Check for updates

Comparing clinical decision-making of AI technology to a multi-professional care team in an electronic cognitive behavioural therapy program for depression: protocol

Callum Stephenson[1†], Jasleen Jagayat[1,2†], Anchan Kumar[1†],
Paniz Khamooshi[1†], Jazmin Eadie[1,3†], Amrita Pannu[1†],
Dekel Meartsi[1†], Eileen Danaee[1†], Gilmar Gutierrez[1†],
Ferwa Khan[1†], Tessa Gizzarelli[1†], Charmy Patel[1†],
Elnaz Moghimi[1†], Megan Yang[1†], Amirhossein Shirazi[4†],
Mohsen Omrani[1,4†], Archana Patel[1†] and Nazanin Alavi[1,2*†]

[1]Department of Psychiatry, Faculty of Health Sciences, Queen's University, Kingston, ON,
Canada, [2]Centre for Neuroscience Studies, Faculty of Health Sciences, Queen's University,
Kingston, ON, Canada, [3]Department of Psychology, Faculty of Arts and Sciences, Queen's
University, Kingston, ON, Canada, [4]OPTT Inc., Toronto, ON, Canada

Introduction: Depression is a leading cause of disability worldwide, affecting up to 300 million people globally. Despite its high prevalence and debilitating effects, only one-third of patients newly diagnosed with depression initiate treatment. Electronic cognitive behavioural therapy (e-CBT) is an effective treatment for depression and is a feasible solution to make mental health care more accessible. Due to its online format, e-CBT can be combined with variable therapist engagement to address different care needs. Typically, a multi-professional care team determines which combination therapy most benefits the patient. However, this process can add to the costs of these programs. Artificial intelligence (AI) has been proposed to offset these costs.

Methods: This study is a double-blinded randomized controlled trial recruiting individuals experiencing depression. The degree of care intensity a participant will receive will be randomly decided by either: (1) a machine learning algorithm, or (2) an assessment made by a group of healthcare professionals. Subsequently, participants will receive depression-specific e-CBT treatment through the secure online platform. There will be three available intensities of therapist interaction: (1) e-CBT; (2) e-CBT with a 15–20-min phone/video call; and (3) e-CBT with pharmacotherapy. This approach aims to accurately allocate care tailored to each patient's needs, allowing for more efficient use of resources.

Discussion: Artificial intelligence and providing patients with varying intensities of care can increase the efficiency of mental health care services. This study aims to determine a cost-effective method to decrease depressive symptoms and increase treatment adherence to online psychotherapy by

allocating the correct intensity of therapist care for individuals diagnosed with depression. This will be done by comparing a decision-making machine learning algorithm to a multi-professional care team. This approach aims to accurately allocate care tailored to each patient's needs, allowing for more efficient use of resources with the convergence of technologies and healthcare.

Ethics: The study received ethics approval and began participant recruitment in December 2022. Participant recruitment has been conducted through targeted advertisements and physician referrals. Complete data collection and analysis are expected to conclude by August 2024.

Clinical trial registration: ClinicalTrials.Gov, identifier NCT04747873.

1 Introduction

Depression is a leading cause of disability, affecting approximately 3.8% of the population worldwide (1, 2). Despite the high prevalence and negative consequences of depression, only one-third of individuals receive treatment, and three in five receive sufficient care (3, 4). In 2018, mental health care needs for 2.3 million Canadians were reported to be insufficient with 78.2% indicating reasons relating to accessibility (e.g., lack of knowledge on how to get help, financial issues, and time restraints) to be the top barrier to receiving mental health care (5). Another major hindering factor in patients' access to effective mental health support is the cost of care (6). This outlines the importance of developing evidence-based and cost-effective solutions to address these issues on mental health care accessibility and efficiency.

Electronic cognitive behavioural therapy (e-CBT) is a good candidate to make mental health care more accessible and is effective in treating various mental health issues including mood and anxiety disorders (7–11). However, this internet-based treatment presents problems of high dropout rates and non-adherence (12). Treatment compliance is an important factor for positive psychotherapy treatment outcomes (13, 14). Drop-out rates for in-person psychotherapy vary from 20–47% depending on the study design (15–17), whereas a systematic review of 29 studies revealed the dropout rate for e-CBT for depression can range from 0–63%, with an average dropout rate of 32% (18). Multiple factors can affect patient compliance such as patient age, education, treatment engagement, and clinician involvement (12, 19, 20).

Furthermore, while pre-designed e-CBT content can make mental health care more scalable and affordable and address the general concerns of most individuals with mild to moderate symptoms (8, 9, 21, 22), individuals with severe symptoms may require further intensive clinical interventions and clinician involvement (22, 23). Stratified care begins treatment by matching care intensity (i.e., low vs. high intensity) based on the patient's needs. Stratified psychiatry is beneficial because it can increase treatment response and remission rates by allocating treatments effectively (24). An important challenge in using a stratified care strategy is recognizing which group each

patient belongs to and which resources are required to address their needs.

Currently, clinical decision-making is conducted by a multi-professional care team which can consist of a psychiatrist, residents, nurses, clinical psychologist, social worker, counsellor, and support worker (25). Although this collaboration is beneficial in many respects (26), is typically costly and results in long wait times (6). A previous framework using real patient data has demonstrated the cost-effectiveness of using machine learning in clinical decision-making (27). The use of machine learning in healthcare has significantly increased over the past few years. In the context of clinical decision-making, machine learning algorithms use large data sets from an array of sources to assist healthcare providers in making rapid and informed decisions (28). To further understand symptom severity and increase treatment efficacy, many researchers have tried to implement supervised machine-learning approaches to identify patient characteristics that may be associated with poor treatment outcomes and develop a treatment for patients (29–32). These novel decision-making algorithms aim to mimic human decision-making to help with objective and accurate evaluation of each patient's needs (33). This technology can make personalized care a possibility, particularly in cases where clinical decision-making by a care team is not feasible or results in prolonged wait times. Overall, the novelty of machine learning in healthcare demands greater exploration and understanding of responsible use (34).

The goal of this study is to develop a machine learning algorithm that can identify the intensity of care an individual needs based on their probability of dropping out from treatment. This algorithm is expected to be comparable to the clinical decision-making process. It is expected to indicate the appropriate level of care for an individual experiencing depression to make treatment efficient and effective. We believe that the integration of AI-driven clinical decision-making would make mental health care delivery more scalable, accessible, and affordable while remaining highly personalized and effective. To make this possible, we need to (1) make objective and quantified evaluations of patients' mental status and their needs, (2) utilize a decision-making algorithm that allocates the right level of care for each patient, and (3) demonstrate that this method of care delivery can enhance the quality of mental health care and reduce costs.

2 Objectives

The primary objective of this study is to evaluate the effectiveness of the suggested level of care made by the AI compared to the healthcare team and the effectiveness of the stratified care model. This will be assessed using the change in the severity of depressive symptoms using the Patient Health Questionnaire – 9 Item (PHQ-9) (35) and Quick Inventory of Depressive Symptoms (QIDS) (36) scores, and the quality of life and functioning using Assessment of Quality of Life (AQoL-8D) (37) scores between the start and the end of the 13-session e-CBT program.

The secondary objectives include evaluating compliance and treatment completion using the number of sessions completed by each participant and evaluating the time and cost commitment of initial assessments (i.e., manual participant stratification by the healthcare team vs. the AI) and the different care intensities (i.e., only e-CBT, e-CBT + phone calls, or e-CBT + video calls). We will also evaluate the functional consequence of our interventions on participant quality of life via Quality-adjusted life years (QALY)s (38). Throughout the study, QALY estimates will be derived from AQoL-8D scores using the area under the curve method (39, 40). Outlined below are the research questions and the corresponding study hypotheses:

1 Will participants with major depressive disorder (MDD) assigned to an e-CBT program by an AI show similar outcomes in depressive symptoms at 3 months to those allocated to an e-CBT program decided by a team of healthcare professionals?
- We hypothesize that the outcomes in depressive symptoms of treatment arm 1 will be comparable to the outcomes of treatment arms 2 and 3 following e-CBT treatment.
2 How does AI-based clinical decision-making compare to that of a multi-professional healthcare team when allocating individuals with depression to an e-CBT program with different degrees of care intensity?
- We hypothesize that AI-based clinical decision-making (treatment arm 2) will provide suggestions comparable to that of a multi-professional care team (treatment arm 1).
3 Is this novel AI approach to e-CBT time and cost-efficient and comparable to the standard multi-professional healthcare decision-making team?
- We hypothesize that the AI approach will decrease the overall time and cost commitment of providing e-CBT.

3 Methods and analysis

3.1 Participants

Participants ($n = 186$: $n = 31$ per e-CBT group * 2 arms) will be recruited at Queen's University from outpatient psychiatry clinics at both Kingston Health Sciences Centre sites (Hotel Dieu Hospital and Kingston General Hospital), as well as Providence Care Hospital in Kingston, Ontario. Additionally, self-referrals and referrals from family doctors, physicians, and clinicians across Ontario will be accepted. After obtaining informed consent from the participant, the participant will be evaluated using the Mini International Psychiatric Assessment (MINI) through a secure video appointment to confirm a diagnosis of MDD using the Diagnostic and Statistical

Manual of Mental Disorders, Fifth Edition (DSM-5) (41), by a trained professional on the research team.

All eligible participants will be randomized to receive a treatment plan based on the decision of either the healthcare team (Arm 1) or the Triage Module using an AI algorithm (Arm 2). Participants will be randomly allocated to one of the two arms of the study by a research assistant on the team who will also balance the group based on demographic variables (i.e., sex, gender, age, and income). Participants and therapists in the study will be blinded to which treatment arm the participant belongs to. By the nature of this study, participants and therapists will not be blinded to which treatment intensity the participant will receive since it will be evident whether the participant is receiving a phone/video call in addition to usual e-CBT care or pharmacotherapy. Each participant will be provided with an effective form of treatment (i.e., e-CBT) regardless of which group they will be allocated to. Participants will be informed that there is no incentive to join the program and that joining or withdrawing at any point will not affect them negatively. It will also be explained to the participants that the program is not a crisis resource and that they will not always have access to their therapists. In the case of an emergency, participants will be directed to proper resources, and this event will be reported to the study's lead psychiatrist (principal investigator). All data will be anonymized and analyzed by research team members who are not directly involved in the patient's care.

3.2 Inclusion and exclusion criteria

The inclusion criteria for this study are at least 18 years of age, diagnosed with MDD according to the Mini International Neuropsychiatric Interview 7.0.2 (MINI) that will be conducted by a trained research assistant. The MINI is a diagnostic interview that assesses 17 common mental disorders by following the diagnostic criteria of the Diagnostic and Statistical Manual of Mental Disorders, 5th edition (DSM-5) (41), ability to provide informed consent, ability to speak and read English, and having consistent and reliable access to the internet. The exclusion criteria include active psychosis, acute mania, severe alcohol, or substance use disorder, and active suicidal or homicidal ideation. Individuals with these disorders are less prone to having good insight into their thoughts; thus, e-CBT on their own may not be fit for them and would be better suited with therapist support throughout therapy. As for individuals with active suicidal or homicidal ideation, this program is not a crisis resource and individuals will be redirected to immediate support and resources that will better assist them. If a participant is receiving or has received CBT or e-CBT in the past year, they will also be excluded from the study to avoid confounding effects on the efficacy of this e-CBT program. If interested, participants excluded from the study will be linked to the appropriate resources.

3.3 Study design

If eligible for this randomized controlled trial, participants ($n = 186$) will be randomized (Figure 1) to receive an e-CBT treatment recommended by a multi-professional healthcare team consisting of

FIGURE 1
Study design flowchart for enrolment, arm allocation, and treatment intensity decision.
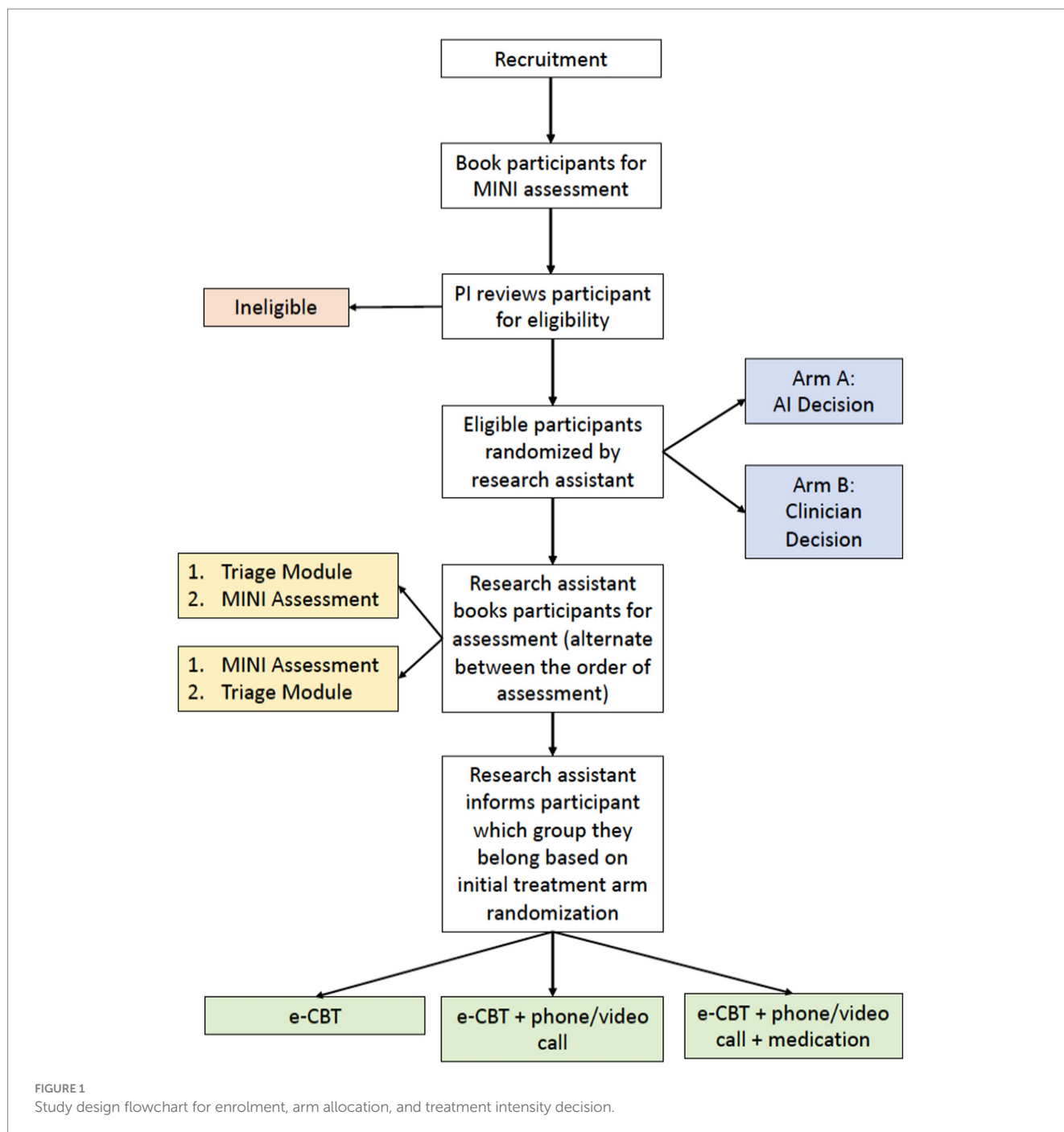
a psychiatrist, psychiatric medical resident, and a trained research assistant (Arm 1, control group; $n = 93$), or the AI machine learning algorithm (Arm 2, experimental group; $n = 93$; Figure 1). To ensure blinding, all participants will complete the intake assessment by the healthcare team (Arm 1) and the Triage Module (Arm 2). Only the relevant data (i.e., Arm 1: intake assessment vs. Arm 2: Triage Module) will be analyzed depending on the treatment arm that the participant is randomly assigned to.

3.3.1 Treatment arm 1: healthcare team allocation

Allocation of treatment intensity by the multi-professional healthcare team will be based on the following criteria:

1 The severity of symptoms/disability (DSM-5 criteria).
2 Mental health factors (prior treatments and responses, current and past psychotic/manic episodes, current and past suicidal/homicidal ideation/attempts, family mental health history, past psychiatric history, and hospital admissions).
3 Medical factors (current medical conditions and medications, personal and family medical history).
4 Social factors (support system and living situation, and occupational, social, and personal functional impairment).

To assess the severity of MDD symptoms and the functional impairments, participants will complete the PHQ-9 and Sheehan

Disability Scale (SDS) (42, 43) before the assessment appointment. The trained research assistant on the multidisciplinary team will conduct the assessment appointment and will relay the information to the rest of the team later to deliberate on treatment intensity allocation. All assessments will occur virtually through phone and video calls. Together, the healthcare team will decide whether the participant should be assigned to the e-CBT-only treatment, e-CBT treatment with weekly phone/video calls, or e-CBT treatment with pharmacotherapy. This process mimics the current triage process in clinical settings. To track cost-effectiveness, the trained research assistant will track the total duration of the individual assessment and team deliberation meetings for analysis of the total time commitment per patient.

3.3.2 Treatment arm 2: AI algorithm allocation

Allocation of treatment intensity by the proposed AI algorithm will be based on the machine learning and natural language processing (NLP) of textual data provided by participants and their PHQ-9 score collected through a pre-treatment screening module called the Triage Module. This module, developed by the research team, (1) provides psychoeducation on the effects of psychotherapy, (2) collects PHQ-9 scores, and (3) asks participants six open-ended questions regarding their mental health history, their experiences with mental health disorders, and what mental health difficulties they are currently facing. Based on the participant's answers to the open-ended questions, a variable called "Symptomatic Score" will be calculated using the NLP algorithm. If the PHQ-9 score < 19 and the Symptomatic Score > 0.75, the participant will be assigned to the e-CBT-only treatment group. However, if either the PHQ-9 score is >19 or the Symptomatic Score is <0.75, the participants will be assigned to the e-CBT treatment with weekly phone/video calls. If both scenarios occur and the PHQ-9 >19 and Symptomatic Score < 0.75, then the participant will be assigned to the e-CBT treatment with pharmacotherapy. This NLP algorithm is patented by OPTT Inc. (International Patent System, PCT/US22/43514).

To gather the relevant data (i.e., participant compliance and change in depression severity, as evaluated by the PHQ-9), the Triage Module was designed. As previously explained, NLP of the participants' written accounts of their challenges with depression in the Triage Module will be used to calculate a Symptomatic Score. To verify the AI's treatment allocation logic, completion rate and the change in PHQ-9 scores in a sample of participants ($n = 190$) who were previously enrolled in e-CBT-only treatment was assessed. The decision-making algorithm determined that the e-CBT-only program was suitable for 62 of the 190 participants (32.63%). Within these 62 participants, 53.22% ($n = 33$) had completed the e-CBT-only program in its entirety and only 20% ($n = 12$) had a final PHQ-9 score > 14. Furthermore, the algorithm indicated that e-CBT with telephone calls would be suitable for 100 out of the 190 participants (52.63%). Of the 100 participants, 41.0% ($n = 41$) completed the whole round of e-CBT-only therapy and 31.0% ($n = 31$) had a final PHQ-9 score > 14. Lastly, the algorithm indicated that e-CBT with video call was appropriate for 28 out of 190 participants (14.74%). Of these 28 participants, 35% ($n = 10$) completed the whole round of e-CBT-only therapy and 40% ($n = 11$) had a final PHQ-9 score > 14. The logic of the AI's decision is therefore justified as those participants allocated to the e-CBT-only group had the highest percentage of completion and lowest percentage of final PHQ-9 scores >14 when completing e-CBT-only. Therefore,

minimal therapist intensity is required for these individuals and e-CBT-only is sufficient. Conversely, participants allocated to the e-CBT with video call had the lowest completion rates and highest rates of final PHQ-9 scores >14 when enrolled in e-CBT-only. These findings justify the AI's logic that greater therapist interaction is required. It is also important to note that demographic factors like age (below or above 40 years), biological sex (male or female) and income (less or more than CAD 50,000) did not have any significant effects on the number of sessions completed by participants ($p = 0.92$, 0.18, 0.90 for age, sex, and income respectively). The demographic factors did not affect the change in PHQ-9 score (i.e., the difference between the beginning and end of treatment scores) either ($p = 0.20$, 0.46, 0.39 for age, sex, and income respectively).

3.4 Treatment intensities

The three treatment intensities decided by the treatment arms (i.e., Arm 1: intake assessment or Arm 2: triage module using an AI algorithm) are:

1 e-CBT Program: The participant will submit their weekly homework and receive personalized feedback from their assigned therapist on the Online Psychotherapy Tool (OPTT). The feedback adds customization by acknowledging the participant's experiences in the past week and ensures the participant has understood the CBT concepts.
2 e-CBT Program + Call: In addition to the e-CBT program (see 1 above), the participant will receive a weekly phone/video call from their assigned therapist. The goal is to build on the therapeutic relationship and to add personalization with direct verbal encouragement. This phone call is limited to a one-time, 15–20-min call each intervention week (44). The purpose is to check with the patient on their treatment progress. The call will either be a secure phone or a video call, depending on the preference of the patient.
3 e-CBT Program + Pharmacotherapy: In addition to the e-CBT program (see 1 above), the participant will receive standard pharmacotherapy following DSM-5 guidelines. A pharmacotherapy allocation system has been developed (Figures 2, 3) that follows clinical guidelines.

While pharmacotherapy in isolation has shown efficacy for the treatment of depression, this study focused on an e-CBT intervention with augmentation options (i.e., e-CBT vs. e-CBT + Call, vs. e-CBT + Call + Pharmacotherapy). Hence, no treatment arm with just pharmacotherapy was included in the treatment intensities.

3.5 e-CBT platform

The web-based platform used for the study, the OPTT, is a secure, cloud-based, digital mental health platform (44). It complies with the Health Insurance Portability and Accountability Act, Personal Information Protection and Electronic Documents Act, and Service Organisation Control-2. In addition, all servers and databases are hosted in the Amazon Web Service Canada cloud infrastructure, which is managed by Medstack to ensure that all provincial and

FIGURE 2
The first phase of the drug-allocation process.
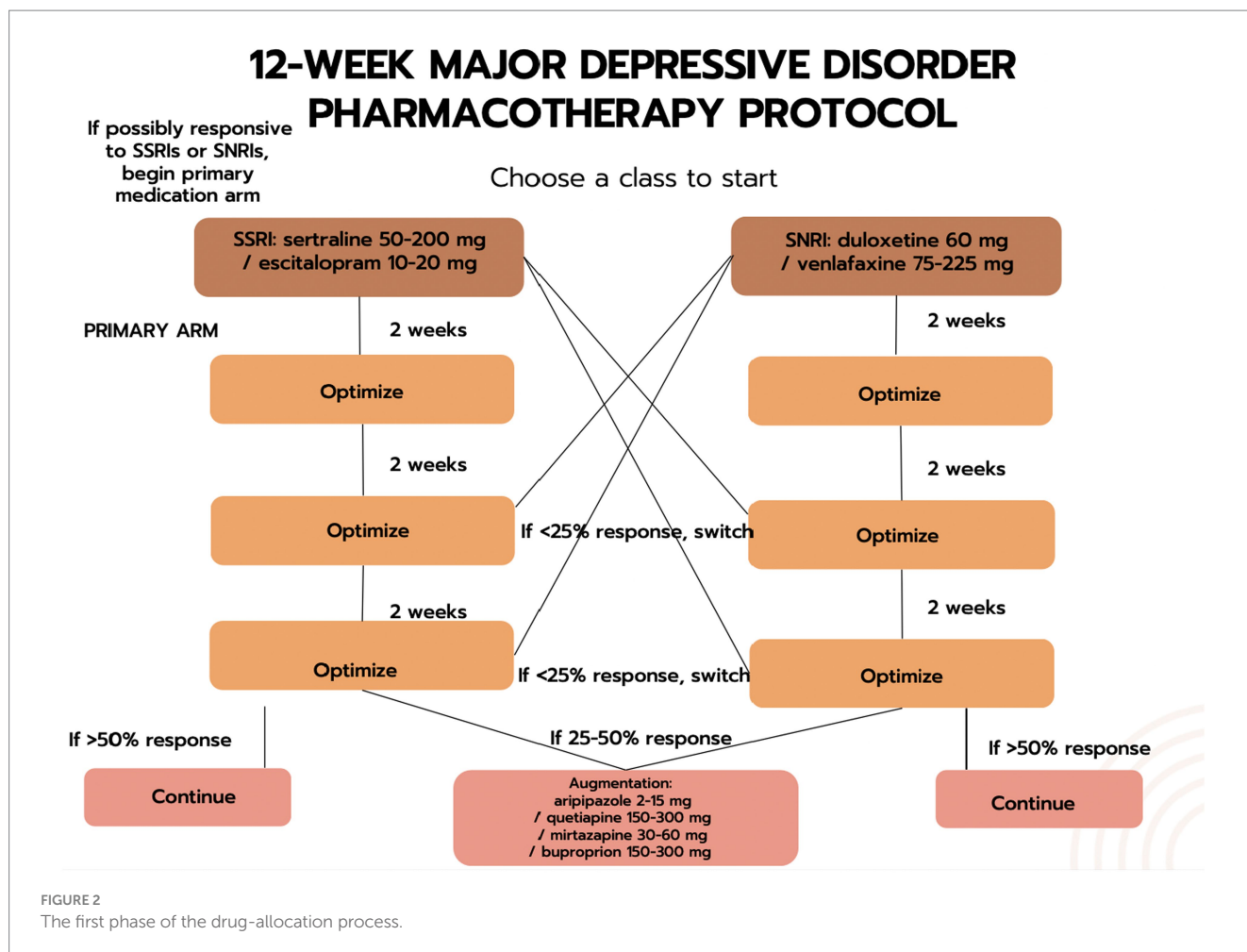
federal privacy and security regulations are met. OPTT does not collect any identifiable personal information or internet protocol addresses for privacy purposes. OPTT only collects anonymized metadata to improve its service quality and provide advanced analytics to the clinician team. OPTT encrypts all data, and no employees have direct access to participants' data. All encrypted backups are kept in the S3 storage that is dedicated to Queen's University, located in Kingston, Ontario, Canada.

3.6 e-CBT program

The e-CBT sessions used in this study include content based on cognitive restructuring and behavioural activation techniques (45). The purpose of the sessions is to help participants become aware of inaccurate or negative thinking patterns so that they can view challenging situations more clearly and respond to them effectively. The sessions prompt participants to understand their situation/environment and the resulting thoughts, behaviours, physical reactions, and feelings. This program aims to help change participants' negative and/or ineffective thoughts to more effective ways of thinking. As expressed in CBT, changing thoughts can subsequently affect feelings, behaviours, and physical reactions to stressful situations. The 13 e-CBT sessions are outlined below (Table 1).

3.7 Care provider

Each participant will be assigned a care provider who will provide feedback for their weekly sessions before the start of their next session. The assigned care provider will be independent of the multi-professional healthcare team that conducted the intake assessment. All care providers are trained in psychotherapy and have experience delivering electronic psychotherapy. They will be informed of each therapeutic session's aim and content. They will also continue receiving specialized training through webinars, workshops and exercises with feedback provided by the lead psychiatrist on the research team, a trained and licensed psychotherapist (22, 46). All care providers will be supervised by a trained psychotherapist and the lead psychiatrist, and all feedback will be reviewed before submission to the participants.

3.8 e-CBT weekly feedback

Weekly homework is reviewed by the independent care provider assigned to the participant, who will provide text-based personalized feedback on OPTT before the next weekly session. Participants and care providers can also communicate asynchronously on OPTT to relay any questions or concerns. The care providers will be provided with sample feedback templates and scripts for the telephone and video call sessions. Templates and scripts will be adapted from previous studies conducted

FIGURE 3
The second phase of the drug-allocation process.
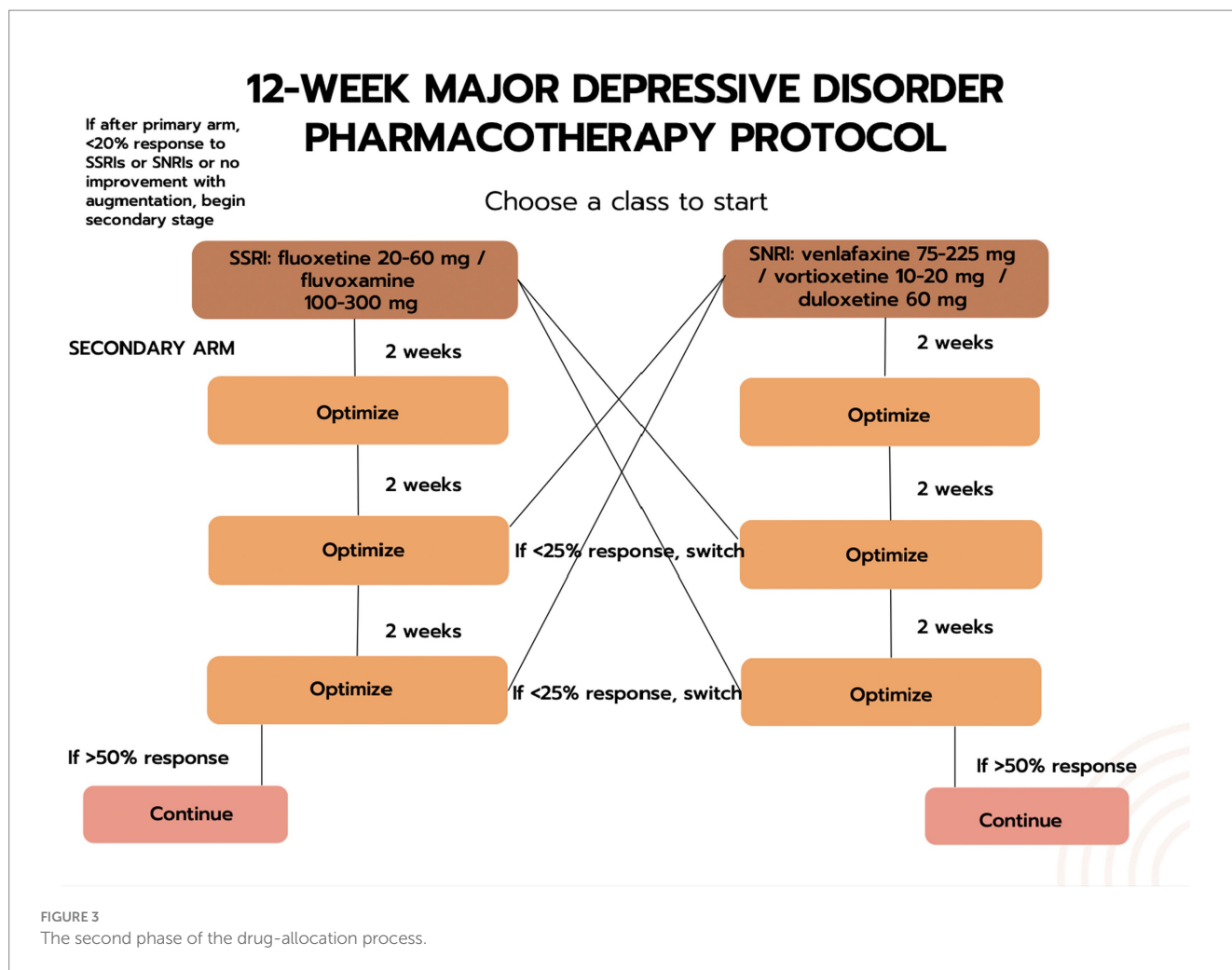
by the research team. Feedback templates and scripts will vary between sessions, and care providers will personalize them for each patient. The feedback templates follow a generic structure starting with, acknowledging the participant's time and effort since the last session, summarising the CBT concepts taught in the previous session, reviewing the event they explained in their homework, validating the participant's experience(s), and encouraging the participant to keep up with the sessions. The feedback is written in a letter format to increase personalization and build rapport with the participants.

3.9 Questionnaires & follow-up

Participants will complete a series of validated questionnaires at baseline including the Patient Health Questionnaire – 9 Item (PHQ-9), Sheehan Disability Scale (SDS), Quick Inventory of Depressive Symptoms (QIDS), Assessment of Quality of Life (AQoL-8D), and a demographic questionnaire. The PHQ-9 and QIDS questionnaire will be used to assess depressive symptoms and the AQoL-8D questionnaire will assess the participants' quality of life. These three questionnaires provide insight into the participant's mental status and perceived quality of life in the context of their depression. Additionally, the SDS will be used to assess the patient's functional impairment for the clinician group assessment during baseline. During the program, participants will complete the PHQ-9 every 3 weeks (i.e., weeks 4, 7, 10, and 13), and

the QIDS and AQoL-8D at weeks 7 and 13. All questionnaires (i.e., PHQ-9, QIDS, and AQoL-8D questionnaires) will be filled at the 3, 6, and 12-month follow-up periods to measure post-study outcomes.

3.10 Sample size determination

In our previous clinical trials and gathered data (47), the average PHQ-9 score changed from 16.2 before e-CBT to 11.48 after 12 sessions of e-CBT (joint standard deviation; SD = 5.45). Based on these numbers, the effect size (Hedges' G) is equal to 0.86. Given the effect size and a power of 0.8, we would need 14 participants to observe a significant effect in a paired sample t-test. Considering the completion rate of 45%, we would require 31 participants in each group to observe significant results. Given the two treatment arms and three treatment intensities, we would require 186 participants to detect significant clinical change across all groups. We believe this sample size is large enough to perform further analyses on factors such as the role of sex and gender in treatment efficacy and completion rate.

To follow up on this sample size determination, we will use previously collected data as mentioned above to simulate the population and use a random sampling method to determine the effect size for our power calculation. We will do this by simulating a population based on the mean and standard deviation of our previous e-CBT studies and collecting data, then randomly sampling

TABLE 1 Session titles and descriptions for the e-CBT program.

Session	Description
1 – What is Depression?	Provides expectations for the course and introduces CBT and depression.
2 – The 5-Part Model	Introduces the concept of the 5-Part Model and how a situation, thoughts, feelings, physical reactions, and behaviours are connected and how they interact.
3 – Sleep Hygiene	Focuses on sleep habits and provides a variety of tips and strategies to use to increase sleep hygiene and get better rest.
4 – Strategies for Stressful Situations	Provides an overview of helpful strategies that can be used in stressful situations including pleasurable activities and helpful breathing techniques.
5 – Thoughts, Feelings, Behaviour, Physical Reactions & Environment	Provides a further detailed exploration of the 5 Part Model and how changes in one area can affect the other 4 parts.
6 – The Thought Record	Highlights the first three columns of the Thought Record; a tool used to help understand the connection between feelings, behaviours, and thoughts. The first three columns include the situation, followed by the feelings and automatic thoughts associated with the situation.
7 – Automatic Thoughts	This delves into the role of automatic thoughts and how they influence feelings. The focus of this session is to understand how to identify automatic thoughts and specifically identify the most dominant idea, or "hot thought" when presented with a stressful situation. Common thinking errors are also discussed in this session.
8 – Activity Scheduling	Provides a break from learning about the Thought Record and instead explains how to use an Activity Record; a tool designed to record and plan weekly activities. This session focuses on how tracking activities can inform mood changes and reinforce the scheduling of pleasurable activities.
9 – Evidence	Focuses on the fourth and fifth columns of the Thought Record, which is designed to help gather the information that supports or does not support the identified hot thought.
10 – Alternative & Balanced Thinking	Focuses on the final two columns of the Thought Record which reflects on the evidence columns to help find an alternative or balanced view of the situation. The last column invites the viewer to re-rate their feelings based on the completion of the Thought Record.
11 – Experiments	Explains the importance of conducting experiments to start believing alternative or balanced thoughts from the Thought Record and begin initiating changes in ineffective thinking patterns.
12 – Action Plans	Centred around identifying a problem that needs to be solved and provides a framework on how to create a plan for solving the problem.
13 – Review	The final session is a review of the course and summarizes the main CBT concepts and tools that have been taught throughout the program.

participants to each treatment arm. This simulation will be developed using Python code. Based on this random sampling, we will calculate the effect size, set the power of 0.8, and consider the completion rate of 45% to determine the sample size required to detect a significant clinical change across the groups.

4 Statistical analysis

Initially, all data will be examined for missing, nonsensical, and outlying variables. Missing data will not be attributed and will be treated as missing. All statistical analyses will be performed at the end of the trial and will consider a significance level of 0.05 and 95% confidence intervals. Analyses will be conducted to account for and assess differences across sex and gender identifications, and how these factors affect the adaptability and efficacy of e-CBT.

4.1 Hypothesis 1: outcomes in depressive symptoms of treatment arm 1 will be comparable to treatment arms 2 and 3

Clinical outcomes of the healthcare teams' decision (arm 1) versus the AI's decision (arm 2), will be compared by assessing the treatment completion rate, number of completed sessions, and change in questionnaire scores. The completion rate across the two study arms and three levels of care will be calculated using a $2 \times 3 \times 2$ contingency matrix (i.e., healthcare team vs. AI $\times$ 3 different treatment intensities $\times$

completers vs. dropouts) and chi-square analysis to compare within and between completion rates in the two arms. We will use a 2×3 (i.e., healthcare team vs. AI $\times$ 3 different levels of care) two-way analysis of variance (ANOVA) to compare the number of completed sessions across participant groups and care levels. Questionnaires are collected at 3 different times; at the start (for all participants), midway (for those who completed at least 6 sessions), and end of the study (for those who completed the treatment program). We will use a $2 \times 3 \times 3$ (i.e., healthcare team vs. AI $\times$ 3 different levels of care $\times$ 3 time points) linear mixed-effects model for each questionnaire to compare clinical outcomes within and between groups and times. A mixed-effects model analysis will account for any missing data caused by participant dropout. Finally, we will assess if demographic covariables affect clinical outcomes using multiple $2 \times 2 \times 3 \times 3$ (demographic factors (e.g., male vs. female) $\times$ healthcare team vs. AI $\times$ 3 different levels of care $\times$ 3-time points for the questionnaire's score) mixed-effects model to account for each demographic variable (i.e., sex, gender, age, and income). Bonferroni corrections will follow the ANOVA analyses in *post hoc* analyses.

4.2 Hypothesis 2: AI-based clinical decision-making will provide suggestions comparable to a multi-professional care team

To compare decision-making between the healthcare team (Arm 1) and AI (Arm 2), we will calculate the percentage of identical choices (e.g., arm 1 e-CBT intensity 1 and arm 2 e-CBT intensity) and use a

3 × 3 crosstab analysis (i.e., 3 e-CBT intensities for treatment arm 1 versus 3 e-CBT intensities for treatment arm 2). A chi-square test will be used to evaluate the significance and whether the two models (i.e., healthcare team vs. AI decision) are related or independent. From a machine learning perspective, we will calculate the model's precision, recall, and F1 score, assuming the healthcare team's performance level is the baseline level. The precision, recall, and F1 scores provide insight into the quality of the machine-learning model (48).

4.3 Hypothesis 3: the AI approach will decrease the time and cost commitment of providing e-CBT

The cost of initial assessments (i.e., manual participant stratification by the healthcare team vs. the AI) and the different care intensities (i.e., only e-CBT, e-CBT + calls, or e-CBT + pharmacotherapy) will also be evaluated. These costs will be evaluated from a health sector and societal perspective (38). For this analysis we will use (1) healthcare costs paid by the government of Ontario as well as healthcare utilization costs self-reported by participants using a resource-use questionnaire (39, 40); (2) medical and productivity costs due to work absence, evaluated using the Questionnaire on Healthcare Consumption and Productivity Losses for patients with a Psychiatric Disorder (TiC-P) (49); (3) intervention costs including staff time-commitment to deliver interventions, development and maintenance costs, and time commitment to assess and enroll participants into the online program. The costs will be expressed in Canadian dollars (CAD). We will use a 2 * 3 (i.e., healthcare team vs. AI * 3 different levels of care) two-way ANOVA to compare costs within and between arms.

5 Discussion

Artificial intelligence and providing patients with varying intensities of care can increase the efficiency of mental health care services. This study aims to determine a cost-effective method to decrease depressive symptoms and increase treatment adherence to online psychotherapy by allocating the correct intensity of therapist care for individuals diagnosed with depression. This will be done by comparing a decision-making machine learning algorithm to a multi-professional care team. This approach aims to accurately allocate care tailored to each patient's needs, allowing for more efficient use of resources with the convergence of technologies and healthcare.

Ethics statement

The studies involving humans were approved by Queen's University Health Science and Affiliated Teaching Hospitals Research Ethics Board. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

Conflict of interest

NA and MO have co-founded the care delivery platform in use (i.e., OPTT) and have ownership stakes in OPTT Inc.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. James SL, Abate D, Abate KH, Abay SM, Abbafati C, Abbasi N, et al. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: a systematic analysis for the global burden of disease study 2017. *Lancet*. (2018) 392:1789–858. doi: 10.1016/S0140-6736(18)32279-7

2. Depression. (2021). World Health Organization, 13. Available at: https://www.who.int/news-room/fact-sheets/detail/depression.

3. Mekonen T, Chan GCK, Connor JP, Hides L, Leung J. Estimating the global treatment rates for depression: a systematic review and meta-analysis. *J Affect Disord*. (2021) 295:1234–42. doi: 10.1016/j.jad.2021.09.038

4. Waitzfelder B, Stewart C, Coleman KJ, Rossom R, Ahmedani BK, Beck A, et al. Treatment initiation for new episodes of depression in primary care settings. *J Gen Intern Med*. (2018) 33:1283–91. doi: 10.1007/s11606-017-4297-2

5. Canada SC G. (2019). Mental health care needs, 2018; 7. Available at: https://www150.statcan.gc.ca/n1/pub/82-625-x/2019001/article/00011-eng.htm.

6. Moroz N, Moroz I, D'Angelo MS. Mental health services in Canada: barriers and cost-effective solutions to increase access. *Health Manage Forum*. (2020) 33:282–7. doi: 10.1177/0840470420933911

7. Luo C, Sanger N, Singhal N, Pattrick K, Shams I, Shahid H, et al. A comparison of electronically delivered and face-to-face cognitive behavioural therapies in depressive disorders: a systematic review and meta-analysis. *EClinicalMedicine*. (2020) 24:100442. doi: 10.1016/j.eclinm.2020.100442

8. Weiner L, Berna F, Nourry N, Severac F, Vidailhet P, Mengin AC. Efficacy of an online cognitive behavioral therapy program developed for healthcare workers during the COVID-19 pandemic: the REduction of STress (REST) study protocol for a randomized controlled trial. *Trials*. (2020) 21:870. doi: 10.1186/s13063-020-04772-7

9. Sztein DM, Koransky CE, Fegan L, Himelhoch S. Efficacy of cognitive behavioural therapy delivered over the internet for depressive symptoms: a systematic review and meta-analysis. *J Telemed Telecare*. (2018) 24:527–39. doi: 10.1177/1357633X17717402

10. Andersson G, Cuijpers P. Internet-based and other computerized psychological treatments for adult Depression: a meta-analysis. *Cogn Behav Ther*. (2009) 38:196–205. doi: 10.1080/16506070903318960

11. Musiat P, Tarrier N. Collateral outcomes in e-mental health: a systematic review of the evidence for added benefits of computerized cognitive behavior therapy interventions for mental health. *Psychol Med*. (2014) 44:3137–50. doi: 10.1017/S0033291714000245

12. Schmidt ID, Forand NR, Strunk DR. Predictors of dropout in internet-based cognitive Behavioral therapy for Depression. *Cogn Ther Res*. (2019) 43:620–30. doi: 10.1007/s10608-018-9979-5

13. Mausbach BT, Moore R, Roesch S, Cardenas V, Patterson TL. The relationship between homework compliance and therapy outcomes: an updated meta-analysis. *Cogn Ther Res*. (2010) 34:429–38. doi: 10.1007/s10608-010-9297-z

14. Serobatse MB, Plessis ED, Koen MP. Interventions to promote psychiatric patients' compliance to mental health treatment: a systematic review. *Health SA Gesondheid*. (2014) 19:1–10. doi: 10.4102/hsag.v19i1.799

15. Olver ME, Stockdale KC, Wormith JS. A meta-analysis of predictors of offender treatment attrition and its relationship to recidivism. *J Consult Clin Psychol*. (2011) 79:6–21. doi: 10.1037/a0022200

16. Swift JK, Greenberg RP. A treatment by disorder meta-analysis of dropout from psychotherapy. *J Psychother Integr*. (2014) 24:193–207. doi: 10.1037/a0037512

17. Wierzbicki M, Pekarik G. A meta-analysis of psychotherapy dropout. *Prof Psychol Res Pract*. (1993) 24:190–5. doi: 10.1037/0735-7028.24.2.190

18. Rost T, Stein J, Löbner M, Kersting A, Luck-Sikorski C, Riedel-Heller SG. User acceptance of computerized cognitive behavioral therapy for depression. *Syst Rev J Med Internet Res*. (2017) 19:e309. doi: 10.2196/jmir.7662

19. Fernandez E, Salem D, Swift JK, Ramtahal N. Meta-analysis of dropout from cognitive behavioral therapy: magnitude, timing, and moderators. *J Consult Clin Psychol*. (2015) 83:1108–22. doi: 10.1037/ccp0000044

20. Haller E, Watzke B. The role of homework engagement, homework-related therapist Behaviors, and their association with depressive symptoms in telephone-based CBT for Depression. *Cogn Ther Res*. (2021) 45:224–35. doi: 10.1007/s10608-020-10136-x

21. Gulliver A, Griffiths KM, Christensen H. Perceived barriers and facilitators to mental health help-seeking in young people: a systematic review. *BMC Psychiatry*. (2010) 10:113. doi: 10.1186/1471-244X-10-113

22. Alavi N, Omrani M. *Online cognitive behavioral therapy: an e-mental health approach to depression and anxiety*. Switzerland: Springer International Publishing (2019).

23. Karyotaki E, Ebert DD, Donkin L. Do guided internet-based interventions result in clinically relevant changes for patients with depression? An individual participant data meta-analysis. *Clin Psychol Rev*. (2018) 63:80–92. doi: 10.1016/j.cpr.2018.06.007

24. Arns M, Dijk H, Luykx JJ, Wingen G, Olbrich S. Stratified psychiatry: tomorrow's precision psychiatry? *Eur Neuropsychopharmacol*. (2022) 55:14–9. doi: 10.1016/j.euroneuro.2021.10.863

25. West M, Almo-Metcalfe B, Dawson J. (2012). Effectiveness of multi-professional team working (MPTW) in mental healthcare. Available at: http://www.netscc.ac.uk/hsdr/files/project/SDO_FR_08-1819-215_V01.pdf.

26. Felker BL, Barnes RF, Greenberg DM. Preliminary outcomes from an integrated mental health primary care team. *Psychiatr Serv*. (2004) 55:442–4. doi: 10.1176/appi.ps.55.4.442

27. Bennett CC, Hauser K. Artificial intelligence framework for simulating clinical decision-making: a Markov decision process approach. *Artif Intell Med*. (2013) 57:9–19. doi: 10.1016/j.artmed.2012.12.003

28. Lysaght T, Lim HY, Xafis V, Ngiam KY. AI-assisted decision-making in healthcare. *Asian Bioeth Rev*. (2019) 11:299–314. doi: 10.1007/s41649-019-00096-0

29. Delgadillo J, Moreea O, Lutz W. Different people respond differently to therapy: a demonstration using patient profiling and risk stratification. *Behav Res Ther*. (2016) 79:15–22. doi: 10.1016/j.brat.2016.02.003

30. Lorenzo-Luaces L, DeRubeis RJ, Straten A, Tiemens B. A prognostic index (PI) as a moderator of outcomes in the treatment of depression: a proof of concept combining multiple variables to inform risk-stratified stepped care models. *J Affect Disord*. (2017) 213:78–85. doi: 10.1016/j.jad.2017.02.010

31. Saunders R, Cape J, Fearon P, Pilling S. Predicting treatment outcome in psychological treatment services by identifying latent profiles of patients. *J Affect Disord*. (2016) 197:107–15. doi: 10.1016/j.jad.2016.03.011

32. Kessler RC, van Loo HM, Wardenaar KJ, Bossarte RM, Brenner LA, Ebert DD, et al. Using patient self-reports to study heterogeneity of treatment effects in major depressive disorder. *Epidemiol Psychiatr Sci*. (2017) 26:22–36. doi: 10.1017/S2045796016000020

33. Cresswell K, Callaghan M, Khan S, Sheikh Z, Mozaffar H, Sheikh A. Investigating the use of data-driven artificial intelligence in computerised decision support systems for health and social care: a systematic review. *Health Informatics J*. (2020) 26:2138–47. doi: 10.1177/1460458219900452

34. Smith H, Fotheringham K. Artificial intelligence in clinical decision-making: rethinking liability. *Med Law Int*. (2020) 20:131–54. doi: 10.1177/0968533220945766

35. Löwe B, Unützer J, Callahan CM, Perkins AJ, Kroenke K. Monitoring depression treatment outcomes with the patient health questionnaire-9. *Med Care*. (2004) 42:1194–201. doi: 10.1097/00005650-200412000-00006

36. Rush AJ, Trivedi MH, Ibrahim HM, Carmody TJ, Arnow B, Klein DN, et al. The 16-item quick inventory of depressive symptomatology (QIDS), clinician rating (QIDS-C), and self-report (QIDS-SR): a psychometric evaluation in patients with chronic major depression. *Biol Psychiatry*. (2003) 54:573–83. doi: 10.1016/S0006-3223(02)01866-8

37. Richardson J, Sinha K, Iezzi A, Khan MA. Modelling utility weights for the assessment of quality of life (AQoL)-8D. *Qual Life Res*. (2014) 23:2395–404. doi: 10.1007/s11136-014-0686-8

38. Touré M, Kouakou CRC, Poder TG. Dimensions used in instruments for QALY calculation: a systematic review. *Int J Environ Res Public Health*. (2021) 18:4428. doi: 10.3390/ijerph18094428

39. Chatterton ML, Mihalopoulos C, O'Neil A. Economic evaluation of a dietary intervention for adults with major depression the "SMILES" trial. *BMC Public Health*. (2018) 18:599. doi: 10.1186/s12889-018-5504-8

40. Glick HA, Doshi JA, Sonnad SS, Polsky D. *Economic evaluation in clinical trials*. Oxford, United Kingdom: OUP Oxford (2014).

41. American Psychiatric Association. *Diagnostic and statistical manual of mental disorders, Fifth Edition (DSM-5)*. 2013.

42. Leon AC, Olfson M, Portera L, Farber L, Sheehan DV. Assessing psychiatric impairment in primary care with the Sheehan disability scale. *Int J Psychiatry Med*. (1997) 27:93–105. doi: 10.2190/T8EM-C8YH-373N-1UWD

43. Sheehan KH, Sheehan DV. Assessing treatment effects in clinical trials with the Discan metric of the Sheehan disability scale. *Int Clin Psychopharmacol*. (2008) 23:70–83. doi: 10.1097/YIC.0b013e3282f2b4d6

44. Digital mental health platform with the human touch. Available at: https://www.optt.health/.

45. Alavi N, Stefanoff M, Hirji A, Khalid-Khan S. Cognitive behavioural therapy through powerpoint: efficacy in an adolescent clinical population with depression and anxiety. *Int J Pediatr*. (2018) 2018 1396216. doi: 10.1155/2018/1396216

46. Alavi N, Hirji A. The efficacy of powerpoint-based CBT delivered through email: breaking the barriers to treatment for generalized anxiety disorder. *J Psychiatr Pract*. (2020) 26:89–100. doi: 10.1097/PRA.0000000000000455

47. Alavi N, Stephenson C, Yang M. Feasibility and efficacy of delivering cognitive behaviour therapy through online psychotherapy tool (OPTT) for depression: protocol. *JMIR Res Protoc*. (2021) 12:e27489. doi: 10.2196/27489

48. Sokolova M, Japkowicz N, Szpakowicz S. Beyond accuracy, F-score and ROC: a family of discriminant measures for performance evaluation In: A Sattar and BH Kang, editors. *AI 2006: advances in artificial intelligence*, vol. *4304*: Springer (2006). 1015–21.

49. Ho FYY, Chung KF, Yeung WF, Ng THY, Cheng SKW. Weekly brief phone support in self-help cognitive behavioral therapy for insomnia disorder: relevance to adherence and efficacy. *Behav Res Ther*. (2014) 63:147–56. doi: 10.1016/j.brat.2014.10.002

Glossary

AI	Artificial intelligence
ANOVA	Analysis of variance
AQoL-8D	Assessment of Quality of Life – 8 Dimension
CAD	Canadian dollars
CBT	Cognitive behavioural therapy
DSM-5	Diagnostic and Statistical Manual of Mental Disorders – 5th Edition
e-CBT	Electronically delivered cognitive behavioural therapy
MDD	Major depressive disorder
MINI	Mini-International Neuropsychiatric Interview
NLP	Natural language processing
OPTT	Online Psychotherapy Tool
PHQ-9	Patient Health Questionnaire – 9 Item
QALY	Quality-adjusted life years
QIDS	Quick Inventory of Depressive Symptoms
SD	Standard deviation
SDS	Sheehan Disability Scale
TiC-P	Questionnaire on Healthcare Consumption and Productivity Loss for Psychiatric Patients