to complete each round of the survey. All correspondence with the participants occurred via email. An initial pilot test of each round was conducted with 5 nonparticipants, which resulted in minor amendments. Round 1 was distributed over eight screen pages, and rounds 2 and 3 were distributed over four pages. The personal information collected was stored separately from the results' data on a secure webserver, and participants were allocated identification numbers, for example, P001, which were used to track responses and duplicates. Only the research team had access to personal and research data. No incentive was offered for participation. Participants were able to change their responses using a back button but could not complete the survey without answering all questions. Round 1 began in November 2019, and round 3 was completed in July 2020.

In round 1, after reading the participant information sheet outlining the length of time of the survey, how data are stored, where, and for how long, and the purpose of the study, participants gave informed consent and provided sociodemographic details including if they had any experience of web-based training (about anything). Participants then rated each training topic and associated definition for importance using a 4-point Likert scale ("0=Not important at all," "1=A bit important," "2=Quite important," and "3=Very important"). A free-text response was available to participants, providing the opportunity to suggest additional training topics or changes to the topic name or definition. Qualitative data were analyzed thematically. Two analysts (AC and NI) independently reviewed and synthesized responses using the following framework:

proposed topic name, change to language or definition, and additional topics to be covered. Agreement on the topics and definitions arising from the first round was determined by further discussion, refinement, and synthesis with a third analyst (MS). The finalized list of training topics was created based on these responses, including the deletion of topics quantitatively rated as not important, refinement of name or definition, and addition of new topics where relevant.

In round 2, participants rated each topic and associated definition for importance using the same rating scale as in round 1, and also rated the extent to which the topic could be delivered on the web using a 4-point Likert scale ("0=No, it can only be delivered through face-to-face training," "1=Partially, eg, as blended learning with some aspects delivered online and some face-to-face," "2=Fully online as a moderated online course, ie, with a peer support worker trainer providing online support and moderation," and "3=It can be fully delivered online as a standalone course without moderation").

In round 3, participants were shown their own round 2 ratings and the mean round 2 ratings from other participants in (1) their group (researcher, PSW, and manager or supervisor), (2) the income level of their country setting (high, middle, and low), and (3) overall. Participants were asked to rerate each component for importance and web-based delivery using the same rating scales as in round 2.

Analysis

For the Delphi consultation, strong consensus was defined as at least 80% of participants in the group with the same rating, and moderate consensus was defined as at least 50% of participants in the group with the same rating. An arbitrary threshold for a high- and moderate-level consensus was implemented.

Results

Overview

The systematized review identified 32 training manuals in English, comprising face-to-face and web-based PSW training from 14 different countries (Argentina, Australia, Brazil, Canada, Chile, Germany, Ireland, the Netherlands, Norway, Scotland, Sweden, Uganda, the United Kingdom, and the United States). Training varied in length from 54 hours to 1 year, and the manuals covered a range of PSW training from working with adults, older adults, adolescents, and young people. A total of 502 topics and 348 learning objectives or definitions were extracted. The coding framework synthesized from the training manuals comprised 18 themes and is shown in [Multimedia Appendix 2](#). The coding framework was used as the basis for developing 18 topics and associated definitions, as shown in the first column of [Multimedia Appendix 3](#).

The characteristics of the Delphi consultation participants, including the number of responses for each round, are shown in [Table 1](#).

Table 1. Characteristics of the Delphi consultation participants (N=110).

Participant characteristics	Value		
	Round 1 (N=110)	Round 2 (n=89)	Round 3 (n=82)
Age (years), n (%)			
21-29	17 (15.4)	13 (14.6)	11 (13.4)
30-39	25 (22.7)	19 (21.3)	21 (25.6)
40-49	28 (25.4)	24 (26.9)	22 (26.8)
50-59	29 (26.3)	24 (26.9)	18 (21.9)
60+	10 (9.0)	9 (10.1)	9 (10.9)
Gender, n (%)			
Female	79 (71.8)	65 (73)	58 (70.7)
Male	29 (26.3)	23 (25.8)	23 (28.0)
Other	2 (1.8)	1 (1.1)	1 (1.2)
Role, n (%)			
PSW ^a	49 (44.5)	38 (42.6)	36 (43.9)
Manager	36 (32.7)	29 (32.5)	24 (29.2)
Researcher	25 (22.7)	22 (24.7)	22 (26.8)
Years of experience in role, n (%)			
<2	13 (11.8)	11 (12.3)	11 (13.4)
2-3	27 (24.5)	18 (20.2)	17 (20.7)
4-6	26 (23.6)	22 (24.7)	20 (24.3)
7-9	21 (19.0)	17 (19.1)	15 (18.2)
≥10	21 (19.0)	19 (21.3)	17 (20.7)
Professional qualification, n (%)			
Yes	42 (38.1)	34 (38.2)	30 (36.5)
No	68 (61.8)	55 (61.7)	52 (63.4)
Lived experience of mental health problems, n (%)			
Yes	82 (74.5)	67 (75.2)	62 (75.6)
No	28 (25.4)	22 (24.7)	20 (24.3)
Experience of initial PSW training, mean (SD)			
Scale: 0 (no experience) to 10 (very experienced)	6.4 (2.9)	6.4 (3.0)	6.3 (3.14)
Experience of web-based PSW training, n (%)			
Yes	25 (22.7)	21 (23.5)	21 (25.6)
Experience of web-based training on any topic, n (%)			
Yes	88 (80.0)	73 (82.0)	67 (81.7)

^aPSW: peer support worker.

Participant Characteristics

The participants came from 21 countries, including higher-income (Australia: n=33; the United Kingdom: n=22; Canada: n=7; Poland: n=7; Germany: n=4; Ireland: n=4; Switzerland: n=4; Israel: n=3; Norway: n=3; Italy: n=2; the United States: n=2; New Zealand: n=1; Belgium: n=1; Singapore: n=1), middle-income (India: n=4; Tunisia: n=2; Brazil: n=1; Egypt: n=1; Argentina: n=1), and lower-income (Uganda: n=6; Tanzania: n=1) countries.

All participants completed round 1 and rated each of the 18 topics for importance. Round 1 ratings and all proposed changes are shown in [Multimedia Appendix 3](#). Analysis of the round 1 free-text responses (n=68) and mean rating scores resulted in multiple refinements to the round 1 topic names and definitions. Although the mean score for *knowledge of mental health* was ranked low compared with other topics, the free-text responses suggested that the definition should be amended to reflect PSW training needs, and this was changed markedly for round 2. Two topics—*role-specific PSW skills and competencies* and *work*

skills—were identified as being relevant in some contexts but not others. These context-specific topics were presented separately in a different format for round 2, along with an explanation for the participants. Three additional topics and definitions were created—*PSW supervision, developing a career as a PSW*, and *role-specific PSW skills and competencies*—that were adapted from the *subpopulation and specialized modules* topic.

The revised list of 20 topics and associated definitions were used for round 2, which are shown in the fifth column of [Multimedia Appendix 3](#). Of 110 participants, a total of 89 (80.9%) participants completed round 2. The round 2 ratings

of importance are shown in [Multimedia Appendix 4](#) and of web-based delivery are shown in [Multimedia Appendix 5](#). Additional comments were received from round 2 participants about the *role-specific PSW skills and competencies* topic, resulting in minor refinements to the definition of *knowledge of mental health* topic. The final list of topics and definitions, which were used in round 3, is shown in [Textbox 1](#).

Of 110 participants, a total of 82 (74.5%) completed round 3. Of these 82 participants, 76 (93%) had completed round 2. The round 3 ratings of importance, ordered by median rating, are shown in [Table 2](#).

Textbox 1. Final list of topics and definitions for initial peer support worker training.

Topics always needing coverage

- Introduction to peer support and peer support worker (PSW)
 - Presenting the local and international history of peer support, survivor or activist grassroots knowledge, and key information on the context of peer support, PSW, principles, and concept of expertise by experience is essential to formal PSWs
- PSW role focus on recovery
 - Teaching about the meaning, stages, and culture of recovery, allowing integration into the PSW's own experiences and practice. Additionally, teaching leadership, supporting informed choice, and working with service users in difficult times
- Approaches, frameworks, and models used in PSWs
 - Familiarizing prospective PSWs with approaches and frameworks underlying which peer support could be practiced. For example, the tree of life, coaching frameworks, strengths-based approach, Intentional Peer Support, and Wellness Recovery Action Planning
- Knowledge of mental health
 - Introducing prospective PSWs to different frames of understanding of mental health, including nonmedical models of understanding mental distress (eg, Hearing Voices, Alternatives to Suicide, or Mad Studies) and medical models (eg, diagnosis or interventions), the different types of service setting (eg, inpatient units), and the mental health needs of different populations (eg, age groups, dual diagnosis, or marginalized and minority groups)
- Human rights and disability legislation
 - Providing training about the meaning and implications of human rights legislation, including regional or national mental health laws and international legislation such as Convention on the Rights of Persons with Disabilities, to inform values-based PSW practice and skills in working within systems to uphold and protect the rights and social justice for people they work with, for example, through advocacy
- Ethics
 - Teaching about PSW values, beliefs, and actions, supporting self-reflection and an understanding about mental health practice and accountability including the importance of boundaries, levels of disclosure, and confidentiality
- Cultural competency
 - Practicing PSWs in a way compatible with the cultural needs, values, background, and context of people using services
- PSW skills and competencies
 - Providing prospective PSWs with essential competencies needed for formal PSWs through an overview of different PSWs' job descriptions; teaching the importance of maintaining role integrity and reflecting on the essential qualities and values of PSWs
- Lived experience as an asset
 - Highlighting how the experience of mental health problems, alongside other peer experiences such as service use, is a central resource for the PSW role; exploring methods and strategies for using lived experience with service users, including the safe, purposeful, and appropriate use of one's story to benefit others
- PSW well-being
 - Supporting self-reflection and offering strategies for PSWs to promote wellness, recovery, and resilience (eg, teaching PSWs about their workplace rights, self-advocacy, stress management techniques, vicarious trauma, self-care, and how to use reflective practice)
- Communication
 - Ensuring prospective PSWs have the fundamental connecting skills (eg, listening skills, use of language, and awareness of verbal and nonverbal cues), which facilitate effective communication with service users in different settings and situations, and helping them develop these skills if necessary
- Trauma-informed peer support practice
 - Offering peer support to understand and respond to the trauma of people using services to help them regain or achieve wellness and healing
- Crisis management
 - Helping PSWs to understand how to respond collaboratively, supportively, respectfully, and empathically to someone in a crisis
- PSWs working with groups

- Training prospective PSWs on the skills needed to start, facilitate, or cofacilitate a peer support group, in addition to understanding the group processes, dynamics, and coproduction practices and addressing any arising issues
- Workplace aspects of PSWs
 - Ensuring PSWs have the skills needed to deal with workplace challenges, including knowledge of support options and training in dealing with work-related pressures, such as working with other professionals with conflicting values, workplace culture, organizational structures, exposure to violence, discrimination, bullying, and managing power dynamics and conflict
- Referral and communication with other services
 - Ensuring prospective PSWs know about local services and community resources and about formal communication or referral processes to other services; ensuring PSWs are sensitive to the balance between helpful referrals and supporting self-management or being heard
- PSW supervision
 - Introducing PSWs to the purpose, types of, and importance of supervision
- Developing a career as a PSW
 - Involves teaching prospective peers about the professionalization of the PSW role, including motivational drivers, career development, training opportunities, and financial management

Context-specific topics

- Role-specific PSW skills and competencies
 - Equipping PSWs with role-specific skills (eg, motivational interviewing, solution-focused thinking, family therapy approach, intentional sharing, and understanding cognitive behavioral therapy and mindfulness), understanding of service settings (eg, inpatient units and community teams) and the mental health needs of different populations (eg, age groups, dual diagnosis, homelessness, and marginalized and minority groups)
- Work skills
 - Teaching the administrative skills of recording and documenting direct mental health care and incidents and other work-related skills, such as time management

Table 2. Delphi Consultation round 3 rating of importance (n=82).

Variables	Total	Participants by role			Participants by income level		
		PSW ^a	Manager	Researcher	High	Middle	Low
Population, n (%)	82 (100)	36 (44)	24 (29)	22 (27)	71 (86)	4 (5)	7 (9)
Lived experience as an asset, median (IQR) ^b	3 (0) ^c	3 (0) ^c	3 (0) ^c	3 (0) ^c	3 (0) ^c	3 (0.25) ^d	3 (0) ^c
Ethics, median (IQR) ^b	3 (0) ^c	3 (0) ^c	3 (0) ^c	3 (0) ^c	3 (0) ^c	3 (0) ^c	3 (0.5) ^d
PSW well-being, median (IQR) ^b	3 (0) ^c	3 (0) ^c	3 (0) ^c	3 (0) ^c	3 (0) ^c	3 (0) ^c	3 (0) ^c
PSW role focus on recovery, median (IQR) ^b	3 (0) ^c	3 (0) ^d	3 (0) ^c	3 (0) ^c	3 (0) ^c	3 (0.25) ^d	3 (0.5) ^d
Communication, median (IQR) ^b	3 (0) ^c	3 (0) ^c	3 (0) ^d	3 (0.75) ^d	3 (0) ^c	3 (0) ^c	3 (1) ^d
Crisis management, median (IQR) ^b	3 (0.75) ^d	3 (1) ^d	3 (0) ^c	3 (1) ^d	3 (0.5) ^d	3 (0.25) ^d	3 (0.5) ^d
Introduction to peer support and PSW, median (IQR) ^b	3 (1) ^d	3 (1) ^d	3 (1) ^d	3 (0) ^d	3 (1) ^d	2.5 (1) ^d	3 (1) ^d
Cultural competency, median (IQR) ^b	3 (1) ^d	3 (1) ^d	3 (1) ^d	3 (1) ^d	3 (1) ^d	2.5 (1) ^d	2 (1)
PSW skills and competencies, median (IQR) ^b	3 (1) ^d	2.5 (1) ^d	3 (0.25) ^d	3 (1) ^d	3 (1) ^d	3 (0.25) ^d	3 (1) ^d
Trauma-informed peer support practice, median (IQR) ^b	3 (1) ^d	3 (1) ^d	3 (0.25) ^d	2 (1)	3 (1) ^d	2 (0.25) ^d	2 (0) ^c
Workplace aspects of PSWs, median (IQR) ^b	3 (1) ^d	3 (1) ^d	3 (1) ^d	2.5 (1) ^d	3 (1) ^d	2.5 (1) ^d	2 (1.5)
PSW supervision, median (IQR) ^b	3 (1) ^d	3 (1) ^d	3 (0) ^d	2 (1) ^d	3 (1) ^d	2 (0.25) ^d	2 (1.5)
PSWs working with groups, median (IQR) ^b	2 (0) ^d	2 (2)	2 (0.25) ^d	2 (0) ^d	2 (0) ^d	2 (0) ^c	2 (1)
Knowledge of mental health, median (IQR) ^b	2 (1)	2.5 (1) ^d	2 (1) ^d	3 (1) ^d	2 (1)	2.5 (1) ^d	2 (1) ^d
Approaches, frameworks, and models used in PSW, median (IQR) ^b	2 (1) ^d	2 (1) ^d	2 (1) ^d	2 (0.75) ^d	2 (1) ^d	2.5 (1) ^d	2 (0) ^c
Human rights and disability legislation, median (IQR) ^b	2 (1) ^d	2 (1) ^d	2 (1) ^d	2 (1) ^d	2 (1) ^d	2.5 (1) ^d	2 (0.5) ^d
Referral and communication with other services, median (IQR) ^b	2 (1) ^d	2 (1.25)	2 (0.25) ^d	2 (0.75) ^d	2 (1) ^d	2.5 (1) ^d	2 (1)
Work skills, median (IQR) ^b	2 (1) ^d	2 (1) ^d	2 (0.25) ^d	2 (1) ^d	2 (0) ^d	3 (0.25) ^d	2 (1)
Developing a career as a PSW, median (IQR) ^b	2 (0.75) ^d	2 (1)	2 (0.25) ^d	2 (0) ^d	2 (0.5) ^d	2 (0.25) ^d	2 (0.5) ^d
Role-specific PSW skills and competencies, median (IQR) ^b	2 (0.75) ^d	2 (1)	2 (0) ^d	2 (0.75) ^d	2 (1)	2.5 (1) ^d	2 (0) ^d

^aPSW: peer support worker.^bScale 0 (low) to 3 (high).^cStrong consensus.^dModerate consensus.

The median rating of importance was “Quite Important” or “Very Important” for all topics. Across all participants, the first five topics in Table 2 reached a strong consensus on importance.

The round 3 ratings for web-based deliverability, ordered by median rating, are shown in Table 3.

The round 3 median ratings for web-based delivery indicated that all topics can be delivered partly or fully on the web with

moderation but none without moderation. No topics reached a strong consensus for the mode of training delivery. The range of median responses relating to web-based delivery was smaller for PSWs (1-1.5) than for managers (0-2) and researchers (1-2), indicating that PSWs were more consistent in placing importance on some face-to-face training contact.

Table 3. Delphi Consultation round 3 rating of web-based delivery (n=82).

Variables	Total	Participants by role			Participants by income level		
		PSW ^a	Manager	Researcher	High	Middle	Low
Population, n (%)	82 (100)	36 (44)	24 (29)	22 (27)	71 (86)	4 (5)	7 (9)
Human rights and disability legislation, median (IQR) ^b	2 (1)	1.5 (1)	2 (2)	2 (1)	2 (1)	3 (0) ^c	2 (1.5)
Developing a career as a PSW, median (IQR) ^b	2 (1) ^d	1.5 (1)	2 (1) ^d	2 (0) ^d	2 (1) ^d	3 (0.25) ^d	2 (1)
Introduction to peer support and PSW, median (IQR) ^b	2 (1)	1 (1) ^d	2 (1)	2 (1.5)	2 (1)	2.5 (1) ^d	1 (1) ^d
Knowledge of mental health, median (IQR) ^b	2 (1)	1 (1)	1.5 (1)	2 (0.75) ^d	2 (1)	3 (0.25) ^d	2 (1)
Role-specific PSW skills and competencies, median (IQR) ^b	1 (0) ^d	1 (0.25) ^d	1 (0.25) ^d	1 (0.75) ^d	1 (0) ^d	1.5 (1) ^d	1 (0) ^c
Referral and communication with other services, median (IQR) ^b	1 (1) ^d	1 (1) ^d	2 (1)	2 (1) ^d	1 (1) ^d	2 (0.25) ^d	1 (1) ^d
Work skills, median (IQR) ^b	1 (1)	1 (1) ^d	1 (1)	2 (0.75) ^d	1 (1)	2 (0.5) ^d	1 (1)
Approaches, frameworks, and models used in PSW, median (IQR) ^b	1 (1)	1 (1) ^d	1 (1) ^d	2 (1) ^d	1 (1)	2 (0.25) ^d	1 (0) ^c
Workplace aspects of PSWs, median (IQR) ^b	1 (1) ^d	1 (1) ^d	1 (1) ^d	2 (1) ^d	1 (1) ^d	2 (0.25) ^d	1 (1) ^d
PSW supervision, median (IQR) ^b	1 (1)	1 (1)	1 (1) ^d	2 (0.75) ^d	1 (1)	2 (0.25) ^d	2 (1)
PSW well-being, median (IQR) ^b	1 (1) ^d	1 (1)	1 (0.25) ^d	1 (1) ^d	1 (1) ^d	2 (0.5) ^d	1 (1)
PSW role focus on recovery, median (IQR) ^b	1 (1) ^d	1 (0) ^d	1 (1) ^d	1 (1) ^d	1 (1) ^d	2 (0.25) ^d	1 (0.5) ^d
Cultural competency, median (IQR) ^b	1 (1)	1 (1)	1 (1) ^d	1 (1) ^d	1 (1) ^d	2 (0) ^c	1 (1.5)
Ethics, median (IQR) ^b	1 (1) ^d	1 (0.5) ^d	1 (0.25) ^d	1 (1) ^d	1 (0) ^d	2.5 (1) ^d	1 (0.5) ^d
PSW skills and competencies, median (IQR) ^b	1 (1) ^d	1 (0) ^d	1 (1)	1 (1) ^d	1 (1) ^d	1 (0.5) ^d	1 (1) ^d
Trauma-informed peer support practice, median (IQR) ^b	1 (1)	1 (1) ^d	1 (1.25)	1 (1) ^d	1 (1)	2.5 (1.25) ^d	1 (0.5) ^d
PSW working with groups, median (IQR) ^b	1 (1) ^d	1 (1) ^d	1 (2)	1 (0) ^d	1 (1) ^d	2.5 (1.25) ^d	1 (0.5) ^d
Crisis management, median (IQR) ^b	1 (1)	1 (1) ^d	0.5 (1) ^d	1 (0.75) ^d	1 (1)	2 (0.5) ^d	1 (0) ^d
Lived experience as an asset, median (IQR) ^b	1 (1)	1 (1)	0 (1) ^d	1 (1)	1 (1)	1 (0.5) ^d	1 (1) ^d
Communication, median (IQR) ^b	1 (1)	1 (1)	0 (1) ^d	1 (1) ^d	1 (1)	1.5 (1.5)	1 (0) ^c

^aPSW: peer support worker.

^bScale 0 (face-to-face) to 3 (fully via the internet).

^cStrong consensus.

^dModerate consensus.

Discussion

Principal Findings

In this 21-country study, 20 topics were identified that can be recommended for inclusion in the curriculum of a PSW initial training program. There was a strong consensus about the high importance of five topics: *lived experience as an asset, ethics, PSW well-being, PSW role focus on recovery, and communication*. There were no substantial differences between role perspectives (PSW, managers, and researchers) and countries with different resource levels relating to importance. All training topics were identified as being partly or fully deliverable on the web, but none could be provided on the web without moderation. There was no consensus about the right balance between face-to-face and web-based training with moderation, even though PSWs were more consistent in identifying the need for a face-to-face training component.

Strengths and Limitations

A strength of this study is the number of participants (N=110) from different countries (n=21) and the low attrition (round 1-2: 19%; round 1-3: 25%) compared with other Delphi studies [46]. Another strength is that the Delphi was reported in line with the Checklist for Reporting Results of Internet e-Surveys checklist [47]. This study has several limitations. First, there is a need for more representation from middle- and lower-income countries, which might have allowed between-setting differences to emerge, which was not achieved despite purposive sampling efforts. Second, participation in a web-based consultation may be more difficult for people in environments with poorer internet access and intermittent electricity, which may disproportionately affect PSWs. Third, participants were asked if they had completed web-based training earlier but not specifically web-based PSW training, which could then have been further explored in the analysis. In addition, web-based deliverability was not defined and was based on participant judgment rather than evidence from experience. Fourth, the use of two web-based platforms that may have confused participants and were not specifically designed for Delphi consultations. Alternative Delphi-specific platforms exist, including ExpertLens [48], Mesydel [49], and Delphi2 [50]. Fifth, PSW training manuals available in languages other than English or Arabic might have identified a wider range of training topics.

Finally, a full systematic review was not conducted, and the limitations associated with implementing a systematic review include the following: (1) lack of patient, population, intervention, comparison, and outcomes criteria; (2) included and excluded manuals were not listed; (3) methodological quality assessment and reliability of manuals were not explored; and (4) discrepancies between reviewers were not reported.

Comparison With Other Work

Achieving an international consensus on topics that are of high importance in PSW training is important for three reasons. First, it offers prospective PSWs a description of the tasks and skills involved in the PSW role, which may inform their decision making about whether to train as a PSW. Second, it provides training providers and organizations with a list of core topics that should be covered in the content of initial PSW training programs in all settings and two additional context-specific topics that may be relevant. Third, it provides an evidence base for developing training curricula and a framework for PSW accreditation. The standardization of PSW training across settings and countries is contentious. On the one hand, an international consortium of peer leaders from 6 continents developed an international charter, which defined peer support and identified key principles and guiding values [51]. In conjunction with this Delphi consultation, a framework is emerging that could underpin the international PSW accreditation process. On the contrary, unintended consequences of institutionalizing the PSW role are emerging, with one qualitative study in the United States concluding that it “has the potential to reduce the very centrality of experiential expertise, reproduce social inequalities, and paradoxically impact stigma” [52].

The identification of training topics relevant to specific contexts reflects cultural and organizational influences on implementation [53]. Identified barriers that may lead to context-specific modifications include the lack of credibility of peer worker roles, professionals’ negative attitudes, tensions with service users, struggles with identity construction, cultural impediments, poor organizational arrangements, and inadequate overarching social and mental health policies [54]. These influences can lead to preplanned modifications implemented during initial PSW training, as well as unplanned extensions to the PSW role [55]. Several studies have found that this role extension can reduce role clarity and integrity, such as by incorporating medical ways of working [56] and creating identity conflict [57]. For example, a 10-site comparative case study across England found that different understandings of professionalism and practice boundaries can erode the distinctiveness of the PSW role [58].

PSW training has evolved in response to organizational needs and more recently the COVID-19 global pandemic. In a recent commentary, barriers to implementing web-based peer support in low- and middle-income countries in the context of COVID-19 were identified [59]. The low-to-moderate consensus about web-based delivery of training found in our study indicates that further work is needed to explore the relative costs and benefits of web-based versus face-to-face training.

All topics were rated as candidates for at least partial web-based delivery, which raises two questions. First, what is the role of web-based moderation? In addition to knowledge and skills

development, an important component of PSW training is ensuring that participants have the ability to maintain role integrity in a context where many will have to deal with microaggressions [60]. Similarly, a recent editorial identified specific contested areas relating to the role of PSW in restraint, administration of medication, and lone working in the community [61]. These may all be sensitive issues for PSWs to explore in training, for example, due to personal experiences, which may be more difficult to explore in moderated web-based discussions. Furthermore, individuals considering PSW training may struggle with motivation [62] and the pressure to succeed [63], and role challenges can include overwork and symptom recurrence [64]. There is some evidence that a therapeutic alliance in digital interventions is possible [65], but the extent to which the requisite resilience and motivation for the PSW role can be fostered through web-based training delivery is an important future focus.

Second, does web-based training prepare recipients better to deliver web-based peer support? Relationships are central to PSWs [66], and one impact of COVID-19 is to increase the use of web-based approaches by trained PSWs as an alternative relationship medium. Combining web-based and offline peer

support has been shown to be a promising concept, with one qualitative Norwegian study of peer support recipients finding it enabled connectedness and allowed individuals to balance anonymity and openness [67]. Web-based training may help future PSWs to have both technological skills and the confidence to engage with PSW recipients on the web. In middle- and low-income countries, this blend of training delivery could also provide an accessible, wide-reaching, and cost-effective approach to increase the availability of PSW training places. A systematic review identified that the role content of PSWs is often underreported [68]. The topics identified in our study can inform the reporting of both the training program and PSW role components in future PSW evaluations.

Conclusions

This study developed a list of training topics for the initial PSW training. One use is to inform PSW training manuals, such as the UPSIDES PSW training program [69], which is being evaluated in Germany, India, Israel, Tanzania, and Uganda [70]. The use of an evidence-based training curriculum will increase the effectiveness of programs to prepare individuals for working as PSWs.

Acknowledgments

The study UPSIDES is a multicenter collaboration between the Department for Psychiatry and Psychotherapy II at Ulm University, Germany (Bernd Puschner, coordinator); the Institute of Mental Health at University of Nottingham, United Kingdom (MS); the Department of Psychiatry at University Hospital Hamburg-Eppendorf, Germany (CM); Butabika National Referral Hospital, Uganda (Juliet Nakku); the Centre for Global Mental Health at London School of Hygiene and Tropical Medicine, United Kingdom (GR); Ifakara Health Institute, Dar es Salaam, Tanzania (Donat Shamba); the Department of Social Work at Ben Gurion University of the Negev, Be'er Sheva, Israel (Galia Moran); and the Centre for Mental Health Law and Policy, Pune, India (Jasmine Kalha). UPSIDES received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant 779263. MS acknowledges the support of the Center for Mental Health and Substance Abuse, University of South-Eastern Norway, the National Institute for Health Research Nottingham Biomedical Research Centre, and research group work support from the Economic and Social Research Council (grant ES/J500100/1 and ES/P000711/1). This publication only reflects the authors' views. The Commission is not responsible for any use that may be made of the information it contains. The funding bodies had no role in the design of the study, writing of the manuscript, or decision to submit the paper for publication.

Authors' Contributions

AC, MS, NI, RN, and CM conceptualized the study. AC, MS, and NI conducted a systematized review, analyzed, and interpreted the data. AC had full access to all the data in the study and had final responsibility for the decision to submit for publication. RN and CM contributed to the design of the Delphi consultation, analysis, and interpretation of data for the work. HN contributed to the design of the Delphi consultation, quality review, and data acquisition. LGM contributed to ethical processes. AC, NI, and MS drafted the manuscript. All authors contributed to the interpretation of the data, critically revised the manuscript, and approved the final submitted draft.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Search strategy for the systematized review.

[DOCX File, 14 KB - [mental_v8i5e25528_app1.docx](#)]

Multimedia Appendix 2

Coding framework developed from the thematic synthesis of peer support worker initial training manuals (n=32).

[DOCX File, 16 KB - [mental_v8i5e25528_app2.docx](#)]

Multimedia Appendix 3

Changes suggested in round 1 of the Delphi consultation (N=110).

[[DOCX File, 23 KB - mental_v8i5e25528_app3.docx](#)]

Multimedia Appendix 4

Round 2 Delphi consultation rating of importance (n=89).

[[DOCX File, 16 KB - mental_v8i5e25528_app4.docx](#)]

Multimedia Appendix 5

Round 2 Delphi consultation rating of web-based delivery (n=89).

[[DOCX File, 16 KB - mental_v8i5e25528_app5.docx](#)]

References

- Shalaby RA, Agyapong VI. Peer support in mental health: literature review. *JMIR Mental Health* 2020 Jun 09;7(6):e15572 [[FREE Full text](#)] [doi: [10.2196/15572](#)] [Medline: [32357127](#)]
- Mead S, Hilton D, Curtis L. Peer support: a theoretical perspective. *Psychiatr Rehabil J* 2001;25(2):134-141. [doi: [10.1037/h0095032](#)] [Medline: [11769979](#)]
- Solomon P. Peer support/peer provided services underlying processes, benefits, and critical ingredients. *Psychiatr Rehabil J* 2004;27(4):392-401. [doi: [10.2975/27.2004.392.401](#)] [Medline: [15222150](#)]
- van Gestel-Timmermans H, Brouwers E, van Assen M, van Nieuwenhuizen C. Effects of a peer-run course on recovery from serious mental illness: a randomized controlled trial. *Psychiatr Serv* 2012 Jan;63(1):54-60. [doi: [10.1176/appi.ps.201000450](#)] [Medline: [22227760](#)]
- Arbour S, Rose B. Improving relationships, lives and systems: the transformative power of a recovery college. *J Recovery Ment Health* 2018;1(3):1-6 [[FREE Full text](#)] [doi: [10.1002/9780470743171.index](#)]
- Mahlke C, Priebe S, Heumann K, Daubmann A, Wegscheider K, Bock T. Effectiveness of one-to-one peer support for patients with severe mental illness - a randomised controlled trial. *Eur Psychiatry* 2017 May;42:103-110. [doi: [10.1016/j.eurpsy.2016.12.007](#)] [Medline: [28364685](#)]
- Schrank B, Bird V, Rudnick A, Slade M. Determinants, self-management strategies and interventions for hope in people with mental disorders: systematic search and narrative review. *Soc Sci Med* 2012 Feb;74(4):554-564. [doi: [10.1016/j.socscimed.2011.11.008](#)] [Medline: [22240450](#)]
- Chinman M, George P, Dougherty R, Daniels A, Ghose S, Swift A, et al. Peer support services for individuals with serious mental illnesses: assessing the evidence. *Psychiatr Serv* 2014 Apr 01;65(4):429-441. [doi: [10.1176/appi.ps.201300244](#)] [Medline: [24549400](#)]
- Rivera J, Sullivan A, Valenti S. Adding consumer-providers to intensive case management: does it improve outcome? *Psychiatr Serv* 2007 Jun;58(6):802-809. [doi: [10.1176/ps.2007.58.6.802](#)] [Medline: [17535940](#)]
- Johnson S, Lamb D, Marston L, Osborn D, Mason O, Henderson C, et al. Peer-supported self-management for people discharged from a mental health crisis team: a randomised controlled trial. *Lancet* 2018 Aug 04;392(10145):409-418 [[FREE Full text](#)] [doi: [10.1016/S0140-6736\(18\)31470-3](#)] [Medline: [30102174](#)]
- Pathare S, Kalha J, Krishnamoorthy S. Peer support for mental illness in India: an underutilised resource. *Epidemiol Psychiatr Sci* 2018 Oct;27(5):415-419 [[FREE Full text](#)] [doi: [10.1017/S2045796018000161](#)] [Medline: [29618392](#)]
- Hall C, Baillie D, Basangwa D, Atukunda J. Brain gain in Uganda: a case study of peer working as an adjunct to statutory mental health care in a low-income country. In: *The Palgrave Handbook of Sociocultural Perspectives on Global Mental Health*. London, United Kingdom: Palgrave Macmillan; 2017:633-655.
- Pathare S, Brazinova A, Levav I. Care gap: a comprehensive measure to quantify unmet needs in mental health. *Epidemiol Psychiatr Sci* 2018 Oct;27(5):463-467 [[FREE Full text](#)] [doi: [10.1017/S2045796018000100](#)] [Medline: [29521609](#)]
- Repper J, Aldridge B, Gilfoyle S, Gillard S, Perkins R, Rennison J. Peer support workers: a practical guide to implementation. ImROC Briefing Paper 7; Centre for Mental Health, London. 2013. URL: <https://imroc.org/wp-content/uploads/2016/09/7-Peer-Support-Workers-a-practical-guide-to-implementation.pdf> [accessed 2021-05-10]
- Utschakowski J. Training programme for people with experience in mental health crisis to work as trainer and peer supporter. 2008. URL: https://www.adam-europe.eu/prj/1871/prd/3/1/Curriculum_english%20i.pdf [accessed 2019-03-16]
- Fortuna KL, Naslund JA, LaCroix JM, Bianco CL, Brooks JM, Zisman-Ilani Y, et al. Digital peer support mental health interventions for people with a lived experience of a serious mental illness: systematic review. *JMIR Ment Health* 2020 Apr 03;7(4):e16460 [[FREE Full text](#)] [doi: [10.2196/16460](#)] [Medline: [32243256](#)]
- Gammon D, Strand M, Eng LS, Børøsund E, Varsi C, Ruland C. Shifting practices toward recovery-oriented care through an e-recovery portal in community mental health care: a mixed-methods exploratory study. *J Med Internet Res* 2017 May 02;19(5):e145 [[FREE Full text](#)] [doi: [10.2196/jmir.7524](#)] [Medline: [28465277](#)]

18. Thomas N, Farhall J, Foley F, Leitan ND, Villagonzalo K, Ladd E, et al. Promoting personal recovery in people with persisting psychotic disorders: development and pilot study of a novel digital intervention. *Front Psychol* 2016;7:196 [FREE Full text] [doi: [10.3389/fpsy.2016.00196](https://doi.org/10.3389/fpsy.2016.00196)] [Medline: [28066271](https://pubmed.ncbi.nlm.nih.gov/28066271/)]
19. Gulliver A, Banfield M, Morse AR, Reynolds J, Miller S, Galati C. A peer-led electronic mental health recovery app in a community-based public mental health service: pilot trial. *JMIR Form Res* 2019 Jun 04;3(2):e12550 [FREE Full text] [doi: [10.2196/12550](https://doi.org/10.2196/12550)] [Medline: [31165708](https://pubmed.ncbi.nlm.nih.gov/31165708/)]
20. Kaplan K, Salzer MS, Solomon P, Brusilovskiy E, Cousounis P. Internet peer support for individuals with psychiatric disabilities: a randomized controlled trial. *Soc Sci Med* 2011 Jan;72(1):54-62. [doi: [10.1016/j.socscimed.2010.09.037](https://doi.org/10.1016/j.socscimed.2010.09.037)] [Medline: [21112682](https://pubmed.ncbi.nlm.nih.gov/21112682/)]
21. Fu Z, Burger H, Arjadi R, Bockting CL. Effectiveness of digital psychological interventions for mental health problems in low-income and middle-income countries: a systematic review and meta-analysis. *Lancet Psychiatry* 2020 Oct;7(10):851-864 [FREE Full text] [doi: [10.1016/S2215-0366\(20\)30256-X](https://doi.org/10.1016/S2215-0366(20)30256-X)] [Medline: [32866459](https://pubmed.ncbi.nlm.nih.gov/32866459/)]
22. Burke EM, Pyle M, Machin K, Morrison AP. Providing mental health peer support 1: a Delphi study to develop consensus on the essential components, costs, benefits, barriers and facilitators. *Int J Soc Psychiatry* 2018 Dec 03;64(8):799-812. [doi: [10.1177/0020764018810299](https://doi.org/10.1177/0020764018810299)]
23. Daniels AS, Bergeson S, Fricks L, Ashenden P, Powell I. Pillars of peer support: advancing the role of peer support specialists in promoting recovery. *J Ment Health Train Edu Practice* 2012 Jun 15;7(2):60-69. [doi: [10.1108/17556221211236457](https://doi.org/10.1108/17556221211236457)]
24. Cronise R, Teixeira C, Rogers ES, Harrington S. The peer support workforce: results of a national survey. *Psychiatr Rehabil J* 2016 Sep;39(3):211-221. [doi: [10.1037/prj0000222](https://doi.org/10.1037/prj0000222)] [Medline: [27618458](https://pubmed.ncbi.nlm.nih.gov/27618458/)]
25. Grant MJ, Booth A. A typology of reviews: an analysis of 14 review types and associated methodologies. *Health Info Libr J* 2009 Jun;26(2):91-108 [FREE Full text] [doi: [10.1111/j.1471-1842.2009.00848.x](https://doi.org/10.1111/j.1471-1842.2009.00848.x)] [Medline: [19490148](https://pubmed.ncbi.nlm.nih.gov/19490148/)]
26. Advanced Course Search (Multiple Criteria). Massive Open Online Courses List. URL: <https://www.mooc-list.com/multiple-criteria> [accessed 2019-03-15]
27. Mahlke C, Nixdorf R, Kalha J, Mpango R, Moran G, Mueller-Stierlin A, et al. A systematic review of influences on implementation of peer support work for adults with mental health problems. *PROSPERO* 2018 CRD42018107772. 2018. URL: https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=107772 [accessed 2019-03-14]
28. Substance Abuse and Mental Health Services Administration. URL: www.samhsa.gov [accessed 2019-03-15]
29. National Mental Health Commission. URL: www.mentalhealthcommission.gov.au [accessed 2019-03-15]
30. Scottish recovery network. URL: www.scottishrecovery.net [accessed 2019-03-16]
31. Mental health America. URL: www.mentalhealthamerica.net [accessed 2019-03-15]
32. Mental health innovation network. URL: www.mhinnovation.net [accessed 2019-03-15]
33. Depression and bipolar support alliance. URL: www.dbsalliance.org [accessed 2019-03-16]
34. REdeAmericas. URL: www.cugmhp.org/research/redeamericas/ [accessed 2019-03-15]
35. Missouri peer specialist. URL: www.mopeer specialist.com [accessed 2019-03-17]
36. Nevada Certification Board. URL: www.nevadacertboard.org [accessed 2019-03-17]
37. National Association of Peer Supporters (iNAPS). URL: www.inaops.org [accessed 2019-03-17]
38. Global Mental Health Peer Network. URL: www.mhinnovation.net/launch-global-mental-health-peer-network [accessed 2019-03-17]
39. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol* 2006 Jan;3(2):77-101. [doi: [10.1191/1478088706qp0630a](https://doi.org/10.1191/1478088706qp0630a)]
40. Hasson F, Keeney S, McKenna H. Research guidelines for the Delphi survey technique. *Journal of Advanced Nursing* 2000 Oct;32(4):1008-1015. [Medline: [11095242](https://pubmed.ncbi.nlm.nih.gov/11095242/)]
41. Donohoe H, Stelfox M, Tennant B. Advantages and limitations of the e-Delphi technique. *Am J Health Educ* 2013 Jan 23;43(1):38-46. [doi: [10.1080/19325037.2012.10599216](https://doi.org/10.1080/19325037.2012.10599216)]
42. Jorm AF. Using the Delphi expert consensus method in mental health research. *Aust N Z J Psychiatry* 2015 Oct;49(10):887-897. [doi: [10.1177/0004867415600891](https://doi.org/10.1177/0004867415600891)] [Medline: [26296368](https://pubmed.ncbi.nlm.nih.gov/26296368/)]
43. Kezar A, Maxey D. The Delphi technique: an untapped approach of participatory research. *Int J Soc Res Methodol* 2014 Jul 21;19(2):143-160. [doi: [10.1080/13645579.2014.936737](https://doi.org/10.1080/13645579.2014.936737)]
44. Williams PL, Webb C. The Delphi technique: a methodological discussion. *J Adv Nurs* 1994 Jan;19(1):180-186. [doi: [10.1111/j.1365-2648.1994.tb01066.x](https://doi.org/10.1111/j.1365-2648.1994.tb01066.x)] [Medline: [8138622](https://pubmed.ncbi.nlm.nih.gov/8138622/)]
45. Recovery Research Network (RRN). URL: <https://www.researchintorecovery.com/rrn> [accessed 2021-05-07]
46. Zelmer J, van Hoof K, Notarianni M, van Mierlo T, Schellenberg M, Tannenbaum C. An assessment framework for e-mental health apps in Canada: results of a modified Delphi process. *JMIR Mhealth Uhealth* 2018 Jul 09;6(7):e10016 [FREE Full text] [doi: [10.2196/10016](https://doi.org/10.2196/10016)] [Medline: [29986846](https://pubmed.ncbi.nlm.nih.gov/29986846/)]
47. Eysenbach G. Improving the quality of Web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). *J Med Internet Res* 2004 Sep 29;6(3):e34 [FREE Full text] [doi: [10.2196/jmir.6.3.e34](https://doi.org/10.2196/jmir.6.3.e34)] [Medline: [15471760](https://pubmed.ncbi.nlm.nih.gov/15471760/)]
48. ExpertLens. URL: www.rand.org/pubs/tools/expertlens.html [accessed 2020-08-10]
49. Mesydel. URL: www.mesydel.com [accessed 2020-08-10]
50. Delphi2. URL: <http://armstrong.wharton.upenn.edu/delphi2> [accessed 2020-08-10]

51. Stratford AC, Halpin M, Phillips K, Skerrett F, Beales A, Cheng V, et al. The growth of peer support: an international charter. *J of Mental Health* 2019 Dec;28(6):627-632. [doi: [10.1080/09638237.2017.1340593](https://doi.org/10.1080/09638237.2017.1340593)] [Medline: [28682640](https://pubmed.ncbi.nlm.nih.gov/28682640/)]
52. Adams WE. Unintended consequences of institutionalizing peer support work in mental healthcare. *Soc Sci Med* 2020 Oct;262:113249. [doi: [10.1016/j.socscimed.2020.113249](https://doi.org/10.1016/j.socscimed.2020.113249)] [Medline: [32768773](https://pubmed.ncbi.nlm.nih.gov/32768773/)]
53. Ibrahim N, Thompson D, Nixdorf R, Kalha J, Mpango R, Moran G, et al. A systematic review of influences on implementation of peer support work for adults with mental health problems. *Soc Psychiatry Psychiatr Epidemiol* 2020 Mar;55(3):285-293. [doi: [10.1007/s00127-019-01739-1](https://doi.org/10.1007/s00127-019-01739-1)] [Medline: [31177310](https://pubmed.ncbi.nlm.nih.gov/31177310/)]
54. Vandewalle J, Debyser B, Beeckman D, Vandecasteele T, Van Hecke A, Verhaeghe S. Peer workers' perceptions and experiences of barriers to implementation of peer worker roles in mental health services: a literature review. *Int J Nurs Stud* 2016 Aug;60:234-250. [doi: [10.1016/j.ijnurstu.2016.04.018](https://doi.org/10.1016/j.ijnurstu.2016.04.018)] [Medline: [27297384](https://pubmed.ncbi.nlm.nih.gov/27297384/)]
55. Charles A, Thompson D, Nixdorf R, Ryan G, Shamba D, Kalha J, et al. Typology of modifications to peer support work for adults with mental health problems: systematic review. *Br J Psychiatry* 2020 Jun;216(6):301-307. [doi: [10.1192/bjp.2019.264](https://doi.org/10.1192/bjp.2019.264)] [Medline: [31992375](https://pubmed.ncbi.nlm.nih.gov/31992375/)]
56. Alberta A, Ploski R. Cooption of peer support staff: quantitative evidence. *Rehab Proc Outcome* 2014 Sep 22;3:RPO.S12343. [doi: [10.4137/RPO.S12343](https://doi.org/10.4137/RPO.S12343)]
57. Walker G, Bryant W. Peer support in adult mental health services: a metasynthesis of qualitative findings. *Psychiatr Rehabil J* 2013 Mar;36(1):28-34. [doi: [10.1037/h0094744](https://doi.org/10.1037/h0094744)] [Medline: [23477647](https://pubmed.ncbi.nlm.nih.gov/23477647/)]
58. Gillard S, Holley J, Gibson S, Larsen J, Luccock M, Oborn E, et al. Introducing new peer worker roles into mental health services in England: comparative case study research across a range of organisational contexts. *Admin Pol Ment Health* 2015 Nov;42(6):682-694. [doi: [10.1007/s10488-014-0603-z](https://doi.org/10.1007/s10488-014-0603-z)] [Medline: [25331447](https://pubmed.ncbi.nlm.nih.gov/25331447/)]
59. Mpango R, Kalha J, Shamba D, Ramesh M, Ngakongwa F, Kulkarni A, et al. Challenges to peer support in low- and middle-income countries during COVID-19. *Glob Health* 2020 Sep 25;16(1):90 [FREE Full text] [doi: [10.1186/s12992-020-00622-y](https://doi.org/10.1186/s12992-020-00622-y)] [Medline: [32977816](https://pubmed.ncbi.nlm.nih.gov/32977816/)]
60. Firmin RL, Mao S, Bellamy CD, Davidson L. Peer support specialists' experiences of microaggressions. *Psychol Serv* 2019 Aug;16(3):456-462 [FREE Full text] [doi: [10.1037/ser0000297](https://doi.org/10.1037/ser0000297)] [Medline: [30382746](https://pubmed.ncbi.nlm.nih.gov/30382746/)]
61. Perkins R, Repper J. Where is peer support going? *Ment Health Soc Incl* 2019 Jun 06;23(2):61-63. [doi: [10.1108/MHSI-05-2019-060](https://doi.org/10.1108/MHSI-05-2019-060)]
62. Moran GS, Russinova Z, Yim JY, Sprague C. Motivations of persons with psychiatric disabilities to work in mental health peer services: a qualitative study using self-determination theory. *J Occup Rehabil* 2014 Mar;24(1):32-41. [doi: [10.1007/s10926-013-9440-2](https://doi.org/10.1007/s10926-013-9440-2)] [Medline: [23576121](https://pubmed.ncbi.nlm.nih.gov/23576121/)]
63. Otte I, Werning A, Nossek A, Vollmann J, Juckel G, Gather J. Challenges faced by peer support workers during the integration into hospital-based mental health-care teams: results from a qualitative interview study. *Int J Soc Psychiatry* 2020 May 12;66(3):263-269. [doi: [10.1177/0020764020904764](https://doi.org/10.1177/0020764020904764)] [Medline: [32046565](https://pubmed.ncbi.nlm.nih.gov/32046565/)]
64. Moran GS, Russinova Z, Gidugu V, Gagne C. Challenges experienced by paid peer providers in mental health recovery: a qualitative study. *Community Ment Health J* 2013 Jun;49(3):281-291. [doi: [10.1007/s10597-012-9541-y](https://doi.org/10.1007/s10597-012-9541-y)] [Medline: [23117937](https://pubmed.ncbi.nlm.nih.gov/23117937/)]
65. Tremain H, McEnery C, Fletcher K, Murray G. The therapeutic alliance in digital mental health interventions for serious mental illnesses: narrative review. *JMIR Mental Health* 2020 Aug 07;7(8):e17204 [FREE Full text] [doi: [10.2196/17204](https://doi.org/10.2196/17204)] [Medline: [32763881](https://pubmed.ncbi.nlm.nih.gov/32763881/)]
66. Gillard S, Gibson S, Holley J, Luccock M. Developing a change model for peer worker interventions in mental health services: a qualitative research study. *Epidemiol Psychiatr Sci* 2015 Oct;24(5):435-445. [doi: [10.1017/S2045796014000407](https://doi.org/10.1017/S2045796014000407)] [Medline: [24992284](https://pubmed.ncbi.nlm.nih.gov/24992284/)]
67. Strand M, Eng L, Gammon D. Combining online and offline peer support groups in community mental health care settings: a qualitative study of service users' experiences. *Int J Ment Health Syst* 2020;14(1):a. [doi: [10.2196/preprints.9804](https://doi.org/10.2196/preprints.9804)]
68. King AJ, Simmons MB. A systematic review of the attributes and outcomes of peer work and guidelines for reporting studies of peer interventions. *Psychiatr Serv* 2018 Sep 01;69(9):961-977. [doi: [10.1176/appi.ps.201700564](https://doi.org/10.1176/appi.ps.201700564)] [Medline: [29962310](https://pubmed.ncbi.nlm.nih.gov/29962310/)]
69. Puschner B, Repper J, Mahlke C, Nixdorf R, Basangwa D, Nakku J, et al. Using Peer Support in Developing Empowering Mental Health Services (UPSIDES): background, rationale and methodology. *Ann Glob Health* 2019 Apr 05;85(1):53 [FREE Full text] [doi: [10.5334/aogh.2435](https://doi.org/10.5334/aogh.2435)] [Medline: [30951270](https://pubmed.ncbi.nlm.nih.gov/30951270/)]
70. Moran G, Kalha J, Mueller-Stierlin A, Kilian R, Krumm S, Slade M, et al. Peer support for people with severe mental illness versus usual care in high-, middle- and low-income countries: study protocol for a pragmatic, multicentre, randomised controlled trial (UPSIDES-RCT). *Trials* 2020 May 01;21(1):371 [FREE Full text] [doi: [10.1186/s13063-020-4177-7](https://doi.org/10.1186/s13063-020-4177-7)] [Medline: [32357903](https://pubmed.ncbi.nlm.nih.gov/32357903/)]

Abbreviations

PSW: peer support worker

UPSIDES: Using Peer Support in Developing Empowering Mental Health Services

Edited by J Torous; submitted 05.11.20; peer-reviewed by K Fortuna, T van Mierlo; comments to author 05.01.21; revised version received 15.01.21; accepted 10.03.21; published 27.05.21.

Please cite as:

Charles A, Nixdorf R, Ibrahim N, Meir LG, Mpango RS, Ngakongwa F, Nudds H, Pathare S, Ryan G, Repper J, Wharrad H, Wolf P, Slade M, Mahlke C

Initial Training for Mental Health Peer Support Workers: Systematized Review and International Delphi Consultation

JMIR Ment Health 2021;8(5):e25528

URL: <https://mental.jmir.org/2021/5/e25528>

doi: [10.2196/25528](https://doi.org/10.2196/25528)

PMID: [34042603](https://pubmed.ncbi.nlm.nih.gov/34042603/)

©Ashleigh Charles, Rebecca Nixdorf, Nashwa Ibrahim, Lion Gai Meir, Richard S Mpango, Fileuka Ngakongwa, Hannah Nudds, Soumitra Pathare, Grace Ryan, Julie Repper, Heather Wharrad, Philip Wolf, Mike Slade, Candelaria Mahlke. Originally published in JMIR Mental Health (<https://mental.jmir.org>), 27.05.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Mental Health, is properly cited. The complete bibliographic information, a link to the original publication on <https://mental.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Learning From Clinical Consensus Diagnosis in India to Facilitate Automatic Classification of Dementia: Machine Learning Study

Haomiao Jin¹, PhD; Sandy Chien¹, MS; Erik Meijer^{1,2}, PhD; Pranali Khobragade¹, MD; Jinkook Lee^{1,2,3}, PhD

¹Center for Economic and Social Research, University of Southern California, Los Angeles, CA, United States

²RAND Corporation, Santa Monica, CA, United States

³Department of Economics, University of Southern California, Los Angeles, CA, United States

Corresponding Author:

Haomiao Jin, PhD

Center for Economic and Social Research

University of Southern California

635 Downey Way, VPD

Los Angeles, CA, 90089

United States

Phone: 1 626 554 3370

Email: haomiaoj@usc.edu

Abstract

Background: The Harmonized Diagnostic Assessment of Dementia for the Longitudinal Aging Study in India (LASI-DAD) is the first and only nationally representative study on late-life cognition and dementia in India (n=4096). LASI-DAD obtained clinical consensus diagnosis of dementia for a subsample of 2528 respondents.

Objective: This study develops a machine learning model that uses data from the clinical consensus diagnosis in LASI-DAD to support the classification of dementia status.

Methods: Clinicians were presented with the extensive data collected from LASI-DAD, including sociodemographic information and health history of respondents, results from the screening tests of cognitive status, and information obtained from informant interviews. Based on the Clinical Dementia Rating (CDR) and using an online platform, clinicians individually evaluated each case and then reached a consensus diagnosis. A 2-step procedure was implemented to train several candidate machine learning models, which were evaluated using a separate test set for predictive accuracy measurement, including the area under receiver operating curve (AUROC), accuracy, sensitivity, specificity, precision, F1 score, and kappa statistic. The ultimate model was selected based on overall agreement as measured by kappa. We further examined the overall accuracy and agreement with the final consensus diagnoses between the selected machine learning model and individual clinicians who participated in the clinical consensus diagnostic process. Finally, we applied the selected model to a subgroup of LASI-DAD participants for whom the clinical consensus diagnosis was not obtained to predict their dementia status.

Results: Among the 2528 individuals who received clinical consensus diagnosis, 192 (6.7% after adjusting for sampling weight) were diagnosed with dementia. All candidate machine learning models achieved outstanding discriminative ability, as indicated by AUROC > .90, and had similar accuracy and specificity (both around 0.95). The support vector machine model outperformed other models with the highest sensitivity (0.81), F1 score (0.72), and kappa (.70, indicating substantial agreement) and the second highest precision (0.65). As a result, the support vector machine was selected as the ultimate model. Further examination revealed that overall accuracy and agreement were similar between the selected model and individual clinicians. Application of the prediction model on 1568 individuals without clinical consensus diagnosis classified 127 individuals as living with dementia. After applying sampling weight, we can estimate the prevalence of dementia in the population as 7.4%.

Conclusions: The selected machine learning model has outstanding discriminative ability and substantial agreement with a clinical consensus diagnosis of dementia. The model can serve as a computer model of the clinical knowledge and experience encoded in the clinical consensus diagnostic process and has many potential applications, including predicting missed dementia diagnoses and serving as a clinical decision support tool or virtual rater to assist diagnosis of dementia.

(JMIR Ment Health 2021;8(5):e27113) doi:[10.2196/27113](https://doi.org/10.2196/27113)

KEYWORDS

dementia; Alzheimer disease; machine learning; artificial intelligence; diagnosis; classification; India; model

Introduction

The World Health Organization estimates that the number of people living with dementia worldwide is approximately 50 million and will almost triple by 2050 [1], with nearly 60% living in low- and middle-income countries like India [2]. Developing effective population-based interventions to address the rising burden of dementia depends on high-quality nationally representative data, which is often scarce in low- and middle-income countries. The Alzheimer's and Related Disorders Society of India estimates that more than 3.7 million Indians have dementia. However, this figure is based on a meta-analysis of prevalence studies with estimated prevalence rates ranging from 0.6% to 10.6% in rural areas and from 0.9% to 7.5% in urban areas [3,4]. The high heterogeneity in reported prevalence could be due to a variety of methodological issues including regional variations and different diagnostic criteria [3].

The Longitudinal Aging Study in India (LASI) is the first and only nationally representative survey of the physical and cognitive health, economic welfare, and social well-being for the country's aging population, with a sample of more than 70,000 individuals aged 45 years and older [3]. The Harmonized Diagnostic Assessment of Dementia for the Longitudinal Aging Study in India (LASI-DAD) further extends the LASI's cognitive data collection by conducting in-depth neuropsychological tests and informant interviews for a subsample of the LASI respondents aged 60 years and older [3]. The design of the LASI-DAD closely follows the Harmonized Cognitive Assessment Protocol (HCAP), which was developed for the assessment of dementia and mild cognitive impairment in the US Health and Retirement Study (HRS) and its associated studies around the world to enable international research collaboration [5].

For conditions such as Alzheimer disease, dementia, and mild cognitive impairment, there is no single definitive diagnostic test. Hence, many clinical researchers rely on a clinical consensus diagnostic process, consisting of data review, adjudication, and consensus by a panel of expert clinicians [6,7]. However, for large population surveys, the gold standard of clinician in-person assessment of respondents and all relevant information from their informants and in-person consensus conference is costly [6,7]. One way to reduce the cost is to replace the in-person consensus conference with a web-based consensus diagnosis approach. This web-based method was implemented first in the Monongahela-Youghiogheny Healthy Aging Team Project [8] and then in the LASI-DAD [9], which developed an online clinical consensus diagnosis platform that provided the detailed information necessary for a clinical assessment [9] and obtained the Clinical Dementia Rating (CDR) for a subsample of the LASI-DAD participants (n=2528).

The objective of this study is to develop a machine learning model that uses information from the clinical consensus diagnosis in the LASI-DAD for classification of dementia. The

resulting machine learning model can serve as a computer model of the clinical knowledge and experience encoded in the clinical consensus diagnostic process. Furthermore, the machine learning model can assist in predicting the dementia status of a subgroup of the LASI-DAD respondents who participate in the extensive cognitive tests and informant interviews but do not obtain the clinical consensus diagnosis due to missing information. The predicted data will become publicly available as a part of the LASI-DAD dataset for potential use in future studies.

This study is, to the best of our knowledge, the first machine learning study on dementia using a nationally representative sample from India. As a part of the LASI-DAD project, this study contributes to a global HCAP-based initiative to advance aging research based on the collection, sharing, and analysis of population data on cognition and dementia [10,11].

Methods

Overall Design

The LASI-DAD data were collected from the larger LASI project between October 2017 and March 2020 and involved a stratified random sample of 4096 individuals aged 60 years and over [3]. All LASI-DAD participants received an extensive cognitive assessment, and interviews were conducted with informants who knew the individual well. The collected data were used in clinical consensus diagnoses by a clinical expert panel to evaluate dementia status based on the CDR. A total of 2528 LASI-DAD participants received clinical consensus diagnoses, while the remaining 1568 individuals did not progress through the diagnostic process. This study developed a machine learning model using the same predictors as the LASI-DAD assessment and informant interview data in the clinical consensus diagnosis. The developed model predicts dementia diagnosis for individuals without consensus diagnosis.

Assessment

The LASI-DAD protocol included a cognitive assessment; self-reported functional difficulties, depression, and anxiety; and an interview of an informant (a relative or friend who knows the individual well) about the respondent's cognitive status and everyday activities. The main LASI collected rich data on sociodemographic status and health history, which were provided to clinicians for evaluation of the CDR. The data presented for the clinical consensus diagnoses were used as predictors in developing the machine learning model.

Cognitive Assessment

The Hindi Mental State Examination [12,13] is an assessment with questions related to tasks, including time orientation, place orientation, 3-word recall, and object naming. Example questions are "What is the year?" and "Can you tell me where we are now? What state? What city?" A summary score is calculated by summing the number of correct answers and ranges from 0 to 30, with a larger number indicating more correct answers.

The Telephone Interview for Cognitive Status (TICS) [14] is a widely used brief questionnaire with 3 questions. An example question is “What do people usually use to cut paper?” (Correct answer: scissors or shears.) The summary score is the total number of correct answers.

The Community Screening Instrument for Dementia (CSID) [15] is a brief assessment with 4 items, including “Where is the local market/store?” and “Point to the window and then the door.” The summary score is the total number of correct answers.

The judgment and problem-solving assessment [16] includes 5 questions, including “What is the difference between a lie and a mistake?” and “What will you do if you find a lost child on the road?” The summary score is the total number of correct answers.

Finally, there are 5 numeracy questions [17], with examples like “How many 25 paisa coins will you give me for one Rupee?” and “If 5 people all have the winning numbers in the lottery and the prize is 1000 Rupees, how much will each of them get?” The summary score is the total number of correct answers.

Self-Reported Functional Difficulties

Activities of daily living (ADLs) [18] assess difficulties in basic self-care tasks including dressing, walking, bathing, eating, getting in or out of bed, and using the toilet. Respondents can choose between yes and no when answering. The summary score is the total number of difficulties, with a higher score indicating more difficulties.

Instrumental activities of daily living [18,19] assess difficulties in daily-life tasks including preparing a meal, shopping for groceries, making phone calls, taking medications, doing housework, managing finances, and getting around or finding an address in an unfamiliar place. The summary score is the total number of difficulties, with a higher score indicating more difficulties.

The LASI mobility module assesses difficulties in 9 tasks such as walking 100 yards, sitting for 2 hours or more, and getting up from a chair after sitting for a long period. The summary score is the total number of difficulties, with a higher score indicating more difficulties.

Depression and Anxiety

Depression was assessed using the 10-item Center for Epidemiological Studies Depression Scale [20], which assesses 10 depressive symptoms in the past week such as having trouble concentrating, feeling depressed, and feeling tired or low in energy. The respondents can choose answers from rarely or never, sometimes, often, or most or all of the time, and scores are coded from 0 to 3. The summary score is calculated by summing the scores of each item and has a range from 0 to 30, with higher scores indicating more depressive symptoms.

Anxiety was assessed using a 5-item scale that is a subset of the Beck Anxiety Inventory [21], which measures anxiety symptoms in the past week, including a fear of the worst happening, being nervous, feeling hands tremble, a fear of dying,

and feeling faint. The respondents can choose answers from never, hardly ever, some of the time, or most or all of the time, and scores are coded from 0 to 3. The summary score is calculated by summing the scores for each item and has a range from 0 to 15, with higher scores indicating more anxiety symptoms.

Informant Interview

The LASI-DAD asked respondents to nominate a close family member or friend as an informant who knows the respondent well, interacts with the respondent frequently, knows the respondent’s daily functions, and can report on the respondent [3]. The informant interview consisted of questions about the respondent’s functional status, social engagement, and memory.

The Informant Questionnaire on Cognitive Decline in the Elderly [22] includes 16 items asking the informant to compare the functional status and memory of the respondent to 10 years ago. Example questions include “How is the respondent at remembering things about family and friends, such as occupations, birthdays, and addresses compared with 10 years ago?” and “How is the respondent at handling money for shopping compared with 10 years ago?” The respondent can choose from much improved, a bit better, not much change, a bit worse, or much worse, and scores are coded from 1 to 5. The summary score is calculated as the average of all item scores.

The Blessed Dementia Scale [23] includes 8 questions for the informant to assess the change in performance and habits of the respondent. An example question is “How well is the respondent able to perform household tasks?” The informant can choose answers from no loss, some loss, and severe loss. If the informant answered some loss or severe loss, a further question is asked: “Is this loss due to physical reasons, mental reasons, or both?” The summary score is calculated by assigning 0 for no loss or loss only due to physical reasons, 0.5 for some loss attributed to mental reasons or both, and 1 for severe loss attributed to mental reasons or both and summing these scores, resulting in a summary score ranging from 0 to 8, with values in multiples of 0.5. There are 3 questions about the habits of the respondent. An example question is “Regarding eating, would you say the respondent feeds himself/herself without assistance, with minor assistance, with much assistance, or has to be fed?” The answer is scored from 1 to 4, with higher scores indicating more difficulties. The summary score is the average score of the 3 items.

Additionally, there are questions for the informant to assess the signs of cognitive change, signs of cognitive impairment, and everyday activities. Example questions include “Does the respondent have difficulty in adjusting to change in the respondent’s daily routine?” for assessing signs of cognitive change, “Has there been a general decline in the respondent’s mental functioning?” for assessing signs of cognitive impairment, and “How often does the respondent go to work or volunteer?” for assessing everyday activities.

Sociodemographic Variables

The sociodemographic variables include age, marital status (married or not), gender, and years of education.

Health History

Health history includes systolic and diastolic blood pressure and previous diagnosis of stroke, heart disease, diabetes, hypertension, depression, dementia, psychiatric problems, neurological problems, vision impairment, and hearing impairment.

Clinical Consensus Diagnosis of Dementia

Obtaining the ground truth is a challenge for all machine learning studies on dementia because there is no single definitive test of the disease. For the basis of the clinical diagnosis of dementia, clinicians used the CDR [16], a global rating device first introduced in a prospective study of patients with dementia [24] that is now widely used to measure dementia severity [25,26]. The CDR comprises 6 cognitive and functional domains [16]: (1) memory, (2) orientation, (3) judgment and problem solving, (4) community affairs, (5) home and hobbies, and (6) personal care. Clinicians complete the CDR ratings based on cognitive test results and informant reports. As noted earlier, the LASI-DAD project built a web-based approach to reach diagnostic consensus [3]. For each individual, at least 3 clinicians were assigned to the first round of review. Each clinician reviewed the case and provided ratings for the subdomains, and based on the subdomain ratings, the CDR algorithm automatically generated a global rating: 0 (normal), 0.5 (very mild dementia), 1 (mild dementia), 2 (moderate dementia), or 3 (severe dementia). For cases where individual global ratings differed, an automatic email was sent to the assigned reviewers, giving them a chance to review the case and read other raters' comments and update their ratings, if desired [3]. This second round of review might reach a consensus for additional cases. For cases where consensus was not reached after the second round of review, a group of clinicians discussed the case through a virtual consensus meeting to determine the global CDR rating. This clinical consensus diagnostic process can surpass the accuracy of individual expert diagnoses and is considered the gold standard for clinical diagnosis of dementia [27]. An individual was classified as having dementia if the global CDR rating from the clinical consensus diagnostic process was equal to or greater than 1.

Statistical Analysis

This study generated descriptive statistics of the data for developing the machine learning model. Available data were then divided into a training set with a random selection of 70% of the sample and a test set involving the remaining 30% of the sample. We trained several candidate machine learning models using the training set, including stochastic gradient boosting, random forest, support vector machine, elastic net, multivariate adaptive regression splines, and multilayer perceptron. Stochastic gradient boosting is an ensemble learning method that produces a prediction model based on weak prediction models, typically decision trees [28]. Random forest constructs a multitude of decision trees at model training and outputs the ultimate prediction that is the mode of the individual trees [29]. Support vector machine constructs hyperplanes to separate different categories of training samples [30]. A radial basis function kernel was used in this study with a support vector machine to construct nonlinear separations [30]. Elastic net is

a regularized regression method that linearly combines the L1 and L2 regularization to achieve improved predictive accuracy [31]. The multivariate adaptive regression splines model is a nonparametric regression technique that automatically models nonlinear relationships [32]. Finally, multilayer perceptron is a type of fully connected artificial neural networks with an input layer, one or more hidden layers, and an output layer [30]. A multilayer perceptron with multiple hidden layers is often known as a type of deep neural network [33]. The model training process as described below tuned the number of hidden layers, number of neurons in each layer, and a weight decay parameter for reducing model overfitting to select the best structure of the neural network.

We trained the candidate machine learning models using a 2-step process. First, we fitted the models based on the training set using repeated cross-validation with 10 repetitions and 10 folds of validation [34]. The objective of this step was to optimize the models' overall discriminative abilities by tuning model meta-parameters, such as the number of decision trees in a random forest and the number of hidden layers in a multilayer perceptron, and using the fitted models to generate predicted risk scores for each training sample (calculated as 100 times the predicted probability of dementia). Whenever possible, a weight inversely proportional to the number of individuals with dementia in the training set was used to account for the imbalance between individuals with versus individuals without dementia in the data. The overall discriminative ability was evaluated by the area under the receiver operating curve (AUROC), which is a measure based on the sensitivity (ie, number of true positive divided by all positive cases) and specificity (ie, number of true negative divided by all negative cases) of different cutoffs for the predicted risk scores. AUROC has a range from 0 to 1, and an AUROC score of more than 0.9 is considered outstanding [35]. In the second step, we trained a majority-voting process that outputted the final classification of dementia by combining 4 weak classifications derived from the predicted risk score. For each individual, the process attached 4 respective group memberships based on the individual's depression and anxiety assessment scores (in the top quartile or not) and whether the individual had vision or hearing impairment. The cutoff scores for each group were selected to maximize the F1 score, which is a summary score calculated from sensitivity and precision (true positive cases divided by the total number of predicted positive cases). Based on the comparisons between the predicted risk score and group-specific cutoffs, each individual received 4 respective weak classifications. An individual was assigned as having dementia if at least 3 weak classifications were positive. Design of this majority-voting process was informed by the clinical evidence that cognitive decline and daily life difficulties may be attributed to alternative conditions other than dementia, such as depression, anxiety, and vision and hearing impairment [36-40]. The above described model training process was implemented by using the R functions `trainControl` and `train` in the `caret` package [41].

We tested the predictive accuracy of the candidate machine learning models using the test set. Predictive accuracy was evaluated by the AUROC of the predicted risk scores and accuracy, sensitivity, specificity, precision, F1 score, and kappa

of the final classifications. We selected the model with the highest kappa as the ultimate prediction model because kappa measures the overall agreement between the predicted classifications and the clinical consensus diagnoses and corrects for the imbalance between positive and negative cases. In general, kappa between .60 and .80 indicates substantial agreement and kappa greater than .80 indicates almost perfect agreement [42]. We further compared the overall accuracy and agreement with the final consensus diagnoses between the selected machine learning model and clinicians who participated in the clinical consensus diagnostic process. Finally, we applied the selected model to predict dementia for individuals without clinical consensus diagnoses. As mentioned earlier, the predicted data will be publicly available as a part of the LASI-DAD dataset.

Results

The sample used to develop the machine learning model included 2528 individuals from the LASI-DAD who received a clinical consensus diagnosis of dementia (Table 1). The sample included 192 individuals living with dementia.

Random selection split the data into a training set with 1770 individuals and a test set with 758 individuals. There were 138 individuals diagnosed with dementia in the training set and 54 individuals diagnosed with dementia in the test set. Evaluation results of the candidate machine learning models on the test set

are shown in Table 2. The tuned multilayer perceptron has one hidden layer with 5 neurons. All candidate models achieved outstanding discriminative ability with AUROC >.90 and similar accuracy and specificity. However, the support vector machine outperformed the other models with the highest sensitivity, F1 score, and kappa and the second highest precision. The support vector machine was selected as the ultimate model since it has the best kappa, indicating the best overall agreement between predicted classifications and clinical consensus diagnoses.

Further examination revealed that accuracy of the selected prediction model (ie, support vector machine) is similar to that of the clinicians who participated in the clinical consensus diagnosis. A total of 12 clinicians participated in the consensus diagnostic process for the 758 individuals in the test set. Compared with the final consensus diagnoses, the average accuracy of clinicians was 0.96 (95% CI 0.94-0.98) and the average kappa was .75 (95% CI 0.61-0.88). There were no significant differences between the selected prediction model and the participating clinicians (accuracy $P=.64$ and kappa $P=.46$).

Application of the selected prediction model to the 1586 individuals without clinical consensus diagnoses results in 127 individuals classified as living with dementia. Hence, the unweighted estimated dementia prevalence in the total sample is $(127+192)/4096=7.8\%$. Applying sampling weights, we can estimate the prevalence in the population as 7.4%.

Table 1. Descriptive statistics of the data used to develop the machine learning model (n=2528).

Characteristics	Unweighted group	Weighted group
Age (years, range 6-103), mean (SD)	68.69 (7.50)	68.54 (7.35)
Married, n (%)	1672 (66.14)	1856 (67.10)
Female, n (%)	1204 (47.63)	1403 (50.70)
Education (years, range 0-20), mean (SD)	3.64 (4.60)	3.40 (4.57)
Blood pressure (systolic, range 76.5-225.0), mean (SD)	138.38 (23.35)	138.06 (23.62)
Blood pressure (diastolic, range 47.5-137.0), mean (SD)	82.73 (12.46)	82.73 (12.46)
Previous diagnosis, n (%)		
Stroke	82 (3.24)	88 (3.18)
Heart disease	155 (6.13)	146 (5.28)
Diabetes	376 (14.87)	352 (12.72)
Hypertension	964 (38.13)	939 (34.12)
Depression	21 (0.83)	23 (0.83)
Dementia	26 (1.03)	25 (0.90)
Psychiatric problems	17 (0.67)	14 (0.51)
Neurologic problems	58 (2.29)	63 (2.28)
Vision impairment	1113 (44.03)	1222 (44.16)
Hearing impairment	712 (28.16)	770 (27.83)
HMSE ^a (range 0-30), mean (SD)	22.11 (5.87)	22.15 (5.73)
TICS ^b (range 0-3), mean (SD)	2.01 (0.92)	1.97 (0.93)
CSID ^c (range 0-4), mean (SD)	3.31 (0.94)	3.29 (0.95)
Judgment and problem solving (range 0-5), mean (SD)	2.21 (1.51)	2.14 (1.52)
Numeracy (range 0-9), mean (SD)	3.85 (2.64)	3.83 (2.61)
ADL ^d (range 0-6), mean (SD)	1.34 (1.76)	1.30 (1.75)
IADL ^e (range 0-7), mean (SD)	2.27 (2.24)	2.27 (2.23)
Difficulties in mobility (range 0-9), mean (SD)	3.72 (2.92)	3.65 (2.94)
Depressive symptoms (range 0-30), mean (SD)	9.92 (5.25)	10.30 (5.17)
Anxiety symptoms (range 0-15), mean (SD)	2.94 (3.30)	3.09 (3.37)
IQCODE ^f (range 1-5), mean (SD)	3.51 (0.56)	3.48 (0.54)
Blessed Dementia Scale—changes in habits (range 1-3.67), mean (SD)	1.09 (0.32)	1.07 (0.28)
Blessed Dementia Scale—changes in performance (range 0-8), mean (SD)	1.26 (1.71)	1.20 (1.63)
Global CDR^g, n (%)		
0 (no dementia)	768 (30.38)	854 (30.86)
0.5 (very mild dementia)	1568 (62.03)	1726 (62.38)
1 (mild dementia)	162 (6.41)	160 (5.78)
2 (moderate dementia)	25 (0.99)	24 (0.87)
3 (severe dementia)	5 (0.20)	2 (0.07)
Diagnosis of dementia (global CDR ≥ 1), n (%)	192 (7.59)	186 (6.72)

^aHMSE: Hindi Mental State Examination.^bTICS: Telephone Interview for Cognitive Status.^cCSID: Community Screening Instrument for Dementia.^dADL: activity of daily living.

^cIADL: instrumental activity of daily living.

^fIQCODE: Informant Questionnaire on Cognitive Decline in the Elderly.

^gCDR: Clinical Dementia Rating.

Table 2. Predictive performance of candidate machine learning models based on evaluation on the test set.

Model	AUROC ^a	Accuracy	Sensitivity	Specificity	Precision	F1	Kappa
Stochastic gradient boosting	.94	.95	.67	.97	.64	.66	.63
Random forest	.95	.95	.67	.98	.68	.67	.65
Support vector machine	.95	.96	.81	.97	.65	.72	.70
Elastic net	.95	.94	.65	.96	.58	.61	.58
Multivariate adaptive regression splines	.94	.94	.69	.96	.60	.64	.61
Multilayer perceptron	.95	.95	.67	.97	.65	.66	.63

^aAUROC: area under the receiver operating curve.

Discussion

Principal Findings

This study developed a machine learning model that uses clinical consensus diagnosis on dementia in a nationally representative survey of individuals aged 60 years and older from India. The ultimate prediction model is a support vector machine model with radial basis function kernel trained on a 2-step process. Validation results suggest that the prediction model has outstanding discriminative ability (AUROC >.90) and substantial agreement with clinical consensus diagnosis (kappa between .60 and .80). Compared with clinicians who participated in the clinical diagnostic process, the machine learning model demonstrates similar overall accuracy and agreement with the final consensus diagnoses. This finding suggests that the prediction model may serve as a decision support tool or even a virtual participating rater in the clinical consensus diagnostic process.

The developed machine learning model has many potential applications. First, as shown in this study, the model can be used to predict dementia diagnosis for individuals without clinical consensus diagnosis in the LASI-DAD project. Future data users may use the predicted data for various purposes, such as estimating the prevalence of dementia in India or examining risk factors of dementia. Second, the prediction model can be built into the online consensus website as a clinical decision support tool or a virtual participating rater to replace one of the clinicians. The next wave of LASI-DAD data collection will implement the machine learning model developed in this paper as a participating virtual rater to replace one of the clinicians in the consensus diagnostic process for a proportion of cases. The implementation data will be used to evaluate whether using the model would impact the accuracy and efficiency of the consensus diagnostic process. Using the machine learning model to replace one of the clinicians has the potential to further reduce the cost associated with implementing the clinical consensus diagnosis while maintaining expert clinicians as the dominating force in the diagnostic process. That is, at least 2 clinicians will still be included in the diagnostic process and any inconsistency between the human and virtual raters will be resolved through the standard consensus process, which involves the meeting of

a group of expert clinicians to discuss cases. Third, since the design of the LASI-DAD closely follows the HCAP to facilitate international collaboration, the developed model may be used in other HCAP-based studies as an external classification tool for dementia in the absence of clinical ratings for those studies. Fourth, the developed machine learning model can serve to capture the significant clinical knowledge and experience encoded in the clinical consensus diagnostic process. Further examination of the computer model using meta-modeling techniques [43] may generate in-depth understanding of the clinical consensus diagnostic process, such as identifying the top influential assessments for making a diagnosis of dementia. Fifth, since there is no definitive diagnostic test of dementia, tracking the misclassified cases by the machine learning model in the next few years may reveal whether those cases are actual false classifications. Finally, the future waves of LASI-DAD data will be used to identify potential improvements and provide further validation of the current model to ensure its predictive accuracy in the long term.

Another important finding of the study is that the 2-step training process may outperform the standard single-step training process for the classification problem of dementia. Our analysis shows that adding the majority-voting as a second step to the training process reduces approximately one-fourth of misclassifications on the test set. The kappa of the model derived from the 2-step training procedure also outperformed the kappa of the model derived from a standard, repeated cross-validation-based training process that directly optimizes the kappa. An important observation of models derived from the standard training process is that these models tend to overfit the training set, as evidenced by the accuracy (>99%) on the training set for most models. Adding the majority-voting process as a second step reduces overfitting, thereby improving predictive performance on the test set.

Comparison With Prior Literature

The techniques of machine learning have been applied to the examination of survey data for predicting a variety of diseases, such as anxiety [44-46], depression [44,46-51], and dementia [52-54]. This study is based on a nationally representative survey involving a clinical consensus diagnosis of dementia. The number of comparable datasets is limited. The Aging,

Demographics, and Memory Study (ADAMS) includes clinical consensus diagnoses for a subsample of 856 individuals aged 70 years and older in the United States from the HRS [55]. Another similar nationally representative dataset is the Hellenic Longitudinal Investigation of Aging and Diet (HELIAD) study with a sample of 1050 individuals aged 65 years and older in Greece [56]. The LASI-DAD dataset used in this study expands the clinical consensus diagnosis to a larger sample with a broader age range than the ADAMS and HELIAD.

Due to the limited number of available nationally representative datasets with a clinical diagnosis of dementia, only a few machine learning studies have used such type of data. Hurd et al [57] developed an ordered probit model to predict the probability of dementia using the ADAMS dataset. The predictive performance of the model is unclear since validation results based on a randomly selected test set were not reported. Nevertheless, such predicted probability of dementia was used in a subsequent machine learning study by de Langavant et al [53] to test the relevance of an unsupervised learning model based on the larger HRS data (HRS is the parent study of ADAMS). de Langavant et al [54] developed a similar unsupervised learning model to assist the estimation of dementia prevalence in 10 nationally representative surveys. Na [52] developed a supervised machine learning model using data from the Korean Longitudinal Study of Aging to facilitate automatic classification of dementia. However, dementia in this paper was classified by the Mini-Mental State Examination scores below one standard deviation of the mean scores of age by educational level stratified groups. Since screening results can misclassify dementia [58], such classification can only serve as a weaker ground truth of dementia than clinical consensus diagnosis.

The majority of existing machine learning studies on dementia are based on neuroimaging data (see systematic review like Pellegrini et al [59]). In contrast, our study relies on cognitive tests and informant reports, which are easier to obtain for a large sample than neuroimaging data. These tests and questionnaires have been carefully selected in a rigorous process developing the multicountry HCAP [5,60] and translated and adapted to fit the Indian context [61]. The psychometric properties of the tests in the LASI-DAD sample have been assessed favorably [62]. However, the measures are not perfect. Specifically, the literature has found that informant reports of individuals' limitations and cognitive decline, while highly correlated with other measures, may differ from the individuals' reports and reports from health professionals, with the discrepancy systematically depending on the type of proxy (eg, whether caregiver or not) [63-66]. Predictive accuracy as assessed by commonly used measures like overall accuracy, AUROC, and kappa of the model developed in this paper is comparable to or, in some cases, outperforms the neuroimaging-based machine learning models. However, since the survey and neuroimaging data are different and there is no definitive diagnosis of dementia, we caution the use of such direct comparison as a criterion to judge the predictive performance of a model.

Limitations

This study has several limitations. Although the dataset for developing the machine learning model more than doubles the

size of similar nationally representative datasets involving a clinical consensus diagnosis of dementia and is much larger than many clinical datasets, the sample size is still limited from a big data perspective. This is evident from the results that the multilayer perceptron with one hidden layer and other so-called shadow learning models like the support vector machine outperform the deep neural network model with multiple hidden layers in this paper. Typically, deep learning techniques outperform shadow learning techniques when the sample size is very large [67]. Second, as mentioned above, informant reports may to some extent systematically vary with the informant type, and neuroimaging data may give a more accurate assessment of the individual than our combination of cognitive tests and informant reports. However, as described above, all measurements used in this study are well validated and widely used worldwide in population-based aging studies. Third, the number of clinicians participating in the diagnostic process for test set samples is 12, which may not be large enough to be representative. Fourth, although the predictive accuracy of the support vector machine model is high, it is difficult to directly interpret the model for generating in-depth understanding of the dementia diagnostic process. As mentioned above, meta-modeling techniques may be useful for improving the interpretability of the model [43]. Finally, even though clinical consensus diagnosis is considered the gold standard in diagnosing dementia [27], it is not without errors. Tracking the misclassified cases in the next few years may reveal whether misclassifications made by the machine learning model are actually false classifications.

Further research will be needed to assess whether the model can be used for individuals from other countries. Comparable data (cognitive tests, informant reports, and online clinical consensus ratings) will be available in the near future for at least 2 other countries (the United States and South Africa, and possibly China later on), which will allow such assessments. An ongoing global initiative known as the Gateway to Global Aging Data is actively working on creating a harmonized multinational dataset, with the LASI-DAD project a part of the initiative [10]. Even if the model developed in this paper does not readily generalize to data from another country without adaptation, the 2-step training process developed in this paper may still be useful for developing similar machine learning models for dementia. In addition, the current model may be recalibrated for a similar population-based dataset from another country by including the predicted risk score or dementia status from the current model as one of the candidate predictors for training a dedicated model for that country [68]. Further improvements in the model may be possible when data from the second wave of LASI-DAD study become available, including online consensus ratings by clinicians who will be able to evaluate data from 2 observations spaced approximately 4 years apart.

Conclusion

This study develops a machine learning model that learns from clinical consensus diagnoses of dementia from a nationally representative survey of the aging population in India to facilitate the automatic classification of dementia. The developed model has outstanding discriminative ability and substantial

agreement with clinical consensus diagnoses of dementia. The model can serve as a computer model of the clinical knowledge and experience encoded in the clinical consensus diagnostic process and has many current and potential applications, including prediction for missing dementia diagnoses and serving as a clinical decision support tool to assist diagnoses of dementia. The predicted missing dementia diagnoses will be

released as a part of the LASI-DAD data for future use in broader aging research. The LASI-DAD study also plans to implement and test the developed model as a participating virtual rater in the consensus diagnostic process in the next wave of data collection. The future implementation data will be valuable for identifying potential further improvements of the model and ensuring its predictive accuracy in the long term.

Acknowledgments

This project is funded by grants R01 AG051125, RF1 AG055273, and U01 AG064948 from the National Institute on Aging and the National Institutes of Health. We thank AB Dey, Joyita Banerjee, Mary Ganguli, Mathew Varghese, Bas Weerman, Kenneth Langa, David Llewellyn, Prasun Chatterjee, Gaurav R Deasi, Krishna Prasad, Sivakumar Thangaraju, Preeti Sinha, Santosh Loganathan, Abhijit Rao, Rishav Bansal, Sunny Singhal, Swaroop Bhatankar, and Swati Bajpai.

Authors' Contributions

SC, PK, and JL collected the data. HJ and EM conducted the data analysis. HJ wrote the first draft of the manuscript. All authors critically reviewed the manuscript.

Conflicts of Interest

None declared.

References

- 10 facts of dementia. Geneva: World Health Organization URL: <https://www.who.int/features/factfiles/dementia/en/> [accessed 2020-12-10]
- Dementia. Geneva: World Health Organization URL: <https://www.who.int/news-room/fact-sheets/detail/dementia> [accessed 2020-12-10]
- Lee J, Banerjee J, Khobragade PY, Angrisani M, Dey AB. LASI-DAD study: a protocol for a prospective cohort study of late-life cognition and dementia in India. *BMJ Open* 2019 Jul 31;9(7):e030300. [doi: [10.1136/bmjopen-2019-030300](https://doi.org/10.1136/bmjopen-2019-030300)]
- Shaji K, Jotheeswaran A, Girish N, Bharath S, Dias A, Pattabiraman M, et al. The dementia India report: prevalence, impact, costs and services for dementia. Alzheimer's and Related Disorders Society of India. 2010. URL: <https://www.mhinnovation.net/sites/default/files/downloads/innovation/reports/Dementia-India-Report.pdf> [accessed 2021-05-01]
- Weir D, Langa K, Ryan L. Harmonized cognitive assessment protocol (hcap): study protocol summary. Ann Arbor: Institute for Social Research, University of Michigan; 2016. URL: https://hrs.isr.umich.edu/sites/default/files/biblio/HRS%202016%20HCAP%20Protocol%20Summary_011619_rev.pdf [accessed 2021-05-01]
- Weir DR, Wallace RB, Langa KM, Plassman BL, Wilson RS, Bennett DA, et al. Reducing case ascertainment costs in U.S. population studies of Alzheimer's disease, dementia, and cognitive impairment—Part 1. *Alzheimers Dement* 2011 Jan;7(1):94-109 [FREE Full text] [doi: [10.1016/j.jalz.2010.11.004](https://doi.org/10.1016/j.jalz.2010.11.004)] [Medline: [21255747](https://pubmed.ncbi.nlm.nih.gov/21255747/)]
- Evans DA, Grodstein F, Loewenstein D, Kaye J, Weintraub S. Reducing case ascertainment costs in U.S. population studies of Alzheimer's disease, dementia, and cognitive impairment—Part 2. *Alzheimers Dement* 2011 Jan 01;7(1):110-123. [doi: [10.1016/j.jalz.2010.11.008](https://doi.org/10.1016/j.jalz.2010.11.008)]
- Ganguli M, Snitz B, Vander Bilt J, Chang CH. How much do depressive symptoms affect cognition at the population level? The Monongahela-Youghiogheny Healthy Aging Team (MYHAT) study. *Int J Geriatr Psychiatry* 2009 Nov;24(11):1277-1284 [FREE Full text] [doi: [10.1002/gps.2257](https://doi.org/10.1002/gps.2257)] [Medline: [19340894](https://pubmed.ncbi.nlm.nih.gov/19340894/)]
- Lee J, Ganguli M, Weerman A, Chien S, Lee D, Varghese M, et al. Online clinical consensus diagnosis of dementia: development and validation. *J Am Geriatr Soc* 2020 Aug;68 Suppl 3:S54-S59. [doi: [10.1111/jgs.16736](https://doi.org/10.1111/jgs.16736)] [Medline: [32815604](https://pubmed.ncbi.nlm.nih.gov/32815604/)]
- Gateway to Global Aging Data. URL: <https://g2aging.org/> [accessed 2020-11-29]
- Lee J, Phillips D, Wilkens J, Chien S, Lin Y, Angrisani M, et al. Cross-country comparisons of disability and morbidity: evidence from the Gateway to Global Aging Data. *J Gerontol A Biol Sci Med Sci* 2018 Oct 08;73(11):1519-1524 [FREE Full text] [doi: [10.1093/gerona/glx224](https://doi.org/10.1093/gerona/glx224)] [Medline: [29211879](https://pubmed.ncbi.nlm.nih.gov/29211879/)]
- Tiwari SC, Tripathi RK, Kumar A. Applicability of the Mini-mental State Examination (MMSE) and the Hindi Mental State Examination (HMSE) to the urban elderly in India: a pilot study. *Int Psychogeriatr* 2009 Feb;21(1):123-128. [doi: [10.1017/S1041610208007916](https://doi.org/10.1017/S1041610208007916)] [Medline: [18983719](https://pubmed.ncbi.nlm.nih.gov/18983719/)]
- Tsolaki M, Iakovidou V, Navrozidou H, Aminta M, Pantazi T, Kazis A. Hindi Mental State Examination (HMSE) as a screening test for illiterate demented patients. *Int J Geriatr Psychiatry* 2000 Jul;15(7):662-664. [doi: [10.1002/1099-1166\(200007\)15:7<662::aid-gps171>3.0.co;2-5](https://doi.org/10.1002/1099-1166(200007)15:7<662::aid-gps171>3.0.co;2-5)] [Medline: [10918349](https://pubmed.ncbi.nlm.nih.gov/10918349/)]

14. Elliott E, Green C, Llewellyn DJ, Quinn TJ. Accuracy of telephone-based cognitive screening tests: systematic review and meta-analysis. *Curr Alzheimer Res* 2020;17(5):460-471. [doi: [10.2174/1567205017999200626201121](https://doi.org/10.2174/1567205017999200626201121)] [Medline: [32589557](#)]
15. Hall KS, Gao S, Emsley CL, Ogunniyi AO, Morgan O, Hendrie HC. Community screening interview for dementia (CSI 'D'): performance in five disparate study sites. *Int J Geriatr Psychiatry* 2000 Jun;15(6):521-531. [doi: [10.1002/1099-1166\(200006\)15:6<521::aid-gps182>3.0.co;2-f](https://doi.org/10.1002/1099-1166(200006)15:6<521::aid-gps182>3.0.co;2-f)] [Medline: [10861918](#)]
16. Morris JC. The Clinical Dementia Rating (CDR): current version and scoring rules. *Neurology* 1993 Nov;43(11):2412-2414. [doi: [10.1212/wnl.43.11.2412-a](https://doi.org/10.1212/wnl.43.11.2412-a)] [Medline: [8232972](#)]
17. Banks J, Oldfield Z. Understanding pensions: cognitive function, numerical ability and retirement saving. *Fiscal Studies* 2007 Jun;28(2):143-170. [doi: [10.1111/j.1475-5890.2007.00052.x](https://doi.org/10.1111/j.1475-5890.2007.00052.x)]
18. Katz S. Assessing self-maintenance: activities of daily living, mobility, and instrumental activities of daily living. *J Am Geriatr Soc* 1983 Dec;31(12):721-727. [doi: [10.1111/j.1532-5415.1983.tb03391.x](https://doi.org/10.1111/j.1532-5415.1983.tb03391.x)] [Medline: [6418786](#)]
19. Lawton MP, Brody EM. Assessment of older people: self-maintaining and instrumental activities of daily living. *Gerontologist* 1969;9(3):179-186. [Medline: [5349366](#)]
20. Amtmann D, Kim J, Chung H, Bamer AM, Askew RL, Wu S, et al. Comparing CESD-10, PHQ-9, and PROMIS depression instruments in individuals with multiple sclerosis. *Rehabil Psychol* 2014 May;59(2):220-229 [FREE Full text] [doi: [10.1037/a0035919](https://doi.org/10.1037/a0035919)] [Medline: [24661030](#)]
21. Beck AT, Epstein N, Brown G, Steer RA. An inventory for measuring clinical anxiety: psychometric properties. *J Consult Clin Psychol* 1988 Dec;56(6):893-897. [doi: [10.1037//0022-006x.56.6.893](https://doi.org/10.1037//0022-006x.56.6.893)] [Medline: [3204199](#)]
22. Harrison JK, Fearon P, Noel-Storr AH, McShane R, Stott DJ, Quinn TJ. Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) for the diagnosis of dementia within a secondary care setting. *Cochrane Database Syst Rev* 2015 Mar 10(3):CD010772. [doi: [10.1002/14651858.CD010772.pub2](https://doi.org/10.1002/14651858.CD010772.pub2)] [Medline: [25754745](#)]
23. Erkinjuntti T, Hokkanen L, Sulkava R, Palo J. The blessed dementia scale as a screening test for dementia. *Int J Geriatr Psychiatry* 1988 Oct;3(4):267-273. [doi: [10.1002/gps.930030406](https://doi.org/10.1002/gps.930030406)]
24. Hughes CP, Berg L, Danziger WL, Coben LA, Martin RL. A new clinical scale for the staging of dementia. *Br J Psychiatry* 1982 Jun;140:566-572. [doi: [10.1192/bjp.140.6.566](https://doi.org/10.1192/bjp.140.6.566)] [Medline: [7104545](#)]
25. Lowe DA, Balsis S, Miller TM, Bengtson JF, Doody RS. Greater precision when measuring dementia severity: establishing item parameters for the Clinical Dementia Rating Scale. *Dement Geriatr Cogn Disord* 2012;34(2):128-134 [FREE Full text] [doi: [10.1159/000341731](https://doi.org/10.1159/000341731)] [Medline: [23006935](#)]
26. Gross AL, Hassenstab JJ, Johnson SC, Clark LR, Resnick SM, Kitner-Triolo M, et al. A classification algorithm for predicting progression from normal cognition to mild cognitive impairment across five cohorts: the preclinical AD consortium. *Alzheimers Dement (Amst)* 2017;8:147-155 [FREE Full text] [doi: [10.1016/j.dadm.2017.05.003](https://doi.org/10.1016/j.dadm.2017.05.003)] [Medline: [28653035](#)]
27. Gabel MJ, Foster NL, Heidebrink JL, Higdon R, Aizenstein HJ, Arnold SE, et al. Validation of consensus panel diagnosis in dementia. *Arch Neurol* 2010 Dec;67(12):1506-1512 [FREE Full text] [doi: [10.1001/archneurol.2010.301](https://doi.org/10.1001/archneurol.2010.301)] [Medline: [21149812](#)]
28. Friedman JH. Stochastic gradient boosting. *Computat Stat Data Anal* 2002 Feb;38(4):367-378. [doi: [10.1016/S0167-9473\(01\)00065-2](https://doi.org/10.1016/S0167-9473(01)00065-2)]
29. Qi Y. Random forest for bioinformatics. In: Zhang C, editor. *Ensemble Machine Learning: Methods and Applications*. Berlin: Springer; 2012:307-323.
30. James G, Witten D, Hastie T, Tibshirani R. *An Introduction to Statistical Learning*. Berlin: Springer; 2013.
31. 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](https://doi.org/10.1111/j.1467-9868.2005.00503.x)]
32. Friedman JH. Multivariate adaptive regression splines. *Ann Statist* 1991 Mar 01;19(1):1-67. [doi: [10.1214/aos/1176347963](https://doi.org/10.1214/aos/1176347963)]
33. Schmidhuber J. Deep learning in neural networks: an overview. *Neural Netw* 2015 Jan;61:85-117. [doi: [10.1016/j.neunet.2014.09.003](https://doi.org/10.1016/j.neunet.2014.09.003)] [Medline: [25462637](#)]
34. Vanwinckelen G, Blockeel H. On estimating model accuracy with repeated cross-validation. 2012 Presented at: Proceedings of the 21st Belgian-Dutch conference on machine learning; 2012; Ghent p. 39-44.
35. Hosmer JD, Lemeshow S, Sturdivant R. *Applied Logistic Regression*. Hoboken: John Wiley & Sons; 2013.
36. Steffens DC, Otey E, Alexopoulos GS, Butters MA, Cuthbert B, Ganguli M, et al. Perspectives on depression, mild cognitive impairment, and cognitive decline. *Arch Gen Psychiatry* 2006 Feb;63(2):130-138. [doi: [10.1001/archpsyc.63.2.130](https://doi.org/10.1001/archpsyc.63.2.130)] [Medline: [16461855](#)]
37. Beaudreau SA, O'Hara R. Late-life anxiety and cognitive impairment: a review. *Am J Geriatr Psychiatry* 2008 Oct;16(10):790-803. [doi: [10.1097/JGP.0b013e31817945c3](https://doi.org/10.1097/JGP.0b013e31817945c3)] [Medline: [18827225](#)]
38. Steffens DC, Potter GG. Geriatric depression and cognitive impairment. *Psychol Med* 2008 Feb;38(2):163-175. [doi: [10.1017/S003329170700102X](https://doi.org/10.1017/S003329170700102X)] [Medline: [17588275](#)]
39. Amieva H, Ouvrard C, Giulioi C, Meillon C, Rullier L, Dartigues J. Self-reported hearing loss, hearing aids, and cognitive decline in elderly adults: a 25-year study. *J Am Geriatr Soc* 2015 Oct;63(10):2099-2104. [doi: [10.1111/jgs.13649](https://doi.org/10.1111/jgs.13649)] [Medline: [26480972](#)]

40. Lin MY, Gutierrez PR, Stone KL, Yaffe K, Ensrud KE, Fink HA, Study of Osteoporotic Fractures Research Group. Vision impairment and combined vision and hearing impairment predict cognitive and functional decline in older women. *J Am Geriatr Soc* 2004 Dec;52(12):1996-2002. [doi: [10.1111/j.1532-5415.2004.52554.x](https://doi.org/10.1111/j.1532-5415.2004.52554.x)] [Medline: [15571533](#)]
41. Kuhn M. caret: classification and regression training. 2020. URL: <https://CRAN.R-project.org/package=caret> [accessed 2021-05-01]
42. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977 Mar;33(1):159-174. [Medline: [843571](#)]
43. Ermakov S, Melas V. Design and Analysis of Simulation Experiments. Berlin: Springer Science & Business Media; 1995.
44. Richter T, Fishbain B, Markus A, Richter-Levin G, Okon-Singer H. Using machine learning-based analysis for behavioral differentiation between anxiety and depression. *Sci Rep* 2020 Oct 02;10(1):1-12. [doi: [10.1038/s41598-020-72289-9](https://doi.org/10.1038/s41598-020-72289-9)]
45. Wang C, Zhao H, Zhang H. Chinese college students have higher anxiety in new semester of online learning during covid-19: a machine learning approach. *Front Psychol* 2020;11:587413 [FREE Full text] [doi: [10.3389/fpsyg.2020.587413](https://doi.org/10.3389/fpsyg.2020.587413)] [Medline: [33343461](#)]
46. Tennenhouse LG, Marrie RA, Bernstein CN, Lix LM, CIHR Team in Defining the Burden and Managing the Effects of Psychiatric Comorbidity in Chronic Immunoinflammatory Disease. Machine-learning models for depression and anxiety in individuals with immune-mediated inflammatory disease. *J Psychosom Res* 2020 Jul;134:110126 [FREE Full text] [doi: [10.1016/j.jpsychores.2020.110126](https://doi.org/10.1016/j.jpsychores.2020.110126)] [Medline: [32387817](#)]
47. Jin H, Wu S, Di CP. Development of a clinical forecasting model to predict comorbid depression among diabetes patients and an application in depression screening policy making. *Prev Chronic Dis* 2015 Sep 03;12:E142 [FREE Full text] [doi: [10.5888/pcd12.150047](https://doi.org/10.5888/pcd12.150047)] [Medline: [26334714](#)]
48. Jin H, Wu S. Developing depression symptoms prediction models to improve depression care outcomes: preliminary results. 2014 Presented at: Proceedings of the 2nd International Conference on Big Data and Analytics in Healthcare; 2014; Singapore p. 22-24.
49. Jin H, Wu S. Use of patient-reported data to match depression screening intervals with depression risk profiles in primary care patients with diabetes: development and validation of prediction models for major depression. *JMIR Form Res* 2019 Oct 01;3(4):e13610 [FREE Full text] [doi: [10.2196/13610](https://doi.org/10.2196/13610)] [Medline: [31573900](#)]
50. Jin H, Wu S, Vidyanti I, Di Capua P, Wu B. Predicting depression among patients with diabetes using longitudinal data. a multilevel regression model. *Methods Inf Med* 2015;54(6):553-559. [doi: [10.3414/ME14-02-0009](https://doi.org/10.3414/ME14-02-0009)] [Medline: [26577265](#)]
51. Kessler RC, van Loo HM, Wardenaar KJ, Bossarte RM, Brenner LA, Cai T, et al. Testing a machine-learning algorithm to predict the persistence and severity of major depressive disorder from baseline self-reports. *Mol Psychiatry* 2016 Oct;21(10):1366-1371 [FREE Full text] [doi: [10.1038/mp.2015.198](https://doi.org/10.1038/mp.2015.198)] [Medline: [26728563](#)]
52. Na K. Prediction of future cognitive impairment among the community elderly: a machine-learning based approach. *Sci Rep* 2019 Mar 04;9(1):1-9 [FREE Full text] [doi: [10.1038/s41598-019-39478-7](https://doi.org/10.1038/s41598-019-39478-7)] [Medline: [30833698](#)]
53. de Langavant LC, Bayen E, Yaffe K. Unsupervised machine learning to identify high likelihood of dementia in population-based surveys: development and validation study. *J Med Internet Res* 2018 Dec 09;20(7):e10493 [FREE Full text] [doi: [10.2196/10493](https://doi.org/10.2196/10493)] [Medline: [29986849](#)]
54. de Langavant LC, Bayen E, Bachoud-Lévi A, Yaffe K. Approximating dementia prevalence in population-based surveys of aging worldwide: an unsupervised machine learning approach. *Alzheimers Dement (N Y)* 2020;6(1):e12074 [FREE Full text] [doi: [10.1002/trc2.12074](https://doi.org/10.1002/trc2.12074)] [Medline: [32885026](#)]
55. Langa KM, Plassman BL, Wallace RB, Herzog AR, Heeringa SG, Ofstedal MB, et al. The Aging, Demographics, and Memory Study: study design and methods. *Neuroepidemiology* 2005;25(4):181-191. [doi: [10.1159/000087448](https://doi.org/10.1159/000087448)] [Medline: [16103729](#)]
56. Dardiotis E, Kosmidis MH, Yannakoulia M, Hadjigeorgiou GM, Scarmeas N. The Hellenic Longitudinal Investigation of Aging and Diet (HELIAD): rationale, study design, and cohort description. *Neuroepidemiology* 2014;43(1):9-14. [doi: [10.1159/000362723](https://doi.org/10.1159/000362723)] [Medline: [24993387](#)]
57. Hurd MD, Martorell P, Delavande A, Mullen KJ, Langa KM. Monetary costs of dementia in the United States. *N Engl J Med* 2013 Apr 04;368(14):1326-1334. [doi: [10.1056/nejmsa1204629](https://doi.org/10.1056/nejmsa1204629)]
58. Ranson JM, Kuźma E, Hamilton W, Muniz-Terrera G, Langa KM, Llewellyn DJ. Predictors of dementia misclassification when using brief cognitive assessments. *Neurol Clin Pract* 2018 Nov 28;9(2):109-117. [doi: [10.1212/cpj.0000000000000566](https://doi.org/10.1212/cpj.0000000000000566)]
59. Pellegrini E, Ballerini L, Hernandez MDCV, Chappell FM, González-Castro V, Anblagan D, et al. Machine learning of neuroimaging for assisted diagnosis of cognitive impairment and dementia: a systematic review. *Alzheimers Dement (Amst)* 2018;10:519-535 [FREE Full text] [doi: [10.1016/j.dadm.2018.07.004](https://doi.org/10.1016/j.dadm.2018.07.004)] [Medline: [30364671](#)]
60. Weir D, McCammon R, Ryan L, Langa K. Cognitive test selection for the Harmonized Cognitive Assessment Protocol (HCAP). Ann Arbor: Institute for Social Research, University of Michigan; 2014. URL: https://hrs.isr.umich.edu/sites/default/files/biblio/HCAP_testselection.pdf [accessed 2021-05-03]
61. Banerjee J, Jain U, Khobragade P, Weerman B, Hu P, Chien S, et al. Methodological considerations in designing and implementing the harmonized diagnostic assessment of dementia for longitudinal aging study in India (LASI-DAD). *Biodemography Soc Biol* 2020;65(3):189-213. [doi: [10.1080/19485565.2020.1730156](https://doi.org/10.1080/19485565.2020.1730156)] [Medline: [32727279](#)]

62. Gross A, Khobragade P, Meijer E, Saxton J. Measurement and structure of cognition in the Longitudinal Aging Study in India: diagnostic assessment of dementia. *J Am Geriatrics Soc* 2020;68:S11-S19. [doi: [10.1093/geroni/igaa057.2280](https://doi.org/10.1093/geroni/igaa057.2280)]
63. Neumann PJ, Araki SS, Gutterman EM. The use of proxy respondents in studies of older adults: lessons, challenges, and opportunities. *J Am Geriatr Soc* 2000 Dec;48(12):1646-1654 [FREE Full text] [doi: [10.1111/j.1532-5415.2000.tb03877.x](https://doi.org/10.1111/j.1532-5415.2000.tb03877.x)] [Medline: [11129756](https://pubmed.ncbi.nlm.nih.gov/11129756/)]
64. Lum TY, Lin W, Kane RL. Use of proxy respondents and accuracy of minimum data set assessments of activities of daily living. *J Gerontol A Biol Sci Med Sci* 2005 May;60(5):654-659. [doi: [10.1093/gerona/60.5.654](https://doi.org/10.1093/gerona/60.5.654)] [Medline: [15972620](https://pubmed.ncbi.nlm.nih.gov/15972620/)]
65. Li M, Harris I, Lu ZK. Differences in proxy-reported and patient-reported outcomes: assessing health and functional status among medicare beneficiaries. *BMC Med Res Methodol* 2015 Aug 12;15:1-10 [FREE Full text] [doi: [10.1186/s12874-015-0053-7](https://doi.org/10.1186/s12874-015-0053-7)] [Medline: [26264727](https://pubmed.ncbi.nlm.nih.gov/26264727/)]
66. Howland M, Allan K, Carlton C, Tatsuoaka C, Smyth K, Sajatovic M. Patient-rated versus proxy-rated cognitive and functional measures in older adults. *PROM* 2017 Mar;8:33-42. [doi: [10.2147/prom.s126919](https://doi.org/10.2147/prom.s126919)]
67. Ciaburro G, Venkateswaran B. *Neural Networks With R: Smart Models Using CNN, RNN, Deep Learning, and Artificial Intelligence Principles*. Birmingham: Packt Publishing Ltd; 2017.
68. King M, Bottomley C, Bellón-Saameño J, Torres-Gonzalez F, Švab I, Rotar D, et al. Predicting onset of major depression in general practice attendees in Europe: extending the application of the predictD risk algorithm from 12 to 24 months. *Psychol Med* 2013 Jan 04;43(9):1929-1939. [doi: [10.1017/s0033291712002693](https://doi.org/10.1017/s0033291712002693)]

Abbreviations

ADAMS: Aging, Demographics, and Memory Study

ADL: activity of daily living

AUROC: area under receiver operating curve

CDR: Clinical Dementia Rating

CSID: Community Screening Instrument for Dementia

HCAP: Harmonized Cognitive Assessment Protocol

HELIAD: Hellenic Longitudinal Investigation of Aging and Diet

HRS: Health and Retirement Study

LASI: Longitudinal Aging Study in India

LASI-DAD: Harmonized Diagnostic Assessment of Dementia for the Longitudinal Aging Study in India

TICS: Telephone Interview for Cognitive Status

Edited by G Eysenbach; submitted 11.01.21; peer-reviewed by KL Ong, V Franzoni; comments to author 25.01.21; revised version received 11.03.21; accepted 17.04.21; published 10.05.21.

Please cite as:

Jin H, Chien S, Meijer E, Khobragade P, Lee J

Learning From Clinical Consensus Diagnosis in India to Facilitate Automatic Classification of Dementia: Machine Learning Study
JMIR Ment Health 2021;8(5):e27113

URL: <https://mental.jmir.org/2021/5/e27113>

doi: [10.2196/27113](https://doi.org/10.2196/27113)

PMID: [33970122](https://pubmed.ncbi.nlm.nih.gov/33970122/)

©Haomiao Jin, Sandy Chien, Erik Meijer, Pranali Khobragade, Jinkook Lee. Originally published in *JMIR Mental Health* (<https://mental.jmir.org>), 10.05.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Mental Health*, is properly cited. The complete bibliographic information, a link to the original publication on <https://mental.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

A Web-Based Group Cognitive Behavioral Therapy Intervention for Symptoms of Anxiety and Depression Among University Students: Open-Label, Pragmatic Trial

Jason Bantjes¹, PhD; Alan E Kazdin², PhD; Pim Cuijpers³, PhD; Elsie Breet¹, PhD; Munita Dunn-Coetzee⁴, PhD; Charl Davids⁴, MA; Dan J Stein⁵, MD, PhD; Ronald C Kessler⁶, PhD

¹Institute for Life Course Health Research, Department of Global Health, Faculty of Medicine and Health Sciences, Stellenbosch University, Stellenbosch, South Africa

²Department of Psychology, Yale University, New Haven, CT, United States

³Department of Clinical, Neuro and Developmental Psychology, Faculty of Behavioural and Movement Sciences, Vrije Universiteit, Amsterdam, Netherlands

⁴Centre for Student Counselling and Development, Student Affairs, Stellenbosch University, Stellenbosch, South Africa

⁵Department of Psychiatry and Mental Health, University of Cape Town, South African Medical Research Council Unit on Risk & Resilience in Mental Disorders, Cape Town, South Africa

⁶Department of Health Care Policy, Harvard Medical School, Boston, MA, United States

Corresponding Author:

Jason Bantjes, PhD

Institute for Life Course Health Research

Department of Global Health, Faculty of Medicine and Health Sciences

Stellenbosch University

Private Bag X1

Matieland

Stellenbosch, 7602

South Africa

Phone: 27 832345554

Fax: 27 832345554

Email: jbantjes@sun.ac.za

Abstract

Background: Anxiety and depression are common among university students, and university counseling centers are under pressure to develop effective, novel, and sustainable interventions that engage and retain students. Group interventions delivered via the internet could be a novel and effective way to promote student mental health.

Objective: We conducted a pragmatic open trial to investigate the uptake, retention, treatment response, and level of satisfaction with a remote group cognitive behavioral therapy intervention designed to reduce symptoms of anxiety and depression delivered on the web to university students during the COVID-19 pandemic.

Methods: Preintervention and postintervention self-reported data on anxiety and depression were collected using the Generalized Anxiety Disorder-7 and Patient Health Questionnaire-9. Satisfaction was assessed postintervention using the Client Satisfaction with Treatment Questionnaire.

Results: A total of 175 students were enrolled, 158 (90.3%) of whom initiated treatment. Among those initiating treatment, 86.1% (135/158) identified as female, and the mean age was 22.4 (SD 4.9) years. The mean number of sessions attended was 6.4 (SD 2.8) out of 10. Among participants with clinically significant symptoms at baseline, mean symptom scores decreased significantly for anxiety ($t_{56}=11.6$; $P<.001$), depression ($t_{61}=7.8$; $P<.001$), and composite anxiety and depression ($t_{60}=10.7$; $P<.001$), with large effect sizes ($d=1-1.5$). Remission rates among participants with clinically significant baseline symptoms were 67.7%-78.9% and were not associated with baseline symptom severity. High overall levels of satisfaction with treatment were reported.

Conclusions: The results of this study serve as a proof of concept for the use of web-based group cognitive behavioral therapy to promote the mental health of university students.

KEYWORDS

anxiety; cognitive behavioral therapy; depression; e-intervention; group therapy; web-based; university students; South Africa

Introduction

Background

Depression is the most common mental disorder and the leading cause of mental health-related disease burden worldwide, affecting more than 300 million people globally and representing a major barrier to sustainable development in all regions of the world [1,2]. Depression is strongly associated with anxiety disorders. If left untreated, these disorders compromise psychosocial function and physical health and are associated with reductions in life expectancy of 11-20 years [3,4]. Both anxiety and depression are common among university students, with a 12-month prevalence for generalized anxiety disorder (GAD) and major depressive episode (MDE) of 16.7% and 18.5%, respectively [5]. In student populations, mood disorders are associated with severe role impairment [6], academic failure [7], and suicide [8]. Despite the availability of viable treatments, the majority of anxious and depressed students, like others in the general population, do not seek care [9-11]. One large cross-national study found that only 25.3%-36.3% of students with mental health problems received treatment [12]. There are a number of reasons that students do not access treatment, including a reluctance to receive help from mental health professionals [13], inability to recognize psychopathology [14], practical issues (eg, time constraints and scheduling problems), and psychological factors (eg, stigma and perceptions of therapy's effectiveness) [15,16]. High rates of attrition among students who access campus-based treatment for anxiety and depression also contribute to the treatment gap [17].

Two key challenges faced by university student counseling centers are how to respond to the large number of students with common mental disorders and how to engage and retain students in effective treatments once they reach out for professional help [13]. Many student counseling centers are underresourced and have difficulty in reaching students in need even when they are operating at full capacity [18], with the situation being more dire in low- and middle-income countries where student counseling services are often nonexistent [19]. Self-guided and guided digital interventions are potentially scalable and cost-effective ways to close the mental health treatment gap and have been shown to be effective in treating university students with anxiety and depression [20-22], although many of these interventions have high rates of attrition and low rates of engagement [20]. Group interventions may be a more viable alternative [19], given that group therapy has been shown to be effective and to have better retention rates than digital interventions [23]. Another appeal of group interventions is that they can be offered remotely using web-based video conferencing platforms, thereby providing greater availability than traditional psychotherapy. Remote group interventions also have the potential to provide greater anonymity than conventional psychotherapy, addressing a major barrier to treatment typically faced by students [13,14]. Another appeal

of web-based groups is that they can enable remote treatment when it is not possible for therapists and students to meet face to face, as was the case during the global COVID-19 pandemic [24].

Group Cognitive Behavioral Therapy

Cognitive behavioral therapy (CBT) is a widely used evidence-based treatment for anxiety and depression [25] and can be effectively delivered via individual and group therapies [26,27], self-help interventions [28], and digital programs [29,30]. The digital revolution has precipitated the development of CBT mobile apps [31], machine learning-based chatbots that provide real-time CBT coaching [32], and guided web-based CBT group courses [33]. A meta-analysis of 35 group cognitive behavioral therapy (GCBT) clinical trials in which a protocol-based intervention that included a minimum of cognitive restructuring and behavioral activation was delivered face to face to adults in 5-16 sessions concluded that GCBT is an effective alternative to individual therapy and compares favorably with other middle-intensity interventions [27]. Evidence also exists that GCBT is effective among students both as a treatment and as an indicated prevention for anxiety and depression [34-36] and that such group therapies can be delivered effectively on the web [33]. Indeed, web-based GCBT might be particularly appropriate for university students given the openness of students to digital interventions and the potential of this modality to overcome barriers to care [37,38], although evidence is needed to show that GCBT is effective and satisfies students' needs.

For web-based GCBT to be meaningfully integrated into student counseling services, it is necessary to not only establish that this modality is effective but also understand which students respond well to it. Adopting a precision medicine (personalized medicine) approach to predict treatment effects is integral to designing evidence-based, efficient, and effective student mental health services [39]. In addition, web-based GCBT could be integrated into stepped care treatments if we understand how symptom severity predicts treatment response [40]. Identifying predictors of treatment response to GCBT could also aid the development of machine learning prediction models that use patient-level data to individualize student mental health care [41]. Predictors of treatment response for traditional interventions for mood disorders include sociodemographic variables (eg, age and sex) and clinical factors (eg, symptom profiles, comorbidities, and symptom severity) [41-43].

Before attempting to conduct controlled trials to document aggregate treatment effects for web-based GCBT or studies designed to develop precision treatment algorithms, it is important to establish whether this modality would engage students and satisfy their needs. Ensuring satisfaction with treatment is an integral component of engaging and retaining students in mental health interventions, and it enables a person-centered approach to student wellness services [44].

Evidence suggests that students are satisfied with web-based CBT interventions irrespective of whether they are delivered by therapist or are self-administered [45], but it remains unclear whether web-based GCBT could satisfy students' needs.

Objectives

The aim of this study is to present the results of a pragmatic open trial to address these uncertainties. We investigate the uptake, attendance, treatment responses, predictors of treatment response, and satisfaction with a 10-session web-based GCBT intervention for anxiety and depression. This pragmatic trial was implemented in South Africa during the global COVID-19 pandemic, when access to traditional campus-based psychotherapy was restricted. A pragmatic design was used for this trial because we wanted to establish whether web-based GCBT could be implemented under real-world conditions to promote mental health in a broad population of students [46]. Unlike explanatory trials that seek to test whether interventions work under optimal conditions with a clearly defined population, pragmatic trials evaluate intervention effects in routine practice conditions [47]. This research is a part of the ongoing work to identify effective and sustainable interventions to promote student mental health by the World Health Organization World Mental Health Surveys International College Student Initiative [48].

Methods

Recruitment

Information about the intervention was posted only once on a student affairs Facebook page at Stellenbosch University in South Africa (N=32,600 students, approximately) near the start of the COVID-19 pandemic when South Africa was in lockdown and access to campus-based counseling was restricted. The posting explained that web-based groups were being offered to help students learn psychological skills to reduce symptoms of anxiety and depression. Although it is not possible to ascertain how many students would have seen the post or how many students who did see it would have shared the information with friends, we know that this Facebook page was followed by 3344 people at the time of recruitment. The 175 students who responded within 24 hours to the notice completed a baseline assessment and provided informed consent before being randomized across 15 groups (to keep group sizes to between 10 and 15 participants). No student who wanted to participate was excluded, and no incentives were given to participate.

Group Content

The intervention was delivered via Microsoft Teams (a secure web-based video conferencing platform) in 10 weekly workshops of 60-75 minutes. The content was drawn from common elements identified from GCBT interventions that were shown to be effective for university students [35,49-52]. The content (Textbox 1) was organized into 5 themes, with each theme spanning 2 workshops. To make the intervention more engaging and relevant to the target population, we consulted a group of 4 student advisors regularly over the development period to select suitable examples, plan activities, and inform the layout and design of materials. Consultation with the student advisors took place in an iterative process over the course of a year while we developed and refined the intervention materials. We also consulted 3 psychologists working in student counseling centers, and the intervention materials were critically reviewed by 2 CBT experts.

Participants were provided with electronic interactive PDF workbooks (Adobe Acrobat) consisting of exercises and brief summaries of the main ideas and skills for each session before each workshop. Participants were given the option to remain anonymous by keeping their web cameras off and/or by using pseudonyms, although they were encouraged to use web cameras to show their faces during the workshops. Participants were also invited to use the web-based chat function to type comments, questions, or responses during the sessions if they felt uncomfortable speaking in the group. The intervention was offered in partnership with the student counseling center and was positioned as a university-endorsed program integrated into the routine clinical care offered to students. Strategies used to improve retention included giving participants permission to miss sessions but encouraging attendance at each new session, sending follow-up emails to students who missed sessions prompting them to join the following week, and giving a brief recap of the previous workshop at the start of each new session.

Sessions were facilitated by registered counselors (the equivalent of psychological technicians with 4 years of university training) and clinical psychology master's students under the supervision of a registered PhD psychologist. A facilitators' handbook, with detailed descriptions of the content of each session and facilitation guidelines, was provided, and facilitators were trained on group therapy facilitation. Web-based sessions were recorded so that they could be reviewed under supervision.

Textbox 1. Overview of the intervention content.**Theme 1: You feel the way you think**

- Workshop 1: Emotional triggers and automatic thoughts
 - How to identify activating events (emotional triggers) and recognize how automatic thoughts contribute to the way you feel
- Workshop 2: Challenging automatic thoughts and core beliefs
 - How to identify and challenge unhelpful core beliefs and automatic thoughts

Theme 2: Planning to succeed

- Workshop 1: Getting on top of problems before they get you down
 - How to recognize stressors and use strategies to solve interpersonal and emotional problems
- Workshop 2: Goal setting and planning
 - How to set goals and plan for behavior change

Theme 3: Hacks to boost your mood

- Workshop 1: Avoiding thinking traps
 - How to identify and modify dysfunctional patterns of thinking
- Workshop 2: Overcoming rumination and guilt
 - How to use strategies to overcome rumination and guilt

Theme 4: Building mastery

- Workshop 1: Behavior activation
 - How to identify and increase activities that promote feelings of well-being and reduce stress
- Workshop 2: Behaviors that matter
 - How to identify unhealthy habits and develop health-promoting behaviors

Theme 5: Avoiding meltdowns

- Workshop 1: Understanding the body's stress response
 - Understanding the body's stress response and how to use strategies to regulate physiological arousal
- Workshop 2: Managing stress and overcoming avoidance
 - How to manage stress and overcome avoidance

Measures

Self-reported preintervention and postintervention surveys were administered on the web using Qualtrics software to record the following sociodemographic and clinical characteristics of participants.

Demographics

In the preintervention survey, participants were asked about their self-identified population group, gender, age, and current year of registration.

Symptoms of Anxiety and Depression

The Generalized Anxiety Disorder-7 (GAD-7) scale [53] and Patient Health Questionnaire-9 (PHQ-9) [54] were administered preintervention and 1-week postintervention. The GAD-7

consists of 7 items assessing the frequency of core symptoms of anxiety disorders over the past 2 weeks, scored using a response scale from 0 (not at all) to 3 (nearly every day) and yielding a total score ranging from 0 to 21. A cutoff of 10 is used to identify clinically significant symptoms (ie, individuals likely to meet the diagnostic criteria for GAD) [53]. Psychometric studies have documented good GAD-7 internal consistency reliability, validity, and sensitivity to changes in the severity of anxiety symptoms in a range of settings, including in Africa [55,56]. The PHQ-9 is a self-reported instrument consisting of 9 items assessing the frequency of core symptoms of an MDE over the past 2 weeks on a 0-4 response scale, yielding a total score ranging from 0 to 27. A cutoff of 10 is used to identify clinically significant symptoms (ie, individuals likely to meet the diagnostic criteria for MDE) [53,57]. The PHQ-9 has good internal consistency reliability, validity, and

sensitivity to changes in the severity of depressive symptoms in a range of settings [58,59], including in South Africa [60]. Cutoffs of 5, 10, and 20 on the GAD-7 and PHQ-9 were used to identify individuals with none, mild, moderate, or severe symptoms.

Client Satisfaction With Treatment

The postintervention survey included the Client Satisfaction with Treatment Questionnaire-8, an 8-item self-reported questionnaire with high internal consistency reliability [61] that provides an efficient, sensitive, and reasonably comprehensive measure of patient satisfaction with services [61,62].

The primary outcomes were symptoms of anxiety (ie, GAD-7 scores), symptoms of depression (ie, PHQ-9 scores), and symptoms of comorbid anxiety and depression (using combined GAD-7 and PHQ-9 scores as proposed by the developers of the instruments to yield a composite anxiety and depression score) [63].

Data Analysis

The data were cleaned and analyzed using SPSS, version 27 (IBM Corporation). Descriptive statistics were used to summarize participant characteristics, attendance rates, primary outcomes, and satisfaction with treatment. A repeated-measures analysis of variance (ANOVA) was used to measure changes in symptoms scores from baseline to follow-up, with effect size determined by η^2 . The mean symptom scores at baseline and follow-up were also compared using a paired-sample t test, with effect sizes determined by Cohen d calculated using Comprehensive Meta-Analysis (CMA) software and the procedures described by Borenstein et al [64]. The McNemar test was used to evaluate the significance of changes in proportions, and multiple regression analysis was used to identify predictors of treatment response and satisfaction with treatment. The analysis was conducted on the subsample of all participants who started the intervention, irrespective of whether they completed the intervention (ie, an intention-to-treat analysis). Participants who did not complete the follow-up assessments were excluded from the analysis via listwise deletion. The results of all regression analyses are reported as adjusted odds ratios (aORs) for logistic regression models and standardized partial regression coefficients for linear regression models, with 95% CI. Statistical significance was evaluated using $P=.05$ -level 2-sided tests.

Ethical Considerations

Ethical approval was obtained from the Health Science Research Ethics Committee (N19/10/145), and institutional permission

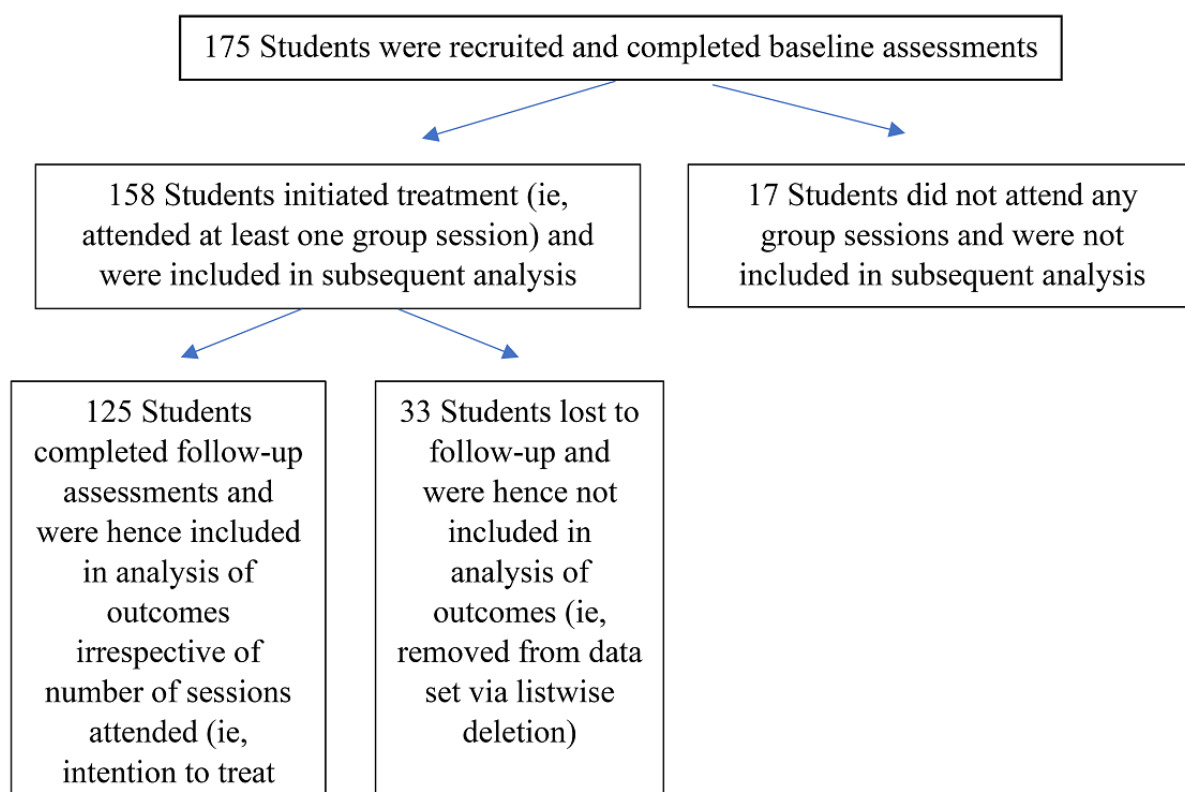
was secured before recruitment. Informed consent was obtained electronically. The participants were given an opportunity to participate in the intervention without being enrolled in the study, although none of them opted for this arrangement. The deidentified data were stored on a password-protected computer. All participants were provided with contact details of the 24-hour campus crisis service, and students who reported suicidal ideation were contacted via email to prompt them to seek individual care.

Results

Sample Characteristics

Of the 175 students who were enrolled and completed baseline assessments, 158 (90.3%) initiated treatment and 125 (71.4%) completed follow-up assessments (Figure 1). No significant baseline differences were found between participants who initiated treatment and those who did not in terms of gender ($P=.30$), age ($P=.76$), undergraduate status ($P=.24$), or baseline severity of anxiety ($P=.72$) or depression ($P=.06$). Participants who were lost to follow-up were not significantly different from participants who were followed up with respect to gender ($P=.91$), age ($P=.29$), undergraduate status ($P=.16$), or baseline severity of anxiety ($P=.24$) or depression ($P=.27$; see Tables S1-S4 in Multimedia Appendix 1). However, attendance was lower among participants who were lost to follow-up (mean number of sessions 4.8 vs 6.8; $P=.02$).

The mean age of the sample of students for whom treatment was initiated ($n=158$) was 22.2 (SD 4.9) years, with a range of 18-54 years (median 21 years, mode 19 years). Among them, 75.9% (120/158) were undergraduates and 24.1% (38/158) were postgraduates. Overall, 85.4% (135/158) self-identified as female, 13.9% (22/158) as male, and 1 (0.6%) student identified as gender nonbinary. Of these, 20.9% (33/158) self-identified as Black African, 19.6% (31/158) as colored (an official term used in South Africa for population classification), 7% (11/158) as Asian, and 52.5% (83/158) as White. The mean baseline GAD-7 and PHQ-9 scores were 9.5 (SD 5.5) and 10.2 (SD 5.4), with 21.5% (34/158) of students scoring less than 5 on the GAD-7 and 14.6% (23/158) scoring less than 5 on the PHQ-9 preintervention survey. The mean number of sessions attended was 6.4 (SD 2.8), with 72.8% (115/158) of participants attending half or more of the sessions and 46.2% (73/158) attending 80% or more (Table 1). Only 10.1% (16/158) of participants attended all 10 sessions. The reasons for nonattendance included competing academic commitments, internet connectivity problems, power failures, and insufficient internet data.

Figure 1. Flowchart of recruitment and assessment process.**Table 1.** Frequency and rate of attendance by number of sessions (n=158).

Number of sessions attended	Participants, n (%)	Participants (cumulative), n (%)
10	16 (10.1)	16 (10.1)
9	31 (19.6)	47 (29.7)
8	26 (16.5)	73 (46.2)
7	21 (13.3)	94 (59.5)
6	13 (8.2)	107 (67.7)
5	8 (5.1)	115 (72.8)
4	7 (4.4)	122 (77.2)
3	14 (8.9)	136 (86.1)
2	8 (5.1)	144 (91.1)
1	14 (8.9)	158 (100)

Improvements in Aggregate Symptoms Scores

To assess intervention effect sizes, we conducted a repeated measures ANOVA comparing symptom scores from baseline to follow-up and calculated the changes in mean symptom scores from baseline to follow-up (with associated effect sizes) for anxiety, depression, and composite anxiety and depression (Table 2). To establish if the effect sizes varied according to symptom severity, we investigated changes in mean scores among the subset of participants with clinically significant symptoms and among subsets with mild, moderate, and severe

symptoms. The repeated measures ANOVA showed significant symptom reductions for anxiety ($F_{1,124}=66.2$; $P<.001$) and depression ($F_{1,123}=45.4$; $P<.001$), with large effect sizes for anxiety ($\eta^2=0.4$) and depression ($\eta^2=0.27$). Among participants with clinically significant symptoms at baseline, the mean symptom scores decreased significantly for anxiety ($t_{56}=11.6$; $P<.001$), depression ($t_{61}=7.8$; $P<.001$), and composite anxiety and depression ($t_{60}=10.7$; $P<.001$), with large effect sizes ($d=1-1.5$).

Table 2. Comparison of mean symptom scores at baseline and follow-up (paired-sample *t* test with effect size; *n*=125).

Symptom severity	Baseline, mean (SD)	Follow-up, mean (SD)	<i>t</i> test (<i>df</i>)	<i>P</i> value	Correlation	Cohen <i>d</i>	Effect size
						<i>d</i>	SE
Symptoms of anxiety (GAD^a-7 scores)							
Clinically significant symptoms ^b	14.1 (3.5)	6.9 (4.2)	11.6 (56)	<.001 ^c	0.3	1.9	0.3 Large
Mild symptoms ^d	7.3 (1.5)	5.6 (4.3)	2.3 (36)	.03 ^c	0.1	0.5	0.2 Medium
Moderate symptoms ^e	13.5 (3.1)	6.5 (3.7)	11.2 (51)	<.001	0.2	2.0	0.3 Large
Severe symptoms ^f	20.2 (0.4)	11.0 (6.9)	3.0 (4)	.04 ^c	0.5	1.4	0.6 Large
Symptoms of depression (PHQ-9^g scores)							
Clinically significant symptoms	14.0 (3.6)	8.5 (5.0)	7.8 (61)	<.001	0.2	1.3	0.2 Large
Mild symptoms	7.1 (1.3)	4.8 (3.1)	4.8 (42)	<.001	0.2	0.9	0.2 Large
Moderate symptoms	13.2 (2.4)	8.1 (4.6)	7.6 (56)	<.001	0.1	1.4	0.3 Large
Severe symptoms	23.0 (2.3)	12.4 (7.8)	2.6 (4)	.06	−0.5	2.0	1.4 Large
Symptoms of anxiety and depression (GAD-7+PHQ-9 scores)							
Clinically significant symptoms ^h	27.2 (5.7)	15.4 (7.7)	10.7 (60)	<.001	0.2	1.7	0.3 Large
Mild symptoms ⁱ	14.5 (2.7)	10.0 (9.2)	3.1 (38)	<.001	0.2	0.6	0.2 Medium
Moderate symptoms ^j	26.7 (5.1)	15.0 (7.0)	10.8 (58)	<.001	0.1	1.9	0.3 Large
Severe symptoms ^k	41.5 (0.7)	28.0 (19.8)	0.9 (1)	.53	−1.0	0.1	0.1 Small

^aGAD: generalized anxiety disorder.^bGeneralized Anxiety Disorder-7/Patient Health Questionnaire-9 score ≥10.^c*P*<.05.^dPatient Health Questionnaire-9/Generalized Anxiety Disorder-7 score 5-9.^ePatient Health Questionnaire-9/Generalized Anxiety Disorder-7 score 10-19.^fPatient Health Questionnaire-9/Generalized Anxiety Disorder-7 score ≥20.^gPHQ-9: Patient Health Questionnaire-9.^hPatient Health Questionnaire-9+Generalized Anxiety Disorder-7 score ≥20.ⁱPatient Health Questionnaire-9+Generalized Anxiety Disorder-7 10-19.^jPatient Health Questionnaire-9+Generalized Anxiety Disorder-7 20-39.^kPatient Health Questionnaire-9+Generalized Anxiety Disorder-7 score ≥40.

The mean symptom scores decreased comparably for participants with moderate and severe symptoms at baseline. The mean symptom scores also decreased among participants with mild symptoms of anxiety ($t_{36}=2.3$; $P=.03$), depression ($t_{42}=4.8$; $P<.001$), and the anxiety and depression composite ($t_{38}=3.1$; $P<.001$) at baseline, with a large effect size for depression ($d=0.9$) and medium effect sizes for anxiety ($d=0.5$) and composite anxiety depression ($d=0.6$). The mean symptom scores decreased significantly, irrespective of symptom severity at baseline, although the effects were largest for moderate and severe symptoms (Table 2).

Treatment Response and Remission Rates

The proportion of participants with clinically significant symptoms decreased significantly from 45.6% (SE 4.5) to 16.9%

(SE 3.4) for anxiety ($\chi^2_1=22.7$; $P<.001$), from 49.6% (SE 4.5) to 8.4% (SE 3.5) for depression ($\chi^2_1=22.7$ -32; $P<.001$), and from 48.8% (SE 4.5) to 15.2% (SE 3.2) for the anxiety and depression composite ($\chi^2_1=33.6$; $P<.001$). These changes represent 54.1%-68.9% proportional reductions in prevalence. The response and remission rates were calculated across all levels of symptom severity to demonstrate how many participants reported changes in their symptom profiles (Table 3), showing significant changes for anxiety ($\chi^2_3=26.3$; $P<.001$), depression ($\chi^2_3=30.6$; $P<.001$), and composite anxiety and depression ($\chi^2_3=35.3$; $P<.001$).

Table 3. Prevalence of symptoms at baseline and follow-up (n=125).

Symptoms	No symptoms ^a , % (SE)	Mild symptoms ^b , % (SE)	Moderate symptoms ^c , % (SE)	Severe symptoms ^d , % (SE)	χ^2 (df)	P value
Symptoms of anxiety ^e at baseline	24.8 (3.9)	29.6 (4.1)	25.6 (3.9)	20.0 (3.6)	N/A ^f	N/A
Symptoms of anxiety ^e at follow-up	46.8 (4.5)	36.3 (4.3)	10.5 (2.8)	6.5 (2.2)	26.3 (3)	<.001 ^g
Symptoms of depression ^h at baseline	16.0 (3.3)	34.4 (4.3)	45.6 (4.5)	4.0 (1.8)	N/A	N/A
Symptoms of depression ^h at follow-up	37.6 (4.3)	44.0 (4.5)	16.0 (3.3)	2.4 (1.4)	30.6 (3)	<.001
Composite anxiety and depression ⁱ at baseline	20.0 (3.6)	31.2 (4.2)	29.6 (4.1)	19.2 (3.5)	N/A	N/A
Composite anxiety and depression ⁱ at follow-up	43.2 (4.4)	41.6 (4.4)	11.2 (2.8)	4.0 (1.8)	35.3 (3)	<.001

^aPatient Health Questionnaire-9/Generalized Anxiety Disorder-7 score <5.

^bPatient Health Questionnaire-9/Generalized Anxiety Disorder-7 score 5-9.

^cPatient Health Questionnaire-9/Generalized Anxiety Disorder-7 score 10-19.

^dPatient Health Questionnaire-9/Generalized Anxiety Disorder-7 score ≤20.

^eGeneralized Anxiety Disorder-7 score.

^fN/A: not applicable.

^g $P < .05$.

^hPatient Health Questionnaire-9 score.

ⁱPatient Health Questionnaire-9+Generalized Anxiety Disorder-7 score.

The proportion of participants with severe symptoms (ie, scores of 20 or more on the GAD-7 or PHQ-9), reduced significantly among those with symptoms of anxiety (20.0% vs 6.5%; $P < .001$), depression (4.0% vs 2.4%; $P < .001$), and composite anxiety and depression (19.2% vs 4.0%; $P < .001$). The proportions of participants with moderate symptoms (ie, scores of 10-19 on the GAD-7 or PHQ-9) decreased significantly for anxiety (25.6% vs 10.5%; $P < .001$), depression (45.6% vs 6.0%; $P < .001$) and composite anxiety and depression (29.6% vs 11.2%; $P < .001$). As expected, the proportion of patients with mild symptoms increased from between 28.8% and 34.4% at baseline to 36.3% and 44.0% across all outcomes at follow-up, and the proportion of asymptomatic patients increased from 12%-24.8% at baseline to 32.0%-46.8% across all outcomes at follow-up.

To provide a more nuanced perspective on response to treatment, we calculated rates of remission (defined as a reduction in

GAD-7/PHQ-9 scores below 10), treatment response (defined as a 50% or greater reduction in baseline symptom scores), and clinical deterioration (defined as a 50% or greater increase in baseline symptom scores; Table 4). Among participants reporting clinically significant baseline symptoms, remission was in the range 67.7%-78.9% across outcomes, and clinically significant deterioration was uncommon, ranging from 1.8% to 3.2% across outcomes. Changes were also observed among participants who at baseline had mild symptoms, among whom treatment response was in the range of 34.9%-43.6% across outcomes and clinically significant deterioration was in the range of 4.7%-10.8% across outcomes. The data indicate the potential for the intervention to reduce the absolute number of students with clinically significant symptoms and demonstrate the low numbers expected to deteriorate over the course of the intervention.

Table 4. Individual-level response (n=125).

Symptoms	Among participants with clinically significant symptoms at baseline ^a		Among participants with mild symptoms at baseline	
	Remission ^b	Deterioration ^c	Treatment response ^d	Deterioration
Symptoms of anxiety^e				
% (SE)	67.7 (6.0)	3.2 (2.3)	34.9 (7.4)	4.7 (3.2)
n (%)	62 (49.6)	62 (49.6)	43 (34.4)	43 (34.4)
Symptoms of depression^f				
% (SE)	78.9 (5.4)	1.8 (1.8)	35.1 (8.0)	10.8 (5.2)
n (%)	57 (45.6)	57 (45.6)	37 (29.6)	37 (29.6)
Composite anxiety and depression^g				
% (SE)	75.4 (5.6)	3.3 (1.6)	43.6 (8.0)	7.7 (4.3)
n (%)	61 (48.8)	61 (48.8)	39 (31.2)	39 (31.2)

^aGeneralized Anxiety Disorder-7/Patient Health Questionnaire-9 score ≥ 10 .

^bRemission: reduction of symptoms so that Patient Health Questionnaire-9/Generalized Anxiety Disorder-7 score < 10 .

^cDeterioration: a 50% increase in symptoms and a change from no to mild symptoms, from mild to moderate symptoms, or from moderate to severe symptoms.

^dTreatment response: a reduction in symptoms of 50% or more.

^eGeneralized Anxiety Disorder-7 score ≥ 10 .

^fPatient Health Questionnaire-9 score ≥ 10 .

^gPatient Health Questionnaire-9+Generalized Anxiety Disorder-7 ≥ 20 .

Predictors of Treatment Response and Remission

We investigated whether information collected from participants at baseline could be used to identify individuals who were likely to respond to the intervention. Multiple logistic regression analysis was used to identify baseline predictors of remission among participants with clinically significant symptoms at baseline and among participants with only mild baseline symptoms (Table 5). Predictors of treatment response among participants with clinically significant symptoms were not examined because of the rarity of response without remission, and predictors of symptom deterioration were not examined because of the small number of participants who deteriorated.

Among participants with clinically significant baseline symptoms, baseline variables were significant in predicting remission from composite anxiety and depression ($\chi^2_5=17.1$; $P=.004$) but not in predicting remission from anxiety ($\chi^2_5=9.3$; $P=.10$) or depression ($\chi^2_5=8.8$; $P=.12$). Among participants with clinically significant composite anxiety and depression, identifying as female was associated with substantially increased odds of remission (aOR 9.6, 95% CI 1.2-78.5; $P=.04$), whereas

baseline depression symptom severity was associated with a modestly decreased odds of remission (aOR 0.8, 95% CI 0.7-1; $P=.03$). No other baseline variables were associated with remission among the participants with clinically significant symptoms. It is noteworthy that the substantial aOR associated with female sex predicting remission from anxiety was, to some extent, because of the comparatively small proportion of male participants, as an investigation of zero-order before-and-after changes in prevalence found little evidence of differences between women (44.9% before to 15.0% after) and men (50.0% before to 27.8% after; Tables S5 and S6 in Multimedia Appendix 1).

Among participants with mild baseline symptoms, baseline variables were significant in predicting treatment response for anxiety ($\chi^2_5=12.3$; $P=.03$) but not in predicting treatment response for depression ($\chi^2_5=5.8$; $P=.33$) or composite anxiety and depression ($\chi^2_5=5.6$; $P=.35$). The only variable associated with treatment response was baseline depression symptom severity, which was associated with decreased odds of treatment response for anxiety (aOR 0.6, 95% CI 0.4-0.9; $P=.02$).

Table 5. Predictors of treatment response (n=125).

Symptoms	Female sex		Age		Number of sessions		Baseline PHQ ^a		Baseline GAD ^b		χ^2 (df)	<i>P</i> value
	OR ^c (95% CI)	<i>P</i> value	OR (95% CI)	<i>P</i> value	OR (95% CI)	<i>P</i> value	OR (95% CI)	<i>P</i> value	OR (95% CI)	<i>P</i> value		
Remission ^d among participants with clinically significant symptoms at baseline												
Symptoms of anxiety ^e , (n=57)	8.1 (1.0-68.0)	.05	1.1 (0.9-1.4)	.43	1.3 (0.9-1.7)	.11	0.9 (0.8-1.1)	.29	0.9 (0.7-1.1)	.16	9.3 (5)	.10
Symptoms of depression ^f , (n=62)	1.9 (0.3-11.5)	.49	1.1 (0.9-1.4)	.21	1.1 (0.9-1.4)	.32	0.9 (0.8-1)	.11	1.1 (0.9-1.3)	.14	8.8 (5)	.12
Composite anxiety and depression ^g , (n=61)	9.6 ^h (1.2-78.5)	.03 ^h	1.3 (0.9-1.8)	.07	1.2 (0.9-1.6)	.13	0.8 ^h (0.7-1)	.03	1.1 (0.9-1.3)	.28	17.1 (5)	.004 ^h
Treatment response ⁱ among participants with mild symptoms at baseline												
Symptoms of anxiety, (n=37)	0.2 (0.0-2.2)	.19	1.0 (0.7-1.3)	.89	0.8 (0.5-1.1)	.19	0.6 ^h (0.4-0.9)	.02 ^h	2.1 (1.0-4.3)	.06	12.3 (5)	.03
Symptoms of depression, (n=43)	2.6 (0.4-19.0)	.34	0.9 (0.7-1.1)	.21	0.8 (0.6-1.1)	.14	0.9 (0.5-1.6)	.78	1.1 (0.9-1.4)	.28	5.8 (5)	.33
Composite anxiety and depression, (n=39)	2.4 (0.3-20.4)	.42	0.9 (0.8-1.1)	.58	0.9 (0.6-1.2)	.33	0.7 (0.5-1.1)	.10	1.2 (0.8-1.6)	.37	5.6 (5)	.35

^aPHQ: Patient Health Questionnaire.^bGAD: generalized anxiety disorder.^cOR: odds ratio.^dRemission: reduction of symptoms so that Patient Health Questionnaire-9/Generalized Anxiety Disorder-7 score <10.^eGeneralized Anxiety Disorder-7 score.^fPatient Health Questionnaire-9 score.^gPatient Health Questionnaire-9+Generalized Anxiety Disorder-7.^hP<.05.ⁱTreatment response: a reduction in symptoms of 50% or more.

Satisfaction With Treatment

Most of the participants (n=125) rated intervention quality as good or excellent (114/125, 91.1%), were satisfied with the kind (108/125, 86.1%) and amount (108/125, 86.4%) of help received, reported being better able to deal effectively with their problems following the intervention (112/125, 89.6%), felt that the intervention met all or most of their needs (93/125, 74.4%), said that they would recommend the intervention to friends (119/125, 95.2%), and were satisfied overall with the intervention (113/125, 90.4%; Table S7 in [Multimedia Appendix 1](#)). Multiple linear regression analysis was used to identify baseline predictors of total satisfaction with treatment ([Multimedia Appendix 1](#), Table S7). Female participants reported significantly higher levels of satisfaction than male participants ($\beta=2$; 95% CI 0.6-4.7; $P=.01$; [Multimedia Appendix 1](#), Table S8). Interestingly, satisfaction was not associated with symptom improvement ($\beta=-.1$; 95% CI -1.3 to 0.7; $P=.57$), indicating that participants who showed less improvement were

as satisfied with the intervention as those who showed higher levels of improvement ([Multimedia Appendix 1](#), Table S8).

Discussion

Principal Findings

Students participating in this 10-week web-based intervention reported significant reductions in anxiety and depression symptom scores, with large effect sizes. Furthermore, there was a significant reduction in the proportion of participants with clinically significant symptoms from baseline to follow-up, with improvements across all levels of symptom severity. Participants reported high levels of satisfaction. This trial, which was conducted under real-world conditions in response to a global crisis, serves as a proof of concept for the use of web-based GCBT, showing that it is possible to recruit, retain, and improve the outcomes of university students with web-based GCBT.

It is noteworthy that our structured GCBT intervention was delivered by graduate clinical psychology students (ie, trainee psychologists) and counselors (with 4 years of training). This group intervention enabled us to reach a larger number of students than would otherwise have been the case using the same resources to deliver individual therapy. These findings speak to the sustainability and cost-effectiveness of the intervention, making it potentially well suited for resource-constrained environments such as South Africa, where the 12-month prevalence of common mental disorders among university students is as high as 31.5% [65] and where even at the most well-resourced institutions, only 28.9% of students with mental disorders receive treatment [19]. Session-by-session fiscal analysis has consistently shown that group therapy is less costly than individual therapy, highlighting the potential for group interventions to be scalable [66]. Efforts to expand the use of group therapy have hitherto been hampered by clinicians' concerns that groups may not be as effective as individual psychotherapy [66], which appears not to be the case for our intervention. Subsequent studies should establish how treatment outcomes are affected by the level of facilitator training and whether the intervention could be delivered by nonprofessionals within a peer-to-peer model. Consistent with the latter possibility, evidence suggests that peer-to-peer interventions can be effective in promoting student mental health [67,68]. Peer-to-peer interventions might also be appropriate given students' reluctance to seek help from mental health professionals [13].

Although attendance rates were relatively good, with a mean attendance of 6 sessions, only 10.1% of the students attended all sessions. This pattern is typical for students. For example, a trial of a 12-session group dialectical behavior therapy intervention for university students reported a 30.0% dropout rate with a mean attendance rate of approximately 60.0%. One survey of counseling center utilization over a 5-year period at a large US university reported that the average number of counseling sessions attended was 4 (SD 3.8, median 3, mode 1), with 18.0% of treatment seekers attending only 1 session [69]. In another study, 51.7% of students who accessed counseling attended 5 or fewer sessions [70]. However, crucially, our retention rates are markedly higher than those typically reported in self-guided or guided individualized internet-based interventions with students [71], suggesting that web-based GCBT might be more engaging than other digital interventions. The use of simple strategies (such as a flexible approach to attendance, encouraging participants to return after they miss a session, and ensuring sufficient repetition and continuity so that missing sessions does not disrupt care) seems to have been effective in retaining students, although it would be valuable to experiment with alternative retention strategies in future implementations.

The remission rates found here (67.7%-78.9%) compare favorably with those reported in clinical trials for other treatments of anxiety and depression. For example, trials of antidepressant treatment typically report response rates in the order of 70%-80% [72], and a trial of GCBT and duloxetine for GAD reported a remission rate of 21.5% [73]. The marked improvements we observed in symptom scores at an aggregate

level are particularly remarkable given that we did not exclude students with mild or moderate symptoms, as would typically be the case in controlled trials [46]. The fact that individuals with mild and moderate symptoms showed significant improvements is important because this population also seeks treatment and is at risk for eventually developing clinical disorders.

Notably, it is possible that our treatment outcomes are strong because we recruited participants into this intervention during the hard lockdown in South Africa at the start of the COVID-19 pandemic, and restrictions were starting to ease as participants completed the intervention. It is also important to note that we only recruited for 24 hours, which means that the intervention took only students very eager to obtain help. By only including very eager students, we have biased our outcomes by including only highly motivated and responsive participants. Before-and-after differences in symptoms might be smaller for more representative samples of students. Nonetheless, our ability to respond quickly to a need for services and to accommodate so many students is a clear strength of this intervention. It is also noteworthy that easing of restrictions on movement occurred over the period of the intervention, undoubtedly contributing to reduced levels of anxiety and depression, although the country was still in a state of national emergency and students were still making stressful transitions to web-based learning and assessment when follow-up data were collected at the end of the intervention. It is also possible that the positive effects we observed for mean symptom scores are a result of statistical regression to the mean (ie, the tendency of extreme measures to move closer to the mean overtime). However, this is less likely than would normally be the case in clinical trials, given that we recruited participants with a wide range of symptoms, including some students with minimal symptom scores at baseline. It will be important to establish in future replications if these favorable outcomes are also observed when students are not confined to their homes and when they have access to other therapies. To this end, follow-up studies including rigorous controlled trials are needed to test the effectiveness of the intervention against other standard treatments.

Although the good outcomes we observed suggest that web-based groups may be an effective and sustainable way to increase students' access to mental health services, it is important to remember that there are significant barriers to providing internet-based services that make them inaccessible to some students. These barriers include the need for students to have appropriate technology, access to broadband internet services and adequate data, a stable internet connection, uninterrupted power supply, and a private space in which to participate [74]. Many of these barriers will be particularly marked for students with limited financial resources, and providing web-based services may thus further exacerbate the inequality in access to treatments in low- and middle-income countries [19].

We found high levels of satisfaction with the intervention, highlighting the potential for web-based GCBT to meet the needs of anxious and depressed students. It is significant that most participants rated the quality of the intervention as good

or excellent and said that they were satisfied with the amount and kind of help they received. This finding is consistent with studies reporting students' positive attitudes toward internet-based mental health interventions [75] and provides further evidence to support the use of digital interventions to promote student wellness [38]. However, it would be helpful if subsequent studies in this area could collect more detailed qualitative data to provide a rich description of students' lived experience of participating in web-based interventions of this kind.

It will be important for future studies in this area to shed more light on factors that predict a good treatment response to these kinds of interventions. Given that this is a group intervention, it also seems appropriate to investigate how group membership and group dynamics (such as group cohesion) may influence treatment outcomes.

Limitations

There are 2 limitations that are worth noting. First, the study did not include a control group, as would be the case in a randomized controlled trial. As a result, we cannot infer that the observed symptom improvements were a result of the intervention. It is well established that even control groups (with no interventions) show improvements over time [76]. However, it was not our intention to undertake a controlled trial; instead, our challenge was to devise a treatment delivered in a novel digital format and tested in a real-world setting where treatment could be applied on a larger scale than is possible in individual therapy. Meeting these challenges along with the positive treatment results makes a randomized controlled trial a logical

next step, as our findings clearly demonstrated that such a trial is both warranted and feasible.

Second, a relatively high number of participants dropped out, with 28.4% (45/158) of participants attending fewer than half of the sessions. Dropout rates in psychotherapy research vary widely, ranging from 0.0% to 83.0% for e-interventions [77], and it is not clear how dropping out should be interpreted, given that many clients show gains early in treatment, which may contribute to *premature* termination. Despite these limitations, our study is an important initial step toward establishing that GCBT can be effectively delivered on the web to students under real-world conditions.

Conclusions

In conclusion, our results suggest that web-based GCBT holds promise as an effective and sustainable intervention for anxiety and depression and that university students participating in this intervention are satisfied with this modality. We demonstrated that it is possible as part of routine care in university counseling centers to recruit and retain students in web-based GCBT and deliver a remote intervention via video conferencing under real-world conditions. This novel intervention could have important implications for increasing access to psychotherapy in low- and middle-income countries. These findings indicate that larger-scale controlled trials of this modality are warranted, particularly trials that expand recruitment to include a wider range of students other than those most eager to participate, especially reaching out for more male students; investigate how retention can be improved; and examine factors that predict treatment response.

Acknowledgments

Development and design of intervention materials: Caitlin Grace Briedenhann, Philip Slabbert and Ryan Van Der Poll. Group facilitators: Dionne Jivan, Sanja Killian, Boitumelo Tebogo Mamiane, Natsai Mano, Nomawethu Matyobeni, Michelle Prevost, Judith Weis, Lisa Wolfaardt, Deviné Kamalie, Vastrohiette Gilbert, Lamees Abrahams, and Alene Smith. CBT experts who reviewed intervention materials: Meyer D. Glantz, Aimee Karam and Chelsey Wilks.

This work was supported with funding from the South African Medical Research Council (SAMRC) under its mid-career scientist development programme (awarded to Jason Bantjes). The views expressed in this manuscript do not represent those of the SAMRC. The research was also supported by a VR(RIPS) Special Covid-19 Research Grant received from the Division of Research Development at Stellenbosch University

Conflicts of Interest

DJS has received research grants and/or consultancy honoraria from Johnson & Johnson, Lundbeck, Servier, and Takeda. In the past 3 years, RCK was a consultant for Datastat, Inc; RallyPoint Networks, Inc; Sage Pharmaceuticals; and Takeda. The other authors have no conflicts of interest to declare.

Multimedia Appendix 1

Detailed statistical analysis to support the findings.

[DOCX File, 52 KB - [mental_v8i5e27400_app1.docx](#)]

References

1. Lund C, Brooke-Sumner C, Baingana F, Baron EC, Breuer E, Chandra P, et al. Social determinants of mental disorders and the Sustainable Development Goals: a systematic review of reviews. *Lancet Psychiatry* 2018 Apr;5(4):357-369. [doi: [10.1016/S2215-0366\(18\)30060-9](https://doi.org/10.1016/S2215-0366(18)30060-9)] [Medline: [29580610](https://pubmed.ncbi.nlm.nih.gov/29580610/)]

2. Herrman H, Kieling C, McGorry P, Horton R, Sargent J, Patel V. Reducing the global burden of depression: a Lancet-World Psychiatric Association Commission. *Lancet* 2019 Jun 15;393(10189):42-43. [doi: [10.1016/S0140-6736\(18\)32408-5](https://doi.org/10.1016/S0140-6736(18)32408-5)] [Medline: [30482607](https://pubmed.ncbi.nlm.nih.gov/30482607/)]
3. Chesney E, Goodwin GM, Fazel S. Risks of all-cause and suicide mortality in mental disorders: a meta-review. *World Psychiatry* 2014 Jun;13(2):153-160 [FREE Full text] [doi: [10.1002/wps.20128](https://doi.org/10.1002/wps.20128)] [Medline: [24890068](https://pubmed.ncbi.nlm.nih.gov/24890068/)]
4. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders, 5th Edition: DSM-5*. Washington, DC: American Psychiatric Publishing; 2013:1-991.
5. Auerbach RP, Mortier P, Bruffaerts R, Alonso J, Benjet C, Cuijpers P, WHO WMH-ICS Collaborators. WHO World Mental Health Surveys International College Student Project: prevalence and distribution of mental disorders. *J Abnorm Psychol* 2018 Oct;127(7):623-638 [FREE Full text] [doi: [10.1037/abn0000362](https://doi.org/10.1037/abn0000362)] [Medline: [30211576](https://pubmed.ncbi.nlm.nih.gov/30211576/)]
6. Alonso J, Vilagut G, Mortier P, Auerbach RP, Bruffaerts R, Cuijpers P, WHO WMH-ICS Collaborators. The role impairment associated with mental disorder risk profiles in the WHO World Mental Health International College Student Initiative. *Int J Methods Psychiatr Res* 2019 Jun 06;28(2):e1750 [FREE Full text] [doi: [10.1002/mpr.1750](https://doi.org/10.1002/mpr.1750)] [Medline: [30402985](https://pubmed.ncbi.nlm.nih.gov/30402985/)]
7. Bruffaerts R, Mortier P, Kiekens G, Auerbach R, Cuijpers P, Demyttenaere K, et al. Mental health problems in college freshmen: prevalence and academic functioning. *J Affect Disord* 2018 Jan 01;225:97-103 [FREE Full text] [doi: [10.1016/j.jad.2017.07.044](https://doi.org/10.1016/j.jad.2017.07.044)] [Medline: [28802728](https://pubmed.ncbi.nlm.nih.gov/28802728/)]
8. Mortier P, Auerbach R, Alonso J, Bantjes J, Benjet C, Cuijpers P, WHO WMH-ICS Collaborators. Suicidal thoughts and behaviors among first-year college students: results from the WMH-ICS Project. *J Am Acad Child Adolesc Psychiatry* 2018 Apr;57(4):263-273 [FREE Full text] [doi: [10.1016/j.jaac.2018.01.018](https://doi.org/10.1016/j.jaac.2018.01.018)] [Medline: [29588052](https://pubmed.ncbi.nlm.nih.gov/29588052/)]
9. Trivedi MH, Rush AJ, Wisniewski SR, Nierenberg AA, Warden D, Ritz L, et al. Evaluation of outcomes with citalopram for depression using measurement-based care in STAR*D: implications for clinical practice. *Am J Psychiatry* 2006 Jan;163(1):28-40. [doi: [10.1176/appi.ajp.163.1.28](https://doi.org/10.1176/appi.ajp.163.1.28)] [Medline: [16390886](https://pubmed.ncbi.nlm.nih.gov/16390886/)]
10. Tolin DF. Is cognitive-behavioral therapy more effective than other therapies? A meta-analytic review. *Clin Psychol Rev* 2010 Aug;30(6):710-720. [doi: [10.1016/j.cpr.2010.05.003](https://doi.org/10.1016/j.cpr.2010.05.003)] [Medline: [20547435](https://pubmed.ncbi.nlm.nih.gov/20547435/)]
11. Wilson CJ, Deane FP. Help-negation and suicidal ideation: the role of depression, anxiety and hopelessness. *J Youth Adolesc* 2010 Mar;39(3):291-305. [doi: [10.1007/s10964-009-9487-8](https://doi.org/10.1007/s10964-009-9487-8)] [Medline: [19957103](https://pubmed.ncbi.nlm.nih.gov/19957103/)]
12. Bruffaerts R, Mortier P, Auerbach RP, Alonso J, De la Torre AE, Cuijpers P, WHO WMH-ICS Collaborators. Lifetime and 12-month treatment for mental disorders and suicidal thoughts and behaviors among first year college students. *Int J Methods Psychiatr Res* 2019 Jun 20;28(2):e1764 [FREE Full text] [doi: [10.1002/mpr.1764](https://doi.org/10.1002/mpr.1764)] [Medline: [30663193](https://pubmed.ncbi.nlm.nih.gov/30663193/)]
13. Prince JP. University student counseling and mental health in the United States: trends and challenges. *Ment Health Prev* 2015 May;3(1-2):5-10. [doi: [10.1016/J.MHP.2015.03.001](https://doi.org/10.1016/J.MHP.2015.03.001)]
14. Eisenberg D, Hunt J, Speer N. Help seeking for mental health on college campuses: review of evidence and next steps for research and practice. *Harv Rev Psychiatry* 2012;20(4):222-232. [doi: [10.3109/10673229.2012.712839](https://doi.org/10.3109/10673229.2012.712839)] [Medline: [22894731](https://pubmed.ncbi.nlm.nih.gov/22894731/)]
15. Collins KA, Westra HA, Dozois DJ, Burns DD. Gaps in accessing treatment for anxiety and depression: challenges for the delivery of care. *Clin Psychol Rev* 2004 Sep;24(5):583-616. [doi: [10.1016/j.cpr.2004.06.001](https://doi.org/10.1016/j.cpr.2004.06.001)] [Medline: [15325746](https://pubmed.ncbi.nlm.nih.gov/15325746/)]
16. Fergusson DM, Horwood LJ, Ridder EM, Beautrais AL. Subthreshold depression in adolescence and mental health outcomes in adulthood. *Arch Gen Psychiatry* 2005 Jan;62(1):66-72. [doi: [10.1001/archpsyc.62.1.66](https://doi.org/10.1001/archpsyc.62.1.66)] [Medline: [15630074](https://pubmed.ncbi.nlm.nih.gov/15630074/)]
17. Hunt C, Andrews G. Drop-out rate as a performance indicator in psychotherapy. *Acta Psychiatr Scand* 1992 Apr;85(4):275-278. [doi: [10.1111/j.1600-0447.1992.tb01469.x](https://doi.org/10.1111/j.1600-0447.1992.tb01469.x)] [Medline: [1595361](https://pubmed.ncbi.nlm.nih.gov/1595361/)]
18. Xiao H, Carney DM, Youn SJ, Janis RA, Castonguay LG, Hayes JA, et al. Are we in crisis? National mental health and treatment trends in college counseling centers. *Psychol Serv* 2017 Nov;14(4):407-415. [doi: [10.1037/ser0000130](https://doi.org/10.1037/ser0000130)] [Medline: [29120199](https://pubmed.ncbi.nlm.nih.gov/29120199/)]
19. Bantjes J, Saal W, Lochner C, Roos J, Auerbach RP, Mortier P, et al. Inequality and mental healthcare utilisation among first-year university students in South Africa. *Int J Ment Health Syst* 2020 Jan 25;14(1):5 [FREE Full text] [doi: [10.1186/s13033-020-0339-y](https://doi.org/10.1186/s13033-020-0339-y)] [Medline: [31998406](https://pubmed.ncbi.nlm.nih.gov/31998406/)]
20. 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](https://doi.org/10.2196/12869)] [Medline: [31333198](https://pubmed.ncbi.nlm.nih.gov/31333198/)]
21. Kählke F, Berger T, Schulz A, Baumeister H, Berking M, Auerbach RP, et al. Efficacy of an unguided internet-based self-help intervention for social anxiety disorder in university students: a randomized controlled trial. *Int J Methods Psychiatr Res* 2019 Jun;28(2):e1766. [doi: [10.1002/mpr.1766](https://doi.org/10.1002/mpr.1766)] [Medline: [30687986](https://pubmed.ncbi.nlm.nih.gov/30687986/)]
22. Farrer L, Gulliver A, Chan JK, Batterham PJ, Reynolds J, Caltar A, et al. Technology-based interventions for mental health in tertiary students: systematic review. *J Med Internet Res* 2013;15(5):e101 [FREE Full text] [doi: [10.2196/jmir.2639](https://doi.org/10.2196/jmir.2639)] [Medline: [23711740](https://pubmed.ncbi.nlm.nih.gov/23711740/)]
23. Barkowski S, Schwartze D, Strauss B, Burlingame GM, Rosendahl J. Efficacy of group psychotherapy for anxiety disorders: a systematic review and meta-analysis. *Psychother Res* 2020 Nov;30(8):965-982. [doi: [10.1080/10503307.2020.1729440](https://doi.org/10.1080/10503307.2020.1729440)] [Medline: [32093586](https://pubmed.ncbi.nlm.nih.gov/32093586/)]

24. Dores AR, Geraldo A, Carvalho IP, Barbosa F. The use of new digital information and communication technologies in psychological counseling during the COVID-19 pandemic. *Int J Environ Res Public Health* 2020 Oct 21;17(20):7663 [FREE Full text] [doi: [10.3390/ijerph17207663](https://doi.org/10.3390/ijerph17207663)] [Medline: [33096650](https://pubmed.ncbi.nlm.nih.gov/33096650/)]
25. Andersen P, Toner P, Bland M, McMillan D. Effectiveness of transdiagnostic cognitive behaviour therapy for anxiety and depression in adults: a systematic review and meta-analysis. *Behav Cogn Psychother* 2016 Nov;44(6):673-690. [doi: [10.1017/S1352465816000229](https://doi.org/10.1017/S1352465816000229)] [Medline: [27301967](https://pubmed.ncbi.nlm.nih.gov/27301967/)]
26. Whitfield G. Group cognitive-behavioural therapy for anxiety and depression. *Adv Psychiatr Treat* 2018 Jan 02;16(3):219-227. [doi: [10.1192/apt.bp.108.005744](https://doi.org/10.1192/apt.bp.108.005744)]
27. Okumura Y, Ichikura K. Efficacy and acceptability of group cognitive behavioral therapy for depression: a systematic review and meta-analysis. *J Affect Disord* 2014 Aug;164:155-164. [doi: [10.1016/j.jad.2014.04.023](https://doi.org/10.1016/j.jad.2014.04.023)] [Medline: [24856569](https://pubmed.ncbi.nlm.nih.gov/24856569/)]
28. Coull G, Morris PG. The clinical effectiveness of CBT-based guided self-help interventions for anxiety and depressive disorders: a systematic review. *Psychol Med* 2011 Nov;41(11):2239-2252. [doi: [10.1017/S0033291711000900](https://doi.org/10.1017/S0033291711000900)] [Medline: [21672297](https://pubmed.ncbi.nlm.nih.gov/21672297/)]
29. Karyotaki E, Riper H, Twisk J, Hoogendoorn A, Kleiboer A, Mira A, et al. Efficacy of self-guided internet-based cognitive behavioral therapy in the treatment of depressive symptoms: a meta-analysis of individual participant data. *JAMA Psychiatry* 2017 Apr 01;74(4):351-359. [doi: [10.1001/jamapsychiatry.2017.0044](https://doi.org/10.1001/jamapsychiatry.2017.0044)] [Medline: [28241179](https://pubmed.ncbi.nlm.nih.gov/28241179/)]
30. 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](https://doi.org/10.1016/j.cpr.2012.02.004)] [Medline: [22466510](https://pubmed.ncbi.nlm.nih.gov/22466510/)]
31. 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 [FREE Full text] [doi: [10.1371/journal.pone.0154248](https://doi.org/10.1371/journal.pone.0154248)] [Medline: [27135410](https://pubmed.ncbi.nlm.nih.gov/27135410/)]
32. Sachan D. Self-help robots drive blues away. *Lancet Psychiatry* 2018 Jul;5(7):547. [doi: [10.1016/S2215-0366\(18\)30230-X](https://doi.org/10.1016/S2215-0366(18)30230-X)] [Medline: [29941139](https://pubmed.ncbi.nlm.nih.gov/29941139/)]
33. van der Zanden R, Kramer J, Gerrits R, Cuijpers P. Effectiveness of an online group course for depression in adolescents and young adults: a randomized trial. *J Med Internet Res* 2012;14(3):e86 [FREE Full text] [doi: [10.2196/jmir.2033](https://doi.org/10.2196/jmir.2033)] [Medline: [22677437](https://pubmed.ncbi.nlm.nih.gov/22677437/)]
34. Peden AR, Rayens MK, Hall LA, Beebe LH. Preventing depression in high-risk college women: a report of an 18-month follow-up. *J Am Coll Health* 2001 May;49(6):299-306. [doi: [10.1080/07448480109596316](https://doi.org/10.1080/07448480109596316)] [Medline: [11413947](https://pubmed.ncbi.nlm.nih.gov/11413947/)]
35. Peden AR, Hall LA, Rayens MK, Beebe LL. Reducing negative thinking and depressive symptoms in college women. *J Nurs Scholarsh* 2000 Jun;32(2):145-151. [doi: [10.1111/j.1547-5069.2000.00145.x](https://doi.org/10.1111/j.1547-5069.2000.00145.x)] [Medline: [10887713](https://pubmed.ncbi.nlm.nih.gov/10887713/)]
36. Vázquez FL, Torres A, Blanco V, Díaz O, Otero P, Hermida E. Comparison of relaxation training with a cognitive-behavioural intervention for indicated prevention of depression in university students: a randomized controlled trial. *J Psychiatr Res* 2012 Nov;46(11):1456-1463. [doi: [10.1016/j.jpsychires.2012.08.007](https://doi.org/10.1016/j.jpsychires.2012.08.007)] [Medline: [22939979](https://pubmed.ncbi.nlm.nih.gov/22939979/)]
37. Holtz BE, McCarroll AM, Mitchell KM. Perceptions and attitudes toward a mobile phone app for mental health for college students: qualitative focus group study. *JMIR Form Res* 2020 Aug 07;4(8):e18347 [FREE Full text] [doi: [10.2196/18347](https://doi.org/10.2196/18347)] [Medline: [32667892](https://pubmed.ncbi.nlm.nih.gov/32667892/)]
38. Lattie EG, Lipson SK, Eisenberg D. Technology and college student mental health: challenges and opportunities. *Front Psychiatry* 2019 Apr 15;10:246 [FREE Full text] [doi: [10.3389/fpsy.2019.00246](https://doi.org/10.3389/fpsy.2019.00246)] [Medline: [31037061](https://pubmed.ncbi.nlm.nih.gov/31037061/)]
39. Lenze EJ, Rodebaugh TL, Nicol GE. A framework for advancing precision medicine in clinical trials for mental disorders. *JAMA Psychiatry* 2020 Jul 01;77(7):663-664. [doi: [10.1001/jamapsychiatry.2020.0114](https://doi.org/10.1001/jamapsychiatry.2020.0114)] [Medline: [32211837](https://pubmed.ncbi.nlm.nih.gov/32211837/)]
40. Cornish PA, Berry G, Benton S, Barros-Gomes P, Johnson D, Ginsburg R, et al. Meeting the mental health needs of today's college student: reinventing services through Stepped Care 2.0. *Psychol Serv* 2017 Nov;14(4):428-442. [doi: [10.1037/ser0000158](https://doi.org/10.1037/ser0000158)] [Medline: [29120201](https://pubmed.ncbi.nlm.nih.gov/29120201/)]
41. Gillett G, Tomlinson A, Efthimiou O, Cipriani A. Predicting treatment effects in unipolar depression: a meta-review. *Pharmacol Ther* 2020 Aug;212:107557. [doi: [10.1016/j.pharmthera.2020.107557](https://doi.org/10.1016/j.pharmthera.2020.107557)] [Medline: [32437828](https://pubmed.ncbi.nlm.nih.gov/32437828/)]
42. Patel V, Weiss H, Mann A. Predictors of outcome in patients with common mental disorders receiving a brief psychological treatment: an exploratory analysis of a randomized controlled trial from Goa, India. *Afr J Psychiatry (Johannesbg)* 2010 Sep;13(4):291-296. [doi: [10.4314/ajpsy.v13i4.61879](https://doi.org/10.4314/ajpsy.v13i4.61879)] [Medline: [20957329](https://pubmed.ncbi.nlm.nih.gov/20957329/)]
43. Deckert J, Erhardt A. Predicting treatment outcome for anxiety disorders with or without comorbid depression using clinical, imaging and (epi)genetic data. *Curr Opin Psychiatry* 2019 Jan;32(1):1-6. [doi: [10.1097/YCO.0000000000000468](https://doi.org/10.1097/YCO.0000000000000468)] [Medline: [30480619](https://pubmed.ncbi.nlm.nih.gov/30480619/)]
44. Hatchett GT. Reducing premature termination in university counseling centers. *J College Stud Psychother* 2004 Oct 14;19(2):13-27. [doi: [10.1300/j035v19n02_03](https://doi.org/10.1300/j035v19n02_03)]
45. Richards D, Timulak L. Satisfaction with therapist-delivered vs. self-administered online cognitive behavioural treatments for depression symptoms in college students. *Br J Guid Counc* 2013 Apr;41(2):193-207. [doi: [10.1080/03069885.2012.726347](https://doi.org/10.1080/03069885.2012.726347)]
46. Patsopoulos NA. A pragmatic view on pragmatic trials. *Dialogues Clin Neurosci* 2011;13(2):217-224 [FREE Full text] [Medline: [21842619](https://pubmed.ncbi.nlm.nih.gov/21842619/)]

47. Schwartz D, Lellouch J. Explanatory and pragmatic attitudes in therapeutical trials. *J Chronic Dis* 1967 Aug;20(8):637-648. [doi: [10.1016/0021-9681\(67\)90041-0](https://doi.org/10.1016/0021-9681(67)90041-0)] [Medline: [4860352](#)]
48. Cuijpers P, Auerbach RP, Benjet C, Bruffaerts R, Ebert D, Karyotaki E, et al. Introduction to the special issue: the WHO World Mental Health International College Student (WMH-ICS) initiative. *Int J Methods Psychiatr Res* 2019 Jun;28(2):e1762 [FREE Full text] [doi: [10.1002/mpr.1762](https://doi.org/10.1002/mpr.1762)] [Medline: [30623516](#)]
49. Hamamci Z. Integrating psychodrama and cognitive behavioral therapy to treat moderate depression. *Arts Psychother* 2006 Jan;33(3):199-207. [doi: [10.1016/j.aip.2006.02.001](https://doi.org/10.1016/j.aip.2006.02.001)]
50. Zemestani M, Davoodi I, Honarmand MM, Zargar Y, Ottaviani C. Comparative effects of group metacognitive therapy versus behavioural activation in moderately depressed students. *J Ment Health* 2016 Dec;25(6):479-485. [doi: [10.3109/09638237.2015.1057326](https://doi.org/10.3109/09638237.2015.1057326)] [Medline: [26439473](#)]
51. Takagaki K, Okamoto Y, Jinnin R, Mori A, Nishiyama Y, Yamamura T, et al. Behavioral activation for late adolescents with subthreshold depression: a randomized controlled trial. *Eur Child Adolesc Psychiatry* 2016 Nov;25(11):1171-1182. [doi: [10.1007/s00787-016-0842-5](https://doi.org/10.1007/s00787-016-0842-5)] [Medline: [27003390](#)]
52. Yang X, Zhao J, Chen Y, Zu S, Zhao J. Comprehensive self-control training benefits depressed college students: a six-month randomized controlled intervention trial. *J Affect Disord* 2018 Jan 15;226:251-260. [doi: [10.1016/j.jad.2017.10.014](https://doi.org/10.1016/j.jad.2017.10.014)] [Medline: [29017069](#)]
53. Spitzer RL, Kroenke K, Williams JB, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch Intern Med* 2006 May 22;166(10):1092-1097. [doi: [10.1001/archinte.166.10.1092](https://doi.org/10.1001/archinte.166.10.1092)] [Medline: [16717171](#)]
54. Kroenke K, Spitzer RL. The PHQ-9: a new depression diagnostic and severity measure. *Psychiatr Ann* 2002 Sep 01;32(9):509-515. [doi: [10.3928/0048-5713-20020901-06](https://doi.org/10.3928/0048-5713-20020901-06)]
55. Mughal A, Devadas J, Ardman E, Levis B, Go V, Gaynes B. A systematic review of validated screening tools for anxiety disorders and PTSD in low to middle income countries internet. *Res Square* 2020. [doi: [10.21203/rs.3.rs-15809/v1](https://doi.org/10.21203/rs.3.rs-15809/v1)]
56. Beard C, Björgvinsson T. Beyond generalized anxiety disorder: psychometric properties of the GAD-7 in a heterogeneous psychiatric sample. *J Anxiety Disord* 2014 Aug;28(6):547-552. [doi: [10.1016/j.janxdis.2014.06.002](https://doi.org/10.1016/j.janxdis.2014.06.002)] [Medline: [24983795](#)]
57. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med* 2001 Sep;16(9):606-613 [FREE Full text] [doi: [10.1046/j.1525-1497.2001.016009606.x](https://doi.org/10.1046/j.1525-1497.2001.016009606.x)] [Medline: [11556941](#)]
58. Wittkampf KA, Naeije L, Schene AH, Huyser J, van Weert HC. Diagnostic accuracy of the mood module of the Patient Health Questionnaire: a systematic review. *Gen Hosp Psychiatry* 2007;29(5):388-395. [doi: [10.1016/j.genhosppsych.2007.06.004](https://doi.org/10.1016/j.genhosppsych.2007.06.004)] [Medline: [17888804](#)]
59. Kroenke K, Spitzer RL, Williams JB, Löwe B. The patient health questionnaire somatic, anxiety, and depressive symptom scales: a systematic review. *Gen Hosp Psychiatry* 2010;32(4):345-359. [doi: [10.1016/j.genhosppsych.2010.03.006](https://doi.org/10.1016/j.genhosppsych.2010.03.006)] [Medline: [20633738](#)]
60. Botha MN. Validation of the Patient Health Questionnaire (PHQ-9) in an African context. North-West University. 2011. URL: <http://hdl.handle.net/10394/4647> [accessed 2020-12-15]
61. Attkisson CC, Greenfield TK. Client satisfaction questionnaire-8 and service satisfaction scale-30. In: Maruish ME, editor. *The Use of Psychological Testing for Treatment Planning and Outcomes Assessment*. New Jersey, United States: Lawrence Erlbaum Associates, Inc; 2007:402-420.
62. Attkisson CC, Greenfield TK. The UCSF client satisfaction scales: I. The client satisfaction questionnaire-8. *The Use of Psychological Testing for Treatment Planning and Outcomes Assessment* (3rd Ed). New Jersey: Lawrence Erlbaum Associates; 2004. URL: https://www.academia.edu/13293847/The_UCSF_Client_Satisfaction_Scales_I_The_Client_Satisfaction_Questionnaire_8 [accessed 2021-04-29]
63. Kroenke K, Wu J, Yu Z, Bair MJ, Kean J, Stump T, et al. Patient health questionnaire anxiety and depression scale: initial validation in three clinical trials. *Psychosom Med* 2016;78(6):716-727 [FREE Full text] [doi: [10.1097/PSY.0000000000000322](https://doi.org/10.1097/PSY.0000000000000322)] [Medline: [27187854](#)]
64. Borenstein M, Hedges LV. *Introduction to Meta-Analysis*. Chichester, UK: John Wiley & Sons, Ltd; 2009.
65. Bantjes J, Lochner C, Saal W, Roos J, Taljaard L, Page D, et al. Prevalence and sociodemographic correlates of common mental disorders among first-year university students in post-apartheid South Africa: implications for a public mental health approach to student wellness. *BMC Public Health* 2019 Jul 10;19(1):922 [FREE Full text] [doi: [10.1186/s12889-019-7218-y](https://doi.org/10.1186/s12889-019-7218-y)] [Medline: [31291925](#)]
66. McDermut W, Miller I, Brown R. The efficacy of group psychotherapy for depression: a meta-analysis and review of the empirical research. *Clin Psychol Sci Pract* 2001;8(1):98-116. [doi: [10.1093/clipsy.8.1.98](https://doi.org/10.1093/clipsy.8.1.98)]
67. Kalkbrenner MT. Peer-to-peer mental health support on community college campuses: examining the factorial invariance of the mental distress response scale. *Community Coll J Res Pract* 2019 Jul 30;44(10-12):757-770. [doi: [10.1080/10668926.2019.1645056](https://doi.org/10.1080/10668926.2019.1645056)]
68. Byrom N. An evaluation of a peer support intervention for student mental health. *J Ment Health* 2018 Jun;27(3):240-246. [doi: [10.1080/09638237.2018.1437605](https://doi.org/10.1080/09638237.2018.1437605)] [Medline: [29451411](#)]
69. Kim JE, Park SS, La A, Chang J, Zane N. Counseling services for Asian, Latino/a, and White American students: initial severity, session attendance, and outcome. *Cultur Divers Ethnic Minor Psychol* 2016 Jul;22(3):299-310 [FREE Full text] [doi: [10.1037/cdp0000069](https://doi.org/10.1037/cdp0000069)] [Medline: [26390372](#)]

70. Huenergarde M. College students' well-being: use of counseling services. *Am J Undergrad Res* 2018 Dec 19;15(3):41 [FREE Full text] [doi: [10.33697/ajur.2018.023](https://doi.org/10.33697/ajur.2018.023)]
71. Becker TD, Torous JB. Recent developments in digital mental health interventions for college and university students. *Curr Treat Options Psych* 2019 Jun 14;6(3):210-220. [doi: [10.1007/s40501-019-00178-8](https://doi.org/10.1007/s40501-019-00178-8)]
72. Tranter R, O'Donovan C, Chandarana P, Kennedy S. Prevalence and outcome of partial remission in depression. *J Psychiatry Neurosci* 2002 Jul;27(4):241-247 [FREE Full text] [Medline: [12174733](https://pubmed.ncbi.nlm.nih.gov/12174733/)]
73. Xie Z, Han N, Law S, Li Z, Chen S, Xiao J, et al. The efficacy of group cognitive-behavioural therapy plus duloxetine for generalised anxiety disorder versus duloxetine alone. *Acta Neuropsychiatr* 2019 Dec 03;31(6):316-324. [doi: [10.1017/neu.2019.32](https://doi.org/10.1017/neu.2019.32)] [Medline: [31405402](https://pubmed.ncbi.nlm.nih.gov/31405402/)]
74. Kruse C, Fohn J, Wilson N, Patlan EN, Zipp S, Mileski M. Utilization barriers and medical outcomes commensurate with the use of telehealth among older adults: systematic review. *JMIR Med Inform* 2020 Aug 12;8(8):e20359 [FREE Full text] [doi: [10.2196/20359](https://doi.org/10.2196/20359)] [Medline: [32784177](https://pubmed.ncbi.nlm.nih.gov/32784177/)]
75. Montagni I, Tzourio C, Cousin T, Sagara JA, Bada-Alonzi J, Horgan A. Mental health-related digital use by university students: a systematic review. *Telemed J E Health* 2020 Feb;26(2):131-146. [doi: [10.1089/tmj.2018.0316](https://doi.org/10.1089/tmj.2018.0316)] [Medline: [30888256](https://pubmed.ncbi.nlm.nih.gov/30888256/)]
76. Cuijpers P, Karyotaki E, Weitz E, Andersson G, Hollon SD, van Straten A. The effects of psychotherapies for major depression in adults on remission, recovery and improvement: a meta-analysis. *J Affect Disord* 2014 Apr;159:118-126. [doi: [10.1016/j.jad.2014.02.026](https://doi.org/10.1016/j.jad.2014.02.026)] [Medline: [24679399](https://pubmed.ncbi.nlm.nih.gov/24679399/)]
77. Donkin L, Christensen H, Naismith SL, Neal B, Hickie IB, Glozier N. A systematic review of the impact of adherence on the effectiveness of e-therapies. *J Med Internet Res* 2011 Aug 05;13(3):e52 [FREE Full text] [doi: [10.2196/jmir.1772](https://doi.org/10.2196/jmir.1772)] [Medline: [21821503](https://pubmed.ncbi.nlm.nih.gov/21821503/)]

Abbreviations

ANOVA: analysis of variance
aOR: adjusted odds ratio
CBT: cognitive behavioral therapy
GAD: generalized anxiety disorder
GAD-7: Generalized Anxiety Disorder-7
GCBT: group cognitive behavioral therapy
MDE: major depressive episode
PHQ-9: Patient Health Questionnaire-9

Edited by J Torous; submitted 23.01.21; peer-reviewed by K Cohen, A Graham; comments to author 19.02.21; revised version received 02.03.21; accepted 15.03.21; published 27.05.21.

Please cite as:

Bantjes J, Kazdin AE, Cuijpers P, Breet E, Dunn-Coetzee M, Davids C, Stein DJ, Kessler RC
A Web-Based Group Cognitive Behavioral Therapy Intervention for Symptoms of Anxiety and Depression Among University Students: Open-Label, Pragmatic Trial
JMIR Ment Health 2021;8(5):e27400
URL: <https://mental.jmir.org/2021/5/e27400>
doi: [10.2196/27400](https://doi.org/10.2196/27400)
PMID: [34042598](https://pubmed.ncbi.nlm.nih.gov/34042598/)

©Jason Bantjes, Alan E Kazdin, Pim Cuijpers, Elsie Breet, Munita Dunn-Coetzee, Charl Davids, Dan J Stein, Ronald C Kessler. Originally published in *JMIR Mental Health* (<https://mental.jmir.org/>), 27.05.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Mental Health*, is properly cited. The complete bibliographic information, a link to the original publication on <https://mental.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Social Representations of e-Mental Health Among the Actors of the Health Care System: Free-Association Study

Margot Morgiève^{1,2,3*}, PhD; Pierre Mesdjian¹, MD; Olivier Las Vergnas^{4,5}, PhD; Patrick Bury⁶, PhD; Vincent Demassiet¹, MA; Jean-Luc Roelandt^{1,7}, MD; Déborah Sebbane^{1,7,8*}, MD

¹WHO Collaborating Centre for Research and Training in Mental Health, EPSM Lille Metropole, Hellemmes, France

²Cermes3, Centre de Recherche Médecine, Sciences, Santé, Santé Mentale et Société, Paris, France

³Department of Emergency Psychiatry and Acute Care, Lapeyronie Hospital, CHU Montpellier, Montpellier, France

⁴University of Lille, EA 4354, Centre Interuniversitaire de Recherche en Education de Lille, Lille, France

⁵UFR Sciences Psychologiques & de l'Éducation, University Paris-Nanterre, Nanterre, France

⁶Cleverside, Courbevoie, France

⁷Inserm, Épidémiologie clinique, évaluation économique appliquées aux populations vulnérables, UMR 1123, Paris, France

⁸University Hospital of Lille, Lille, France

*these authors contributed equally

Corresponding Author:

Déborah Sebbane, MD

WHO Collaborating Centre for Research and Training in Mental Health

EPSM Lille Metropole

211 rue Roger Salengro

Hellemmes, 59260

France

Phone: 33 320437100

Email: deborah.sebbane@ghpty-npdc.fr

Abstract

Background: Electronic mental (e-mental) health offers an opportunity to overcome many challenges such as cost, accessibility, and the stigma associated with mental health, and most people with lived experiences of mental problems are in favor of using applications and websites to manage their mental health problems. However, the use of these new technologies remains weak in the area of mental health and psychiatry.

Objective: This study aimed to characterize the social representations associated with e-mental health by all actors to implement new technologies in the best possible way in the health system.

Methods: A free-association task method was used. The data were subjected to a lexicometric analysis to qualify and quantify words by analyzing their statistical distribution, using the ALCESTE method with the IRaMuTeQ software.

Results: In order of frequency, the terms most frequently used to describe e-mental health in the whole corpus are: “care” (n=21), “internet” (n=21), “computing” (n=15), “health” (n=14), “information” (n=13), “patient” (n=12), and “tool” (n=12). The corpus of text is divided into 2 themes, with technological and computing terms on one side and medical and public health terms on the other. The largest family is focused on “care,” “advances,” “research,” “life,” “quality,” and “well-being,” which was significantly associated with users. The nursing group used very medical terms such as “treatment,” “diagnosis,” “psychiatry,” and “patient” to define e-mental health.

Conclusions: This study shows that there is a gap between the representations of users on e-mental health as a tool for improving their quality of life and those of health professionals (except nurses) that are more focused on the technological potential of these digital care tools. Developers, designers, clinicians, and users must be aware of the social representation of e-mental health conditions uses and intention of use. This understanding of everyone's stakes will make it possible to redirect the development of tools to adapt them as much as possible to the needs and expectations of the actors of the mental health system.

(JMIR Ment Health 2021;8(5):e25708) doi:[10.2196/25708](https://doi.org/10.2196/25708)

KEYWORDS

e-mental health; social representations; free association task; psychiatry; mental health; mental health service users; technology; digital health

Introduction

Context

Mental health care continues to face many challenges such as cost, accessibility, and the stigma associated with mental health. This results in inequalities and inadequacies in the treatment of many people with lived experiences of mental health problems [1]. The field of eHealth offers an opportunity to overcome these structural and personal barriers to seeking help [2]. Electronic mental health (or digital mental health) includes teleservice or telemedicine, interoperability repositories, shared medical records, mobile applications (mHealth), e-learning, online information searches and sharing, and others. Psychiatry, more than any other discipline, will be able to benefit from these new technologies. During the COVID-19 pandemic crisis, rapid virtualization demonstrated that clinicians, mental health service users, and health care systems were able to quickly adapt to telepsychiatry, overcoming previous obstacles including regulatory constraints, system inertia, and general resistance to telepsychiatry [3,4].

The use of technology is exponentially increasing in our society, especially the use of smartphones (more than 3.8 billion users worldwide) [5] to travel, communicate, work, manage one's finances, or to have social relations. Recently, the World Health Organization (WHO) announced its intention to use "new opportunities, creativity, learning and technology [...] to ensure the health and well-being of everyone" [6]. The use of information and communication technologies (ICTs) in health care since the 2000s is already improving access to care by strengthening communication between health service users and providers and by making health systems and decisions more efficient and cost-effective [7]. Indeed, in addition to the provision of direct service delivery, eHealth enables people with lived experiences of mental health problems to access their shared medical records and receive medical advice and information directly on their computers, tablets, or smartphones.

However, the use of these new technologies in the area of health and mental health remains weak. In France, only 6% of the population have already experienced a teleconsultation, and 9% of health care professionals have already done (at least) a teleconsultation with one of their patients [8]. Also, nearly two-thirds of the French population report they are not ready to use connected objects in the future in the health care field [9]. On the contrary, people with lived experiences of mental health problems are more and more connected, and most are in favor of using applications and websites to manage their mental health problems [10].

This study explored the social representations of e-mental health with the actors of the mental health system with the hypothesis that these social representations can help to understand and characterize the intentions of use.

Social Representations in eHealth

Several questionnaires were created in order to get an understanding of the barriers to the use of new technology in general. Those Technology Acceptance Models (TAMs), created in the 1980s, were used to better target the eHealth expectations of users [11] and professionals [12] based on 2 main questions: "Is this new technology useful for me?" and "Is this technology easy to use?" Some researchers [13,14] have highlighted the need to broaden this questioning to include environmental factors of individuals, including the social influence between subjects but also between the tool and the subject. Indeed, this very logical-scientific approach to TAMs must be supplemented by a vision, certainly more subjective, but which directly questions social cognitions referring to the "object of e-mental health." In order to understand the place of the individual in relation to this object in society and the socioeconomic power issues that emerge from it, it seems essential to question the mental image of e-mental health according to the beliefs and attitudes about it [15]. According to Jodelet [16], it is from this singular mental representation that a form of "knowledge is constructed, socially elaborated and shared, having a practical aim and contributing to the constitution of a reality common to a social group." Thus, the social representation creates a link between the individual and the feeling of belonging to a group in society with the same interpretations and uses of e-mental health.

Objective

This study aimed to characterize the social representations associated with e-mental health by all actors in order to implement new technologies in the best possible way within the health system.

Methods

Study Design

A qualitative study (EQUME) was conducted by the WHO Collaborating Centre of Lille (France) in order to assess the social representations and norms of 10 typologies of actors involved in the health care system. These 10 categories were chosen in order to have access to different professional groups with different references and practices (general practitioners, psychiatrists, social workers, psychologists, occupational therapists, and nurses); to service users and family carers; to user representatives, who have a discourse significantly different from the users; and to the general public. Participants were recruited through announcements to various professional networks, peer support groups, user and carer representatives (mainly posted on their respective websites), and by word of mouth. The inclusion criteria were as follows: (1) belong to one of these 10 categories of actors, (2) speak the French language, (3) be of legal age, (4) agree to participate in this study. There were no other criteria for noninclusion. The data were collected during focus groups (moderated by a social sciences researcher

and an assistant psychiatric moderator), which took place in 2 French cities.

The first part of the study, based on data collected in focus groups, revealed a heterogeneous and unstable definition of e-mental health with regard to the different groups of actors concerned as well as within each group [17]. The second part of the study, presented here, is based on the free-association task method.

Each focus group was initiated with a sociodemographic questionnaire collecting the following variables: age, gender, and profession. Each group was then asked to complete a self-reported familiarity scale ranging from 0 to 10 with e-mental health devices. Finally, a free-association task — detailed in subsequent sections — was conducted to collect words related to e-mental health.

The EQUIME study was the subject of a declaration of compliance with reference methodology at the French National Agency for Medicines and Health Products Safety (N°2040798 v 0, March 3, 2017). All participants were asked to sign a consent form.

Procedures and Methods

The free-association method was chosen to study the social representations of e-mental health. This method is based on a question of evocation (or word associations) with the following written instructions: “Quote three words related to ‘e-mental health,’ then three more words related to these words” (see [Multimedia Appendix 1](#)). This exercise will result in having 3 words at the first level, then 9 words at the second level as each word from the first level will be associated with 3 other words. This makes a total of 12 words or expressions per participant. The free-association method is a classic tool in studies on social representations [18-20]. It calls upon the latent content of representation [15,21] opening a path to the semantic field of the social object studied through the spontaneity and projective dimension of the method of free associations [15]. According to Abric [15], social representation is composed of a content (eg, information, opinions, beliefs, attitudes) and a structure. The structure consists of a central system (or central core) and a peripheral system, each of which is composed of the beliefs of the same name. The central elements have “evidential status” and help to “provide a framework for interpreting and categorizing new information” [15]. The peripheral system links the central core of the representation to the reality of the moment for individuals. For example, if we consider “knowledge acquisition” as the central core of the object “study,” for some, “the library” will be a peripheral element, while for others, “the scholarship” will be an entirely different one (considering that knowledge acquisition would allow one to obtain a scholarship related to further study).

Data Analysis

We used several types of text data analysis (TDA) in this study. TDA corresponds to a set of methods that aim to analyze the information contained in a text. Two of the authors (OLV and PB) who specialize in the statistical analysis of textual data conducted the technical analyses, guided by a social sciences researcher (MM) and mental health clinicians (PM and DS).

They use categories to qualify elements of the text and quantify them by analyzing their statistical distribution.

Lexical Analysis

The data were subjected to a lexical or lexicometric analysis: the ALCESTE method. It was developed by Reinert [22] on the basis of the work by Benzécri and the textual statistics of Lebart and Salem. We used the IRaMuTeQ software, which is open source, is free, uses the R language [23], and was developed by Ratinaud and Dejean.

Text segments have been created from each “level 1 word” and the three “level 2 words” associated with them (equivalent to a branch of the tree structure of the free-association diagram, see [Multimedia Appendix 1](#)). This makes 3 sentences or text segments called B1, B2, B3 (in the order of the word branches quoted from left to right on the diagram) per participant. In order to identify groups of words often together in these text segments, the analysis performed is mainly a Hierarchical Descending Classification (HDC). The software builds a tree structure, and a classification is proposed grouping the words most often used together in the same sentences or segments.

Still using the IRaMuTeQ software, we obtained a visualization of the relations between the word clusters and the variables studied (age, sex, familiarity with e-mental health, categories of actors, order of text segments [ie, B1, B2, and B3]) with the corresponding chi-squared value ($P=.05$). We defined the significance threshold at 5%. Based on correspondence factor analysis (CFA) applied to the center of the clusters, this visualization provides pairs of images that can be stacked together. One of the images represents the relative proximity of words, and the other represents the types of text segments concerned, around the centers of these lexicons. The central words are the most common, and the distance from the center indicates the specificity of one or the other word. The axes mathematically maximize the visibility of specificities, but their orientation on the page (top/bottom and right/left) is arbitrary.

Thematic or Categorical Analysis

To deepen the links that exist between the different terms, we used a graph representation tool (Neovis) in order to visually understand the different word associations. For this purpose, 3 researchers independently classified the 180 terms in level 1 into 24 categories.

On a technical level, the graphs were built from the Excel file resulting from the encoding of the responses that we injected into a dedicated database (neo4j) using a python script. An HTML page connecting to this database was then built; its role was to retrieve the relationship of interest and represent it using Neovis.

Results

Baseline Assessment

The sample comprised a total of 70 people (37 women and 33 men) between 24 and 77 years old (average age of 44 years; [Table 1](#)). They correspond to 10 categories of actors: general practitioners, psychiatrists, user representatives, general public,

family carers, social workers, psychologists, service users, occupational therapists, and nurses.

Self-reported familiarity with e-mental health ranged from 0 to 9/10 but was on average, very low for all groups. The occupational therapists report the lowest level of familiarity

(1.1), while general practitioners report the highest level of familiarity (4.5).

Responses to the free-association questionnaire had 167 words missing for 828 possible answers (20.2%). The user group had the highest rate of missing words (46%), as it was the group with the most participants.

Table 1. Participants' characteristics and self-assessment of eHealth knowledge.

Categories of actors	Participants, n			Age of participants (years), mean (range)	Knowledge of e-mental health tools, mean (range)
	Men	Women	Total		
General practitioners	4	1	5	48.4 (40-59)	4.5 (3-5)
Psychiatrists	3	2	5	43.6 (25-62)	3.2 (0-8.5)
User representatives	2	1	3	54.3 (29-77)	3.3 (1-6)
General public	0	6	6	38.5 (29-53)	3.2 (1-7)
Family carers	6	3	9	62.2 (48-74)	1.8 (0-4)
Social workers	0	5	5	43.2 (29-57)	1.6 (0-5)
Psychologists	1	6	7	35.7 (25-59)	1.7 (0-5)
Service users	11	1	12	42 (30-59)	3.7 (0-9)
Occupational therapists	2	7	9	38.4 (24-56)	1.1 (0-4)
Nurses	4	5	9	36.7 (25-48)	2.6 (0-6)
Total group	33	37	70	44.3 (24-77)	2.2 (0-9)

Lexical Analysis

In order of frequency, the terms most frequently used to describe e-mental health in the whole corpus were “care” (n=21), “internet” (n=21), “computing” (n=15), “health” (n=14), “information” (n=13), “patient” (n=12), and “tool” (n=12).

The terms that were cited several times together (co-occurrences) throughout the corpus were “internet” and “information” (n=6), “internet” and “computer” (n=6), “activity” and “workshop” (n=5), “activity” and “care” (n=5), “hope” and “activity” (n=5), “carefree” and “activity” (n=5), and “will” and “activity” (n=5).

Analysis of the similarities and differences between the terms used by the participants (Figure 1) shows 5 clusters of word characteristics in the main themes addressed. The percentages represent the number of times the words are cited together throughout the corpus.

The corpus of text is divided in 2, with technological and computing terms on one side and medical and public health terms on the other.

As for the medico-social terms, the largest family (cluster 5) is focused on “care,” “advances,” “research,” “life,” “quality,” and “well-being.” It is related to 2 families (clusters 2 and 3), which also include health-related terms but differ from them by more general terms. These 2 other clusters are distinguished by more specific terms related to psychiatry and preventive medicine (eg, “psychiatry,” “diagnosis,” “prevention,” “information”) and by access terms related to public health and the direct environment of the user of the health system (eg, “health,” “public,” “share,” “user,” “family”).

During the task of free association, we can see that the participants very frequently quoted in the first line of the response ($\chi^2_1=4.93$, $P=.02$) terms associated with the lexical fields of technology and computer science (B1, Figure 2) overlapping with cluster 1 (Figure 3).

Participants over 61 years of age-related e-mental health to terms in the fields of “health,” “public,” “professional,” “medical,” “accessibility,” “family,” “user,” and “network” ($\chi^2_1=3.93$, $P=.04$).

The CFA (Figures 2 and 3) shows that the group of users consist of those who use the terms focused on “care,” “progress,” “research,” “life,” “quality,” and “well-being” the most ($\chi^2_1=11.16$, $P<.001$). It appears that this group of participants would make very little use of the other families of words, and almost none of them used terms related to technology or computing (clusters 1 and 4).

The nursing group used very medical terms such as “treatment,” “diagnosis,” “psychiatry,” and “patient” to define e-mental health ($\chi^2_1=4.8$, $P=.02$). They also used more global words focusing on “quality,” “care,” “progress,” and “well-being” as well as “users.” They did not associate e-mental health at all with the terms in the public health-oriented family (cluster 3).

The general public group associated terms such as “application,” “technology,” “digital,” “web,” “monitoring,” “computer,” “site,” and “knowledge” with e-mental health ($\chi^2_1=4.63$, $P=.03$), as did the group of user representatives ($\chi^2_1=3.11$, $P=.07$).

The general public, psychiatrists, occupational therapists, user representatives, and general practitioners used very little or no medico-social vocabulary (from clusters 2, 3, and 5) within their representations of e-mental health. These groups are more likely to use terms focused on computing and technology.

Figure 1. Classification of words used according to their frequency of co-citation within the same text segment by all participants in the 5 clusters, with $P<.05$ for all words.

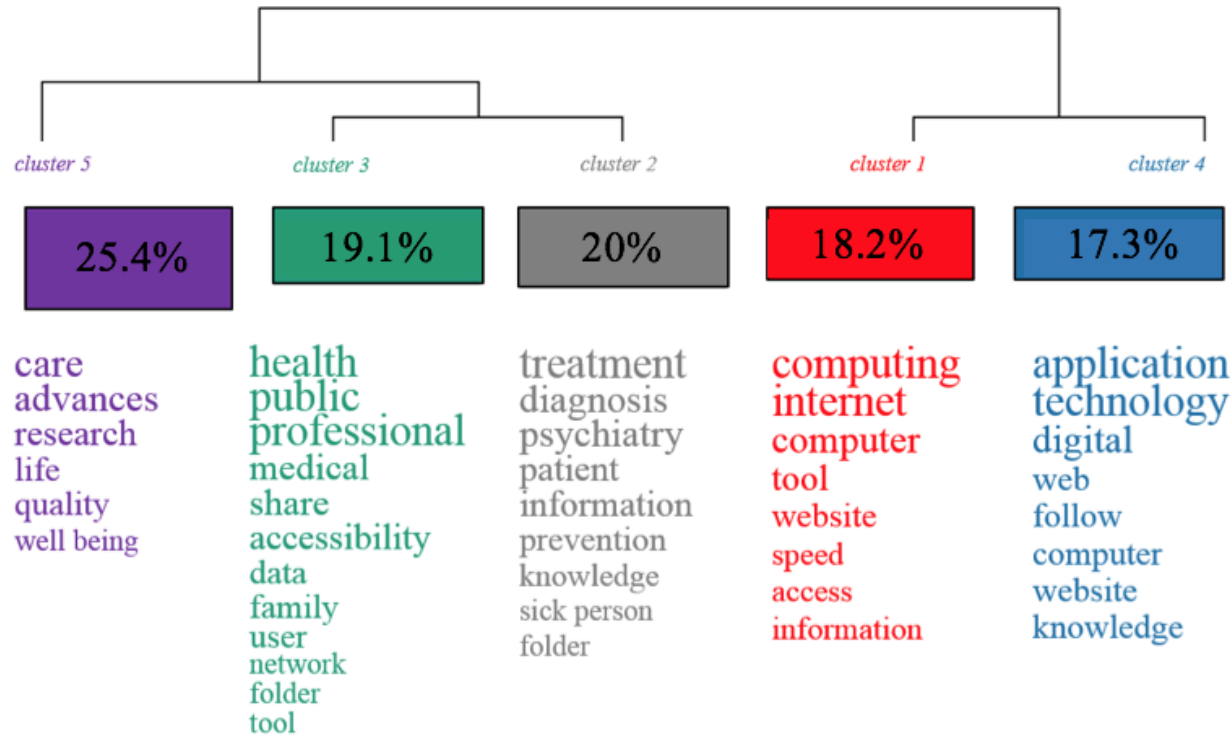


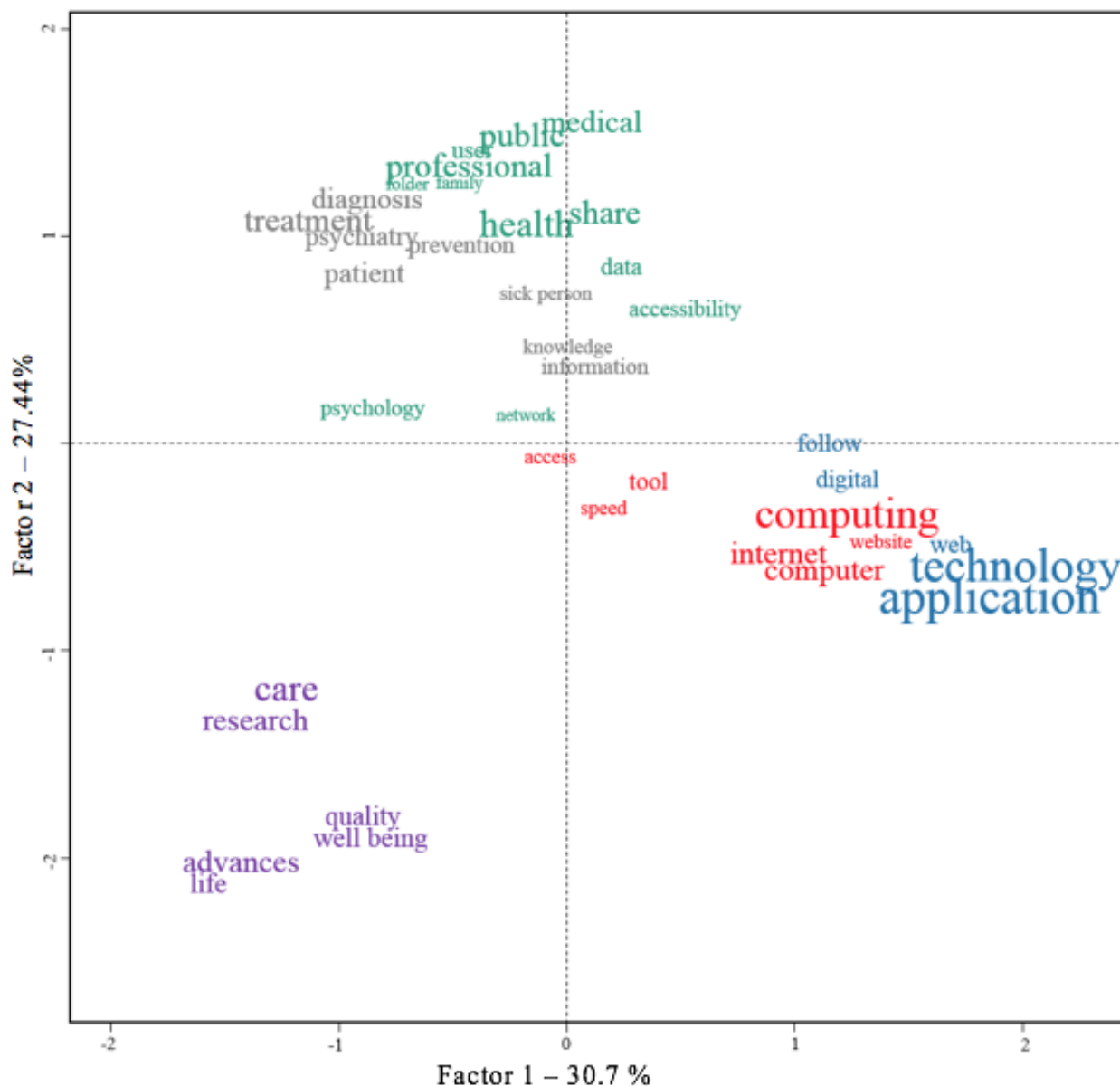
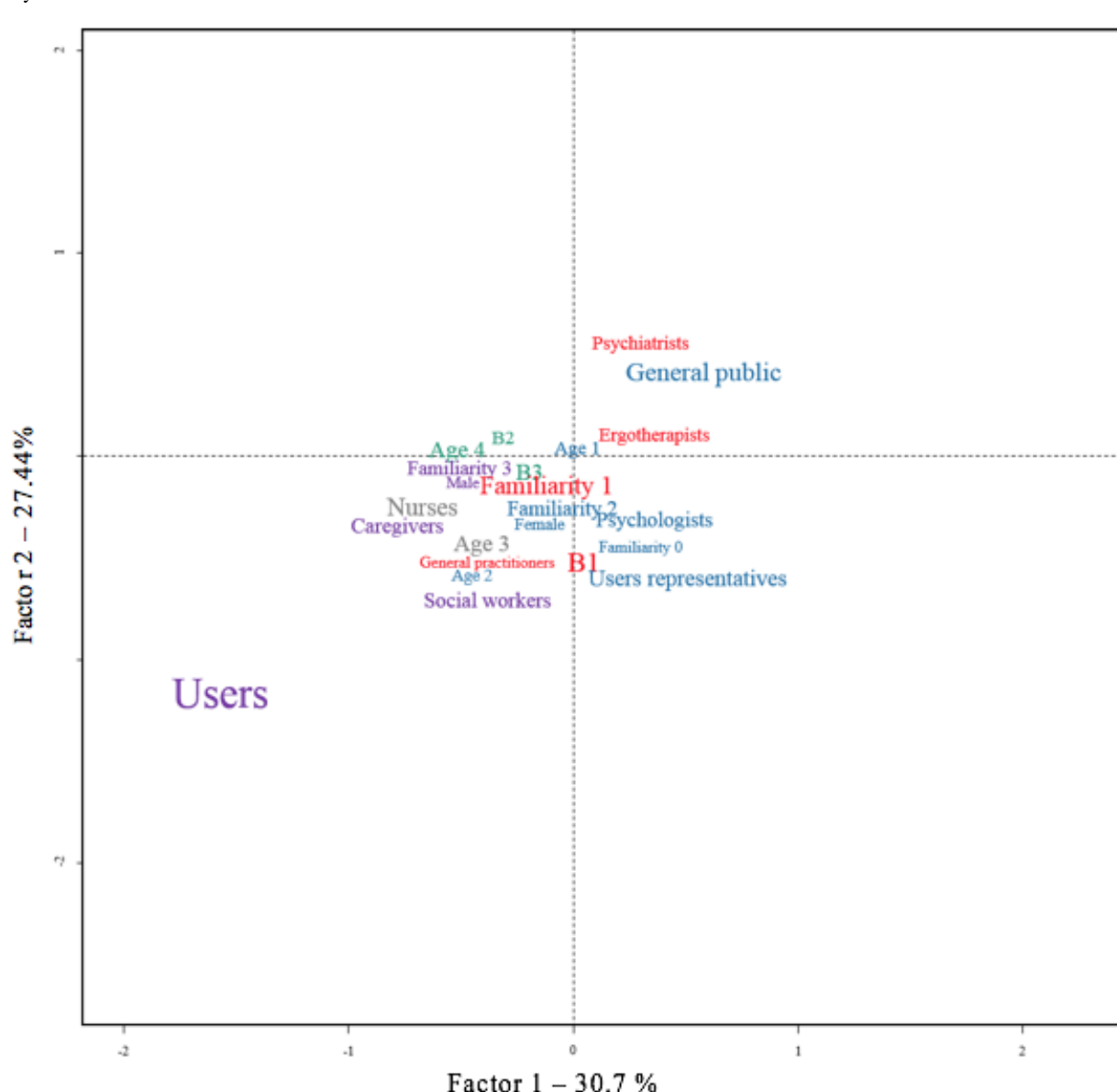
Figure 2. Correspondence factor analysis of the free-word association about e-mental health.

Figure 3. Correspondence factor analysis with the different variables studied. B1: text segment 1; B2: text segment 2, B3; text segment 3; Age 1: 20-30 years old; Age 2: 31-40 years old; Age 3: 41-60 years old; Age 4: ≥ 61 years old; Familiarity 0: no answer; Familiarity 1: 0-2; Familiarity 2: 3-5; Familiarity 3: 6-9.



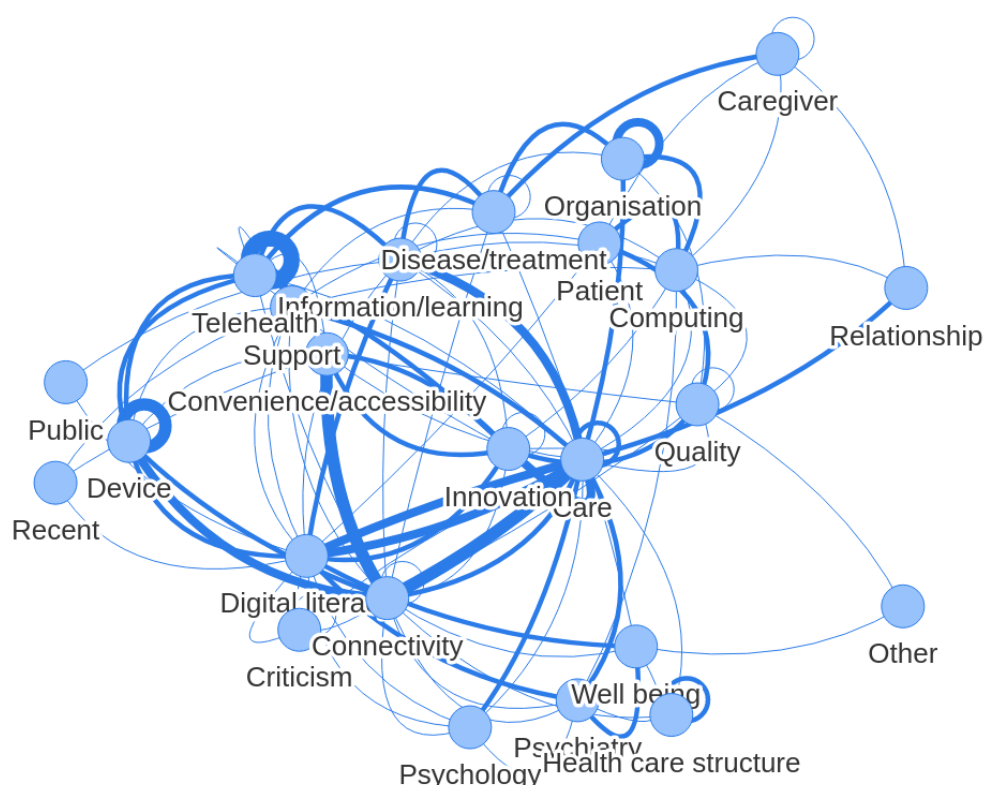
Semantic Analysis

The 24 ad hoc categories constituted by the investigators and based on level 1 terms are illustrated in Figure 4. This graph summarizes the main corpus by lexical fields. It introduces a dynamic dimension by adding links between the different categories, which can be compared to more or less “stretched springs” depending on the number of relationships between groups of words.

As shown in the figure, it is possible to notice that the central place of the category “care” has a link with almost all the other categories. The terms constituting this category are therefore related at least once to the words used in the other categories.

The other graphs that were made according to the different variables (age, gender, familiarity with e-mental health, except groups of participants) show a core of close relationships between the categories “care,” “connectivity,” “pathology/treatment,” “device,” “telehealth,” “computing,” “information/training,” “digital literacy,” “practicality/accessibility,” and “innovation.” There was no clear structural difference in the graph (Figure 4). It is possible to observe differences between the groups of participants depicted in the graphs; however, it is not possible to provide a clear conclusion because of the low number of participants per group.

Figure 4. Links between ad hoc categories of level 1 words. The stroke thickness and node distances represent the frequencies of word co-occurrences between categories.



Discussion

Main Findings

The scores of the self-reported familiarity scales are generally below average and are opposed to the richness of the words and lexical fields mobilized by the participants during the task of free association. This highlights a necessary distinction between daily digital use and access to digital health literacy that is controlled [24]. It is the responsibility of the state to set up an education system at school that allows future e-citizens to know how to use these tools in an informed way and to manage their digital identities and a digital infrastructure in order to avoid the “digital divide” as well as digital health illiteracy [7].

In the main corpus, a homogeneous and frequent vocabulary field relating to health care and ICTs (Figure 4) allows us to formulate the hypothesis of the centrality of these lexical fields illustrating the social representations of e-mental health of the participants. The absence of terms with positive or negative valences is to be noted. In addition, a very consensual and materialistic definition characterizes the central system of social representation. e-Mental health is considered a new technological, computerized, and medical tool that would be able to offer a diagnosis or treatment to people with a lived experience of mental health problems. These tools are at the service of information and training. Data from the free-association task suggest a relative openness or at least a lack of aversion to the mental health of participants. The subsequent focus group discussions also point in this direction, but nevertheless highlighted fears linked to “dehumanization” or the replacement of humans by technological tools [17]. The

peripheral elements are linked to the structural and organizational dimensions of e-mental health (ie, “structure of care,” “organization”).

The group of service users of the mental health system is clearly distinguished by a specific vocabulary. It differs from the words most found in the main corpus but also from the other groups of participants. These discrepancies evoke the nuance between users’ expectations of improving their quality of life in the first place and that of health professionals (except nurses), which focused more on the potential of new digital tools to perform repetitive tasks for them, allowing them to refocus their practice on what makes (clinical) sense.

“Psychiatry” and “psychology” are also peripheral elements of the representation of e-mental health. While psychiatry has established itself as the “normal practice” that has regulated the conception of disorders and their treatments for many years, it may seem “natural” that it now extends its jurisdiction to the field of mental health. This extension could thus announce the renewal of psychiatric practices, as well as their social role. Current frameworks guiding clinical practice in psychiatry and psychology are limited because they do not address the complex reality of people with lived experiences of mental health problems. They project on them a predefined reading grid and neglect the dynamic interaction between their real, lived experiences, which are inextricably linked to social, psychological, and biological contexts [25]. The mental health care system thus does not consider the fundamental realities of people with lived experiences of mental health problems in their daily life. We therefore urgently need new paradigms of clinical practice to effectively treat these people in vivo, in which what

matters the most for them — loss of meaning, impoverishment, social isolation, and/or disability associated with symptoms — is also what matters the most to clinicians [26,27]. However, new digital tools can precisely enable people to be observed and treated in vivo, by integrating a stream of ecological and multidimensional data. These developments require theorizing methodological approaches to guide the design of new digital tools adapted to the challenges of a digital clinic.

This integration of digital tools in the daily practice can thus become part of a “professional project” in order to gain status and expand territories [27].

“Well-being” is also a peripheral element associated with “care” for users specifically. This representation illustrates the process of gradually extending psychiatry to “mental health” and even happiness since the 1980s. This extension is based on the redefinition of health by the WHO, no longer as the absence of disease but as “complete physical, mental and social well-being.” “Mental health” has become ubiquitous in public health discourse and more broadly throughout the social landscape since the early 2000s. Many actors in the field see it as a form of injunction to happiness and well-being beyond the scope of psychiatrists’ interventions. This presence of “well-being” in the discourse of users can also be explained by the fact that the current technological tools are not necessarily medical tools but common objects (eg, connected watches, actimetry bracelets, smartphone applications) that have been designed according to ways of thinking about the world from fields other than psychiatry, in particular the well-being and quantification of oneself.

“Relationship” is also one of the peripheral representations associated with e-mental health. New technologies are changing the relationships between caregivers and people with lived experiences of mental health problems, enabling new forms of digital intimacy thanks to a new form of continuity of care. According to Fairhurst and May [28], abstract medical knowledge (“knowing the patient”) can thus be supplemented and enriched by personal exchanges that can help the clinician to “know about the patient.” The most recent example comes from the COVID-19 health care crisis during which telepsychiatry allowed the maintenance of social links through digital health despite the need for physical distancing [29]. Although technology has enabled the maintenance of caregivers’ and patients’ “connection,” experts recommend the complementary use of telepsychiatry with face-to-face interviewing [3,29]. It is a question of finding the balance of these new hybrid relationship modalities within the patient-caregiver-technology triad. Foucault et al [30] raised the following question: “Is there a virtual saturation point at which the benefits of a virtual relationship diminish, or patients demand more face-to-face interactions?” The relationship between service users and professionals seems to be evolving towards a rebalancing of each other’s roles and is being profoundly transformed under the effect of new technologies; the e-citizen user is thus becoming an informed actor of his or her health, expert, and partner in an increasingly digitalized ecosystem.

Although the processes of “autonomy” and empowerment are recommended by public health authorities and that these terms are increasingly present in the contemporary discourses of patient and service user associations as well as more widely disseminated in society, it is surprising that they are totally absent from the task of free association. However, the discursively configured involvement of patients in their care through technology is a matter of debate: advocated by some as a means of horizontalizing the caregiver-patient relationships and contested by others as a social injunction and the sign of the expression of a Foucauldian bio-power [30]. The promising discourse of digital health policy positions citizens as objects of political intervention but neglect the many social, political, cultural, and economic inequalities that specifically prevent engagement in digital health [31].

Similarly, “data” is absent from free associations. Participants thus did not seem to question the place of their own data in the mental health ecosystem or to be concerned about the use that private lobbies can make of it. This absence of “data” from the discourses of all the typologies of actors can be interpreted at different levels. Users of the health system may seem to “not care” about the confidentiality and security of their data. There might be a few reasons for that: They may trust the e-mental health ecosystem, it may seem that the benefit-risk balance makes it preferable for them to use these digital tools, or they might not have mastered the issues related to the use and circulation of their data. This could be explained by the difficult acquisition of health literacy for users of the health care system as well as for professionals, not because of lack of interest but rather by the complexity of the health ecosystem (ie, the lack of resources and reliable evaluation of medical information on the internet). To use the concept by Petersen et al [32], “cartographies of trust” has now become extremely complex and follows tortuous and emotionally charged paths that require navigating between online and offline resources. Health literacy is evolving; it requires medical, informational, and more recently, digital skills [24]. Considered increasingly civic and social [33], it is now part of the community with, on the one hand, the need for an awareness of “self-concern” at the individual level [31] and, on the other hand, the need for optimal organization of the health ecosystem managed by the guardianships.

As Henwood and Marent [31] rightly pointed out, at the level of the individual, the ways people “make sense with numbers” and numbers “make sense of people” interact so finely that it is extremely complex to determine towards or tilt the balance between freedom and power, determining and being determined, acting and being acted upon; in such a way, it is urgent to expand our sociological imaginations of the “reflexive” patient or citizen.

Limitations

The aim of this study was to “photograph” the social representations of e-mental health from the different typologies of actors in the health care system. The speed of development of new devices implies new uses will likely have a retroactive impact on users’ representations, making it difficult to capture these constantly evolving representations. Also, one of the main

limitations of our work is related to the small number of participants present in each group. Our material has a certain number of nonresponses to the task of free association without the possibility of exchange with participants. More qualitative studies, using narrative content, interviews, focus groups, and field observation methodologies, are needed to further explore the social representations of e-mental health among different actors.

Conclusions

The rise of e-mental health in our health systems is both a challenge and an opportunity for mental health. This study showed that the social representations of e-mental health differ according to the social group to which participants belong. It conditions an intention of use that developers, designers, clinicians, and users must be aware of. This better understanding

of everyone's stakes will make it possible to redirect the development of tools and adapt them as well as possible to the needs and expectations of the actors of the mental health system. In this process of listening and horizontalization of the relationships between actors, the aim was to harmonize the contribution of digital tools and enable their appropriation by all users, as well as facilitating equal access to care by bridging the digital divide. In order to do so, the guardianships must ensure the deployment process of the tools. If all user citizens have to be concerned by these policies and if they are to remain committed to a better knowledge of themselves and their health, these reflections must be participatory and collaborative. In this sense, the improvement of the components relating to the training of actors through the acquisition of digital skills and the increase of literacy e-mental health is at the dawn of a successful implementation of digital mental health.

Acknowledgments

The authors would like to thank the French Ministry of Health for funding the study, as well as the different institutions that supported this work and especially the Interreg North-West Europe Program, Psycos, the Sainte-Anne Hospital of Paris, the "Users' House" in the Sainte-Anne Hospital, the National Union of Family or Friends of Sick and Disabled Psychic Persons (Union Nationale de Familles ou Amis de Personnes Malades et Handicapées Psychiques- Unafam), and the Health Cooperation Group for the research and training in mental health (Groupement de Coopération Sanitaire- GCS).

We would like to thank Lisa Aissaoui for her careful proofreading of the article.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Free association task instructions.

[PNG File, 117 KB - [mental_v8i5e25708_app1.png](#)]

References

1. Ngui EM, Khasakhala L, Ndeti D, Roberts LW. Mental disorders, health inequalities and ethics: A global perspective. *Int Rev Psychiatry* 2010 Jun 09;22(3):235-244 [FREE Full text] [doi: [10.3109/09540261.2010.485273](https://doi.org/10.3109/09540261.2010.485273)] [Medline: [20528652](#)]
2. Shaw T, McGregor D, Brunner M, Keep M, Janssen A, Barnet S. What is eHealth (6)? Development of a Conceptual Model for eHealth: Qualitative Study with Key Informants. *J Med Internet Res* 2017 Oct 24;19(10):e324 [FREE Full text] [doi: [10.2196/jmir.8106](https://doi.org/10.2196/jmir.8106)] [Medline: [29066429](#)]
3. La télé-psychiatrie en période COVID-19?: un outil d'appui aux soins en psychiatrie publique. Conférence nationale des présidents de CME/CHS Groupe ressource PSY/COVID-19. 2020 Apr 24. URL: <https://fedepsychiatrie.fr/wp-content/uploads/2020/04/T%C3%A9l%C3%A9consultation-en-période-COVID19-le-24042020.pdf> [accessed 2021-04-10]
4. Khanna R, Forbes M. Telepsychiatry as a public health imperative: Slowing COVID-19. *Aust N Z J Psychiatry* 2020 Jul 04;54(7):758-758. [doi: [10.1177/0004867420924480](https://doi.org/10.1177/0004867420924480)] [Medline: [32363911](#)]
5. O'Dea S. Number of smartphone users worldwide from 2016 to 2023. Statista. 2021 Mar 31. URL: <https://www.statista.com/statistics/330695/number-of-smartphone-users-worldwide/> [accessed 2021-04-10]
6. Kluge H. Invitation to share feedback on European Programme of Work. World Health Organization Regional Office for Europe. 2020 Jul 27. URL: <https://www.euro.who.int/en/about-us/regional-director/news/news/2020/06/have-your-say-on-whoeuropes-proposed-european-programme-of-work/invitation-message-from-dr-hans-henri-p.-kluge.-who-regional-director-for-europe> [accessed 2021-04-10]
7. D'un système de santé curatif à un modèle préventif grâce aux outils numériques. *Renaiss Numérique*. 2014. URL: https://www.renaissancenumerique.org/ckeditor_assets/attachments/55/lb_sante_preventive_renaissance_numerique_1.pdf [accessed 2021-04-10]
8. Agence du Numérique en Santé. Le Baromètre de La Télémédecine. Agence du Numérique en Santé. 2020 Oct 22. URL: https://esante.gouv.fr/sites/default/files/media_entity/documents/odoxa-pour-lans-et-le-mag-de-la-sante---barometre-telemedecine-vague-2-publie-le-22-octobre-2020.pdf [accessed 2021-04-11]

9. Baromètre du numérique 2019: enquête sur la diffusion des technologies de l'information et de la communication dans la société française 2019. Conseil Général de l'Economie, de l'Industrie, de l'Energie et des Technologies (CGE), Autorité de Régulation des Communications Electroniques et des Postes (ARCEP), Agence du numérique. 2019 Nov 27. URL: <https://www.blogdumoderateur.com/barometre-numerique-2019/> [accessed 2021-04-10]
10. Torous J, Chan SR, Yee-Marie Tan S, Behrens J, Mathew I, Conrad EJ, et al. Patient Smartphone Ownership and Interest in Mobile Apps to Monitor Symptoms of Mental Health Conditions: A Survey in Four Geographically Distinct Psychiatric Clinics. *JMIR Ment Health* 2014 Dec;1(1):e5 [FREE Full text] [doi: [10.2196/mental.4004](https://doi.org/10.2196/mental.4004)] [Medline: [26543905](https://pubmed.ncbi.nlm.nih.gov/26543905/)]
11. Safi S, Danzer G, Schmailzl KJ. Empirical Research on Acceptance of Digital Technologies in Medicine Among Patients and Healthy Users: Questionnaire Study. *JMIR Hum Factors* 2019 Nov 29;6(4):e13472 [FREE Full text] [doi: [10.2196/13472](https://doi.org/10.2196/13472)] [Medline: [31782741](https://pubmed.ncbi.nlm.nih.gov/31782741/)]
12. Yarbrough AK, Smith TB. Technology acceptance among physicians: a new take on TAM. *Med Care Res Rev* 2007 Dec 05;64(6):650-672. [doi: [10.1177/1077558707305942](https://doi.org/10.1177/1077558707305942)] [Medline: [17717378](https://pubmed.ncbi.nlm.nih.gov/17717378/)]
13. Venkatesh V, Morris MG, Davis GB, Davis FD. User Acceptance of Information Technology: Toward a Unified View. *MIS Quarterly* 2003;27(3):425. [doi: [10.2307/30036540](https://doi.org/10.2307/30036540)]
14. Tonn P, Reuter SC, Kuchler I, Reinke B, Hinkelmann L, Stöckigt S, et al. Development of a Questionnaire to Measure the Attitudes of Laypeople, Physicians, and Psychotherapists Toward Telemedicine in Mental Health. *JMIR Ment Health* 2017 Oct 03;4(4):e39 [FREE Full text] [doi: [10.2196/mental.6802](https://doi.org/10.2196/mental.6802)] [Medline: [28974485](https://pubmed.ncbi.nlm.nih.gov/28974485/)]
15. Jean-Claude A. Pratiques Sociales et Représentations Social Practices and Representations. Paris, France: Presses Universitaires de France; 2011.
16. Jodelet D. Représentations sociales: un domaine en expansion. In: Jodelet D, editor. *Les Représentations Sociales*. Paris, France: Presses Universitaires de France; 2003:45-78.
17. Morgiève M, Sebbane D, De Rosario B, Demassiet V, Kabbaj S, Briffault X, et al. Analysis of the Recomposition of Norms and Representations in the Field of Psychiatry and Mental Health in the Age of Electronic Mental Health: Qualitative Study. *JMIR Ment Health* 2019 Oct 09;6(10):e11665 [FREE Full text] [doi: [10.2196/11665](https://doi.org/10.2196/11665)] [Medline: [31356151](https://pubmed.ncbi.nlm.nih.gov/31356151/)]
18. Wachelke J. Relationship between response evocation rank in social representations associative tasks and personal symbolic value. *Revue Internationale de Psychologie Sociale* 2008;21(3):113-126.
19. Bergamaschi A. Attitudes et représentations sociales. *ress* 2011 Dec 15(2):93-122 [FREE Full text] [doi: [10.4000/ress.996](https://doi.org/10.4000/ress.996)]
20. Fontaine S, Hamon JF. La représentation sociale de l'école des parents et des enseignants à La Réunion. *Les Cahiers Internationaux de Psychologie Sociale* 2010;85(1):69-109. [doi: [10.3917/cips.085.0069](https://doi.org/10.3917/cips.085.0069)]
21. Salès-Wuillemin E, Galand C, Cabello S, Folcher V. Validation d'un modèle tri-composantiel pour l'étude des représentations sociales à partir de mesures issues d'une tâche d'association verbale. *Les Cahiers Internationaux de Psychologie Sociale* 2011 Mar;91(3):231-252. [doi: [10.3917/cips.091.0231](https://doi.org/10.3917/cips.091.0231)]
22. Reinert M. Quel Objet Pour une Analyse Statistique du Discours?. 1998. URL: <http://lexicometrica.univ-paris3.fr/jadt/jadt1998/reinert.htm> [accessed 2021-04-10]
23. Fallery B, Rodhain F. Quatre approches pour l'analyse de données textuelles. 2007 Presented at: XVIème Conférence Internationale de Management Stratégique; June 6-9, 2007; Montreal, Canada p. 3-17.
24. Le Deuff O. La littératie digitale de santé: un domaine en émergence. *Les écosystèmes numériques et la démocratisation informationnelle : Intelligence collective, Développement durable, Interculturalité, Transfert de connaissances*. 2015 Nov. URL: <https://hal.univ-antilles.fr/hal-01258315/document> [accessed 2021-04-10]
25. Braslow JT, Brekke JS, Levenson J. Psychiatry's Myopia-Reclaiming the Social, Cultural, and Psychological in the Psychiatric Gaze. *JAMA Psychiatry* 2020 Sep 09;1. [doi: [10.1001/jamapsychiatry.2020.2722](https://doi.org/10.1001/jamapsychiatry.2020.2722)] [Medline: [32902598](https://pubmed.ncbi.nlm.nih.gov/32902598/)]
26. Bolton D. Clinical significance, disability, and biomarkers: Shifts in thinking between DSM-IV and DSM-5. In: Kendler KS, Parnas J, editors. *Philosophical Issues in Psychiatry IV: Classification of Psychiatric Illness*. Oxford, England: Oxford University Press; 2017.
27. Abbott A. *The System of Professions: An Essay on the Division of Expert Labor*. Chicago, IL: University of Chicago Press; 1988.
28. Fairhurst K, May C. Knowing patients and knowledge about patients: evidence of modes of reasoning in the consultation? *Fam Pract* 2001 Oct 01;18(5):501-505. [doi: [10.1093/fampra/18.5.501](https://doi.org/10.1093/fampra/18.5.501)] [Medline: [11604371](https://pubmed.ncbi.nlm.nih.gov/11604371/)]
29. Anderson RM, Heesterbeek H, Klinkenberg D, Hollingsworth TD. How will country-based mitigation measures influence the course of the COVID-19 epidemic? *The Lancet* 2020 Mar;395(10228):931-934 [FREE Full text] [doi: [10.1016/s0140-6736\(20\)30567-5](https://doi.org/10.1016/s0140-6736(20)30567-5)]
30. Foucault M, Fontana A, Gros F. *L'Herméneutique du sujet*. Cours au Collège de France. Paris, France: Seuil; 2001.
31. Henwood F, Marent B. Understanding digital health: Productive tensions at the intersection of sociology of health and science and technology studies. *Sociol Health Illn* 2019 Oct 10;41 Suppl 1(S1):1-15. [doi: [10.1111/1467-9566.12898](https://doi.org/10.1111/1467-9566.12898)] [Medline: [31599984](https://pubmed.ncbi.nlm.nih.gov/31599984/)]
32. Petersen A. *Digital health and technological promise: a sociological inquiry*. Newbury Park, CA: Routledge Publishing; 2019.
33. Health literacy: The solid facts. World Health Organization Regional Office for Europe. 2013. URL: <https://apps.who.int/iris/bitstream/handle/10665/128703/e96854.pdf> [accessed 2021-04-10]

Abbreviations

CFA: correspondence factor analysis
HDC: Hierarchical Descending Classification
ICT: information and communication technology
TAMs: Technology Acceptance Models
TDA: text data analysis
WHO: World Health Organization

Edited by J Torous; submitted 12.11.20; peer-reviewed by N Schulze, P Tonn; comments to author 08.01.21; revised version received 29.01.21; accepted 19.02.21; published 27.05.21.

Please cite as:

Morgiève M, Mesdjian P, Las Vergnas O, Bury P, Demassiet V, Roelandt JL, Sebbane D
Social Representations of e-Mental Health Among the Actors of the Health Care System: Free-Association Study
JMIR Ment Health 2021;8(5):e25708
URL: <https://mental.jmir.org/2021/5/e25708>
doi: [10.2196/25708](https://doi.org/10.2196/25708)
PMID: [34042591](https://pubmed.ncbi.nlm.nih.gov/34042591/)

©Margot Morgiève, Pierre Mesdjian, Olivier Las Vergnas, Patrick Bury, Vincent Demassiet, Jean-Luc Roelandt, Déborah Sebbane. Originally published in JMIR Mental Health (<https://mental.jmir.org>), 27.05.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Mental Health, is properly cited. The complete bibliographic information, a link to the original publication on <https://mental.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Problematic Social Media Use in Sexual and Gender Minority Young Adults: Observational Study

Erin A Vogel¹, PhD; Danielle E Ramo^{2,3}, PhD; Judith J Prochaska¹, MPH, PhD; Meredith C Meacham³, MPH, PhD; John F Layton³, BA; Gary L Humfleet³, PhD

¹Stanford Prevention Research Center, Department of Medicine, Stanford University, Stanford, CA, United States

²Hopelab, San Francisco, CA, United States

³Department of Psychiatry and Behavioral Sciences, University of California, San Francisco, CA, United States

Corresponding Author:

Erin A Vogel, PhD

Stanford Prevention Research Center

Department of Medicine

Stanford University

1265 Welch Road

X3C16

Stanford, CA, 943050000

United States

Phone: 1 6507243608

Email: eavogel@stanford.edu

Abstract

Background: Sexual and gender minority (SGM) individuals experience minority stress, especially when they lack social support. SGM young adults may turn to social media in search of a supportive community; however, social media use can become problematic when it interferes with functioning. Problematic social media use may be associated with experiences of minority stress among SGM young adults.

Objective: The objective of this study is to examine the associations among social media use, SGM-related internalized stigma, emotional social support, and depressive symptoms in SGM young adults.

Methods: Participants were SGM young adults who were regular (≥ 4 days per week) social media users ($N=302$) and had enrolled in Facebook smoking cessation interventions. As part of a baseline assessment, participants self-reported problematic social media use (characterized by salience, tolerance, and withdrawal-like experiences; adapted from the Facebook Addiction Scale), hours of social media use per week, internalized SGM stigma, perceived emotional social support, and depressive symptoms. Pearson correlations tested bivariate associations among problematic social media use, hours of social media use, internalized SGM stigma, perceived emotional social support, and depressive symptoms. Multiple linear regression examined the associations between the aforementioned variables and problematic social media use and was adjusted for gender identity.

Results: A total of 302 SGM young adults were included in the analyses (assigned female at birth: 218/302, 72.2%; non-Hispanic White: 188/302, 62.3%; age: mean 21.9 years, SD 2.2 years). The sexual identity composition of the sample was 59.3% (179/302) bisexual and/or pansexual, 17.2% (52/302) gay, 16.9% (51/302) lesbian, and 6.6% (20/302) other. The gender identity composition of the sample was 61.3% (185/302) cisgender; 24.2% (73/302) genderqueer, fluid, nonbinary, or other; and 14.6% (44/302) transgender. Problematic social media use averaged 2.53 (SD 0.94) on a 5-point scale, with a median of 17 hours of social media use per week (approximately 2.5 h per day). Participants with greater problematic social media use had greater internalized SGM stigma ($r=0.22$; $P<.001$) and depressive symptoms ($r=0.22$; $P<.001$) and lower perceived emotional social support ($r=-0.15$; $P=.007$). Greater internalized SGM stigma remained significantly associated with greater problematic social media use after accounting for the time spent on social media and other correlates ($P<.001$). In addition, participants with greater depressive symptoms had marginally greater problematic social media use ($P=.05$). In sum, signs of problematic social media use were more likely to occur among SGM young adults who had internalized SGM stigma and depressive symptoms.

Conclusions: Taken together, problematic social media use among SGM young adults was associated with negative psychological experiences, including internalized stigma, low social support, and depressive symptoms. SGM young adults experiencing minority stress may be at risk for problematic social media use.

KEYWORDS

sexual and gender minorities; social media; Facebook; internet; social stigma; mobile phone

Introduction

Background

Despite sweeping social change in many parts of the United States, many sexual and gender minority (SGM) youth and young adults face stigma, prejudice, and discrimination [1]. With social media's broad reach and its potential for supportive interactions, SGM individuals—especially SGM youth and young adults—may seek connection, validation, and community-building on social media. Indeed, SGM individuals use social media more frequently than the average US adult [2]. However, social media use can become problematic by interfering with functioning in daily life [3]. SGM young adults experiencing processes and effects of minority stress (ie, internalized stigma, depressive symptoms, and low emotional social support) may also be at risk for problematic social media use.

As a first step toward informing recommendations for health social media use for SGM young adults, it is important to identify those who may be at risk for problematic social media use (eg, preoccupation with social media use and difficulty with reducing or abstaining from use). In a sample of SGM young adults participating in clinical trials of 2 Facebook-delivered interventions for smoking cessation, this study examined the associations between problematic social media use and internalized SGM stigma, emotional social support, and depressive symptoms. We first review problematic social media use, including its theoretical underpinnings and correlates, among young adults. We then review the minority stress experienced by SGM young adults, social media use among SGM young adults, and how minority stress may influence the development of problematic social media use. Finally, we present potential correlates of problematic social media use in SGM young adults, informed by theoretical models of problematic social media use and minority stress.

Problematic Social Media Use

Social media use is almost ubiquitous among young adults, with 91% of young adults in the United States reporting the use of at least one social media platform in a 2019 survey [4]. In some social media users, habitual use leads to problematic use, characterized by interference with functioning, negative feelings when unable to access social media, inability to control one's social media use, and dominance of social media use over other thoughts and behaviors [5]. Young adults are particularly vulnerable to problematic social media use compared with older adults [5]. Numerous studies have examined the predisposing risk factors for problematic social media use among young adults. According to the Interaction of Person-Affect-Cognition-Execution (I-PACE) model, problematic social media use occurs when individuals with predisposing risk factors experience gratifications (eg, stress relief, positive mood) from social media use [6]. Such

gratifications reinforce the use of social media, which can then escalate until it becomes problematic (ie, becomes difficult to control and creates negative consequences) [6].

In the general young adult population, predisposing risk factors associated with vulnerability to problematic social media use include certain personality traits, mental health symptoms, and experiences. First, high extraversion [3,7,8] and low conscientiousness [3,7-9] are consistently associated with problematic social media use across studies. Social media offers an outlet for connecting with others in a way that requires little planning and can be an escape from other daily tasks [7]. Second, young adults who experience depressive symptoms are at risk of problematic social media use. A nationally representative study of young adults in the United States found that problematic social media use was associated with a 9% increase in the odds of experiencing depressive symptoms [10]. Individuals with depressive symptoms may use social media to relieve stress and seek social support, thereby developing problematic social media use patterns [11]. Third, young adults who frequently experience fear of missing out (FOMO) are more likely to exhibit problematic social media use [8,12-14]. FOMO refers to an uncomfortable feeling that others are having positive experiences in one's absence and is associated with higher engagement in social media use [15]. In line with the I-PACE model, these factors may interact to influence problematic social media use. For example, depressive symptoms may interfere with an individual's ability to socialize and enjoy experiences, thereby producing FOMO. Consistent with the I-PACE model, a three-wave longitudinal study found that Chinese young adults who experienced greater depressive symptoms subsequently experienced more FOMO, followed by more problematic smartphone use [16]. In sum, predisposing factors to problematic social media use in the general young adult population include personality traits, mental health concerns, and social cognition.

Minority Stress Among SGM Young Adults

SGM young adults face unique challenges and may have different predispositions than their non-SGM peers because of their experiences of minority stress. The minority stress model, as applied to sexual and gender minorities, posits that being a minority leads to chronic experiences of stigma, prejudice, and discrimination as well as more proximal stress processes, including vigilance toward future rejection, the concealment of one's sexual or gender identity, and the internalization of negative societal attitudes toward SGM individuals (ie, internalized stigma) [17]. Both proximal and distal stress processes can lead to negative mental health outcomes, including depressive symptoms [17-19]. Receiving adequate social support may ameliorate the negative effects of minority stress on SGM youth and young adults' depressive symptoms [20]. In other words, SGM individuals who experience negative effects of minority stress may report internalized SGM stigma, depressive symptoms, and low emotional social support. These processes

and outcomes of minority stress may contribute to problematic social media use, especially if SGM individuals turn to social media for social support but still feel unsupported.

Social Media Use Among Sexual and Gender Minorities

Social media may help undersupported individuals compensate for their lack of in-person emotional social support [21]. In a qualitative study, transgender adolescents reported using social media to obtain emotional support and inspiration from other transgender individuals and to obtain information about gender-affirming therapy [22]. Alternatively, social media use may have an overall negative effect on the mental health of SGM young adults. Transgender adolescents also report seeing hurtful content related to their gender identity [22]. SGM social media users face the unique challenge of managing disclosure of their SGM identity on social media, where different members of their social networks (eg, friends, family, and coworkers) may be able to see their posts [23]. A systematic review of the association between social media use and depressive symptoms among SGM individuals concluded that social media had both positive effects (eg, connection with the SGM community and self-expression) and negative effects (eg, being a victim of cyberbullying and experiencing the stress of hiding one's SGM identity from certain others) [24]. A study conducted in China found that among SGM individuals aged 13-58 years, problematic social media use was associated with greater depressive symptoms, time spent on social media, and web-based social support seeking [25]. However, these specific relationships have not yet been studied among young adults in the United States, and there may be important cross-cultural differences in social media use and in the social experience of identifying as SGM. In addition, Han et al [25] did not measure offline emotional social support; therefore, it is unclear whether SGM individuals who lacked offline emotional social support engaged in more problematic social media use.

This Research and Hypotheses

SGM young adults are vulnerable to internalized stigma, low emotional social support, and depressive symptoms as part of minority stress. The I-PACE model suggests that these predispositions may lead to problematic social media use in response to stress. Although research has examined predisposing factors for problematic social media use in the general young adult population, the effects and processes of minority stress as potential predispositions for problematic social media use in SGM young adults have not been explored.

This study is largely exploratory, testing the associations between problematic social media use and individual differences (informed by the I-PACE and minority stress models) in a convenience sample of SGM young adults enrolled in clinical trials of 2 smoking cessation interventions delivered on Facebook. Nonetheless, the extant literature and the minority stress model provided a framework for testing relevant factors that may predispose SGM young adults to problematic social media use. Specifically, we predicted that problematic social media use would be associated with greater internalized SGM stigma, lesser emotional social support, and greater depressive symptoms. Problematic social media use was expected to be

positively associated, but not synonymous, with the time spent on social media [25-27]. Overall, this study assessed the relative strength of the associations among problematic social media use, time spent on social media, internalized SGM stigma, social support, and depressive symptoms, in a cross-sectional study of SGM young adults enrolled in 2 smoking cessation intervention trials on Facebook.

Methods

Participants and Procedures

Data were taken from the baseline assessments of 2 randomized controlled trials that tested the efficacy of SGM-tailored Facebook smoking cessation interventions compared with similar, nontailored interventions. Participants (N=302) were aged between 18 and 25 years, identifying as SGM (ie, nonheterosexual and/or noncisgender), having smoked at least 100 cigarettes in their lifetime, using Facebook for at least 4 days per week, living in the United States, and were well-versed in English. Participants in both trials were recruited between April and December 2018 with an SGM-tailored Facebook advertising campaign [28]. Clicking on an ad directed the participants to a confidential screener. After verifying their age and identity, eligible participants received a baseline web-based survey. Relevant measures from the baseline web-based survey are described in the *Measures* section. Informed consent was obtained from all participants, and all research activities were approved by the University of California, San Francisco, Institutional Review Board.

Measures

Problematic Social Media Use

Problematic social media use was measured using the Bergen Social Media Addiction Scale [29]. Participants rated, on a 1-5 Likert-type scale, the degree to which 6 statements about their social media use were true for them (1=very rarely; 2=rarely; 3=sometimes; 4=often; 5=very often). The possible range of scores was 1-5. Sample items included "You spend a lot of time thinking about social media or planning how to use it" and "You become restless or troubled if you are prohibited from using social media."

Time Spent on Social Media

Participants answered, "Approximately how many hours per week do you spend using social media?" The item was a free response; the natural log of responses was used in analyses.

Frequency of Social Media Use

A total of 9 items assessed the frequency of use of 8 social media platforms (Facebook, Instagram, Snapchat, Twitter, Pinterest, LinkedIn, Reddit, and Tumblr), plus a write-in "other" option, using 1-5 Likert-type scales (1=never; 2=monthly; 3=weekly; 4=once a day; 5=multiple times a day). The possible scores for each platform ranged from 1-5.

Internalized SGM Stigma

To measure internalized SGM stigma, participants completed the Revised Internalized Homophobia Scale, a 5-item measure adapted for all SGM identities [30] by rating their agreement

with 5 statements on a 1-5 Likert-type scale (1=strongly disagree; 2=disagree; 3=neutral; 4=agree; 5=strongly agree), with possible average scores ranging from 1-5. Sample items included “I wish I weren’t LGBTQ+ (lesbian, gay, bisexual, transgender, queer+)” and “I have tried to stop being LGBTQ+.”

Perceived Emotional Social Support

The 8-item Emotional Support subscale of the National Institutes of Health (NIH) Toolbox for the assessment of neurological and behavioral functions measured the perceived emotional social support [31]. Sample items include, “I have someone who understands my problems” and “I feel there are people I can talk to if I am upset” rated on a 1-5 Likert-type scale (1=never; 5=always), with total scores ranging from 8-40.

Depressive Symptoms

Depressive symptoms were measured using the 2-item Patient Health Questionnaire, a validated screener assessing the frequency of depressed mood and anhedonia in the past 2 weeks, with possible scores ranging from 0-6 [32]. The 2-item Patient Health Questionnaire is sensitive to changes in depressive symptoms and has been found to detect symptom improvement during depression treatment [33].

Demographic Characteristics

The participants reported their age and sex assigned at birth (male or female). Participants selected all applicable terms to describe their gender identity (male/man, female/woman, trans male/trans man, trans female/trans woman, genderqueer/gender nonconforming, genderfluid, agender, nonbinary, or different identity), sexual identity (straight [heterosexual], lesbian/gay [homosexual], bisexual, pansexual, or not listed), and race and ethnicity (American Indian/Alaska Native, Asian, Black, Hispanic, Pacific Islander/Native Hawaiian, White, or other). Household income was presented in 7 ranges from “less than US \$20,000” to “over US \$200,000.” Subjective social status was measured on a 1-10 scale from “worst off” to “best off” compared with others in the United States and in their communities [34]. Years of education were measured with “How many total years of school have you completed?”

Statistical Analysis

The analyses were conducted in 3 steps. First, we examined the distributions and reliability of the measures before creating composite scores. Second, we examined bivariate Pearson correlations between problematic social media use, social media use in hours per week, internalized SGM stigma, social support, and depressive symptoms. Third, significant correlates of problematic social media use identified in bivariate screening were entered into a multiple linear regression analysis as independent variables, with problematic social media use as the outcome dependent variable. As gender identity was significantly associated with problematic social media use, we adjusted for gender identity by entering it in step 1. The main correlates of interest were entered as independent variables in step 2. In the final model, gender identity (control variable) was entered in step 1; time spent on social media, internalized SGM stigma, emotional social support, and depressive symptoms (independent variables) were entered in step 2, and problematic social media use was entered as the dependent variable.

Results

Sample Characteristics

The sample characteristics are listed in Table 1. Most participants (218/302, 72.2%) were assigned female at birth; however, they varied in their current gender identity and sexual identity. The sample comprised 61.3% (185/302) cisgender participants; 24.2% (73/302) genderqueer, fluid, and/or nonbinary participants; and 14.6% (44/302) transgender participants. Most participants were bisexual and/or pansexual (179/302, 59.3%), followed by gay (52/302, 17.2%), lesbian (51/302, 16.9%), or other (20/302, 6.6%) participants. The sample majority (188/302, 62.3%) was non-Hispanic White. Participants perceived themselves as below average in social standing compared with others in the United States (mean 3.91, SD 1.61) and their communities (mean 4.32, SD 1.96), where the scales ranged from 1-10.

Table 1. Participant characteristics (N=302).

Characteristics	Value
Age (years), mean (SD)	21.9 (2.2)
Sex assigned at birth (female), n%	218 (72.2)
Gender identity^a, n (%)	
Cisgender	185 (61.3)
Transgender	44 (14.6)
Genderqueer, fluid, and/or nonbinary	73 (24.2)
Sexual identity^b, n (%)	
Gay	52 (17.2)
Lesbian	51 (16.9)
Bisexual/pansexual	179 (59.3)
Other	20 (6.6)
Race and ethnicity, n (%)	
Non-Hispanic White	188 (62.3)
Native American/Alaskan Native	3 (1)
African American/Black	3 (1)
Asian	10 (3.3)
Hispanic/Latinx	29 (9.6)
Pacific Islander/Hawaiian Native	2 (0.7)
More than one race and ethnicity	59 (19.5)
Other	8 (2.6)
Household income, n (%)	
Less than US \$20,000	75 (24.8)
US \$21,000-US \$60,000	155 (51.3)
US \$61,000-US \$100,000	52 (17.2)
More than US \$100,000	20 (6.6)
Subjective social status (out of 10), mean (SD)	
Compared with others in the United States	3.9 (1.6)
Compared with others in your community	4.3 (2.0)
Years of education, mean (SD)	13.5 (2.1)

^aParticipants selected all applicable terms from straight (heterosexual), lesbian/gay (homosexual), bisexual, pansexual, and not listed (please specify). Sexual identity was recoded into the categories presented here. When multiple identities were selected, “Gay/lesbian” took priority in coding, followed by “bisexual/pansexual” then “other” [28].

^bParticipants selected all applicable terms from male/man, female/woman, trans male/trans man, trans female/trans woman, genderqueer/gender nonconforming, genderfluid, agender, nonbinary, or different identity (please specify). Gender identity was recoded into the categories presented here. When multiple identities were selected, “transgender” took priority, followed by “cisgender” then “other” [28].

Distributions of Measures

The distributions of the key measures are presented in Table 2. Composite scores were computed using the mean of the internalized homophobia items ($\alpha=.74$), sum of social support

items ($\alpha=.96$), sum of depressive symptoms items ($r=.075$), and mean of problematic social media use items ($\alpha=.84$), according to the recommended scoring guidelines for each measure. Owing to a skewed distribution, the natural log of social media use frequency was used in the analyses.

Table 2. Distributions of key measures.

Measure	Respondents, n (%)	Mean (SD) or median (IQR)	Observed range	Possible range	Scale
Problematic social media use (score)	302 (100)	2.5 (0.9)	1-5	1-5	1=very rarely; 2=rarely; 3=sometimes; 4=often; 5=very often
Social media (hours per week)	300 (99.3)	17.0 (10-30)	1-120	≥0	Measured continuously
Natural log (Ln) of hours per week	300 (99.3)	2.8 (0.7)	0.7-4.8	≥0	Measured continuously
Frequency of platform use (score)					1=never; 2=monthly; 3=weekly; 4=once/day; 5=multiple times/day
Facebook	302 (100)	4.7 (0.6)	2-5	1-5	
Instagram	301 (99.7)	4.0 (1.3)	1-5	1-5	
Snapchat	300 (99.3)	3.7 (1.4)	1-5	1-5	
Twitter	297 (98.3)	2.5 (1.6)	1-5	1-5	
Pinterest	292 (96.7)	1.8 (1.1)	1-5	1-5	
LinkedIn	293 (97)	1.3 (0.7)	1-4	1-5	
Reddit	292 (96.7)	1.9 (1.3)	1-5	1-5	
Tumblr	297 (98.3)	2.6 (1.5)	1-5	1-5	
Other	237 (78.5)	1.2 (0.7)	1-5	1-5	
Internalized SGM ^a stigma (score)	299 (99)	1.7 (0.8)	1-5	1-5	1=strongly disagree; 2=disagree; 3=neutral; 4=agree; 5=strongly agree
Emotional social support (score)	302 (100)	30.7 (6.9)	11-40	8-40	1=never; 2=rarely; 3=sometimes; 4=usually; 5=always (sum of 8 items)
Depressive symptoms (score)	302 (100)	3.5 (1.8)	0-6	0-6	0=not at all and 3=nearly every day (sum of 2 items)

^aSGM: sexual and gender minority.

On average, problematic social media use was around the midpoint of the 1-5 scale (mean 2.5, SD 0.9). Median use of social media was 17 hours per week (an average of approximately 2.5 h per day). The 3 most frequently used social media platforms were Facebook, Instagram, and Snapchat. Internalized SGM stigma scores, which can range from 1-5, averaged 1.7 (SD 0.8), which was comparable with previous research (mean 1.5, SD 0.7 [30]). Perceived emotional social support was moderate (mean 30.7, SD 6.9), on a scale from 8-40; depressive symptoms averaged 3.5 (SD 1.8) on a scale from 0-6.

Bivariate Correlates of Problematic Social Media Use

Problematic social media use was significantly greater among participants who spent more time on social media ($r=0.24$; $P<.001$). Participants with greater problematic social media use also reported greater internalized stigma about their SGM identities ($r=0.22$; $P<.001$), perceived less emotional social support ($r=-0.16$; $P=.007$), and had greater depressive symptoms ($r=0.22$; $P<.001$) compared with those with lower problematic social media use. Participants who spent more time on social media had significantly greater depressive symptoms ($r=0.15$; $P=.009$); however, time spent on social media was not associated with internalized SGM stigma ($r=-0.05$; $P=.41$) or emotional social support ($r=0.02$; $P=.78$). The correlations between all of the measures are presented in Table 3.

Table 3. Pearson correlations between key variables (N=302)^a.

Measure	Problematic social media use	Social media hours per week (natural log [Ln])	Internalized sexual and gender minority stigma	Emotional social support	Depressive symptoms
Problematic social media use					
<i>r</i>	1	0.24 ^b	0.22 ^b	−0.16 ^c	0.22 ^b
<i>P</i> value	— ^d	<.001	<.001	.007	<.001
Social media hours per week (natural log [Ln])					
<i>r</i>	0.24 ^b	1	−0.05	0.02	0.15 ^c
<i>P</i> value	<.001	—	.41	.78	.009
Internalized sexual and gender minority stigma					
<i>r</i>	0.22 ^b	−0.05	1	−0.23 ^b	0.13 ^e
<i>P</i> value	<.001	.41	—	<.001	.02
Emotional social support					
<i>r</i>	−0.16 ^c	0.02	−0.23 ^b	1	−0.28 ^b
<i>P</i> value	.007	.78	<.001	—	<.001
Depressive symptoms					
<i>r</i>	−0.22 ^b	0.15 ^c	0.13 ^e	−0.28 ^b	1
<i>P</i> value	<.001	.009	.02	<.001	—

^an=300 for social media hours per week; n=299 for internalized sexual and gender minority stigma.

^bSignificant at the $P<.001$ level.

^cSignificant at the $P<.01$ level.

^dNot applicable.

^eSignificant at the $P<.05$ level.

Multivariable Model of Problematic Social Media Use

Time spent on social media ($\beta=.24$; $P<.001$) and internalized SGM stigma ($\beta=.20$; $P<.001$) remained significantly and positively associated with problematic social media use in the

adjusted multiple linear regression analysis (presented in [Table 4](#)). The positive association between depressive symptoms and problematic social media use was marginally significant in the adjusted multiple linear regression analysis ($\beta=.11$; $P=.05$).

Table 4. Adjusted multiple linear regression analysis with problematic social media use as the outcome variable.

Variable	β	<i>P</i> value
Step 1		
Gender identity		
Transgender (reference: cisgender)	.04	.48
Nonbinary and/or other gender (reference: cisgender)	.18	.002
Step 2		
Gender identity		
Transgender (reference: cisgender)	.02	.66
Nonbinary and/or other gender (reference: cisgender)	.15	.009
Social media (hours per week; natural log [Ln])	.24	<.001
Internalized stigma	.20	<.001
Emotional social support	−.06	.28
Depressive symptoms	.11	.05

Discussion

Principal Findings

In this study of SGM young adults enrolled in 2 Facebook smoking cessation intervention trials, those with greater problematic social media use had greater internalized SGM stigma and depressive symptoms and lower perceived emotional social support. Greater internalized SGM stigma remained significantly associated with greater problematic social media use after accounting for the time spent on social media and other correlates. In addition, participants with greater depressive symptoms had marginally greater problematic social media use. Overall, signs of problematic social media use were more likely to occur among SGM young adults who had internalized SGM stigma and depressive symptoms.

Comparison With Prior Work

These results are consistent with our hypothesis that SGM young adults with greater internalized SGM stigma would be more likely to exhibit problematic social media use than those with lower internalized SGM stigma. According to the I-PACE model, personal characteristics can predispose individuals to problematic social media use. Internalized SGM stigma may be one such characteristic for SGM young adults. Individuals with poor self-concepts often seek social support on social media, and these attempts are often unsuccessful. Specifically, previous research found that adults with low self-esteem primarily posted negative thoughts [35] and sought reassurance [36]. Such posts are less likely to garner social support than positive posts [37]. Similarly, SGM young adults who feel marginalized and different from their peers may find that these feelings are exacerbated by the positive self-presentation bias on social media [38]. Seeing others appear to be happy and comfortable with themselves may be harmful to SGM young adults who internalize negative messages about their sexual and gender identities. Therefore, problematic social media use may contribute to further internalization of stigma. The I-PACE model posits bidirectional relationships between predispositions to problematic social media use and social media use behavior. Although this cross-sectional analysis cannot determine causality, the results are consistent with the I-PACE model and suggest that internalized SGM stigma may be a predisposing factor for problematic social media use.

SGM young adults with low emotional social support were more likely to have problematic social media use in bivariate analyses, thereby corroborating the associations between low social support and problematic social media use found in the general population [39] and partially supporting our hypothesis. The association did not remain significant in the multivariable model, suggesting that internalized SGM stigma accounted for more variance in problematic social media use than did emotional social support. SGM young adults who experienced lower emotional social support also reported greater internalized SGM stigma. According to the minority stress model, internalized stigma and low emotional social support can result from experiencing prejudice and discrimination [18]. Given the frequency with which SGM individuals experience prejudice and discrimination [1], assessing SGM individuals' emotional

support is critical. Low emotional social support may be an additional risk factor for problematic social media use, albeit a weaker predictor than internalized SGM stigma.

Problematic social media use was associated with depressive symptoms in bivariate analyses, partially supporting our hypothesis. This is consistent with the I-PACE model and with extant research examining the link between depressive symptoms and problematic social media use in the general population [11,40].

Consistent with extant literature involving young adult social media users [41,42] and SGM social media users in China [25], SGM young adults in this study who spent more time on social media experienced more depressive symptoms. Frequent social media use may lead to depressive symptoms through social mechanisms such as FOMO [43]. Experiencing depression may also increase the risk of frequent social media use, as browsing social media may provide a temporary escape that requires little energy [27]. SGM individuals are at a high risk of depression [17]. Problematic social media use should be assessed as both a contributing factor to an individual's depressive symptoms and the result of depressive symptoms.

This pattern of results suggests multiple predisposing factors for problematic social media use among SGM young adults, especially internalized SGM stigma. Social media may appeal to young adults with poor self-concepts and insufficient emotional support, as social media presents opportunities to connect with others. However, social media may not be an adequate source of support and connection for SGM young adults with poor self-views. Our results suggest that SGM young adults who used social media intensively, to the point of problematic use, still had high internalized SGM stigma, high depressive symptoms, and low emotional support. Moreover, SGM young adults with problematic social media use may withdraw from offline social opportunities and thus receive less support. Importantly, the relationship between problematic social media use and social support was significant only in bivariate analyses, suggesting that internalized SGM stigma is more strongly associated with problematic use. Problematic social media use was also associated with depressive symptoms in bivariate analyses. Seeking social support and not receiving it may exacerbate depressive symptoms [44].

Conversely, spending more time on social media was not associated with internalized stigma or social support. SGM young adults who experience distress around their identities may be more vulnerable to problematic social media use, regardless of how frequently they use social media. Consistent with previous research, time spent on social media and problematic social media use were modestly correlated but distinct [25-27]. Spending more time on social media did not appear to account for the relationships between problematic social media use, internalized SGM stigma, and depressive symptoms. SGM young adults who spent more time on social media but did not experience impairment or distress from their social media use may have used social media differently than those with greater problematic use. Problematic social media use may be a stronger indicator or consequence of underlying

concerns (eg, internalized stigma and low social support) than the time spent on social media.

Implications

This study supports and extends prior research on SGM individuals' social media use by demonstrating associations between problematic use and common challenges faced by SGM young adults. When used in moderation, social media offers benefits such as self-expression and connection with others [24]. Although this study cannot confirm the direction of the associations among problematic social media use, internalized SGM stigma, low social support, and depressive symptoms, the results suggest that intense social media use may not meet the psychosocial needs of SGM young adults. Importantly, the time spent on social media was not significantly associated with internalized SGM stigma ($P=.411$) or low emotional social support ($P=.775$). SGM young adults with problematic social media use may represent a high-risk subset of social media users who are likely to experience social and mental health challenges. SGM young adults are vulnerable to problematic social media use, internalized stigma, low emotional support, and depressive symptoms. Clinicians working with SGM young adults should discuss social media use with clients and explore how social media may reflect or affect clients' mental health. Social media may be an effective intervention platform for helping young adults improve their well-being and engage in health-promoting behaviors (eg, smoking cessation [45]).

Limitations

First, this cross-sectional, correlational study design did not confirm causal inferences about the role of social media use in mental and physical health outcomes. Longitudinal data would enable the use of more sophisticated statistical modeling to understand the interplay among experiences of stigma, social support, and depressive symptoms. This study underscores the

need for future research on the social media use of SGM young adults. Second, all participants were young adults who self-identified as SGM and smoked cigarettes. Other dimensions of sexual orientation (ie, sexual attraction and sexual behavior) were not assessed. SGM young adults who smoke may be a particularly vulnerable subset of SGM young adults, and it is unclear whether the results can be generalized to SGM individuals who do not smoke. Future research should aim to oversample non-White young adults and those assigned male at birth, as the sample in this study was mainly White (188/302, 62.3%) and assigned female at birth (218/302, 72.2%). Third, all participants were frequent Facebook users. Future research could aim to recruit SGM young adults from other social media platforms. Finally, this study was a secondary analysis that captured only a few potential correlates of problematic social media use. Measures of some constructs (eg, depressive symptoms and time spent on social media) were brief mental health symptoms aside from depression (eg, anxiety symptoms) were not measured, and "social media" was not defined. Future research could use more comprehensive measures to further our understanding of problematic social media use among SGM young adults.

Conclusions

Although social media use can be beneficial for SGM young adults [24], it may not be an adequate source of social support and can become problematic. Problematic social media use is distinct from the time spent on social media and involves feelings of distress when disconnected from social media [5]. SGM young adults experiencing minority stress may be at risk for problematic social media use. Taken together, these results suggest that problematic social media use among SGM young adults is associated with internalized stigma, low social support, and depressive symptoms.

Acknowledgments

This work was supported by the National Institute on Minority Health and Health Disparities of the NIH (grant R21 MD011765), the National Institute on Drug Abuse of the NIH (grant K01 DA046697), and the Tobacco-Related Disease Research Program (grants 26IR-0004 and 28FT-0015). Sponsors had no role in the study design; collection, analysis, or interpretation of data; writing of the report; or decision to submit the paper for publication. The authors have no relevant financial interests to disclose.

Authors' Contributions

EAV was involved in the conceptualization, formal analysis, investigation, and writing of the original draft. Investigation; writing, reviewing, and editing; supervision; and funding acquisition was conducted by DER. JJP and MCM were involved in conceptualization; investigation; and writing, reviewing, and editing. Conceptualization; investigation; data curation; and writing, reviewing, and editing was done by JFL. GLH was involved in conceptualization; investigation; writing, reviewing, and editing; and supervision.

Conflicts of Interest

None declared.

References

1. Burton CM, Marshal MP, Chisolm DJ, Sucato GS, Friedman MS. Sexual minority-related victimization as a mediator of mental health disparities in sexual minority youth: a longitudinal analysis. *J Youth Adolesc* 2013 Mar 5;42(3):394-402 [[FREE Full text](#)] [doi: [10.1007/s10964-012-9901-5](https://doi.org/10.1007/s10964-012-9901-5)] [Medline: [23292751](https://pubmed.ncbi.nlm.nih.gov/23292751/)]

2. A survey of LGBT Americans: attitudes, experiences, and values in changing times. Pew Research Center. 2013. URL: https://www.pewsocialtrends.org/wp-content/uploads/sites/3/2013/06/SDT_LGBT-Americans_06-2013.pdf [accessed 2021-05-15]
3. Andreassen CS, Torsheim T, Brunborg GS, Pallesen S. Development of a Facebook addiction scale. *Psychol Rep* 2012 Apr 01;110(2):501-517. [doi: [10.2466/02.09.18.pr0.110.2.501-517](https://doi.org/10.2466/02.09.18.pr0.110.2.501-517)]
4. Perrin A, Anderson M. Share of U.S. adults using social media, including Facebook, is mostly unchanged since. Pew Research Center. 2019. URL: <https://www.pewresearch.org/fact-tank/2019/04/10/share-of-u-s-adults-using-social-media-including-facebook-is-mostly-unchanged-since-2018/> [accessed 2021-05-15]
5. Turel O, Serenko A. The benefits and dangers of enjoyment with social networking websites. *Eur J Inf Sys* 2017 Dec 19;21(5):512-528. [doi: [10.1057/ejis.2012.1](https://doi.org/10.1057/ejis.2012.1)]
6. Brand M, Young KS, Laier C, Wölfling K, Potenza MN. Integrating psychological and neurobiological considerations regarding the development and maintenance of specific Internet-use disorders: an Interaction of Person-Affect-Cognition-Execution (I-PACE) model. *Neurosci Biobehav Rev* 2016 Dec;71:252-266 [FREE Full text] [doi: [10.1016/j.neubiorev.2016.08.033](https://doi.org/10.1016/j.neubiorev.2016.08.033)] [Medline: [27590829](https://pubmed.ncbi.nlm.nih.gov/27590829/)]
7. Wilson K, Fornasier S, White KM. Psychological predictors of young adults' use of social networking sites. *Cyberpsychol Behav Soc Netw* 2010 Apr;13(2):173-177. [doi: [10.1089/cyber.2009.0094](https://doi.org/10.1089/cyber.2009.0094)] [Medline: [20528274](https://pubmed.ncbi.nlm.nih.gov/20528274/)]
8. Moore K, Craciun G. Fear of missing out and personality as predictors of social networking sites usage: the Instagram case. *Psychol Rep* 2020 Jul 12:A. [doi: [10.1177/0033294120936184](https://doi.org/10.1177/0033294120936184)] [Medline: [32659168](https://pubmed.ncbi.nlm.nih.gov/32659168/)]
9. Kircaburun K, Griffiths MD. Instagram addiction and the Big Five of personality: the mediating role of self-liking. *J Behav Addict* 2018 Mar 01;7(1):158-170 [FREE Full text] [doi: [10.1556/2006.7.2018.15](https://doi.org/10.1556/2006.7.2018.15)] [Medline: [29461086](https://pubmed.ncbi.nlm.nih.gov/29461086/)]
10. Shensa A, Escobar-Viera CG, Sidani JE, Bowman ND, Marshal MP, Primack BA. Problematic social media use and depressive symptoms among U.S. young adults: a nationally-representative study. *Soc Sci Med* 2017 Jun;182:150-157 [FREE Full text] [doi: [10.1016/j.socscimed.2017.03.061](https://doi.org/10.1016/j.socscimed.2017.03.061)] [Medline: [28446367](https://pubmed.ncbi.nlm.nih.gov/28446367/)]
11. Brailovskaia J, Velten J, Margaf J. Relationship between daily stress, depression symptoms, and Facebook addiction disorder in Germany and in the United States. *Cyberpsychol Behav Soc Netw* 2019 Sep 01;22(9):610-614. [doi: [10.1089/cyber.2019.0165](https://doi.org/10.1089/cyber.2019.0165)] [Medline: [31397593](https://pubmed.ncbi.nlm.nih.gov/31397593/)]
12. Dempsey AE, O'Brien KD, Tiamiyu MF, Elhai JD. Fear of missing out (FoMO) and rumination mediate relations between social anxiety and problematic Facebook use. *Addict Behav Rep* 2019 Jun;9:A [FREE Full text] [doi: [10.1016/j.abrep.2018.100150](https://doi.org/10.1016/j.abrep.2018.100150)] [Medline: [31193746](https://pubmed.ncbi.nlm.nih.gov/31193746/)]
13. Liu C, Ma J. Adult attachment orientations and social networking site addiction: the mediating effects of online social support and the fear of missing out. *Front Psychol* 2019 Nov 26;10:A [FREE Full text] [doi: [10.3389/fpsyg.2019.02629](https://doi.org/10.3389/fpsyg.2019.02629)] [Medline: [32038342](https://pubmed.ncbi.nlm.nih.gov/32038342/)]
14. Sheldon P, Antony MG, Sykes B. Predictors of problematic social media use: personality and life-position indicators. *Psychol Rep* 2021 Jun 24;124(3):1110-1133. [doi: [10.1177/0033294120934706](https://doi.org/10.1177/0033294120934706)] [Medline: [32580682](https://pubmed.ncbi.nlm.nih.gov/32580682/)]
15. Przybylski AK, Murayama K, DeHaan CR, Gladwell V. Motivational, emotional, and behavioral correlates of fear of missing out. *Comput Hum Behav* 2013 Jul;29(4):1841-1848. [doi: [10.1016/j.chb.2013.02.014](https://doi.org/10.1016/j.chb.2013.02.014)]
16. Yuan G, Elhai JD, Hall BJ. The influence of depressive symptoms and fear of missing out on severity of problematic smartphone use and Internet gaming disorder among Chinese young adults: a three-wave mediation model. *Addict Behav* 2021 Jan;112:A. [doi: [10.1016/j.addbeh.2020.106648](https://doi.org/10.1016/j.addbeh.2020.106648)] [Medline: [32977268](https://pubmed.ncbi.nlm.nih.gov/32977268/)]
17. Meyer IH. Prejudice, social stress, and mental health in lesbian, gay, and bisexual populations: conceptual issues and research evidence. *Psychol Bull* 2003 Sep;129(5):674-697 [FREE Full text] [doi: [10.1037/0033-2909.129.5.674](https://doi.org/10.1037/0033-2909.129.5.674)] [Medline: [12956539](https://pubmed.ncbi.nlm.nih.gov/12956539/)]
18. Meyer I, Dean L. Internalized homophobia, intimacy, sexual behavior among gay/bisexual men. In: Herek GM, editor. *Stigma and Sexual Orientation: Understanding Prejudice Against Lesbians, Gay Men, and Bisexuals*. Thousand Oaks, CA: Sage Publications Inc; 1997:160-186.
19. Gonzales G, Henning-Smith C. Health disparities by sexual orientation: results and implications from the behavioral risk factor surveillance system. *J Community Health* 2017 Dec 2;42(6):1163-1172. [doi: [10.1007/s10900-017-0366-z](https://doi.org/10.1007/s10900-017-0366-z)] [Medline: [28466199](https://pubmed.ncbi.nlm.nih.gov/28466199/)]
20. McConnell EA, Birkett MA, Mustanski B. Typologies of social support and associations with mental health outcomes among LGBT youth. *LGBT Health* 2015 Mar;2(1):55-61 [FREE Full text] [doi: [10.1089/lgbt.2014.0051](https://doi.org/10.1089/lgbt.2014.0051)] [Medline: [26790019](https://pubmed.ncbi.nlm.nih.gov/26790019/)]
21. Rains SA, Tsetsi E. Social support and digital inequality: does internet use magnify or mitigate traditional inequities in support availability? *Commun Monogr* 2016 Sep 09;84(1):54-74. [doi: [10.1080/03637751.2016.1228252](https://doi.org/10.1080/03637751.2016.1228252)]
22. Selkie E, Adkins V, Masters E, Bajpai A, Shumer D. Transgender adolescents' uses of social media for social support. *J Adolesc Health* 2020 Mar;66(3):275-280. [doi: [10.1016/j.jadohealth.2019.08.011](https://doi.org/10.1016/j.jadohealth.2019.08.011)] [Medline: [31690534](https://pubmed.ncbi.nlm.nih.gov/31690534/)]
23. McConnell EA, Clifford A, Korpak AK, Phillips G, Birkett M. Identity, victimization, and support: Facebook experiences and mental health among LGBTQ youth. *Comput Human Behav* 2017 Nov;76:237-244 [FREE Full text] [doi: [10.1016/j.chb.2017.07.026](https://doi.org/10.1016/j.chb.2017.07.026)] [Medline: [29225412](https://pubmed.ncbi.nlm.nih.gov/29225412/)]

24. Escobar-Viera CG, Whitfield DL, Wessel CB, Shensa A, Sidani JE, Brown AL, et al. For better or for worse? A systematic review of the evidence on social media use and depression among lesbian, gay, and bisexual minorities. *JMIR Ment Health* 2018 Jul 23;5(3):e10496 [FREE Full text] [doi: [10.2196/10496](https://doi.org/10.2196/10496)] [Medline: [30037786](https://pubmed.ncbi.nlm.nih.gov/30037786/)]
25. Han X, Han W, Qu J, Li B, Zhu Q. What happens online stays online? — Social media dependency, online support behavior and offline effects for LGBT. *Comput Hum Behav* 2019 Apr;93:91-98. [doi: [10.1016/j.chb.2018.12.011](https://doi.org/10.1016/j.chb.2018.12.011)]
26. Kanat-Maymon Y, Almog L, Cohen R, Amichai-Hamburger Y. Contingent self-worth and Facebook addiction. *Comput Hum Behav* 2018 Nov;88:227-235. [doi: [10.1016/j.chb.2018.07.011](https://doi.org/10.1016/j.chb.2018.07.011)]
27. Brailovskaia J, Margraf J, Köllner V. Addicted to Facebook? Relationship between Facebook Addiction Disorder, duration of Facebook use and narcissism in an inpatient sample. *Psychiatry Res* 2019 Mar;273:52-57. [doi: [10.1016/j.psychres.2019.01.016](https://doi.org/10.1016/j.psychres.2019.01.016)] [Medline: [30639564](https://pubmed.ncbi.nlm.nih.gov/30639564/)]
28. Vogel EA, Belohlavek A, Prochaska JJ, Ramo DE. Development and acceptability testing of a Facebook smoking cessation intervention for sexual and gender minority young adults. *Internet Interv* 2019 Mar;15:87-92 [FREE Full text] [doi: [10.1016/j.invent.2019.01.002](https://doi.org/10.1016/j.invent.2019.01.002)] [Medline: [30792958](https://pubmed.ncbi.nlm.nih.gov/30792958/)]
29. Andreassen CS, Billieux J, Griffiths MD, Kuss DJ, Demetrovics Z, Mazzoni E, et al. The relationship between addictive use of social media and video games and symptoms of psychiatric disorders: a large-scale cross-sectional study. *Psychol Addict Behav* 2016 Mar;30(2):252-262. [doi: [10.1037/adb0000160](https://doi.org/10.1037/adb0000160)] [Medline: [26999354](https://pubmed.ncbi.nlm.nih.gov/26999354/)]
30. Herek GM, Gillis JR, Cogan JC. Internalized stigma among sexual minority adults: insights from a social psychological perspective. *J Couns Psycho* 2009 Jan;56(1):32-43. [doi: [10.1037/a0014672](https://doi.org/10.1037/a0014672)]
31. Cyranowski JM, Zill N, Bode R, Butt Z, Kelly MA, Pilkonis PA, et al. Assessing social support, companionship, and distress: National Institute of Health (NIH) Toolbox Adult Social Relationship Scales. *Health Psychol* 2013 Mar;32(3):293-301 [FREE Full text] [doi: [10.1037/a0028586](https://doi.org/10.1037/a0028586)] [Medline: [23437856](https://pubmed.ncbi.nlm.nih.gov/23437856/)]
32. Kroenke K, Spitzer RL, Williams JB. 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](https://doi.org/10.1097/01.MLR.0000093487.78664.3C)] [Medline: [14583691](https://pubmed.ncbi.nlm.nih.gov/14583691/)]
33. 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](https://doi.org/10.1016/j.genhosppsych.2018.11.003)] [Medline: [30508772](https://pubmed.ncbi.nlm.nih.gov/30508772/)]
34. Singh-Manoux A, Adler NE, Marmot MG. Subjective social status: its determinants and its association with measures of ill-health in the Whitehall II study. *Soc Sci Med* 2003 Mar;56(6):1321-1333. [doi: [10.1016/s0277-9536\(02\)00131-4](https://doi.org/10.1016/s0277-9536(02)00131-4)]
35. Forest AL, Wood JV. When social networking is not working: individuals with low self-esteem recognize but do not reap the benefits of self-disclosure on Facebook. *Psychol Sci* 2012 Mar 07;23(3):295-302. [doi: [10.1177/0956797611429709](https://doi.org/10.1177/0956797611429709)] [Medline: [22318997](https://pubmed.ncbi.nlm.nih.gov/22318997/)]
36. Clerkin EM, Smith AR, Hames JL. The interpersonal effects of Facebook reassurance seeking. *J Affect Disord* 2013 Nov;151(2):525-530. [doi: [10.1016/j.jad.2013.06.038](https://doi.org/10.1016/j.jad.2013.06.038)] [Medline: [23850160](https://pubmed.ncbi.nlm.nih.gov/23850160/)]
37. Vogel EA, Rose JP, Crane C. "Transformation Tuesday": temporal context and post valence influence the provision of social support on social media. *J Soc Psychol* 2018 Nov 10;158(4):446-459 [FREE Full text] [doi: [10.1080/00224545.2017.1385444](https://doi.org/10.1080/00224545.2017.1385444)] [Medline: [29023225](https://pubmed.ncbi.nlm.nih.gov/29023225/)]
38. Vogel EA, Rose JP. Self-reflection and interpersonal connection: making the most of self-presentation on social media. *Transl Issues Psychol Sci* 2016 Sep;2(3):294-302. [doi: [10.1037/tps0000076](https://doi.org/10.1037/tps0000076)]
39. Brailovskaia J, Rohmann E, Bierhoff H, Schillack H, Margraf J. The relationship between daily stress, social support and Facebook Addiction Disorder. *Psychiatry Res* 2019 Jun;276:167-174. [doi: [10.1016/j.psychres.2019.05.014](https://doi.org/10.1016/j.psychres.2019.05.014)] [Medline: [31096147](https://pubmed.ncbi.nlm.nih.gov/31096147/)]
40. Shensa A, Sidani JE, Escobar-Viera CG, Switzer GE, Primack BA, Choukas-Bradley S. Emotional support from social media and face-to-face relationships: associations with depression risk among young adults. *J Affect Disord* 2020 Jan 01;260:38-44 [FREE Full text] [doi: [10.1016/j.jad.2019.08.092](https://doi.org/10.1016/j.jad.2019.08.092)] [Medline: [31493637](https://pubmed.ncbi.nlm.nih.gov/31493637/)]
41. Steers MN, Wickham RE, Acitelli LK. Seeing everyone else's highlight reels: how Facebook usage is linked to depressive symptoms. *J Soc Clin Psychol* 2014 Oct;33(8):701-731. [doi: [10.1521/jscp.2014.33.8.701](https://doi.org/10.1521/jscp.2014.33.8.701)]
42. Feinstein BA, Hershenberg R, Bhatia V, Latack JA, Meuwly N, Davila J. Negative social comparison on Facebook and depressive symptoms: rumination as a mechanism. *Psychol Pop Media Cult* 2013 Jul;2(3):161-170. [doi: [10.1037/a0033111](https://doi.org/10.1037/a0033111)]
43. Baker ZG, Krieger H, LeRoy AS. Fear of missing out: relationships with depression, mindfulness, and physical symptoms. *Transl Issues Psychol Sci* 2016 Sep;2(3):275-282. [doi: [10.1037/tps0000075](https://doi.org/10.1037/tps0000075)]
44. Frison E, Eggermont S. The impact of daily stress on adolescents' depressed mood: the role of social support seeking through Facebook. *Comput Hum Behav* 2015 Mar;44:315-325. [doi: [10.1016/j.chb.2014.11.070](https://doi.org/10.1016/j.chb.2014.11.070)]
45. Vogel E, Ramo D, Meacham M, Prochaska J, Delucchi K, Humfleet G. The Put It Out Project (POP) Facebook intervention for young sexual and gender minority smokers: outcomes of a pilot, randomized, controlled trial. *Nicotine Tob Res* 2020 Aug 24;22(9):1614-1621 [FREE Full text] [doi: [10.1093/ntr/ntz184](https://doi.org/10.1093/ntr/ntz184)] [Medline: [31562765](https://pubmed.ncbi.nlm.nih.gov/31562765/)]

Abbreviations

FOMO: fear of missing out

I-PACE: Interaction of Person-Affect-Cognition-Execution

LGBTQ+: lesbian, gay, bisexual, transgender, queer+

NIH: National Institutes of Health

SGM: sexual and gender minority

Edited by J Torous; submitted 19.08.20; peer-reviewed by C Escobar-Viera, M Gansner; comments to author 01.10.20; revised version received 23.11.20; accepted 18.12.20; published 28.05.21.

Please cite as:

Vogel EA, Ramo DE, Prochaska JJ, Meacham MC, Layton JF, Humfleet GL

Problematic Social Media Use in Sexual and Gender Minority Young Adults: Observational Study

JMIR Ment Health 2021;8(5):e23688

URL: <https://mental.jmir.org/2021/5/e23688>

doi: [10.2196/23688](https://doi.org/10.2196/23688)

PMID:

©Erin A Vogel, Danielle E Ramo, Judith J Prochaska, Meredith C Meacham, John F Layton, Gary L Humfleet. Originally published in JMIR Mental Health (<https://mental.jmir.org>), 28.05.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Mental Health, is properly cited. The complete bibliographic information, a link to the original publication on <https://mental.jmir.org/>, as well as this copyright and license information must be included.

Original Paper

Knowledge-Infused Abstractive Summarization of Clinical Diagnostic Interviews: Framework Development Study

Gaur Manas¹, MSc; Vamsi Aribandi², BSc; Ugur Kursuncu¹, PhD; Amanuel Alambo², BSc, MSc; Valerie L Shalin³, PhD; Krishnaprasad Thirunarayan², PhD; Jonathan Beich⁴, MD, MSEE; Meera Narasimhan⁵, MD; Amit Sheth¹, PhD

¹Artificial Intelligence Institute, University of South Carolina, Columbia, SC, United States

²Kno.e.sis Center, Department of Computer Science and Engineering, Wright State University, Dayton, OH, United States

³Department of Psychology, Wright State University, Dayton, OH, United States

⁴Department of Psychiatry, Wright State University, Dayton, OH, United States

⁵Department of Neuropsychiatry & Behavioral Science, School of Medicine, Prisma Health, University of South Carolina, Columbia, SC, United States

Corresponding Author:

Gaur Manas, MSc
Artificial Intelligence Institute
University of South Carolina
Room 513, 1112 Greene St
Columbia, SC, 29208
United States
Phone: 1 5593879476
Email: mgaur@email.sc.edu

Abstract

Background: In clinical diagnostic interviews, mental health professionals (MHPs) implement a care practice that involves asking open questions (eg, “What do you want from your life?” “What have you tried before to bring change in your life?”) while listening empathetically to patients. During these interviews, MHPs attempted to build a trusting human-centered relationship while collecting data necessary for professional medical and psychiatric care. Often, because of the social stigma of mental health disorders, patient discomfort in discussing their presenting problem may add additional complexities and nuances to the language they use, that is, hidden signals among noisy content. Therefore, a focused, well-formed, and elaborative summary of clinical interviews is critical to MHPs in making informed decisions by enabling a more profound exploration of a patient’s behavior, especially when it endangers life.

Objective: The aim of this study is to propose an unsupervised, knowledge-infused abstractive summarization (KiAS) approach that generates summaries to enable MHPs to perform a well-informed follow-up with patients to improve the existing summarization methods built on frequency heuristics by creating more informative summaries.

Methods: Our approach incorporated domain knowledge from the Patient Health Questionnaire-9 lexicon into an integer linear programming framework that optimizes linguistic quality and informativeness. We used 3 baseline approaches: extractive summarization using the SumBasic algorithm, abstractive summarization using integer linear programming without the infusion of knowledge, and abstraction over extractive summarization to evaluate the performance of KiAS. The capability of KiAS on the Distress Analysis Interview Corpus-Wizard of Oz data set was demonstrated through interpretable qualitative and quantitative evaluations.

Results: KiAS generates summaries (7 sentences on average) that capture informative questions and responses exchanged during long (58 sentences on average), ambiguous, and sparse clinical diagnostic interviews. The summaries generated using KiAS improved upon the 3 baselines by 23.3%, 4.4%, 2.5%, and 2.2% for thematic overlap, Flesch Reading Ease, contextual similarity, and Jensen Shannon divergence, respectively. On the Recall-Oriented Understudy for Gisting Evaluation-2 and Recall-Oriented Understudy for Gisting Evaluation-L metrics, KiAS showed an improvement of 61% and 49%, respectively. We validated the quality of the generated summaries through visual inspection and substantial interrater agreement from MHPs.

Conclusions: Our collaborator MHPs observed the potential utility and significant impact of KiAS in leveraging valuable but voluminous communications that take place outside of normally scheduled clinical appointments. This study shows promise in generating semantically relevant summaries that will help MHPs make informed decisions about patient status.

KEYWORDS

knowledge-infusion; abstractive summarization; distress clinical diagnostic interviews; Patient Health Questionnaire-9; healthcare informatics; interpretable evaluations

Introduction

Background

The diagnosis of mental illness is unique to medicine. Although other specialties can rely on physical examinations, imaging, and laboratory tests for diagnosis and ongoing assessment, psychiatry often relies on only a patient's narrative. An accurate assessment, diagnosis, and treatment hinge on the ability of a trained mental health professional (MHP) to elicit not only information but also subtle indicators of human emotions that portend clues to severely life-threatening situations [1]. Although MHPs might find a "second set of eyes or ears" valuable, it is generally impractical and costly to hire additional personnel for this purpose. The shortage of qualified MHPs and the increasing amount of clinical data dictate novel approaches in the diagnosis and treatment processes. Summarizing patients' relevant electronic health records, including clinical diagnostic interview logs between clinicians and patients, has emerged as a novel method. Simultaneously, the techniques and tools require rigorous evaluation by domain experts [2,3]. As accuracy and false-negative rate are crucial metrics for the success of the deployment of such tools, we leveraged knowledge-infused learning to achieve this goal, as described in recent studies [4-9]. Sheth et al [6,8] define knowledge-infused learning as "the exploitation of domain knowledge and application semantics to enhance existing artificial intelligence methods by infusing relevant conceptual information into a statistical and data-driven computational approach," which in this study is *integer linear programming* (ILP). A paper on "knowledge infusion" from Valiant et al [10] theoretically assesses the importance of teaching materials (eg, lexicons) in reducing prediction errors and making the model robust. This study, theoretically, quantitatively, and qualitatively evaluates the knowledge-infusion paradigm in improving the outcomes of recent artificial intelligence algorithms, specifically in the context of deep-learning algorithms, as defined in Sheth et al [11,12] on knowledge-infused learning and for achieving explainability.

Building upon the recent efforts in knowledge-infused learning, we propose an end-to-end summarization framework, called KiAS (knowledge-infused abstractive summarization) for clinical diagnostic interviews, using domain knowledge derived from the Patient Health Questionnaire-9 (PHQ-9). Informative summaries capture insightful questions from the interviewer and relevant patient responses that best express intent and expected behavior. This system concisely encapsulates major themes and clarifies patient concerns toward the goal of focused treatment. In addition, summaries for individual patients will provide a new type of historical record of MHP-patient interactions that was not previously possible. This allows for more quantifiable measures for the program of a patient of mental health. We validated our approach by using the *Distress*

Analysis Interview Corpus-Wizard of Oz (DAIC-WoZ) data set comprising recorded interviews between a patient (participant) and a computerized animated virtual interviewer *Ellie* [13]. Our MHP coauthors analyzed the clinical intricacies in the data set (described in the *Data Set and Analysis* section). Our analysis of the corpus revealed a key finding: a clinical interview's patient response is not always specific to the previous question. Instead, a meaningful response from a patient may be semantically linked to other earlier questions. Furthermore, the patient's responses can be ambiguous, be redundant, and, with distant anaphora, challenge the summarization task.

Previous work has attempted to summarize structured interviews or meeting logs where every response to a question is known to be informative and nonredundant [14]. The dialogues in clinical interviews are ambiguous and, following a simple filtering or preprocessing process, leads to the loss of critical pieces of information that might be relevant to MHPs. Therefore, in our task, we did not consider redundant responses or ambiguous responses as noise. Our aim is to prefer recall over precision because MHPs will eventually decide the action plan and reduce their manual labor.

Abstraction-based summarization is generally more complicated than extractive summarization (ES), as it involves context understanding, content organization, rephrasing, and relevance-based matching of sentences to form coherent summaries. In addition, the challenges in the problem domain increase the complexity of the abstractive summarization (AS) task. Previous studies by Clarke and Lapata [15,16] provided the first attempt for AS using sentence compression techniques (eg, tree based [17] and sentence based [eg, lexicalization or markovization]). However, these approaches rely on syntactic parsing using a part-of-speech tagger, which relies on quality annotation, a process that is both knowledge-intensive and time-consuming [18]. Instead, a direct word graph (WG)-based approach was used employing TextRank to generate compressed sentences by finding the k-shortest paths [19,20]. Filippova used TextRank over LexRank [21] because of the use of a cosine similarity, which is not semantically preserved, a property required in meaningful summarization. However, linguistic quality is sacrificed while improving the informativeness of the summaries. Banerjee et al [22] and Nayeem et al [23] developed an AS scheme using a skip-gram word-embedding model and ILP to summarize multiple documents. Our proposed method optimizes the grammaticality and informativeness constrained by the length of the summaries [22,23].

Recently, researchers have employed a neural network-based framework to address the summarization problem. Li et al [24] described an attention-based encoder and decoder component to find words or phrases that would guide ILP procedures to generate concise and meaningful summaries.

See et al [25] summarized news articles using a supervised sequence-to-sequence approach that uses human-written summaries to learn model parameters. They evaluated their approach using the CNN or Daily Mail data set, which contains news articles and their corresponding human-written abstracts. This method outperformed the state-of-the-art solution by at least two Recall-Oriented Understudy for Gisting Evaluation (ROUGE) points [26]. However, it is not appropriate for summarizing dialogues in diagnostic interviews in an unsupervised setting.

Shang et al [27] designed an unsupervised, end-to-end AS architecture to summarize the meeting speech. However, a meeting structure is different from a clinical interview structure, mostly because it is centered around diagnosing patients. Redundancy in a corpus of interviews is semantic and not merely lexical. An MHP might ask the following 2 paraphrased questions during an interview: (Q1) “Have you been diagnosed with clinical depression?” or (Q2) “You are showing signs of clinical depression; have you seen an MHP before?” Whereas such paraphrasing is absent in a meeting, these questions require domain knowledge to calculate their semantic proximity. Therefore, the approach proposed by Shang et al [27] does not apply to our problem.

Furthermore, the study by Wang et al [28] composed word embedding and word frequency to calculate a *word attraction force score* to integrate previous knowledge. Such word-based models seldom capture implicit semantics (eg, “difficulty in sleeping at night” can allow an MHP to infer insomnia) in complex discourse such as clinical interviews. Identifying relevant phrases (eg, “feeling hopeless”) and characterizing patient behavior is essential for an MHP to subsequently make informed decisions. In addition, word-embedding models (eg, BERT or Word2Vec) do not conflate and provide robustness against lexical ambiguity [29]. Furthermore, the scarcity of clinical diagnostic interviews restricts the training of problem-specific word-embedding models.

MacAvaney et al [30] designed a method to summarize radiology reports using a neural sequence-to-sequence architecture, infusing previous knowledge in the summarization process through ontology. They compared the use of medical ontology (Quick Unified Medical Language System [QuickUMLS]) with domain-specific ontology (RadLex). They evaluated their approach using human-written summaries and acknowledged that AS methods generate readable summaries. For a comprehensive evaluation, human evaluators employed the criteria of readability, accuracy, and completeness. The current state-of-the-art deep-learning architecture, BART (Bidirectional and Auto-Regressive Transformers) [31], comprises a bidirectional encoder over the document to be summarized and an autoregressive decoder over reference summaries (RS). The model is trained using cosine-similarity loss, resulting in a dense representation, which generates summaries using beam search. The method is useful in question answering, reading comprehension, and summarization with gold standard summaries [32,33]. The complexity of the model requires considerable training time over a sizable domain-specific corpus. Unfortunately, most of the previous works on AS have been developed and tested on benchmark

data sets on news articles and meeting notes, which differ significantly from clinical interviews. Furthermore, domain knowledge inclusion is critical for associating a meaningful response to the relevant question asked by an MHP, which is challenging for current deep-learning architectures for summarization.

To leverage the benefits of domain knowledge, we used the PHQ-9 lexicon to incorporate relevant concepts in a machine-processable manner. Yazdavar et al [34] built a depression lexicon from the established clinical assessment questionnaire, PHQ-9. They divided the lexicon into 9 signals as per PHQ-9, denoted using indicative phrases as follows: decreased pleasure in most activities (S1), feeling down (S2), sleep disorders (S3), loss of energy (S4), a significant change in appetite (S5), feeling worthless (S6), concentration problems (S7), hyper or lower activity (S8), and suicidal thoughts (S9). Karmen et al [35] also proposed a lexicon for depression symptoms but did not categorize them. Neuman et al [36] crawled the web for metaphorical and nonmetaphorical relations that embed the word *depression*, which added noise to the lexicon because of its polysemous nature (eg, great depression, depressing incident, depressing movie, and economic depression). Furthermore, these lexicons captured *transient sadness* instead of clinical symptoms influencing the diagnosis of *major depressive disorder*. The lexicon created by Yazdavar et al [34] was chosen to conduct this study based on the shortcomings of the alternatives. Alhanai et al [37] used the DAIC-WoZ data set to predict the depression of an individual leveraging multimodal data. The study used audio, visual, and textual data recorded during an interview between a patient and a virtual clinician to train 2 sequence models for detecting depression. Furthermore, the data set was annotated with labels that help train supervised models for detecting depression. However, the DAIC-WoZ data set does not have ground truth summaries that support our research.

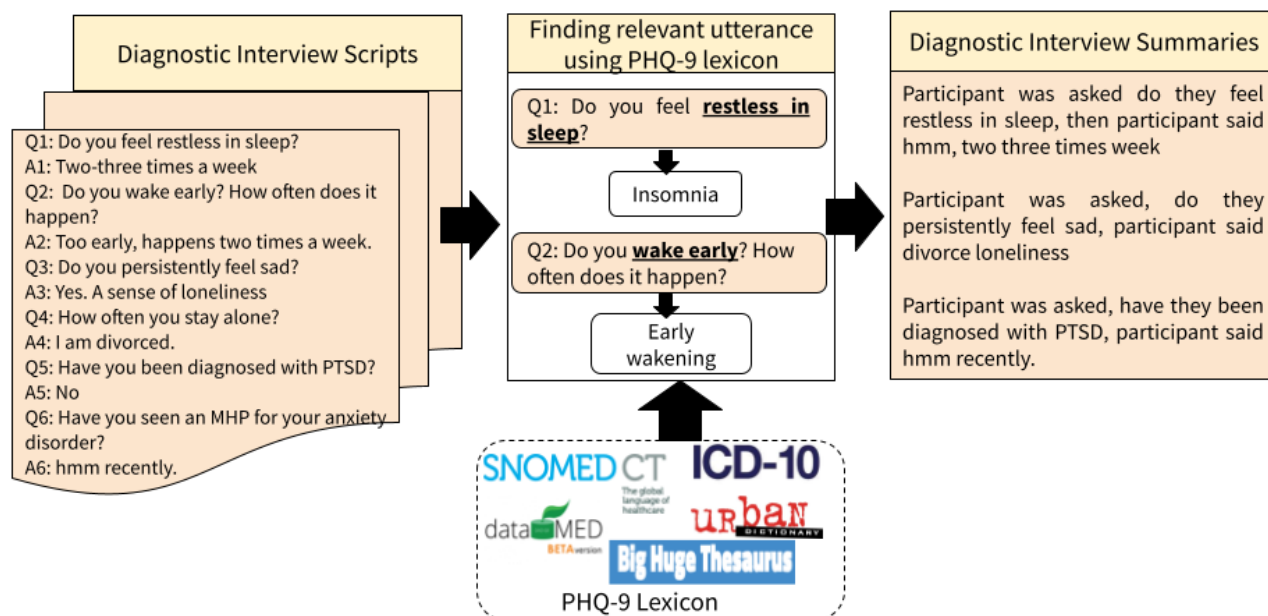
Objective

We seek to improve upon these existing approaches by capturing the critical nuances of clinical diagnostic interviews that will lead to a more robust analysis, paraphrasing, and reorganizing the source content [38]. Our approach uses an ILP framework, exploiting linguistic quality constraints to capture end-user information needs. Furthermore, existing methods seldom leverage domain-specific information to generate summaries by filtering out the noninformative utterances of a patient. For example, the utterance “uh well most recently we went to Israel for a pilgrimage” is informative in everyday conversation but not to an MHP. Furthermore, an isolated utterance “I have trouble sleeping” might not be considered informative by a purely statistical algorithm but is essential information for an MHP. We leveraged a semantic lexicon from a recent study by Yazdavar et al [34] to filter out irrelevant utterances. The lexicon was used to retrofit (contextualize) ConceptNet word embedding and improve the informativeness of summaries [39]. For instance, in Figure 1, the patient replied “No” to the question “Have you been diagnosed with PTSD?” However, the patient replied “hmm recently” to the question “Have you seen an MHP for your anxiety disorder?” During the flattening of the hierarchy of Systematized Nomenclature of Medicine-Clinical Terms

(SNOMED-CT) medical knowledge in the semantic lexicon, we observed anxiety disorder (SNOMED-CT ID: 197480006) and posttraumatic stress disorder (PTSD; SNOMED-CT ID: 47505003) were associated with parent-child relationships. Thus, in the generated summary, for the question “Have you

been diagnosed with PTSD?” the deduced response is “hmm recently,” which could be a limitation of the summarizer; however, the generated summaries are intended for further scrutiny from MHPs. The proposed approach is illustrated in Figure 1.

Figure 1. Overview of knowledge-infused abstractive summarization for an interview snippet of a patient's responses to the question asked by Ellie (virtual interviewer). Phrases relevant to mental health are identified using the PHQ-9 lexicon. The contextual similarity between utterances is calculated through a retrofitted embedding model. The resulting summaries contain relevant questions and meaningful responses. ICD-10: International Statistical Classification of Diseases, Tenth Revision; MHP: mental health professional; PHQ-9: Patient Health Questionnaire-9; PTSD: posttraumatic stress disorder; SNOMED CT: SNOMED Clinical Terms.



A total of 4 professors assessed a set of summaries in neuropsychiatry specializing in rehabilitation counseling and mental health assessment. We avoided crowdsourcing of workers for qualitative evaluation because of issues described in a study by Chandler et al [40]. The key contributions of this study are four-fold:

- We developed an unsupervised framework to generate readable and informative summaries of clinical diagnostic interviews. The proposed KiAS framework does not only rely on target summaries but also uses domain-specific knowledge based on PHQ-9 [4]. The summarization task was formulated to optimize both linguistic quality and informativeness.
- The PHQ-9 lexicon knowledge is infused into KiAS through a trigram language model (LM) and a retrofitted ConceptNet embedding [5,6].
- A quantitative evaluation to assess the efficacy of our approach is based on contextual similarity, Jensen Shannon Divergence (JSD; information entropy), thematic overlap, and readability. As the data set lacks ground truth handwritten summaries, such metrics would evaluate coherence, content-preserving nature, and summaries' understandability.
- A qualitative evaluation was used to assess the usability of the summaries based on context-specific questions and meaningful responses.

The infusion of distress-related knowledge through language modeling and embedding methods has considerably improved

the quality of summaries over comparable baselines. Furthermore, KiAS summarizes long conversations (mean 2061, SD 813 words) in 7-8 sentences (103 words on par with an SD of 52 words) and has the potential to reduce the follow-up time from 3 patients to 5-6 patients in 7 days.

Methods

Data Set and Analysis

We used the DAIC-WoZ data set, which consists of clinical interviews to support the diagnosis of psychological disorders such as anxiety, depression, and PTSD [13]. Data were collected via Wizard-of-Oz interviews, conducted by an animated virtual interviewer called Ellie, and controlled by a human interviewer. It contains data from 189 patient interviews, including transcripts, audio recordings, video recordings, and PHQ-8 depression questionnaire responses [41]. For this study, we used only the transcripts. We did not employ existing annotations in the data set in our unsupervised approach. Incomplete words were replaced with their complete version (eg, a vague word such as *peop* is turned into its full form *people*), and unrecognizable words were annotated as *xxx*. The interviews were generally 7-33 minutes long, with an average length of 16 minutes and 58 statements. The data set was anonymized to comply with privacy and ethical standards. From 189 interviews, 5 were excluded from the study either because of imperfections in data collection or transcription (interruptions during the interview or missing transcription of the interviewer's utterances). Note that we are not predicting depression; instead,

we retrieve utterances that help an MHP make inferences relevant to depression, PTSD, or stress.

Our collaborator, clinical psychiatrists, observed that questions related to mental health conditions varied based on their sequence. The clinical diagnostic interviews were open ended and semistructured (eg, with a clinician improvising questions based on an earlier response of a patient) where items transition from clinically irrelevant to clinically relevant (Textbox 1). The interview started with neutral questions to build a rapport with the participant before moving on to inquiries related to depression, PTSD, or mental health. Finally, it ends with a *cool-down* phase to ensure that the participant leaves the interview in a peaceful state of mind. The DAIC-WoZ interviews were designed to diagnose mental health conditions and their severity by assessing the level of interference in a patient's life. However, much of the meeting is designed to relax the participant by engaging them in neutral conversations. Utterances in a neutral conversation do not provide clinicians with insights about the mental condition; therefore, they can be discarded, and we call this operation *pruning*. Identifying *revealing and diagnostic* utterances is challenging because, without background knowledge, such critical utterances can be

missed by statistical methods that are based on frequency counts. For example, if a patient says “nobody likes me” once, a purely statistical approach may not consider this utterance relevant. However, using background knowledge (in our case, the depression lexicon), we can identify such valuable signals. Similarly, the utterance “I was raised in New York, but I live in L.A. now” is irrelevant to a clinician but may be identified as relevant by a domain-agnostic tool. To prune noisy questions or answers and focus on mental health-related issues, we used the PHQ-9 lexicon for semantic pattern-matching of phrases in a conversation. Specifically, semantic pattern-matching captures PHQ-8 responses implicit in the DAIC-WoZ corpus of interviews [42]. On the other hand, previous work on ES relies on the intrinsic capability to remove noisy utterances from conversations instead of a domain-specific lexicon [43] (we provided our evaluation for comparison in the *Results* section). Finally, we converted the questions and their answers from filtered conversations into one combined statement. For example, the question asked: “How long ago were you diagnosed;” patient's response: “a few years ago”; converted statement: “participant was asked how long ago they were diagnosed,” and the participant replied “a few years ago.”

Textbox 1. Example questions from the interviewer at the start of the interview (top). These questions are casual to make the patient comfortable. Over time, questions become specific to the patient's condition and behavior (middle). At the end of the interview, the questions become less subjective and target the patient's life (bottom).

<p>Start of interview</p> <ul style="list-style-type: none"> • “Okay what'd you study at school” • “That's good where are you from originally” • “How do you like L.A.” <p>Middle of interview</p> <ul style="list-style-type: none"> • “Is there anything you regret” • “Do you feel down” • “Have you been diagnosed with posttraumatic stress disorder” <p>End of interview</p> <ul style="list-style-type: none"> • “Okay when was the last time you felt really happy” • “Cool how would your best friend describe you” • “who's someone that's been a positive influence in your life”
--

We converted the question and answer (Q and A) pairs into sentences to preserve the inherent structure of clinical diagnostic interviews, which is described by the sequence of utterances by the interviewer and the interviewee. For instance, the interviewer asks the same question multiple times by rephrasing it to derive meaningful responses from the patient. Therefore, the answer to a question at the beginning of the interview can be found at the end or later parts of the interview (ie, anaphora [44]). To address this issue, we converted questions and answers into a template “participant was asked X, the participant said Y” so that TextRank can measure the statistical relevance of the response to the question (if the answer is contextually irrelevant to the question, the TextRank algorithm fails to generate a cohesive graph; rather, it generates disjoint graphs—one for the

question, one for the answer—with no common node). Although we optimize using ILP over such disjoint structures, the informativeness score is still very low, as distances are large (a constraint that we minimize in ILP). Therefore, it is essential to develop a summarization model to recognize this structure.

We recognize that, in some cases, this structure might generate longer statements; however, the statement conveys minimal information. There are various forms of AS, which include or exclude paraphrasing, depending on the problem. For this study, we built and improved upon the past work on AS from Banerjee et al [38], Filippova [19], and Tuan et al [44]. In our qualitative evaluation guided by our domain expert, we specifically focused on the ability of the model to select good questions and meaningful responses.

Baseline Summarization Methodologies

We considered 3 baselines for comparison with the proposed approach. First, ES generates human-readable summaries. The greedy nature of ES identifies meaningful utterances from the interview scripts. Second, we use AS, which brings coherence to the summaries. Third, we hybridized *abstraction over extractive summarization* (AoES) to leverage the advantages of ES to improve AS summary quality.

Extractive Summarization

ES generates a subset of sentences from the input document of the corpus as a summary. This approach can be likened to a condensation of the source document according to the *what you see is what you get* paradigm. ES techniques generally guarantee the linguistic quality of the generated summaries, whereas they are not abstractive enough to mirror manual summaries. We used the SumBasic (SB) algorithm to perform ES over the interview transcripts [45]. SB selects important sentences based on word probabilities ($P(w_i)$) [46]. For a given input sentence (S_j), the sentence-importance weight is computed using Equation 1:

$$w_j = \frac{c(w_i)}{N}$$

where $c(w_i)$ is the number of occurrences w_i in the input sentence, and N is the total number of words in the input sentence ($N \gg S_j$). A greedy selection strategy of SB selects the sentence that contains words with the highest probability. This selection strategy embodies the intuition that the words with the highest probabilities represent the document's most important topics. After selecting a sentence, the word probabilities in the selected sentence are updated by squaring word probabilities before the sentence was selected. Such an SB update prevents the selections of the same or similar sentences multiple times, thus creating a diverse sentence summary. However, ES-generated summaries, although short, lack readability and informativeness. These are critical criteria in making understandable abstracts of clinical interviews where understandability takes precedence over summary length.

Abstractive Summarization

Typically, in a clinical diagnostic interview, the patient's response aligns better with the question asked earlier in the interview. Therefore, to obtain a meaningful summary, it is

necessary to consider the informativeness of a response to the question that has been asked earlier. Relating the answer to the most relevant question improves the cohesiveness of the summaries. This thematic rephrasing aspect makes AS superior to ES [47,48]. We implemented the AS method using the ILP optimization framework to optimize linguistic quality and informativeness by leveraging a generic LM [22]. This study contrasts with existing studies investigating supervised AS algorithms leveraging sequence-to-sequence architectures because we do not rely on human-written summaries [49]. A caveat in the use of AS in an unsupervised setting is its susceptibility to generate summaries with inaccurate information because the constraints do not consider domain knowledge. To minimize the impact of unsupervised learning, particularly in the mental health domain, constraints can be modeled using medical domain knowledge (eg, UMLS [Unified Medical Language System] and International Classification of Disease, 10th Edition) to generate outcomes that facilitate reliable decisions to develop our proposed approach, KiAS.

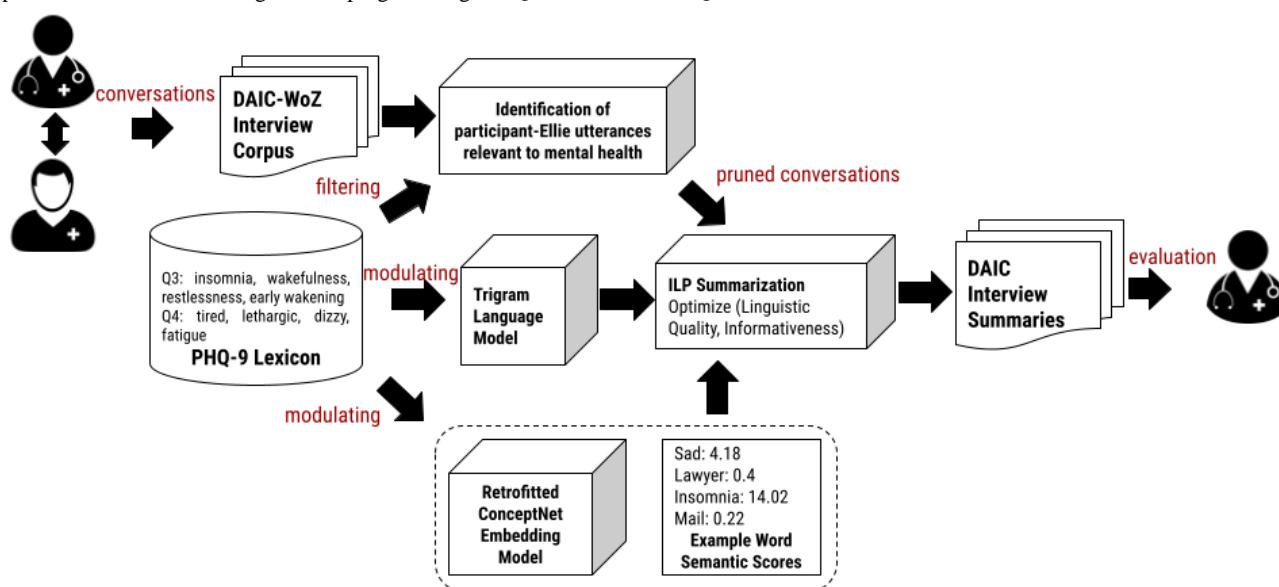
Abstraction Over ES

AoES uses ES as a prefilter so that AS can generate high-quality summaries. However, AoES fails to focus on the domain-specific verbiage implicit in the conversation. For example, in the following conversation piece, "participant was asked what is going on with you, a participant said i am sick and tired of losses," the italicized phrases are essential to an MHP but occur with low frequency and thus get removed by AoES. Another example, "participant was asked that's good where you are from originally, the participant said originally I'm from glendale california," was identified as relevant by AoES. In contrast, it was filtered out using the proposed approach. Although the location in which people live could influence their mental health [50], it is less of a concern in clinical diagnostic interviews, which are face-to-face.

Proposed Approach: KiAS

KiAS has 4 key steps: (1) creation of pruned conversations, (2) tuning generic LM, (3) retrofitting the concept net embedding model, and (4) creating abstractive summaries in an unsupervised manner. Figure 2 illustrates the proposed framework of KiAS for generating knowledge-aware summaries from real-world, simulated clinical diagnostic interviews in the DAIC-WoZ data set.

Figure 2. The overall workflow of our proposed model to generate contextual summaries of clinical diagnostic interviews. First, a PHQ-9 lexicon filters out irrelevant occurrences. Next, the ConceptNet embedding model was modulated using PHQ-9 Lexicon using a retrofitting procedure. The improved model was used to generate Word Semantic Scores. These scores quantify the importance of words in a conversation piece. A unified ILP framework of pruned conversations, trigram language model, and WSS, generates abstractive summaries. Q3 and Q4 in PHQ-9 Lexicon are questions in the PHQ-9 questionnaire. So, our Lexicon has nine categories associated with nine items in the questionnaire. DAIC-WoZ: Distress Analysis Interview Corpus-Wizard of Oz; ILP: integer linear programming; PHQ-9: Patient Health Questionnaire-9.



Creation of Pruned Conversations

In the clinical diagnostic interview, the interaction between the patient and MHP is initially unstructured and noisy but becomes focused and specific to mental health over time. The interview involves exchanging medical analogies and condition-expressive phrases, which implicitly refer to mental health conditions. These phrases were identified using semantic lexicon. Furthermore, to improve the quality of the summaries, initial filtering of noninformative summaries needs to be conducted.

A cleaned set of conversations is termed as *pruned conversations* and is essential to allow KiAS to put more weight on terms that are important to MHPs. In this pruning process, we used the PHQ-9 lexicon [34], which covers concepts related to depressive disorders obtained from human-curated medical knowledge bases (eg, UMLS and SNOMED-CT) and web dictionaries (eg, Urban Dictionary). We extracted phrases (eg, bigrams and trigrams) from the pruned conversations and find their presence in the PHQ-9 lexicon. We used normalized pointwise mutual information (NPMI) to evaluate the quality of n-grams. On the basis of the observation that trigrams NPMI scores (0.78) are higher than those for bigrams (0.72), we used a trigram LM in our summarization process [51].

Tuning the Generic LM

Our primary motivation for using an LM is to improve the linguistic quality of generated summaries. LM implicitly measures the grammatical nature and readability of text. On the other hand, one can either use a pretrained generic LM as-is or tune an LM for a specific domain and task. In this study, we chose to optimize an existing generic LM (we used a trigram model from CMUSphinx [52]) using the PHQ-9 lexicon.

This ensures that depression-related terms are retained when sentences are synthesized. We introduced a semantic score for

each word in the sentence, termed the word semantic score (WSS), to measure how a word relates to the depression symptoms present in the PHQ-9 lexicon. To compute this semantic score, we leveraged ConceptNet embeddings retrofitted with the PHQ-9 lexicon [53].

Retrofitting ConceptNet

This procedure enriches word representations in the generic word-embedding model using a semantic lexicon [53]. In this study, we retrofitted ConceptNet to improve the similarity between the concepts of clinical relevance using the PHQ-9 lexicon. For example, the similarity between the words *feeling* and *lethargic* is 0.39 in ConceptNet, whereas after retrofitting, it is 0.86. As generic LMs generate representations of words in a sentence solely depending on the data, they may not reflect the true meaning of these words. Retrofitting adjusts the representations of words related to depression to have a similarity score to those in the lexicon. With the retrofitted ConceptNet embeddings, a word's semantic score is the maximum cosine similarity the word can have with a term in the depression lexicon. The formulation for WSS is as follows:

$$c(w_l) = \max_{w_k \in \text{lexicon}} \cos(w_l, w_k)$$

Here, $\{c(w_l) = \max_{w_k \in \text{lexicon}} \cos(w_l, w_k) \mid l \in \text{lexicon categories}, k \in l\}$, denotes the maximum cosine similarity of w_l with a term in the depression lexicon. WSS improves the linguistic quality of the summaries by enhancing the probabilities of trigram phrases.

Knowledge-Infused Abstractive Summarization

We input the semantic score of a word, the trigram LM, and the relevant utterances to the ILP framework that will maximize the summaries' informativeness and linguistic quality (Table 1). In the context of clinical diagnostic interviews, where the presence of anaphora and the free-flowing nature of discourse

is profound, a Q and A pair would carry more meaning if previous and subsequent utterances were examined. For example, the Q and A pair—Ellie: “Was that hard for you?” Participant: “Yes”—can be unambiguously decoded if the previous few interactions between the interviewer and the participant are used to provide the necessary context for interpretation. Our domain experts estimated 7 Q and A pairs as an adequate window to capture the context based on a random

sample of 25 patient conversations. The pruned conversations were divided evenly into the maximum number of slices such that each piece was not larger than 7 Q and A pairs. For example, a pruned conversation with 20 Q and A pairs is divided into a set of 3 slices (S) of sizes 7 (s_1), 7 (s_2), and 6 (s_3) Q and A pairs. Furthermore, these fixed-window slices enable grouping the most semantically related sentences, which enhances informativeness.

Table 1. Example dialogues with their respective informativeness (I) and linguistic quality (Q) scores.

Example question and answer pair (path in word graph) ^a	Informativeness score (I)	Linguistic quality score (Q)	Included in summary
Participant was asked have they been diagnosed with depression participant said yeah while ago	0.25	0.1	Yes
Participant was asked uh huh, then participant said pretty easy	0.08	0.04	No

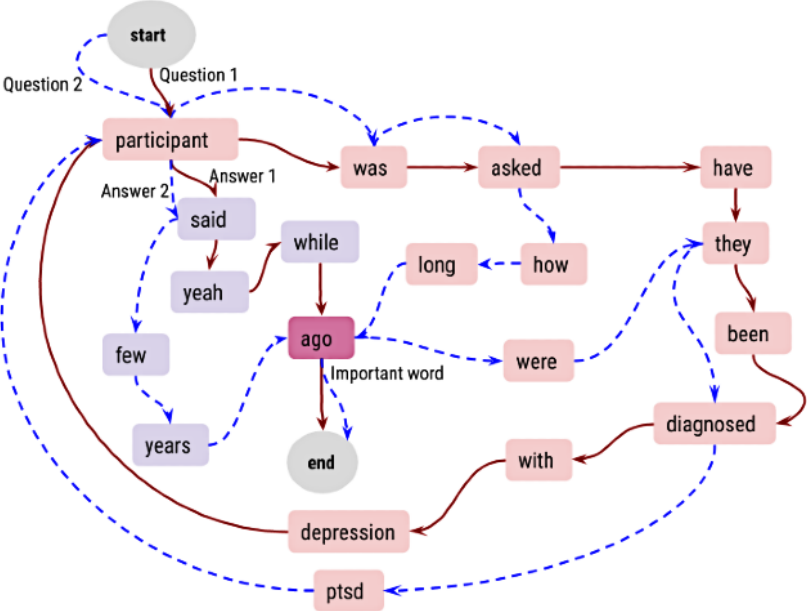
^aSentences containing words that are semantically close to those in the depression lexicon have higher Q than those that do not. The last column says whether the integer linear programming framework selected the sentence or not based on I and Q.

Informativeness (I)

We used TextRank that creates a WG from these slices, with words as vertices along with their scores computed based on in-degree and out-degree metrics (Figure 3) [20]. The frequency of connections between words in a corpus determines the contextual information, thereby factoring in informativeness [54]. TextRank assigns higher scores to vertices (words) with higher degrees in WG. The score of a given vertex v_i in WG is calculated as follows:

where the damping factor d is set to 0.78 (defined empirically), $In(v_i)$ denotes the in-degree of v_i , and $Out(v_j)$ denotes the out-degree of v_j . The information content of a path in a slice $I(p_i^{s_j})$ is calculated as the sum of its words’ scores ($Imp(v_i)$) that represent their importance.

Figure 3. Word Graph (WG) of two Q/A pairs from the Distress Analysis Interview Corpus-Wizard of Oz data set. WG shows “ago” and “diagnosed” as the words with high importance score. WG is created per patient interview and becomes dense as the interview proceeds. This allow the summarizer to locate important words. However, importance scores of domain-specific words “depression” and “ptsd” is elevated using the word semantic score. Start and End are dummy nodes. ptsd: posttraumatic stress disorder.



Linguistic Quality (Q)

Sequential characteristics (ie, using an LM) and contextual coherence (ie, via the WSS of the words in the path) determine the linguistic quality of a sentence. Our approach uses the contextual knowledge using WSS (defined in Equation 2) in the LM, emphasizing the sentences with more frequent trigrams

and words extracted through the mental health lexicon. The modified Q of a path p_i in a slice s_j is defined as:




where p_i is the i^{th} path in the WG formed from q words (w_1, w_2, \dots, w_q). L is the number of conditional probabilities; $P(w_t | w_{t-1}, w_{t-2})$ is calculated using the words in p_i , (w_1, w_2, \dots, w_q). $1 - LL$ (log-likelihood) in Equation 5 maximizes the understanding of the summaries.

ILP Formulation

To simultaneously maximize both Q and I of s_j , we formulated the following objective function:



The ILP framework optimizes $I(p_i^{s_j})$ and $Q(p_i^{s_j})$ of a path p_i and does the same over K paths in the slice $s_j (s_j \in S)$ created from a pruned conversation. The term  ensures that higher weight is given to paths with fewer words for concise summary generation, where $|W(p_i^{s_j})|$ indicates the number of words in the path. To ensure that F chooses one path from WG of paths ($\{p_1, p_2, \dots, p_K\} \subset s_j$), it adheres to the following constraint:



Results

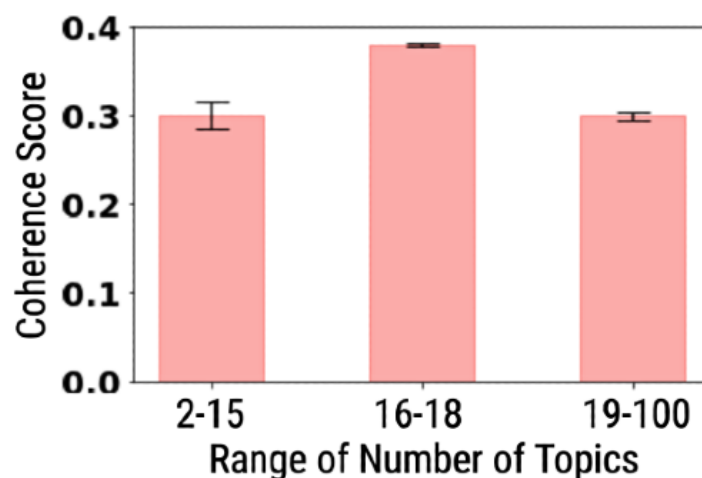
Extrinsic Evaluation

As there are no ground truth summaries of the clinical diagnostic interviews for our task, we considered the pruned conversations (which are the inputs that need to be summarized) as if they were written by humans as reference summaries (RS). For an extrinsic evaluation of the summaries, we adopted JSD, contextual similarity, Flesch Reading Ease (FRE), and Topic Coherence Scores [55]. JSD measures how well the generated

summaries probability distribution approximates the reference probability distribution. We considered the RS to provide the actual probability distribution over mental health concepts (input) and the generated summary as providing an approximation. For JSD, we needed to represent the generated summaries and RS. For this, we trained a topic model and used it to create a probabilistic topical representation of the generated summaries and RS. We trained Latent Dirichlet Allocation (LDA) topic models with varying numbers of topics (2 to 100) and measured coherence scores for each model. From Figure 4, a trained LDA model with 18 topics, having the highest coherence score, was used to generate topical representations of reference and generated summaries. For contextual similarity, we used the cosine-similarity metric to measure the distance between representations of the generated summaries and RS using retrofitted ConceptNet embeddings. A higher contextual similarity score is desirable. We used FRE to measure the readability of the text generated by an automatic summarizer [20]. A higher score reflects the ease of reading. Note that FRE is a validated instrument in the domain of education. Although generated summaries are expected to contain more focused and domain-specific information, they should have a considerable thematic overlap with the pruned conversations (or RS). Topics that describe psychological distress (eg, “I was diagnosed with PTSD a couple of weeks back”) and multiple behavioral dimensions (eg, sad, anger, positive or negative empathy, feeling low) need to be captured in summaries.

We used the LDA model trained over pruned conversations to create topics of summaries generated from different summarization methods. The thematic overlap score was calculated for the generated summaries [56]. We expected the created summaries to provide information on the diagnostic disorder (eg, stress and depression), patient behavior (eg, sad, feeling lonely, and lack of sleep), and time-related information (eg, years, ago, and weeks). As expected, a higher score for thematic overlap was desired.

Figure 4. Coherence scores of latent Dirichlet allocation trained over pruned conversations (or reference summaries) over a number of topics. The topic coherence was measured using the formulation of CV [55].



Extrinsic Evaluation Results

The interviews transcribed in DAIC-WoZ are diverse and noisy discourses because of poor grammar, thereby making the inputs significantly challenging to interpret across different patients.

Moreover, some of the patients' responses were short, with no mental health-related information. We showed these variations and linguistic irregularities using the metrics of contextual

similarity, JSD, readability, and thematic overlap across 184 patients and reported mean scores with SDs for comparison.

In Figure 5, we compare the summaries' representations of each patient with the averaged representation of the RS. We observed that the summaries generated by ES, AS, and AoES, were similar, whereas those produced by KiAS were marginally better. We used information entropy (JSD) to measure the information gain in the summarization approaches (Figure 5 and Table 2). A lower JSD was more desirable. We observed that the JSD of KiAS was relatively small (ES [2.5% decrease], AS [2% decrease], and AoES [3% decrease]), and 2.2% decrease on average, as the use of the PHQ-9 lexicon and retrofitted

ConceptNet enables KiAS to preserve the semantic relationships in the dialogue. For example, "Have you been diagnosed with clinical depression?" and "How often did you feel depressed because of poor sleep and fatigue?" elicit similar information from the patients. The response to either of these questions is related to the mental condition; therefore, it should be informative and preserved. Figures 6 and 7 show a comparison between the baseline summarization approaches and KiAS for readability. We observed that KiAS provides better readability than the baseline methods (ES [5.5% increase], AS [5.3% increase], and AoES [2.3% increase]), whereas all summaries seem to be more readable than the pruned conversations.

Figure 5. The heatmap shows the evaluation of generated summaries using information entropy and contextual similarity metrics. The figure shows the performance of KiAS against three baselines on retaining the content and context in summaries. AoES: abstraction over extractive summarization; AS: abstractive summarization; ES: extractive summarization; KiAS: knowledge-infused abstractive summarization.

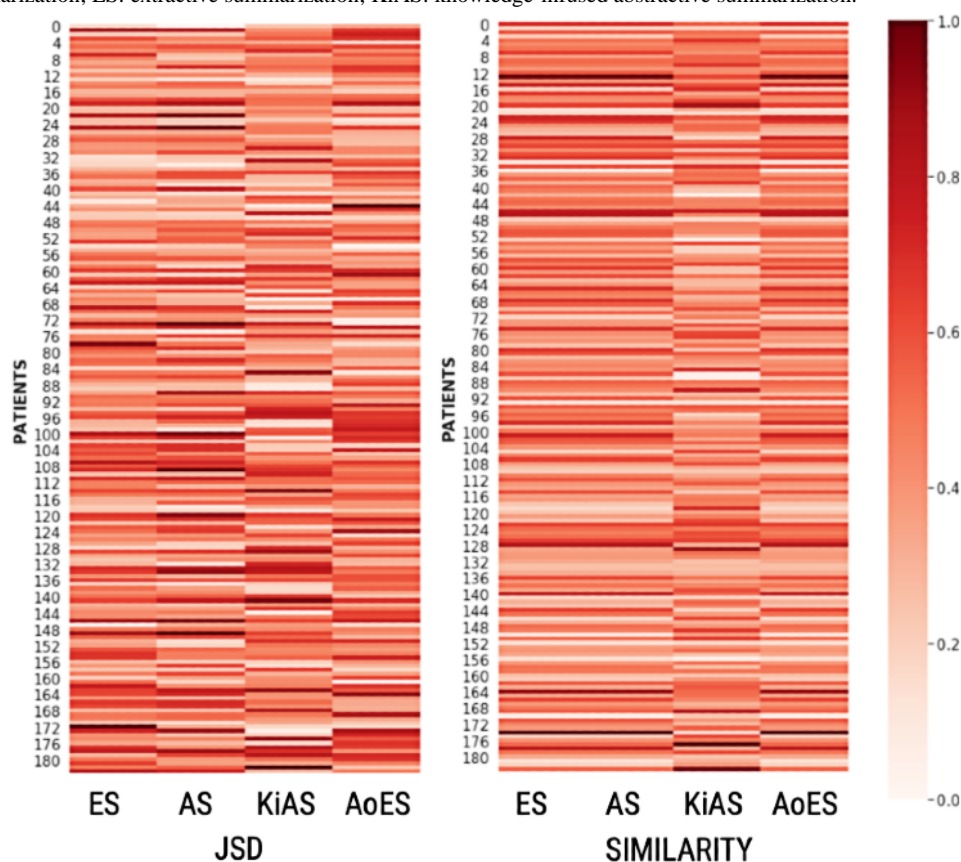


Table 2. Evaluating contextual similarity (higher is better) and information entropy (lower is better) of summaries generated from four different summarization approaches. The reported scores are mean and standard deviation calculated across 184 patient conversations in the Distress Analysis Interview Corpus-Wizard of Oz data set.

Methods	Contextual similarity, mean (SD)	JSD ^a , mean (SD)
Extractive summarization	0.672 (0.076)	0.387 (0.107)
Abstractive summarization	0.676 (0.076)	0.373 (0.102)
Abstractive over extractive summarization	0.669 (0.092)	0.385 (0.103)
Knowledge-infused abstractive summarization	0.689 (0.088)	0.360 (0.104)

^aJSD: Jensen Shannon Divergence.

Figure 6. Heatmap of 184 patient summaries created using the Flesch Reading Ease scale. We compare the readability of summaries created from ES, AS, AoES, and KiAS with reference summaries. Darker patches indicate patient summaries, which are easy to read. AoES: abstraction over extractive summarization; AS: abstractive summarization; ES: extractive summarization; KiAS: knowledge-infused abstractive summarization; RS: reference summaries.

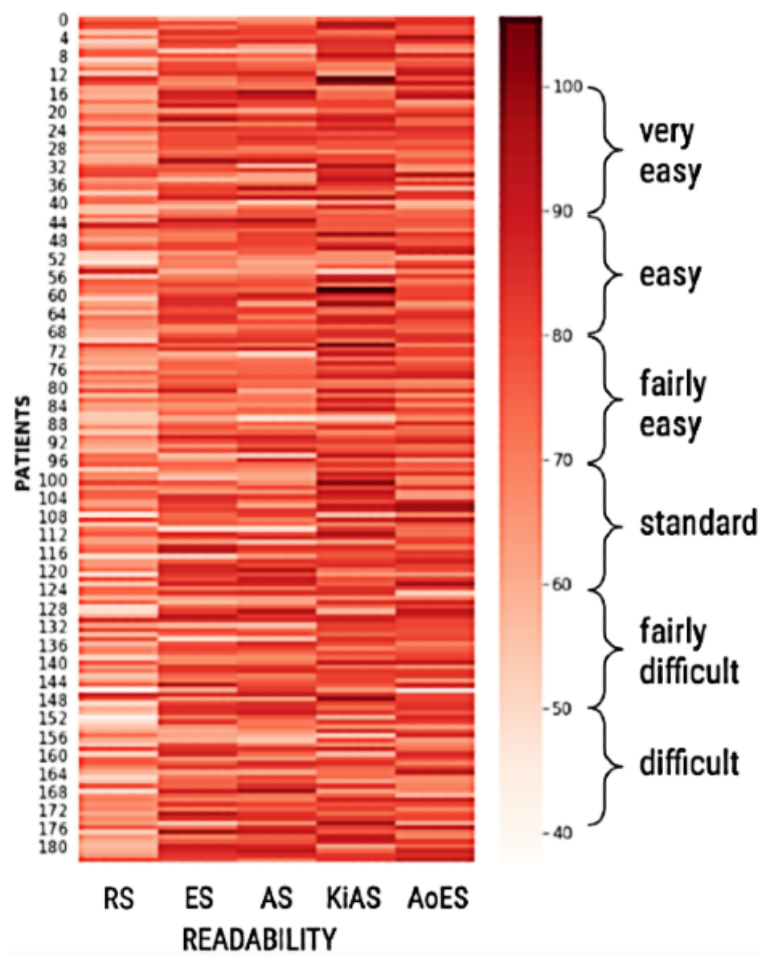
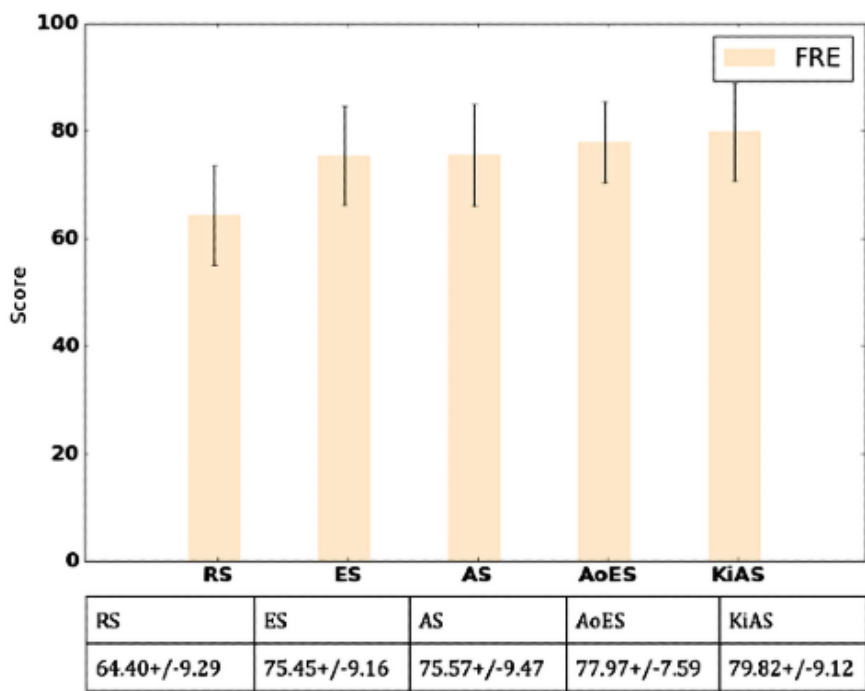


Figure 7. Evaluating readability of the generated summaries from RS. FRE method assesses comprehensibility and engagement of the summaries. AoES: abstraction over extractive summarization; AS: abstractive summarization; ES: extractive summarization; FRE: Flesch Reading Ease; KiAS: knowledge-infused abstractive summarization; RS: reference summaries.



We measured the thematic overlap between the referenced and generated summaries to assess the presence of psychological stressors and behavioral signals, which would be helpful for MHPs. We noticed that ES, AS, and AoES created summaries capturing time-related words (eg, how long, away for a while, about a year, end of the day) and emotional words (eg, annoyed, obnoxious behavior, happy, angry); however, disorder- and response-related words were rarely present. KiAS summaries showed higher occurrences of psychological stress-related words (eg, lack of energy, dizzy whole day, loss of appetite) and behavioral concepts (eg, down in guilt and fear of returning to work). As a result, we observed ~40% thematic overlap between KiAS and reference summaries, which is 17%, 20.5%, 32.5% better than ES (32.8%), AS (31.5%), and AoES (26.65%). On the other hand, KiAS might provide low-quality summaries when the conversation is introductory and does not use mental health-specific words (white and dark patches in [Figures 5 and 7](#)).

In comparing KiAS with ES based on FRE, JSD, and contextual similarity scores, we noticed a significant statistical difference on a two-tailed *t* test at a significance level of .05. Across all 184 patients' summaries, which were quantitatively and qualitatively evaluated, KiAS gave consistently higher scores than ES. Similarly, KiAS summaries were reasonably statistically significant when compared with AS and AoES across the various FRE, JSD, and contextual similarity scores. The main feature of KiAS is to reveal the implicit clinical information from the interviews, specifically the responses of patients to the questions; however, there are few such conversations. Therefore, noticeable differences in KiAS

summaries can be observed by focusing on these samples. Nonetheless, we observed an improvement over AoES and AS, although it was modest.

ROUGE Evaluation

ROUGE is a measure of the informational adequacy of the summaries generated from a summarization system given as input as long text (eg, paragraph, meeting notes, interview logs, or RS). The metric is statistical, as it measures the overlap in *n*-grams in the generated summaries and RS. We reported F_1 scores and Recall of KiAS and compared it with other methods of summarization (ES, AS, and AoES, [Table 3](#)). As KiAS and AS methods are guided by constraints, we considered recall as an appropriate measure, alongside the F_1 score for performance evaluation. With an improvement of approximately 61% in ROUGE-2 (ROUGE for bigrams) recall, KiAS captures more relevant bigrams compared with ES and AS. Most of the clinical terms in mental health care occur as bigrams, and the use of the PHQ-9 lexicon (which contains mostly bigram phrases) enabled the generation of comparatively more relevant summaries (approximately 55% improvement in the F_1 score). In another version of ROUGE-2, ROUGE-L (ROUGE for longest subsequence) calculates the sentence-level structural similarity between the generated summaries and RS. KiAS outperformed ES and AS and AoES with improvements of 48%, 48%, and 53%, respectively, in ROUGE-L recall. We did not see bilingual evaluation understudy as an appropriate metric based on the reasons mentioned in previous literature [[57,58](#)]. We further conducted a human evaluation to analyze the informativeness and fluency (or readability) of the generated summaries [[58,59](#)].

Table 3. Quantitative Recall-Oriented Understudy for Gisting Evaluation (ROUGE)-1, ROUGE-2, and ROUGE-L results on the summarization task of clinical diagnostic interviews. All the scores have a 95% CI of at most ± 0.18 (SD).

Methods	ROUGE ^a -1		ROUGE-2		ROUGE-L	
	Recall	F_1 score	Recall	F_1 score	Recall	F_1 score
KiAS ^b	22.53	30.43	10.65	14.62	24.46	32.57
Abstractive summarization	9.89	15.22	4.09	6.42	12.80	20.53
AoES ^c	8.69	12.37	2.25	3.21	11.51	17.24
Extractive summarization	9.79	14.37	4.18	6.52	12.72	20.47

^aROUGE: Recall-Oriented Understudy for Gisting Evaluation.

^bKiAS: knowledge-infused abstractive summarization.

^cAoES: abstraction over extractive summarization.

Intrinsic Evaluation

Our qualitative evaluation has been designed by 4 practicing psychiatrists who are also the end users of the proposed system. We evaluated the quality of the questions and meaningful responses provided by the patient. Our rubric for domain expert evaluation was as follows:

1. Good question with unclear context: the summary may or may not include context related to mental health, specific to the patient. For example, the patient was asked, "When was the last time that happened?" for which the referent of that is unclear.
2. Good question with clear context: the summary includes context related to only the mental health situation of the patient. These questions are complete, and no inference is required by MHPs. For example, the patient was asked, "Did you ever suffer from PTSD?"
3. Meaningful response: the response is meaningful and understandable concerning the patient and the question asked by the MHP. For example, the patient was asked, "Have you ever been diagnosed with depression?" The patient then responded, "Not really," meaningful only in the context of the preceding question.

Another example: a patient was asked, “How long ago were you diagnosed?” The patient responded, “A few years ago.” In this qualitative evaluation, we only considered summaries generated from KiAS and AS because ES- and AoES-based summaries were not identified as useful by our collaborator MHPs. Randomly selected 25 (179 Q and A sentences) patient summaries were given to the 4 practicing MHPs for expert evaluation. In summary, if a sentence is (1) a good question with an unclear context, (2) a good question with a clear context, or (3) a meaningful response, a +1 score is given. A score of 0 is assigned otherwise. We then totaled the scores for each patient. Considering the varied experiences of MHPs in treating patients, we investigated interrater agreement using Cohen κ [60,61].

Intrinsic Evaluation Results

To evaluate the quality of summaries from the clinicians’ perspective, we randomly selected 25 summaries (179 Q and A pairs) generated from AS and KiAS methods. For intrinsic

evaluation, MHPs would first read the original interview transcript of the patient and then assess the summaries based on (1) the number of good questions and (2) the association of the most relevant patient’s response to the good questions.

Overall, KiAS (1) provides more contextual questions and answers, (2) is more informative even in the absence of context, (3) captures implicit references to medical vocabulary in patients’ answers, and (4) improves upon AS on identifying good questions by 2.6% and meaningful responses by 4.1%. Better Q and A with meaningful responses should improve the follow-up time for MHPs. The 4 MHPs had an estimated reduction of up to 46% in the follow-up time with KiAS than without KiAS, enabling them to see more patients. MHPs mostly agree on their evaluation of KiAS and AS, although MHP 3 is more inclined toward AS than KiAS (Table 4). MHPs provided a substantial agreement that both KiAS and AS provided *good questions*, whereas they showed substantial agreement on KiAS for *meaningful response* compared with a moderate agreement on AS (Table 5).

Table 4. Performance of methods on intrinsic evaluation from 4 mental health professionals gathered after counting the number of good questions with clear context, good questions with unclear context, and meaningful responses in 50 (2 methods) summaries generated.

Domain Experts	MHP ^a 1		MHP 2		MHP 3		MHP 4	
Methods	AS ^b	KiAS ^c	AS	KiAS	AS	KiAS	AS	KiAS
GQCC ^d	71	78	75	74	66	63	59	64
GQUC ^e	86	92	93	95	87	85	82	85
Meaningful responses	60	63	61	65	52	57	72	76

^aMHP: mental health professional.

^bAS: abstractive summarization.

^cKiAS: knowledge-infused abstractive summarization.

^dGQCC: good questions with clear context.

^eGQUC: good questions with unclear context.

Table 5. Interannotator agreement (Cohen κ) calculated over summaries evaluated by 4 mental health professionals.

Method	Good questions ^a , Cohen κ	Meaningful response, Cohen κ
Abstractive summarization	0.63	0.43
KiAS ^b	0.70	0.65

^aGood questions include good questions with unclear context and good questions with clear context.

^bKiAS: knowledge-infused abstractive summarization.

Discussion

Principal Findings

Textbox 2 shows summary snippets generated from AS and KiAS for patient ID 313 (because of a page limit, we have not shown the pruned conversation of patient ID 313, but it can be viewed on a link [62]). The number of Q and A pairs are 51 (number of words=1190). On visual inspection of the summaries by 4 MHPs, we found that the summary provided by the KiAS was more informative and aligned with relevant responses to relevant questions. For example, considering the patient question “How long ago were they diagnosed with depression?” the

response obtained by KiAS was “a year ago,” which is meaningful compared with the response by AS (“they are still depressed”).

Furthermore, an implication that can be drawn from Table 4 is that our summary includes the reason for the patient’s visit, which makes questions on *therapy* and *diagnosis of depression* relevant. However, the *purpose of the visit* is missing in the summary generated by AS. Although the summary created from our approach is longer than that from AS, a recent study by Sun et al [63] illustrated that summary length alone is not a good measure of the summary. To validate the quality of KiAS, we performed an intrinsic evaluation designed to investigate its potential utility in real-world applications.

Textbox 2. Summary generated using abstractive summarization and knowledge-infused abstractive summarization. The proposed approach captures near-exact questions and responses as in Distress Analysis Interview Corpus-Wizard of Oz (DAIC-WoZ) interview scripts compared with AS. Both models received the same pruned conversations of patient ID 313 from the DAIC-WoZ. The tendency to model Q and I as constraints in integer linear programming enhances its capability to generate descriptive summaries.

Summary using abstractive summarization

- *Participant was asked:* What do they do when they are annoying until they stop
- *Participant said:* That they stop talking
- *Participant was asked:* When was the last time they felt really happy
- *Participant said:* A year while ago
- *Participant was asked:* How long ago were they diagnosed depression
- *Participant said:* They are still depressed

Summary using knowledge-infused abstractive summarization

- *Participant was asked:* What do you do when they are annoying
- *Participant said:* She stop talking
- *Participant was asked:* Can you explain with example
- *Participant said:* Yeah
- *Participant was asked:* When was the last time they felt happy
- *Participant said:* A while ago
- *Participant was asked:* What got them to seek help
- *Participant said:* They are still depressed
- *Participant was asked:* Tell me more about that
- *Participant said:* Yeah
- *Participant was asked:* Do they feel like therapy useful
- *Participant said:* Oh yeah definitely
- *Participant was asked:* How long ago were they diagnosed depression
- *Participant said:* A year ago

Conclusions

We experimented with the infusion of knowledge to summarize mental health conversations that contain ambiguous, noisy, and nongrammatical content while preserving the key low-frequency content of clinical significance. Using clinical diagnostic interviews, we proposed a simple and effective summarization strategy called KiAS, which incorporates the knowledge in the PHQ-9 lexicon into an ILP framework to extract and summarize Q and A pairs that describe a patient's condition. We used a sequence of understandable evaluation criteria to test the summarization framework's ability to capture contextual, syntactic, and semantic characteristics that match human-level judgments. Our approach significantly outperforms the baselines on dialogues that had explicit and implicit indicators of mental health conditions. Furthermore, our evaluation showed substantial agreement between MHPs regarding the presence of *good questions* and *meaningful responses* in summaries.

Limitations and Future Work

The result of this study is to inform the development of an automated summarization tool for clinical diagnostic interviews. The clinical interviews were characterized by open-ended conversations comprising questions on diagnosis, symptoms,

medication, and response, which provide implicit references to disorders, symptoms, and medicine. Our analysis of the interview transcripts showed a minimal number of medical concepts, redundancy in questions and responses, and implicit references to mental health conditions. These data characteristics raise challenges in the summarization task in addition to data sparsity because of the limited number of patient interviews. Our primary goal was to demonstrate the feasibility of a method that generates actionable summaries containing meaningful questions and responses for a well-informed follow-up. With guidance from coauthor MHP, we selected the ILP framework, which allows integration of domain knowledge to resolve summarization over clinically sparse data, optimize informativeness in summaries, and generate summaries with good linguistic quality. By optimizing *informativeness* using the PHQ-9 lexicon, our method elicits cues that act as subtle indicators of human emotions. However, we certainly acknowledge that abstractive systems may lose nuance from their rewriting of the text and accept this trade-off to improve recall or reduce information overload. Similarly, social media interaction typically demonstrates sarcasm, irony, and other idiomatic languages that are difficult to capture using deep-learning approaches. Thus, we state that “an accurate assessment, diagnosis, and treatment, hinges on the ability of a

trained MHP to elicit not only information but also extract subtle indicators of human emotions that portend clues to a severely life-threatening situation” with a caution that generated summaries need to be evaluated by domain experts.

A limitation of this study is that we did not explore transformer-based architectures (eg, Medical Bidirectional Encoder Representations from Transformers, Clinical Bidirectional Encoder Representations from Transformers, and BART), fine-tuning procedures, and pretrained models owing to (1) the lack of sufficient patient interviews, (2) no gold standards, and (3) the lack of data from discussion forums that parallel the structure of clinical interviews. Aside from the limitations on the data sets and insufficient exercise of deep-learning methods, a different set of MHPs may have highlighted a different approach or qualitative evaluation strategy. Our research was limited to transcripts from the University of Southern California Institute of Creative Technologies.

Furthermore, we are working toward extending the data set cohort by including clinical interviews recorded by clinicians

in these medical schools. This sample may not represent diagnostic interviews in other clinics, such as the University of California, Los Angeles; University of California San Francisco; Weill Cornell; and Wake Forest School of Medicine, which follow different styles of eliciting information from patients. This can be remedied in future studies by having MHPs at these medical schools. Furthermore, we would extend our qualitative analysis of opinions from MHPs from these medical schools. Furthermore, we envisioned that our approach can be used with guidance from an MHP in a mental health virtual assistant (VA) app to summarize the conversations between the patient and VA.

Reproducibility

We acknowledge the DIAC-WoZ dataset created by the Institute for Creative Technologies at the University of Southern California is available at: <https://dcapswoz.ict.usc.edu/>. The repository containing the models, including the baselines, is made public on: <https://github.com/AmanuelF/DIAC-WoZ-Summarization>.

Acknowledgments

This work was supported in part by the National Science Foundation (NSF) award CNS-1513721 “Context-Aware Harassment Detection on Social Media”; the National Institutes of Health (NIH) award MH105384-01A1 “Modeling Social Behavior for Healthcare Utilization in Depression”; and the National Institute on Drug Abuse (NIDA) grant 5R01DA039454-02 “Trending: Social Media Analysis to Monitor Cannabis and Synthetic Cannabinoid Use.” Any opinions, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect that of the NSF, NIH, or NIDA.

Conflicts of Interest

None declared.

References

1. Park J, Kotzias D, Kuo PF, Logan IV RL, Merced K, Singh S, et al. Detecting conversation topics in primary care office visits from transcripts of patient-provider interactions. *J Am Med Inform Assoc* 2019 Dec 01;26(12):1493-1504 [FREE Full text] [doi: [10.1093/jamia/ocz140](https://doi.org/10.1093/jamia/ocz140)] [Medline: [31532490](https://pubmed.ncbi.nlm.nih.gov/31532490/)]
2. Gigioli P, Sagar N, Rao A, Voyles J. Domain-aware abstractive text summarization for medical documents. In: Proceedings of the IEEE International Conference on Bioinformatics and Biomedicine (BIBM). 2018 Presented at: IEEE International Conference on Bioinformatics and Biomedicine (BIBM); Dec 3-6, 2018; Madrid, Spain p. 2338-2343. [doi: [10.1109/bibm.2018.8621539](https://doi.org/10.1109/bibm.2018.8621539)]
3. Mishra R, Bian J, Fiszman M, Weir CR, Jonnalagadda S, Mostafa J, et al. Text summarization in the biomedical domain: a systematic review of recent research. *J Biomed Inform* 2014 Dec;52:457-467. [doi: [10.1016/j.jbi.2014.06.009](https://doi.org/10.1016/j.jbi.2014.06.009)]
4. Gaur M, Kursuncu U, Alambo A, Sheth A, Daniulaityte R, Thirunarayan K, et al. "Let Me Tell You About Your Mental Health!": contextualized classification of Reddit posts to DSM-5 for web-based intervention. In: Proceedings of the 27th ACM International Conference on Information and Knowledge Management. 2018 Presented at: CIKM '18: The 27th ACM International Conference on Information and Knowledge Management; October, 2018; Torino Italy p. 753-762. [doi: [10.1145/3269206.3271732](https://doi.org/10.1145/3269206.3271732)]
5. Kursuncu U, Gaur M, Sheth A. Knowledge Infused Learning (K-IL): towards deep incorporation of knowledge in deep learning. *AAAI-MAKE Spring Symposium*. 2020. URL: <https://arxiv.org/pdf/1912.00512.pdf> [accessed 2021-03-25]
6. Sheth A, Gaur M, Kursuncu U, Wickramarachchi R, Sheth A. Shades of knowledge-infused learning for enhancing deep learning. *IEEE Internet Comput* 2019 Nov 1;23(6):54-63. [doi: [10.1109/mic.2019.2960071](https://doi.org/10.1109/mic.2019.2960071)]
7. Kapanipathi P, Thost V, Sankalp Patel S, Whitehead S, Abdelaziz I, Balakrishnan A, et al. Infusing knowledge into the textual entailment task using graph convolutional networks. In: Proceedings of the AAAI Conference on Artificial Intelligence. 2020 Apr 03 Presented at: Thirty-Fourth AAAI Conference on Artificial Intelligence; February 7-12, 2020; New York Hilton Midtown, New York p. 8074-8081. [doi: [10.1609/aaai.v34i05.6318](https://doi.org/10.1609/aaai.v34i05.6318)]

8. Sheth A, Perera S, Wijeratne S, Thirunarayan K. Knowledge will propel machine understanding of content : extrapolating from current examples. In: Proceedings of the International Conference on Web Intelligence. 2017 Presented at: WI '17: International Conference on Web Intelligence 2017; August, 2017; Leipzig Germany p. 1-9. [doi: [10.1145/3106426.3109448](https://doi.org/10.1145/3106426.3109448)]
9. Alambo A, Gaur M, Lokala U, Kursuncu U, Thirunarayan K, Gyrard A, et al. Question answering for suicide risk assessment using reddit. In: 2019 IEEE 13th International Conference on Semantic Computing (ICSC). 2019 Presented at: 13th International Conference on Semantic Computing; 30 Jan-Feb 1, 2019; Newport Beach, CA URL: <https://ieeexplore.ieee.org/document/8665525/authors#authors> [doi: [10.1109/ICOSC.2019.8665525](https://doi.org/10.1109/ICOSC.2019.8665525)]
10. Valiant LG. Knowledge infusion. American Association for Artificial Intelligence. 2006. URL: <https://people.seas.harvard.edu/~valiant/AAAI06.pdf> [accessed 2021-03-26]
11. Purohit H, Shalin VL, Sheth AP, Sheth A. Knowledge graphs to empower humanity-inspired AI systems. IEEE Internet Comput 2020 Jul 1;24(4):48-54. [doi: [10.1109/mic.2020.3013683](https://doi.org/10.1109/mic.2020.3013683)]
12. Gaur M, Kursuncu U, Sheth A, Wickramarachchi R, Yadav S. Knowledge-infused deep learning. In: Proceedings of the 31st ACM Conference on Hypertext and Social Media. 2020 Presented at: HT '20: 31st ACM Conference on Hypertext and Social Media; July, 2020; Virtual Event USA p. 309-310. [doi: [10.1145/3372923.3404862](https://doi.org/10.1145/3372923.3404862)]
13. Gratch J, Ron A, Lucas G, Stratou G, Scherer S, Nazarian A, et al. The distress analysis interview corpus of human and computer interviews. Language Resource and Evaluation Conference. 2014. URL: http://www.lrec-conf.org/proceedings/lrec2014/pdf/508_Paper.pdf [accessed 2021-03-25]
14. Wang L, Cardie C. Unsupervised topic modeling approaches to decision summarization in spoken meetings. In: Proceedings of the 13th Annual Meeting of the Special Interest Group on Discourse and Dialogue. 2012 Presented at: 13th Annual Meeting of the Special Interest Group on Discourse and Dialogue; July 5-6, 2012; Seoul National University in Seoul, South Korea p. 40-49 URL: <https://arxiv.org/pdf/1606.07829.pdf>
15. Clarke J, Lapata M. Models for sentence compression: a comparison across domains, training requirements and evaluation measures. In: Proceedings of the 21st International Conference on Computational Linguistics and the 44th annual meeting of the Association for Computational Linguistics. 2006 Presented at: 21st International Conference on Computational Linguistics and the 44th annual meeting of the Association for Computational Linguistics; July, 2006; Sydney, Australia p. 377-384. [doi: [10.3115/1220175.1220223](https://doi.org/10.3115/1220175.1220223)]
16. Clarke J, Lapata M. Global inference for sentence compression: an integer linear programming approach. J Artif Intell Res 2008 Mar 11;31:399-429. [doi: [10.1613/jair.2433](https://doi.org/10.1613/jair.2433)]
17. Galley M, McKeown K. Lexicalized Markov grammars for sentence compression. In: Proceedings of the Conference on Human Language Technologies 2007: The Conference of the North American Chapter of the Association for Computational Linguistics. 2007 Presented at: Human Language Technologies 2007: The Conference of the North American Chapter of the Association for Computational Linguistics; April 2007; Rochester, New York p. 180-187 URL: <https://www.aclweb.org/anthology/N07-1023.pdf>
18. Fan JW, Prasad R, Yabut RM, Loomis RM, Zisook DS, Mattison JE, et al. Part-of-speech tagging for clinical text: wall or bridge between institutions? AMIA Annu Symp Proc 2011;2011:382-391 [FREE Full text] [Medline: [22195091](https://pubmed.ncbi.nlm.nih.gov/22195091/)]
19. Filippova K. Multi-sentence compression: finding shortest paths in word graphs. In: Proceedings of the 23rd International Conference on Computational Linguistics (Coling 2010). 2010 Presented at: 23rd International Conference on Computational Linguistics (Coling 2010); August 2010; Beijing, China p. 322-330 URL: <https://www.aclweb.org/anthology/C10-1037.pdf>
20. Mihalcea R, Tarau P. Textrank: bringing order into text. In: Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing. 2004 Presented at: Conference on Empirical Methods in Natural Language Processing; July, 2004; Barcelona, Spain p. 404-411 URL: <https://www.aclweb.org/anthology/W04-3252.pdf>
21. Erkan G, Radev DR. LexRank: graph-based lexical centrality as salience in text summarization. J Artif Intell Res 2004 Dec 01;22:457-479. [doi: [10.1613/jair.1523](https://doi.org/10.1613/jair.1523)]
22. Banerjee S, Mitra P, Sugiyama K. Multi-document abstractive summarization using ILP based multi-sentence compression. In: Proceedings of the 24th International Joint Conference on Artificial Intelligence. 2015 Presented at: 24th International Joint Conference on Artificial Intelligence, IJCAI 2015; Jul 25-31, 2015; Buenos Aires, Argentina URL: <https://arxiv.org/pdf/1609.07034.pdf>
23. Nayeem MT, Fuad TA, Chali Y. Abstractive unsupervised multi-document summarization using paraphrastic sentence fusion. In: Proceedings of the 27th International Conference on Computational Linguistics. 2018 Presented at: 27th International Conference on Computational Linguistics; August, 2018; Santa Fe, New Mexico, USA p. 1191-1204 URL: <https://www.aclweb.org/anthology/C18-1102.pdf>
24. Li P, Lam W, Bing L, Guo W, Li H. Cascaded attention based unsupervised information distillation for compressive summarization. In: Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing. 2017 Presented at: Conference on Empirical Methods in Natural Language Processing; September, 2017; Copenhagen, Denmark p. 2081-2090. [doi: [10.18653/v1/d17-1221](https://doi.org/10.18653/v1/d17-1221)]
25. See A, Liu PJ, Manning CD. Get to the point: summarization with pointer-generator networks. In: Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). 2017 Presented at: 55th Annual

- Meeting of the Association for Computational Linguistics; July, 2017; Vancouver, Canada p. 1073-1083. [doi: [10.18653/v1/p17-1099](https://doi.org/10.18653/v1/p17-1099)]
26. Nallapati R, Zhou B, Santos CD, Gu'lcehre, C, Xiang B. Abstractive text summarization using sequence-to-sequence RNNs and beyond. In: Proceedings of The 20th SIGNLL Conference on Computational Natural Language Learning. 2016 Presented at: 20th SIGNLL Conference on Computational Natural Language Learning; August, 2016; Berlin, Germany p. 280-290. [doi: [10.18653/v1/k16-1028](https://doi.org/10.18653/v1/k16-1028)]
 27. Shang, Guokan, Wensi D, Zekun Z, Antoine T, Polykarpos M, et al. Unsupervised abstractive meeting summarization with multi-sentence compression and budgeted submodular maximization. In: Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). 2018 Presented at: 56th Annual Meeting of the Association for Computational Linguistics; July, 2018; Melbourne, Australia p. 664-674. [doi: [10.18653/v1/p18-1062](https://doi.org/10.18653/v1/p18-1062)]
 28. Wang R, Liu W, McDonald C. Corpus-independent generic keyphrase extraction using word embedding vectors. In Software Engineering Research Conference. 2014. URL: <https://www.semanticscholar.org/paper/Corpus-independent-Generic-Keyphrase-Extraction-Wang/bd3794c777af5ba363abae5708050ea78ecc97e2> [accessed 2021-03-25]
 29. Wang B, Wang A, Chen F, Wang Y, Kuo CJ. Evaluating word embedding models: methods and experimental results. APSIPA Transactions on Signal and Information Processing 2019 Jul 08;8:1-14. [doi: [10.1017/atsip.2019.12](https://doi.org/10.1017/atsip.2019.12)]
 30. MacAvaney S, Sotudeh S, Cohan A, Goharian N, Talati I, Filice RW. Ontology-aware clinical abstractive summarization. In: Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval. 2019 Presented at: Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval; July, 2019; Paris France p. 1013-1016. [doi: [10.1145/3331184.3331319](https://doi.org/10.1145/3331184.3331319)]
 31. Lewis M, Liu Y, Goyal N, Ghazvininejad M, Mohamed A, Levy O, et al. Bart: denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension. In: Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. 2020 Presented at: 58th Annual Meeting of the Association for Computational Linguistics; July, 2020; ACL, Online p. 7871-7880. [doi: [10.18653/v1/2020.acl-main.703](https://doi.org/10.18653/v1/2020.acl-main.703)]
 32. Gaur M. 313_TRANSCRIPT_CONTINUOUS. Google Docs. URL: https://docs.google.com/spreadsheets/d/17ax_FsLs4Xkb95g4RDWT04g631vciktqH_mwisE6A6s/edit#gid=1305923778 [accessed 2021-04-06]
 33. Gaur M, Faldu K, Sheth A, Sheth A. Semantics of the black-box: can knowledge graphs help make deep learning systems more interpretable and explainable? IEEE Internet Comput 2021 Jan;25(1):51-59. [doi: [10.1109/MIC.2020.3031769](https://doi.org/10.1109/MIC.2020.3031769)]
 34. 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. In: Proc IEEE ACM Int Conf Adv Soc Netw Anal Min. 2017 Presented at: ASONAM '17: Advances in Social Networks Analysis and Mining; July, 2017; Sydney Australia p. 1191-1198 URL: <http://europepmc.org/abstract/MED/29707701> [doi: [10.1145/3110025.3123028](https://doi.org/10.1145/3110025.3123028)]
 35. Karmen C, Hsiung RC, Wetter T. Screening Internet forum participants for depression symptoms by assembling and enhancing multiple NLP methods. Comput Methods Programs Biomed 2015 Jun;120(1):27-36. [doi: [10.1016/j.cmpb.2015.03.008](https://doi.org/10.1016/j.cmpb.2015.03.008)] [Medline: [25891366](https://pubmed.ncbi.nlm.nih.gov/25891366/)]
 36. Neuman Y, Cohen Y, Assaf D, Kedma G. Proactive screening for depression through metaphorical and automatic text analysis. Artif Intell Med 2012 Sep;56(1):19-25. [doi: [10.1016/j.artmed.2012.06.001](https://doi.org/10.1016/j.artmed.2012.06.001)] [Medline: [22771201](https://pubmed.ncbi.nlm.nih.gov/22771201/)]
 37. Alhanai T, Ghassemi M, Glass J. Detecting depression with audio/text sequence modeling of interviews. Interspeech. 2018. URL: https://groups.csail.mit.edu/sls/publications/2018/Alhanai_Interspeech-2018.pdf [accessed 2021-03-25]
 38. Banerjee S, Mitra P, Sugiyama K. Abstractive meeting summarization using dependency graph fusion. In: Proceedings of the 24th International Conference on World Wide Web. 2015 Presented at: 24th International Conference on World Wide Web; May, 2015; Florence Italy p. 5-6. [doi: [10.1145/2740908.2742751](https://doi.org/10.1145/2740908.2742751)]
 39. Speer R, Chin J, Havasi C. ConceptNet 5.5: an open multilingual graph of general knowledge. In: Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence. 2017 Presented at: AAAI'17: Thirty-First AAAI Conference on Artificial Intelligence; February, 2017; San Francisco, California USA p. 4444-4451 URL: <https://arxiv.org/pdf/1612.03975.pdf>
 40. Chandler J, Shapiro D. Conducting clinical research using crowdsourced convenience samples. Annu Rev Clin Psychol 2016 Mar 28;12(1):53-81. [doi: [10.1146/annurev-clinpsy-021815-093623](https://doi.org/10.1146/annurev-clinpsy-021815-093623)] [Medline: [26772208](https://pubmed.ncbi.nlm.nih.gov/26772208/)]
 41. Kroenke K, Strine TW, Spitzer RL, Williams JB, Berry JT, Mokdad AH. The PHQ-8 as a measure of current depression in the general population. J Affect Disord 2009 Apr;114(1-3):163-173. [doi: [10.1016/j.jad.2008.06.026](https://doi.org/10.1016/j.jad.2008.06.026)] [Medline: [18752852](https://pubmed.ncbi.nlm.nih.gov/18752852/)]
 42. Shin C, Lee S, Han K, Yoon H, Han C. Comparison of the usefulness of the PHQ-8 and PHQ-9 for screening for major depressive disorder: analysis of psychiatric outpatient data. Psychiatry Investig 2019 Apr;16(4):300-305 [FREE Full text] [doi: [10.30773/pi.2019.02.01](https://doi.org/10.30773/pi.2019.02.01)] [Medline: [31042692](https://pubmed.ncbi.nlm.nih.gov/31042692/)]
 43. Pérez-Rosas V, Sun X, Li C, Wang Y, Resnicow K, Mihalcea R. Analyzing the quality of counseling conversations: the tell-tale signs of high-quality counseling. In: Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018). 2018 Presented at: Eleventh International Conference on Language Resources and Evaluation; May, 2018; Miyazaki, Japan URL: <https://www.aclweb.org/anthology/L18-1591.pdf>
 44. Tuan DT, Chi NV, Nghiem MQ. Multi-sentence compression using word graph and integer linear programming. In: Advanced Topics in Intelligent Information and Database Systems. Switzerland: Springer; 2017:367-377.

45. Mihalcea R. Language independent extractive summarization. In: Proceedings of the ACL 2005 on Interactive poster and demonstration sessions. 2005 Presented at: ACLdemo '05: ACL 2005 on Interactive poster and demonstration sessions; June 2005; Ann Arbor, Michigan p. 49-52. [doi: [10.3115/1225753.1225766](https://doi.org/10.3115/1225753.1225766)]
46. Nenkova A, McKeown K. A survey of text summarization techniques. In: Mining Text Data. Boston: Springer; 2012:43-76.
47. Celikyilmaz A, Bosselut A, He X, Choi Y. Deep communicating agents for abstractive summarization. In: Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers). 2018 Presented at: Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers); June, 2018; New Orleans, Louisiana p. 1662-1675 URL: <https://arxiv.org/pdf/1803.10357.pdf> [doi: [10.18653/v1/N18-1150](https://doi.org/10.18653/v1/N18-1150)]
48. Yang Z, Zhu C, Gmyr R, Zeng M, Huang X, Darve E. TED: a pretrained unsupervised summarization model with theme modeling and denoising. Findings of the Association for Computational Linguistics: EMNLP 2020 2020:1865-1874 [FREE Full text] [doi: [10.18653/v1/2020.findings-emnlp.168](https://doi.org/10.18653/v1/2020.findings-emnlp.168)]
49. Lin H, Ng V. Abstractive summarization: a survey of the state of the art. In: Proceedings of the AAAI Conference on Artificial Intelligence. 2019 Presented at: AAAI Conference on Artificial Intelligence; January 27 – February 1, 2019; Hilton Hawaiian Village, Honolulu, Hawaii, USA p. 9815-9822. [doi: [10.1609/aaai.v33i01.33019815](https://doi.org/10.1609/aaai.v33i01.33019815)]
50. Weir K. Exploring how location affects mental health. Monitor on Psychology. 2017. URL: <https://www.apa.org/monitor/2017/01/location-health> [accessed 2020-05-22]
51. Bouma G. Normalized (pointwise) mutual information in collocation extraction. Proceedings of German Society for Computational Linguistics & Language Technology. 2009. URL: [https://www.semanticscholar.org/paper/Normalized-\(pointwise\)-mutual-information-in-Bouma/15218d9c029cbb903ae7c729b2c644c24994c201](https://www.semanticscholar.org/paper/Normalized-(pointwise)-mutual-information-in-Bouma/15218d9c029cbb903ae7c729b2c644c24994c201) [accessed 2021-04-09]
52. Lamere P, Kwok P, Gouvêa E, Raj B, Singh R, Walker W, et al. The CMU SPHINX-4 speech recognition system. Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2003). 2003. URL: <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.406.8962&rep=rep1&type=pdf> [accessed 2021-03-25]
53. Faruqi M, Dodge J, Jauhar, SK, Dyer C, Hovy E, Smith NA. Retrofitting word vectors to semantic lexicons. In: Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. 2015 Presented at: Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies; May–June, 2015; Denver, Colorado p. 1606-1615 URL: <https://arxiv.org/pdf/1411.4166.pdf> [doi: [10.3115/v1/n15-1184](https://doi.org/10.3115/v1/n15-1184)]
54. Haveliwala TH. Topic-sensitive pagerank: a context-sensitive ranking algorithm for web search. IEEE Transactions on Knowledge and Data Engineering 2003 Jul;15(4):784-796. [doi: [10.1109/tkde.2003.1208999](https://doi.org/10.1109/tkde.2003.1208999)]
55. Röder M, Both A, Hinneburg A. Exploring the space of topic coherence measures. In: Proceedings of the Eighth ACM International Conference on Web Search and Data Mining. 2015 Presented at: WSDM 2015: Eighth ACM International Conference on Web Search and Data Mining; February, 2015; Shanghai China p. 399-408. [doi: [10.1145/2684822.2685324](https://doi.org/10.1145/2684822.2685324)]
56. Flesch R. The art of readable writing. Stanford Law Review 1950 Apr;2(3):625. [doi: [10.2307/1225957](https://doi.org/10.2307/1225957)]
57. Sulem E, Abend O, Rappoport A. BLEU is not suitable for the evaluation of text simplification. In: Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing. 2018 Presented at: Conference on Empirical Methods in Natural Language Processing; October–November, 2018; Brussels, Belgium p. 738-744. [doi: [10.18653/v1/d18-1081](https://doi.org/10.18653/v1/d18-1081)]
58. Reiter E. A structured review of the validity of BLEU. Computational Linguistics 2018 Sep;44(3):393-401. [doi: [10.1162/coli_a_00322](https://doi.org/10.1162/coli_a_00322)]
59. Novikova J, Dušek O, Curry AC, Rieser V. Why we need new evaluation metrics for NLG. In: Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing. 2017 Presented at: Conference on Empirical Methods in Natural Language Processing; September, 2017; Copenhagen, Denmark p. 2241-2252. [doi: [10.18653/v1/d17-1238](https://doi.org/10.18653/v1/d17-1238)]
60. Cohen J. Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. Psychol Bull 1968 Oct;70(4):213-220. [doi: [10.1037/h0026256](https://doi.org/10.1037/h0026256)] [Medline: [19673146](https://pubmed.ncbi.nlm.nih.gov/19673146/)]
61. Goel S, Gaur M, Jain E. Nature inspired algorithms in remote sensing image classification. Procedia Comput Sci 2015;57:377-384. [doi: [10.1016/j.procs.2015.07.352](https://doi.org/10.1016/j.procs.2015.07.352)]
62. Lassalle E, Denis P. Joint anaphoricity detection and coreference resolution with constrained latent structures. In: Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence. 2015 Presented at: AAAI Conference on Artificial Intelligence (AAAI); January 25–30, 2015; Austin, Texas, USA p. 2274-2280 URL: <https://bit.ly/patid313>
63. Sun S, Shapira O, Dagan I, Nenkova A. How to compare summarizers without target length? Pitfalls, solutions and re-examination of the neural summarization literature. In: Proceedings of the Workshop on Methods for Optimizing and Evaluating Neural Language Generation. 2019 Presented at: Workshop on Methods for Optimizing and Evaluating Neural Language Generation; June, 2019; Minneapolis, Minnesota p. 21-29. [doi: [10.18653/v1/w19-2303](https://doi.org/10.18653/v1/w19-2303)]

Abbreviations

AoES: abstraction over extractive summarization

AS: abstractive summarization
BART: Bidirectional and Auto-Regressive Transformers
ES: extractive summarization
FRE: Flesch Reading Ease
ILP: integer linear programming
JSD: Jensen Shannon Divergence
KiAS: knowledge-infused abstractive summarization
LDA: Latent Dirichlet Allocation
LM: language model
MHP: mental health professional
NIDA: National Institute on Drug Abuse
NIH: National Institutes of Health
NPMI: normalized pointwise mutual information
NSF: National Science Foundation
PHQ-9: Patient Health Questionnaire-9
PTSD: posttraumatic stress disorder
Q and A: question and answer
ROUGE: Recall-Oriented Understudy for Gisting Evaluation
RS: reference summaries
SB: SumBasic
SNOMED-CT: Systematized Nomenclature of Medicine-Clinical Terms
UMLS: Unified Medical Language System
VA: virtual assistant
WG: word graph
WSS: word semantic score

Edited by J Torous; submitted 30.05.20; peer-reviewed by T Baldwin, A Poddar; comments to author 09.07.20; revised version received 05.10.20; accepted 04.03.21; published 10.05.21.

Please cite as:

Manas G, Aribandi V, Kursuncu U, Alambo A, Shalin VL, Thirunarayan K, Beich J, Narasimhan M, Sheth A
Knowledge-Infused Abstractive Summarization of Clinical Diagnostic Interviews: Framework Development Study
JMIR Ment Health 2021;8(5):e20865
URL: <https://mental.jmir.org/2021/5/e20865>
doi: [10.2196/20865](https://doi.org/10.2196/20865)
PMID: [33970116](https://pubmed.ncbi.nlm.nih.gov/33970116/)

©Gaur Manas, Vamsi Aribandi, Ugur Kursuncu, Amanuel Alambo, Valerie L Shalin, Krishnaprasad Thirunarayan, Jonathan Beich, Meera Narasimhan, Amit Sheth. Originally published in JMIR Mental Health (<https://mental.jmir.org>), 10.05.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Mental Health, is properly cited. The complete bibliographic information, a link to the original publication on <https://mental.jmir.org/>, as well as this copyright and license information must be included.

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org

<https://www.jmirpublications.com/>