Can Mental Health Chatbots Help Chronic Disease Populations?

[Evaluating the Usefulness of an Artificial Intelligence (A.I.) Mental Health Chatbot in a Healthcare Setting]

Protocol and Statistical Analysis Plan

NCT04620668

October 15, 2020

Updated February 13, 2024 for title page, formatting

People dealing with chronic health conditions are susceptible to mental health issues such as depression and anxiety (Clarke & Currie, 2009). Providing conventional mental health services to all of these individuals is not practical given the limited resources of the healthcare system. Artificial intelligence (A.I.) mental health chatbots may be an accessible and cost-effective means by which people can receive some degree of mental health support while they cope with their conditions. These automated programs act as a source of virtual support, talking with individuals and providing them with therapeutic exercises to improve their mental wellbeing. Several chatbots have been designed to deliver interventions based on popular psychological therapies (e.g., Wysa, Woebot, and Tess). Research has shown that these programs can reduce symptoms of depression, anxiety, and stress in nonclinical populations (Fitzpatrick et al., 2017; Fulmer et al., 2018; Inkster et al., 2018; Ly et al., 2017). However, the effectiveness of these programs has not been tested in chronic disease populations.

The purpose of the current research is to gain a better understanding of the usefulness of mental health chatbots for chronic disease populations. This research will be guided by two fundamental objectives: (1) to determine whether a mental health chatbot can reduce or prevent negative mental health symptoms in individuals who are dealing with a chronic health condition, and (2) to learn more about how individuals with a chronic health condition view these programs, particularly in terms of their potential benefits or drawbacks when used in healthcare settings. This research will focus on two specific chronic disease populations that are prone to elevated levels of mental health symptoms: people with arthritis and diabetes.

Method

Design

The study objectives will be pursued using a mixed-method research design (Creswell & Plano Clark, 2011), which will involve collecting and analyzing quantitative data on individuals' mental health symptomology and qualitative data on their experiences with the mental health chatbot.

Sample

Participants will be individuals between the ages of 19 and 65 who have a diagnosis of diabetes (type 1 or type 2 diabetes) or arthritis (osteoarthritis, rheumatoid arthritis, or another type of arthritis). Although people of any age might be susceptible to mental health issues when dealing with these conditions, individuals in the young adult and middle adult age groups may be at particularly high risk because they have to navigate certain life obligations while coping with their conditions (e.g., caring for a young family, working a job) that many older and younger individuals do not. They are also more likely to use and be familiar with chat-based programs than young children or older adults, meaning that they will almost certainly have the modest technical knowledge required to complete the study.

A power analysis conducted using G*Power (Heinrich-Heine-Universität, Düsseldorf) indicated that a total of 40 participants would be required for the quantitative portion of the study, assuming a moderate effect size, .80 power, an alpha level of .05, and a moderate correlation between repeated measures (.50). However, a larger number of participants ($N = 60$) will be sought to strengthen the analysis. Individuals will be excluded from participation if they are receiving ongoing treatment from a mental health professional, if they already use a mental health chatbot, or if they started or experienced a dosage change in a psychopharmacological drug within the previous month. Participants will also need to have a phone with an active Internet connection. Participants will be offered a $20 gift card (Canadian funds) to an online retailer of their choice for completing the study. Participants who drop out of the study prior to

completion will be eligible for a $5 gift card. An additional $10 gift card will be offered to participants in the treatment group who complete an optional interview following the main study.

Setting

The study researchers will be based at the Centre for Research in Integrated Care (CRIC) at the University of New Brunswick in Saint John. The centre will provide access to workspace and necessary equipment and software. All participant communication will occur remotely through phone calls, videoconferencing, and/or email. The developer of the chatbot, Touchkin eServices, will provide technical support and assistance to the research team as necessary from their location Bengaluru (Bangalore), India.

Procedure

Participants will be recruited through social media channels (including online groups), newspaper advertisements, and emails and newsletters from relevant organizations (e.g., the Arthritis Society, Diabetes Canada). After volunteering to participate, participants will set up a phone or video conferencing call with the primary investigator to orient them into the study. Participants will be randomly assigned to either a treatment group or control group. Those assigned to the treatment group will download the mental health chatbot Wysa (Touchkin eServices, Bangalore) on their smartphones. They will be encouraged to interact with the chatbot at least two times per week over a period of four weeks. Participants assigned to the control group will receive no chatbot (i.e., they will be in a no-treatment control group).

Regardless of their group assignment, participants will complete online materials via Qualtrics at the outset of the study, two weeks into the study, and four weeks into the study (i.e., the final assessment point). At the outset of the study, participants will fill out an informed consent form, a demographic questionnaire, and four psychological assessments tools: measures of depression (Patient Health Questionnaire; Kroenke et al., 2001), anxiety (Generalized Anxiety Disorder Scale; Spitzer et al., 2006), stress (Perceived Stress Scale; Cohen & Williamson, 1988), and life satisfaction (Satisfaction with Life Scale; Diener, et al., 1985). Two weeks into the study, participants will complete the four psychological assessment tools a second time. Four weeks into the study, participants will complete the four assessment tools a final time, and those in the treatment group will be presented with a post-study questionnaire that contains qualitative questions regarding their experiences with the chatbot. Participants in both groups will be presented with a debriefing form providing more information about the study. Those in the control group will be given the opportunity to download and use the chatbot.

To facilitate participant compliance, email reminders about outstanding surveys will be sent throughout the week as necessary. In addition, participants in the treatment group will be sent a general reminder about the intervention once per week.

After the data from the four-week study are analyzed, a subset of participants from the treatment group may be asked to complete optional phone or video interviews to gain more insight into their experiences with and opinions on the chatbot. Approximately 15 to 20 participants will be sought for the interviews. The questions for these interviews will be developed based on the collective results from the quantitative and qualitative analysis described above.

Data Analysis

Quantitative data will be entered into SPSS version 26 (IBM Corp., Armonk, NY, USA) for statistical analysis. Comparisons between the treatment and control groups will be performed using four mixed analysis of variance (ANOVA) statistical tests, with experimental group and health condition as the between-subjects factors, time point as the within-subjects factor, and

mental health outcomes (i.e., scores on the Patient Health Questionnaire, Generalized Anxiety Disorder Scale, Perceived Stress Scale, Satisfaction with Life Scale) as the dependent variables. Significant interactions will be followed up with repeated measures ANOVAs (one within each experimental group) and dependent t-tests with Bonferroni adjustments, as needed.

Qualitative data will be analyzed using Braun and Clarke's (2006) six phases of thematic analysis: (1) familiarize oneself with the data; (2) generate initial codes; (3) search for themes; (4) review themes; (5) define and name themes; and (6) provide the report. This analysis will be performed on the qualitative data from the post-study questionnaire as well as the phone or video interviews. Qualitative data management will be performed with the assistance of NVivo version 12 (QSR International, Burlington, MA, USA).

Implications

There would be several positive implications if results show that a mental health chatbot can help people with chronic health conditions manage their mental health symptoms. Although these types of programs are not a replacement for more conventional mental health providers, they would be an accessible and cost-effective means by which people could gain some degree of mental health support while they are coping with their chronic health conditions. People who do not have access to a mental health provider (e.g., those who live in rural communities or who have reduced mobility due to their health conditions) and those who lack the means to pay for a mental health provider on their own (e.g., low income individuals) would be able to receive some mental health support within the comfort of their own homes for little or no cost. These programs would also be an invaluable tool for helping people with chronic health conditions cope when experiencing socially mandated restrictions on their mobility (e.g., self-isolation during a pandemic).

Despite this positive potential, participants in the current study will almost certainly identify limitations with these programs as well. These limitations could include technical limitations with the program itself (e.g., conversations may lack a certain degree of naturalness), as well as limitations in terms of the practical utility of these programs (e.g., there is no possibility of interprofessional collaboration in situations where this would be beneficial). This type of feedback could be useful for the future refinement of these programs. It could also help to inform policy decisions on how these programs should be used in a healthcare context, or even whether they should be used at all.

References

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77-101. doi: 10.1191/1478088706qp063oa

Clarke, D.M. and Currie, K.C. (2009). Depression, anxiety and their relationship with chronic diseases: A review of the epidemiology, risk and treatment evidence. *Medical Journal of Australia, 190(*S7), S54-S60. doi:10.5694/j.1326-5377.2009.tb02471.x

Cohen, S., & Williamson, G.M. (1988). Perceived stress in a probability sample of the United States. In S. Spacapan & S. Oskamp (Eds.), *The social psychology of health: Claremont Symposium on Applied Social Psychology* (pp. 31-67). Sage.

Creswell, J.W., & Plano Clark, V.L. (2017). *Designing and conducting mixed methods research (Vol. 3)*. Sage.

Diener, E., Emmons, R.A., Larsen, R.J., & Griffin, S. (1985). The Satisfaction with Life Scale. *Journal of Personality Assessment, 49*(1), 71-75. doi: 10.1207/s15327752jpa4901_13

Fitzpatrick, K.K., Darcy, A., & Vierhile, M. (2017). Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): A randomized controlled trial. *JMIR Mental Health, 4*(2), e19. doi: 10.2196/mental.7785

Fulmer, R., Joerin, A., Gentile, B., Lakerink, L., & Rauws, M. (2018). Using psychological artificial intelligence (Tess) to relieve symptoms of depression and anxiety: Randomized controlled trial. *JMIR Mental Health, 5*(4), e64. doi: 10.2196/mental.9782

Inkster, B., Sarda, S., & Subramanian, V. (2018). An empathy-driven, conversational artificial intelligence agent (Wysa) for digital mental well-being: Real-world data evaluation mixed-methods study. *JMIR Mental Health, 6*(11), e12106. doi: 10.2196/12106

Kroenke, K., Spitzer, R.L., & Williams, J.B. (2001). The PHQ-9: Validity of a brief depression severity measure. *Journal of General Internal Medicine, 16*(9), 606-613. doi: 10.1046/j.1525-1497.2001.016009606.x

Ly, K.H., Ly, A.-M., & Andersson, G. (2017). A fully automated conversational agent for promoting mental well-being: A pilot RCT using mixed methods. *Internet Interventions, 10*, 39-46. doi: 10.1016/j.invent.2017.10.002

Spitzer, R.L., Kroenke, K., Williams, J.B.W., & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: The GAD-7. *Archives of Internal Medicine, 166*(10), 1092–1097. doi: 10.1001/archinte.166.10.1092