ClinicalTrials.gov Cover Page

Project Title
Automated Symptom Tracking for Measurement-Based Care of Depression

NCT Number
NCT03990389

Document Version Date
May 19, 2019

Project Title
Automated Symptom Tracking for Measurement-Based Care of Depression

Principal Investigator
David G. Beiser, MD (University of Chicago)

Background & Significance

Perinatal depression is one of the most common medical complications during pregnancy and the postpartum period.[1] The prevalence of women that experience antenatal and postpartum depression are 12% and 10-15% accordingly.[2] While depression is one of the most treatable mental illnesses, the rates of mothers receiving treatment is far too low.[3] In addition, the annual cost of not treating a mother with depression is $7,200 making the annual cost of untreated depression in mothers $5.7 billion a year.[4]

Gaps in depression treatment

Efficacious treatments exist, yet barely half of Americans, and only 40% of African Americans, with major depressive disorder (MDD) receive treatment.[5,6] Factors contributing to low treatment rates are manifold but are known to be rooted in a national shortage of mental health practitioners that is projected to worsen over the next decade.[7,8] In the face of this shortage, primary care providers, who exclusively manage 73% of all cases of MDD, have become the *de facto* mental health providers for much of society.[9,10] For patients fortunate enough to access treatment, most do not receive guideline-concordant, that is, measurement-based care (MBC).

Measurement-based care is the key – but it is difficult to deliver

MBC, which involves the use of quantitative assessments for depression screening, diagnosis, and symptom monitoring,[11] is recognized as the key to depression management.[12–14] MBC can be thought of as an iterative “treat to target” strategy where a diagnosis is made, treatment is initiated, symptoms are measured over time, and treatments are modified to optimize symptom burden. It is highly effective and is associated with a much lower relapse rate (25%) as compared with usual care (75%).[15]

Providers struggle to deliver guideline-concordant depression care as only 22% of patients who receive care receive what would be considered adequate treatment.[5,6] This rate drops to 14% in African Americans -- highlighting a disparity in MDD treatment effectiveness.[6] Barriers to providing MBC include the additional resources needed for administration, collection, scoring, and recording of patient assessment instruments with some successful models of MBC, often utilizing a dedicated care manager to coordinate and communicate patient results.[10] Lacking the access to MBC, many patients with MDD receive a single prescription for a low-dose antidepressant medication and little follow-up or adjuvant talk therapy.[5,6] It is reasonable to hypothesize that low rates of MBC adoption are responsible for low rates of adequate MDD care. New approaches are needed to address the gap in depression care.

What is a chatbot?

The general definition of a chatbot is a computer program designed to simulate conversation with human users, especially over the Internet. Chatbots vary in design and function. We propose the development and validation of a chatbot mediated approach for implementing MBC that we call CB-MBC. CB-MBC is designed to orient the subject to the study and methods of interacting with the chatbot. Users will respond to the chatbot using reply buttons (not free text) and will have a persistent menu at the bottom of their screen to be used with questions or when further assistance is needed.

Rationale for using Facebook Messenger app

Facebook Messenger was selected as the method of administration for this study as it is extensively used by this study population demographic and it supports chatbot technology. The most popular form of messaging within the U.S. and Canada is overwhelmingly Facebook Messenger.[16] Facebook is most used by people ages 18-44 with the highest usage ages being 25-34).[17] This is the ideal age range for participants in our study population. In 2016 it was reported that Facebook users in the U.S. spent approximately 50 minutes a day on the website, an increase from 40 minutes the previous year.[18] There are several other messaging apps with end-to-end encryption including Signal, VIBER, Telegram and WhatsApp. However, we are restricted to using a platform that supports chatbots thereby eliminating Signal and WhatsApp. Telegram and VIBER are bot friendly but far more popular in Eastern Europe, Iran & Uzbekistan, than the U.S.

Rationale for Studying Perinatal Population

The CB-MBC chatbot is essentially, a depression symptom tracker delivered via a smartphone application called Facebook Messenger. In order to maximize the treatment exposure of this intervention, it is important to apply it within a population that is likely to own a smartphone and use Facebook Messenger in their daily lives. Data from the Pew Research Center suggests that smartphone ownership and social networking application use is highest among people ages 18-29.[19] In addition, women are more likely to use their smartphone for health-related messaging than men.[20] The perinatal population matches this demographic quite closely and also tends to be younger than the general adult psychiatry population and is thus more likely to engage with the CB-MBC intervention. In addition, women often face a myriad of practical barriers related to childcare and transportation when it comes to making follow-up clinic visits.[21] This data point suggests that remote symptom tracking, as enabled by the CB-MBC intervention, might have more value in this less mobile population.

Safety and Efficacy of Smartphone Apps in Mental Health Care Delivery

While it is still very early in the development, evaluation and adoption curves of smartphone-based digital health platforms, there is much interest and enthusiasm for their applications in mental health.[22] Like the proposed CB-MBC application, many of the existing mental health applications involve a component of symptom tracking. At home symptom tracking and journaling are commonly recommended parts of a mental health treatment plan. There is no literature to support the notion that at home symptom tracking or journaling are associated with worsening depression or suicide risk in any population of patients. In fact, it should be noted that the FDA considers mobile applications that enable symptom tracking to be so low risk that it does not intend to regulate such applications as medical devices.[23]

There is a growing literature of scientific studies that supports the safety and efficacy of smartphone mental health applications in non-pregnant populations. For example smartphone-based self-monitoring apps for symptom detection, monitoring and cognitive-based therapy have been previously studied in a variety of patient populations with serious mental illness.[24–28] One fully-remote trial of 626 adults with mild to moderate depression demonstrated differential treatment improvements for patients using an app designed to deliver cognitive behavioral therapy.[29] Another group recently reported substantial improvement in symptoms through the use of interactive apps for treating depression and anxiety.[30] Another study compared the efficacy of using either a mobile phone- or computer-based platform for delivering cognitive behavioral therapy. A Facebook Messenger chatbot, similar to the proposed CB-MBC platform, has also been recently studied in students who self-identify as having depression.[31] None of these studies reported any safety concerns.

The safety and feasibility of using digital health apps have also been previously studied in vulnerable pregnant females. In one observational trial of a smartphone app for tracking gestational hypertension and weight gain patients exhibited high patient engagement with the platform and reported high satisfaction with the experience.[32] Another randomized control trial in a vulnerable population of pregnant women with depressive symptomatology at < 32 weeks gestation, patients randomized to a mood tracking mobile application as compared to a control group receiving either usual care email communication or a patient portal mobile app exhibited greater self-rated ability to manage their own health better than controls.[33] Similar to the literature in non-pregnant populations, neither of these studies reported any safety concerns. In addition, digital health applications that facilitate self-monitoring and empower patients in the doctor-patient relationship are also highly valued and accepted by pregnant women.[34] While much remains to study regarding the efficacy, effectiveness, and cost-effectiveness of digital mental health interventions, we conclude that a symptom tracking intervention, such as CB-MBC, poses an acceptably low risk of physical harm to pregnant and non-pregnant patients with depression.

The proposed trial will bridge a significant "effectiveness" gap in depression treatment

We will test the use of automated symptom tracking using a "chatbot" for delivering MBC via the Facebook Messenger platform. Tapping the pervasive presence of smartphone technology and ubiquity of text-messaging, this approach will reduce the burden of MBC delivery and improve safety.

The proposed trial will improve scientific knowledge and clinical practice

High-frequency depression severity data yielded from this feasibility trial will reveal new insights into the day-to-day mood variations of patients being treated for depression.

The methods and services that drive this field will be changed if our aims are achieved

Data from this pilot trial will provide the basis a multi-center trial for CB-MBC which could change our approach to depression care.

Specific Aims

We propose the development and validation of a chatbot-mediated approach for implementing MBC that we call Chatbot-Measurement Based Care (CB-MBC.) We hypothesize that the use of CB-MBC, which combines a conversational interface with state-of-the-art quantitative assessments, will improve depression symptom severity for patients with depression.

Specific Aim 1: Patient subjects in the intervention group will exhibit reduced depression severity compared to usual care.

Specific Aim 2: Patient subjects in the intervention group, compared to usual care, will report

a. lower side effect burden
b. greater behavioral self-efficacy
c. greater maternal functioning

Specific Aim 3: Obstetric providers will report satisfaction with the weekly email trends on their patients' depression severity.

Methods

Study Sites
University of Chicago is the lead site. This study is also being conducted at Northshore under the guidance of Dr. Richard Silver.

Subject Population
Provider Subject Cohort: Obstetric care providers (obstetricians, nurse midwives, residents and attending physicians) from University of Chicago Medical Center OBGYN.

Patient Subject Cohort: Postnatal women who score 12 or more on the Edinburgh Postnatal Depression Scale at obstetric clinics at University of Chicago Medical Center.

Patient Subject Inclusion and Exclusion Criteria

Table 2. Eligibility criteria	
Inclusion Criteria	**Exclusion Criteria**
(1) Age ≥ 18 years and able to and demonstrate English reading literacy of at least 8th-grade level (REALM-R ≥ 6) (2) Postnatal women diagnosed or receiving treatment for depression with a severity-based entry criterion of moderate to severe (EPDS ≥ 12) (3) Willing to participate and able to give written informed consent (4) Own a smartphone with a data plan (5) Have a Facebook account and Facebook Messaging app on phone (or be willing to create account/download app during enrollment) (6) Sufficient cognitive ability to provide self-report data on a computer touchscreen/ standard computer with minimal assistance	(1) Subjects with documented dysthymia or Axis II diagnoses. (2) Subjects with self-reported or documented history of: • anorexia or bulimia • obsessive compulsive disorder • prior hospitalization for suicidal ideation (3) Active suicidality as determined by clinician† (4) Non-English speakers .

† Suicidality may be incidentally disclosed during screening or identified via Question #10 of the EPDS ("The thought of harming myself has occurred to me:") An answer to Question #10 of "hardly ever", "sometimes", or "yes quite often", will trigger an evaluation for active suicidality by the clinician as described in Section VI (below).

Study-Wide Number of Subjects
Phase 1- Conducted at Univ. of Chicago only: 20 provider subjects and 5 patient subjects for a total enrollment of approximately 30 subjects.

Phase 2- 20 provider subjects and 40 patient subjects from each study site for a total enrollment of approximately 60 subjects per site (120 subject total).

Recruitment Methods
University of Chicago Medical Center – A social worker unaffiliated with the study will notify the research coordinator of any potential study participants for screening. We will recruit outpatients and inpatients.

A research coordinator unaffiliated with the patient's medical care will discuss possible study participation. The research coordinator will then provide information about the study and obtain written consent prior to screening subject. For screening, coordinators will confirm that the subject owns a smartphone in their possession with a data plan. Next, they will administer a health literacy screen (REALM-R), followed by a baseline depression severity measurement (EPDS). Subjects that fit the eligibility criteria will continue further in the study. The research coordinator will record information on the Chatbot Enrollment Log for each patient who is actually approached by the research coordinator for study enrollment. This will enable us to assess the possibility of selection bias. The enrollment log will not collect PHI.

The research coordinator will develop a separate log with patient names, date of birth (and study IDs, when applicable). This log will help to ensure that patients are not approached multiple times without permission, will facilitate future contacts with study participants, etc. This log will be kept in a locked area (e.g. desk drawer) accessible only to approved study personnel.

Provider Subjects: The providers will be reminded that they are not obligated to participate and that declining to participate will not in any impact employment. The research assistant will email or fax the consent form to the providers, depending on their preference. The providers will have the opportunity to consider the study before deciding whether to participate. Written consent will be obtained for interested subjects. Subjects will be asked to sign, date, and securely email the signed consent form to the research assistant.

Study Setting

Subjects will be recruited from the obstetric clinics and as inpatients at the University of Chicago Medical Center.

Trial Design

Phase 1 (Completed): We begin with a feasibility study in which all subjects receive CB-MBC care for 3 weeks. We will review the results and make any necessary changes in the chatbot platform before proceeding with the Phase 2 clinical trial. In Phase 2, subjects are randomized to Group 1 (usual care) or Group 2 (CB-MBC) and followed for 12 weeks, which is the outer-limit of the acute phase of depression recovery.[36] The primary outcome measure is depressive symptom severity. Exploratory secondary outcomes include side effect burden, patient-engagement, patient satisfaction, and provider satisfaction.

Interventions

Phase 1 (Completed) - All subjects will be asked to connect with the chatbot via Facebook Messenger. Subjects will then receive a 10-minute in person orientation to the chatbot and will be told to expect weekly messages from the chatbot asking them to complete a depression severity and side effect burden assessments. After one week, they will receive a check-in call from a study coordinator to answer any questions regarding use of the chatbot. Providers will receive weekly reports summarizing patient-level responses on administered assessments and are encouraged to respond if clinically-indicated. Upon completion of the study, the subject will be asked to participate in a semi-structured debrief interview focused on the user's experience.

Phase 2 - This is an open-label randomized control trial comparing depression severity in subjects entering treatment for major depression. Using a random number generator (www.random.org) subjects will be separated into the following groups:

Group 1: Subjects in this group will receive a monthly phone call reminding them to complete surveys that were securely emailed. The five (5) monthly surveys will ask about mood,

medication adherence if any, side effects from any medication, maternal functioning, and maternal confidence. The surveys could take up to 20 minutes to complete. The survey answers are stored in a secure database. Information collected will not be shared with the obstetric provider.

Group 2: Subjects in group 2 will receive a short explanation on how to use a depression "chatbot." A chatbot is like having a chat with a robot. The chatbot will use Facebook Messenger. The research assistant will discuss FB/FB Messenger privacy settings with each study participant prior to enrollment. This includes verification that all settings related to FB Messenger logs and data sharing (e.g., call and text logs for Android phones) will be switched to "OFF".

Two of the surveys (mood and side effects of medication) will be administered weekly by chatbot, and should take about 5 minutes to complete. The rest of the surveys will be emailed monthly. Those could take up to 20 minutes to complete. The survey answers are stored in a secure database. Subjects will be given the option of receiving daily medication reminders via the chatbot. Following the first week, subjects will receive a phone call from a research assistant to discuss any questions about the chatbot. Feedback on the information collected by the chatbot like graphs and trends will be sent to subjects and their obstetric provider each week. Some providers may choose to contact subjects to adjust treatment based on the email.

Providers will receive weekly secure emails summarizing patient responses on administered scales and encouraged to respond if clinically-indicated.

Subjects in Group 2 will also be provided with feedback within the chatbot messenger interface. Specifically, we plan to provide subjects with a variety of feedback including graphs of depression trends, text-based trend interpretation, or words of encouragement.[37] The feedback will be delivered in a random fashion in order to capitalize on the behavioral work of Skinner and others on variable reward conditioning.[37] This group will also receive additional "gamified" feedback regarding completion streaks (e.g., "Congratulations, you've completed 5 out of 5 depression check ins!"). Subjects in this group will also have their CAT-DI measurements displayed back to them via the CA interface. The results will be displayed using a combination of graphic and story-telling elements (e.g., "Your mood has improved by 10% since we last connected! That's great progress."). Subjects will be oriented to printed examples of these feedback elements and have any of their questions answered by the research coordinator.

Both groups: Upon completion of the study, subjects from both groups will be asked to participate in a semi-structured interview.

Survey Instruments

The following survey instruments will be utilized during the study. The schedule and survey administration method is described in Section 8 below.

Computerized Adaptive Testing: Depression Screen (CAT-DI): (secondary outcome), is a computerized adaptive dimensional severity measure for depression.[38] The CAT-DI utilizes a bank of about 375 depression items whose response patterns are fitted to a multidimensional item response theory model. Adaptive tests are then constructed that use the item parameters to select an optimal small subset of items from the item bank that is tailored to the subject's depression severity, which is dynamically estimated as the subject responds to successive test items. The process continues until the uncertainty in this severity score estimate drops below a

pre-specified level (e.g., 5 points on a 100-point scale). A prior study has shown that an average of 12 items and a median time of 137 seconds had a 95% correlation of with a 389 total item bank score.[38] The resulting severity score maintains a strong correlation with the DSM-IV (SCID) diagnostic categories of none, minor depression (including dysthymia), and MDD. The CAT-DI categorized responses as 1) normal, 2) mild, 3) moderate, or 4) severe depression. The CAT-DI will be administered on a weekly basis via the chatbot interface to Phase 1 and Phase 2 (for Group 2 subjects). Severity results from the CAT-DI will form the basis of feedback provided to subjects in this arm. A refined version of the CAT-DI, optimized for perinatal depression, will be used for this study.[39]

Rapid Estimate of Adult Literacy in Medicine, Revised (REALM-R)[35] is a word recognition test used to identify people at risk for poor health literacy. 11-items, administration time: 3 minutes This survey is administered on paper.

Edinburgh Postnatal Depression Scale (EPDS)[40] provides a discrete measure of perinatal depression. It will be administered by the research coordinator upon entry and then monthly via the chatbot, 10-item, administration time: 5 minutes. At screening this survey is administered on paper. At all following visits, subjects complete the survey via RedCap.

Medication Adherence Question will be used to document medication adherence, 1-item, administration time: 1 minute Subjects complete the survey via RedCap.

The Frequency, Intensity, and Burden of Side Effects Rating (FIBSER)[41] questionnaire uses 3 questions on a 6-point Likert measurement scale to measure 3 side effect domains of impact: frequency, intensity, and burden, administration time: 2 minutes Group 1 subjects complete the survey via RedCap. Group 2 subjects complete the survey via RedCap on some days and complete it using the chatbot interface on other days.

Confidence Questionnaire – to measure behavioral self-efficacy, 7-items, administration time: 5-minutes. This questionnaire has been previously used by our colleagues at NorthShore but not published. Subjects complete the survey via RedCap.

Barkin Index of Maternal Functioning[43] questionnaire designed to measure maternal functioning after childbirth, 20-items, administration time: 7 minutes Subjects complete the survey via RedCap.

A Debrief Tool has been developed to be used as the basis for semi-structured interviews following the trial, administration time: 6 minutes The study team will administer this interview over the phone or in person.

Study Activity Schedule

Table 1: Patient Timeline of Assessments – Phase 1 (Completed)

Timeline: Phase 1 – CB-MBC	Week 1			Week 2				Week 3	
	Days since enrollment								
	1	3	5	8	10	12	15	17	19

Screening: REALM-R & EPDS via interview									
Orientation training (10 minutes)									
Weekly notification to complete CAT-DI & FIBSER via CB-MBC									
User experience check-in phone call									
Debrief Interview									
Compensation ($25)									

Table 2: Patient Timeline of Assessments for Groups 1 & 2 – Phase 2

Timeline: Phase 2 - Group 1: Usual Care (UC)	**Month 1**					**Month 2**				**Month 3**			
	Week												
	Screen	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
Screening: REALM-R & EPDS via interview	X												
Randomize		X											
Reminder phone call from RA		X			X				X				X
EPDS, FIBSER, BIMF, CQ & Medication Adherence via RedCap		X			X				X				X
Debrief Interview													X
Compensation ($25)					X				X				X
Timeline: Phase 2 – Group 2: CB-MBC	**Month 1**					**Month 2**				**Month 3**			
	Week												
	Screen	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
Screening: REALM-R & EPDS	X												
Randomize		X											
Orientation training (10 minutes)		X											
User experience check-in phone call		X											
Reminder phone call from RA to complete questionnaires		X	X	X	X	X	X	X	X	X	X	X	X
EPDS, FIBSER, BIMF, CQ & Medication Adherence via RedCap		X			X				X				X
CAT-DI & FIBSER via CB-MBC		X	X	X	X	X	X	X	X	X	X	X	X

Debrief Interview													**X**
Compensation ($25)					**X**				**X**				**X**

*. Questions related to self-harm asked via CAT-DI will be responded using the method displayed in the Safety Response Algorithm (Appendix 1).

Table 3: Timeline of Provider Involvement – Phase 1 (Completed)

Timeline: Phase 1 – Providers	Week 1			Week 2			Week 3		
	Days since enrollment								
	1	3	5	8	10	12	15	17	19
Receive weekly email report on patients									
Debrief Interview									

Table 4: Timeline of Provider Involvement – Phase 2

Timeline: Phase 2 - Providers	Month 1				Month 2				Month 3			
	Week											
	1	2	3	4	5	6	7	8	9	10	11	12
Receive weekly email report on patients enrolled in chat-bot group	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**
Debrief Interview												**X**

Covariates

Conversational agent metrics (secondary outcomes) will be assessed to better understand the acceptability of CB-MBC. When a subject reads a CB-MBC message, that session will be considered "active". If the subject responds to the message, that session will be considered "engaged". A "Conversation" will be defined as an interchange involving more than 2 subject inputs within a 5-minute interval. Using these definitions, we will track:

- Active rate = (number of active sessions of a user)/(total number of sessions of that subject).
- Engaged rate = (number of engaged sessions of a user)/(total sessions of that subject).
- Average number of conversations/user

Clinical covariates including sociodemographic including subject gender, age, insurance status, and race/ethnicity will be collected from the chart review. Co-morbidity burden will be quantified using the enhanced version of the validated Charlson–Deyo comorbidity index (CCI) for administrative data.[69–71] Other healthcare-related covariates included having a primary care provider; and current use of illicit drugs, alcohol, and tobacco will also be collected.

Statistical methods, sample size, and expected outcomes

The primary outcome of depression severity as measured by longitudinal repeat measurements of EPDS will be compared between usual care and CB-MBC intervention groups. While we anticipate that our sample will be underpowered for formal hypothesis testing, an intention to treat (IIT) analysis with multiple imputation and linear mixed-effects regression will be used to test for differences and examine potential effect size. Similarly, the study results will provide us with estimates odds ratios and 95% CI for discrete secondary endpoints.

While a definitive randomized control trial is not the intent of this trial, a power calculation was performed to understand the limitations of our design and to plan for next steps. Assuming our population exhibits normally distributed depression severity scores, a typical wave-to-wave re-interview attrition rate of 5%, a 1-tailed test, an alpha = 0.05, power = 0.8, and simple residual auto-correlated errors, or rho = 0.3, we would need 56 patient subjects at baseline to detect a large effect size (ES) of 1.0 at the 4th (12 week) time-point. We anticipate that the ES of CB-MBC will be much more modest as the average ES of antidepressants is 0.32 (.11 – 0.69).[44] We expect that intervention group will exhibit a trend towards reduced depression severity as compared to usual care.

Data Storage

Any PHI collected will be directly entered in the HIPAA-compliant REDCap electronic data capture system at the University of Chicago. All information will be treated as confidential material and will be available only to research staff affiliated with the site. All paper-based materials, like consent, HIPAA and data collection forms, will be kept at our ED Research Offices under securely locked conditions with access limited to authorized staff only. Signed and dated consent documents will be maintained for at least seven years after completion of the research.

Data Confidentiality

Research subjects will be assigned a unique study number. No personal health information (PHI) is entered into the database.

All information will be treated as confidential material and will be available only to the designated research and clinical staff at each site. All paper-based materials, like consent forms and data collection forms, will be kept at our ED Research Office under securely locked conditions with access limited to authorized staff only.

In the research records, subjects will be identified by a subject number and the name will be kept separate and only known and accessed by trained research staff members that were given access by the PI or study coordinator. All data stored in RedCap (electronic database) will be accessible only to trained RAs who are given access by the PI or coordinator.

Consent and HIPAA forms will be stored in locked cabinet until the study is closed and no further manuscripts are ongoing. These documents will be accessible only to trained staff. In compliance with HIPAA, data will be kept for 7 years and destroyed afterward.

Chat logs will be maintained in a HIPAA compliant secure cloud database that can only be accessed by the PI and the research coordinator. The servers reside in the University of Chicago data center located at 6045 S. Kenwood Avenue. The security is maintained and monitored by the CRI Systems Team, under the direction of Thornbjorn Axelsson, the CRI Director of Systems and IT Infrastructure. The architecture is monitored through weekly security scans conducted by the Office of Information Security under the direction of Chief Information Security Officer Plamen Martinov. All researchers involved in the study have completed HIPAA

training. All researchers are instructed that the data are to remain on the CRI-hosted secure servers. As an additional layer of safety, chat logs will be reviewed by study personnel each day and any PHI discovered will be deleted.

For subjects using CB-MBC via Facebook Messenger, chat logs are maintained on the Facebook Messenger platform. It should be noted that any results of depression or other assessments will be displayed to the subject in non-machine readable formats such as images. As an extra precaution, Facebook Messenger logs will be deleted each day. It should be noted that Facebook generally already has access to PHI including name, address, and birthdate as part of their Terms of Service. If a breach of data privacy or security is discovered, the PI will report the breach to the IRB and the DSMB at the University of Chicago within 48 hours.

Risks

Risk to privacy. As with all medical research involving human subjects, there is a risk of inadvertent disclosure or theft of protected health information (PHI) from paper enrollment log, REDCap electronic data capture system, the HIPAA-compliant cloud database, or the subject's Facebook Messenger account. We see these risks as minimal and will address ways of reducing these risks below.

Risk of discomfort. Subjects will be able to withdraw from the study at any time. If a subject becomes noticeably distressed, the RA will inform his or her treating physician.

Risk of additional/reduced assessment or treatment. Treating providers will not have access to screening or follow-up depression severity and side effect information obtained from monthly telephone interviews. Assessment results from subjects randomized to CB-MBC will be automatically sent to the provider on a weekly basis. It is a desired outcome that these assessment results motivate the provider to consider contacting the subject and possibly modifying their treatment plan. This type of additional assessment and treatment would be expected to be beneficial goal of this intervention. We believe the risk of an adverse event related to additional assessment by the provider to be minimal.

Conversely, participation in the study will not prohibit involvement in any other type of assessment and/or treatment. All participants will be allowed (and/or encouraged) to participate in whatever type or amount of treatment is thought to be clinically indicated for the subject. Thus, participation in the study will not cause any participant to be deprived of a clinically appropriate assessment and/or treatment.

Risk of disclosure of suicidal ideation. Women with perinatal depression are at higher risk for experiencing suicidal ideation. We will be actively assessing for suicidal ideation via in-clinic screening and telephone assessments. There is also the risk of inadvertent disclosure of self-harm behavior during these contact points. In addition, for patients using the chatbot, the CAT-DI depression item bank includes 11 items that make reference to death, suicide, or self-harm. Thus, there is a risk of increased detection of suicidality or self-harm during the study. Safety procedures will be described in the section below.

Risk of other serious adverse events. The overall potential for serious adverse events leading to physical or psychological harm during the study is minimal as we are testing an information technology intervention. Risk of increased suicidal or self-harm behaviors. We are unaware of any literature to suggest that the act of tracking depression symptoms in depressed patients increases the risk of suicide or worsening depression as compared to usual care. The principal

investigator will report the adverse event to the IRB within 48 hours.

Risk of social harm. There is minimal risk of social harm to the subject.

Medicolegal risk to providers. As part of the Phases 1 & 2, provider subjects will receive a weekly encrypted email report of their patients enrolled in the chatbot group. During the consent process, providers will be oriented to the format of these email reports and the study's intention that they review the report and make treatment changes as necessary. They will be alerted to the potential medicolegal risk associated with ignoring or not responding to these reports. Providers will have the option of not participating in the study if they are uncomfortable with this risk.

Protections Against Risk

Physical harm. In order to manage subject expectations, during orientation we will explain that the CB-MBC is an experimental platform and that no one is actively monitoring their responses on a daily basis. Patients will be told to expect follow-up from the team and/or their providers if they disclose potential thoughts of suicide, self-harm, or harm to others via the chatbot or during phone interviews.

The potential safety mechanisms for detecting and responding to suicidal ideation are outlined below. For all cases, the PI will take whatever actions are judged to be clinically appropriate to protect subject safety per standard of care guidelines:

- **Detection during initial enrollment:** If suicidal ideation or self-harm is disclosed by the subject or detected by formal suicide risk assessment or incidental disclosure during initial enrollment in clinic, the research coordinator will immediately contact the treating provider or other clinical staff. In addition, the research assistant will notify Dr. Beiser to discuss the case. As recommended in the original publication, a patient who scores a 1, 2, or 3 on question #10 of the EPDS ("The thought of harming myself has occurred to me" "1. hardly ever", "2. sometimes", or "3. yes quite often") will trigger an evaluation for active suicidality by the clinician.[40]

- **Detection during telephone follow-up:** If suicidal ideation or self-harm is disclosed by the subject or detected during an EPDS depression severity assessment the research coordinator will ask the subject if he/she feels safe. If the subject does not feel safe, they will be encouraged to call 911 or go to the nearest emergency department They will also be given suicide prevention resources including the Suicide Hotline (available 24 hours a day, 7 days per week) at 1-800-273-TALK (1-800-273-8255) or 1-800-SUICIDE (1-800-784-2433) or the crisis text line (text HOME to 741741). We will not remind patients of the risk involuntary admission to the hospital at this critical juncture as it may represent a disincentive for seeking help. If the patient feels safe, he/she will be told to expect a call from their treating physician or therapist to discuss. The research assistant will then contact Dr. Beiser to discuss the case. Dr. Beiser will then contact the treatment team and discuss the need for immediate follow-up with the patient. The treating physician or therapist will then contact the patient by phone to discuss the thoughts of suicide or self-harm and make a safety plan.

- **Detection during REDCap email EPDS survey**: If the patient scores a 1, 2, or 3 on question #10 of the EPDS, ("The thought of harming myself has occurred to me" "1. hardly ever", "2. sometimes", or "3. yes quite often") the REDCap email survey will automatically branch and

ask if the patient feels safe. If the subject does not feel safe, they will be encouraged to call 911 or go to the nearest emergency department. They will also be given suicide prevention resources including the Suicide Hotline (available 24 hours a day, 7 days per week) at 1-800-273-TALK (1-800-273-8255) or 1-800-SUICIDE (1-800-784-2433) or the crisis text line (text HOME to 741741). Finally, an automated text message that contains a studyID without any PHI will be sent to Dr. Silver indicating that a patient has triggered a suicide alert. At any point, NorthShore's 866-364-MOMS hotline may be used for clinical consultation and immediate patient support

- **Detection via CB-MBC:** If suicidal ideation or self-harm is indicated by a participant's responses to the chatbot, a member of the team will follow-up with them either immediately or within 24 hours by phone depending on their responses to subsequent questions. This safety response algorithm is outlined in Appendix 1. The participant's treatment provider will also be notified within 24 hours or sooner if clinically indicated by PI. In addition, patients will be made aware of suicide prevention resource links they have access to via a persistent menu at the bottom of the chatbot window.

Potential Benefits To Subjects

While we cannot assure subjects that they will personally benefit from participation in our study, subjects will have the opportunity to share their feelings, symptoms, concerns, and sense well-being via self-report measures. Subjects may become more aware of their adjustment and improvements from regular follow-ups.

If our hypothesis is correct, CB-MBC will improve the care of patients with depression and lead to an improvement in depression symptoms and side effect burden in participating subjects and potentially, future patients. We believe the anticipated benefits to subjects and others far outweigh the minimal risks outlined above.

Sharing of Results with Subjects

All participants will receive the standard of care for their presenting problem. Any subjects that screen positive for suicidal ideation or behavior will be reported to their treating physician, and their treating physician will manage their suicidality per standard operating protocols.

The consent process explains this (i.e., that they may be further evaluated by staff and mental health professionals if they are deemed to have any level of suicidality). This evaluation, in rare circumstances, may involve involuntary commitment to protect the individual or someone else if they report that they are imminently suicidal or homicidal. This judgment is based on standard operating protocols and clinical policies.

Monitoring Plan to Ensure the Safety of Participants

A plan to periodically evaluate information collected regarding risks or harms to determine whether participants remain safe is in place. The study will be monitored by the study investigator as well as an individual who is not associated with the study. Principal Investigator and an unaffiliated individual will review the data collected from the study on a monthly basis. These reviews will allow the study personnel to monitor the progress as well as to determine if there are any unanticipated problems occurring.

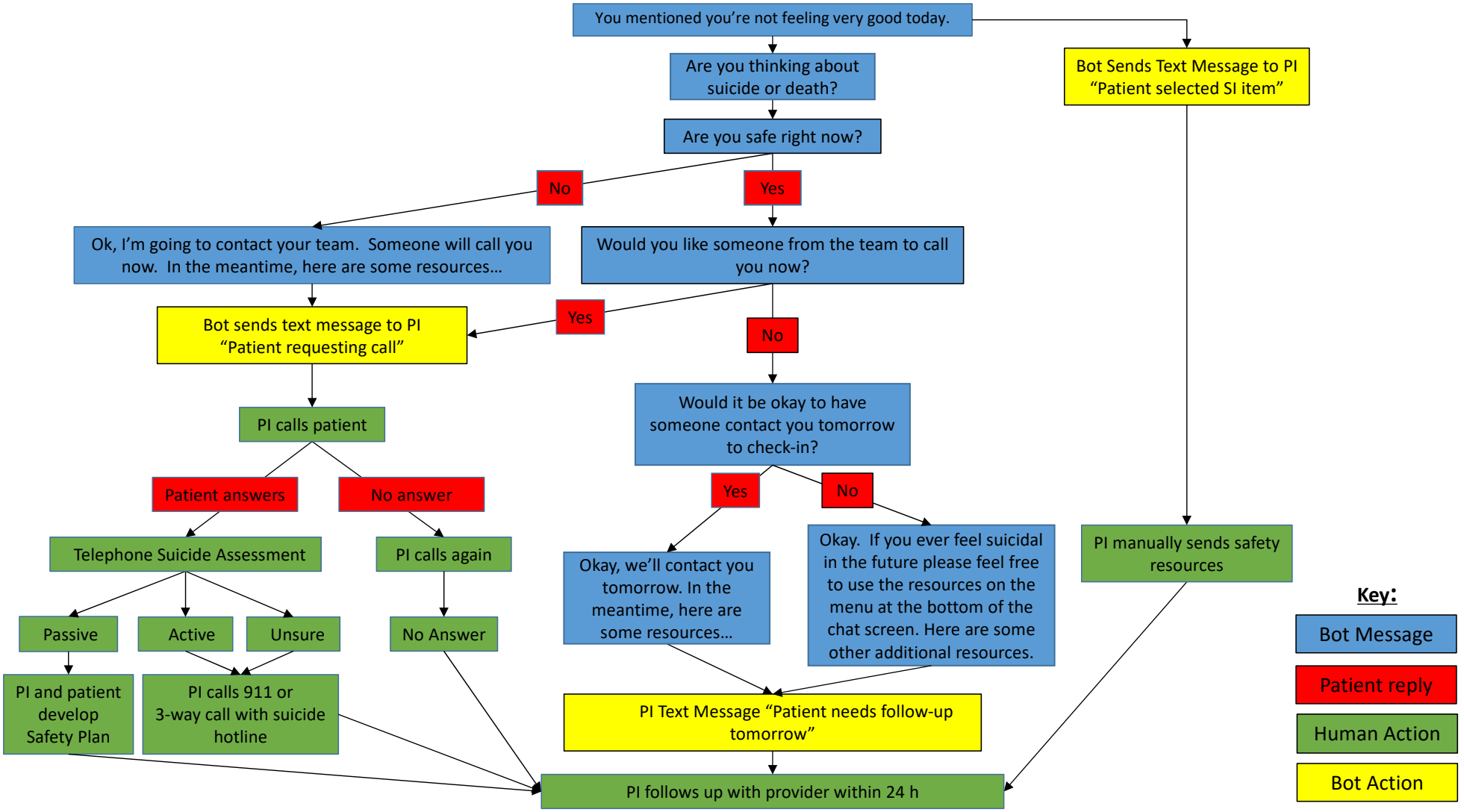
Video Demonstration

A video demonstration of the production version of the chatbot can be viewed below.

https://youtu.be/wFe3-MuIlxk

Appendix 1

Safety Response Algorithm



References

1. Bennett HA, Einarson A, Taddio A, Koren G, Einarson TR. Prevalence of Depression During Pregnancy: Systematic Review. *Obstet Gynecol*. 2004;103(4):698-709. doi:10.1097/01.AOG.0000116689.75396.5f.

2. Ko JY, Rockhill KM, Tong VT, Morrow B, Farr SL. Trends in Postpartum Depressive Symptoms - 27 States, 2004, 2008, and 2012. *MMWR Morb Mortal Wkly Rep*. 2017;66(6):153-158. doi:10.15585/mmwr.mm6606a1.

3. Lowenstein C, Golden O, Emam D, Mcdaniel M. Linking Depressed Mothers to Effective Services and Supports: A Policy and Systems Agenda to Enhance Children's Development and Prevent Child Abuse and Neglect. (May 2013).

4. Diaz JY, Chase R. The Cost of Untreated Maternal Depression. 2010;(October):2700.

5. Kessler RC, Berglund P, Demler O, et al. The epidemiology of major depressive disorder. *JAMA psychiatry*. 2003;289(23):3095-3105. doi:10.1001/jama.289.23.3095.

6. Gonzalez HM, Vega WA, Williams DR, Tarraf W, West BT, Neighbors HW. Depression Care in the United States: Too Little for Too Few. *Arch Gen Psychiatry*. 2010;67(1):37. doi:10.1001/archgenpsychiatry.2009.168.Depression.

7. Services UD of H and H. *National Projections of Supply and Demand for Selected Behavioral Health Practitioners: 2013-2025*.; 2016.

8. *Health Workforce Projections : Mental Health Counselors*.; 2016.

9. Kroenke K, Unutzer J. Closing the False Divide : Sustainable Approaches to Integrating Mental Health Services into Primary Care. *J Gen Intern Med*. 2016;32(4):404-410. doi:10.1007/s11606-016-3967-9.

10. Sanchez K. Collaborative care in real-world settings: barriers and opportunities for sustainability. *Patient Prefer Adherence*. 2017;Volume 11:71-74. doi:10.2147/PPA.S120070.

11. Gelenberg AJ. Using assessment tools to screen for, diagnose, and treat major depressive disorder in clinical practice. *J Clin Psychiatry*. 2010;71 Suppl E:e01. doi:10.4088/JCP.9058se1c.01gry.

12. Harding KJK, Rush AJ, Arbuckle M, Trivedi MH, Pincus HA. Measurement-based care in psychiatric practice: A policy framework for implementation. *J Clin Psychiatry*. 2011;72(8):1136-1143. doi:10.4088/JCP.10r06282whi.

13. Trivedi MH. Tools and strategies for ongoing assessment of depression: A measurement-based approach to remission. *J Clin Psychiatry*. 2009;70(SUPPL. 6):26-31. doi:10.4088/JCP.8133su1c.04.

14. Morris DW, Trivedi MH. Measurement-Based Care for Unipolar Depression. 2011:446-458. doi:10.1007/s11920-011-0237-8.

15. Paykel ES, Ramana R, Cooper Z, Hayhurst H, Kerr J, Barocka A. Residual symptoms after partial remission: an important outcome in depression. *Psychol Med*. 1995;25(6):1171-1180. doi:10.1017/S0033291700033146.

16. Schwartz J. The Most Popular Messaging App in Every Country. The Market Intelligence Blog by Similar Web. https://www.similarweb.com/blog/worldwide-messaging-apps. Published 2016. Accessed December 21, 2017.

17. Statista. Number of Facebook users by age in the U.S. 2017. https://www.statista.com/statistics/398136/us-facebook-user-age-groups/. Published 2017. Accessed December 21, 2017.

18. Stewart JB. Facebook Has 50 Minutes of Your Time Each Day. It Wants More. New York Times. https://www.nytimes.com/2016/05/06/business/facebook-bends-the-rules-of-audience-engagement-to-its-advantage.html?_r=0. Published 2016. Accessed December 21, 2017.

19. Smith A, Mcgeeney K, Rainie L, Keeter S, Page D. U.S. Smartphone Use in 2015. *Smartphone Differ*. 2015:60. doi:10.1017/CBO9781107415324.004.

20. Fox S, Duggan M. *Mobile Health 2012*. Pew Internet & American Life Project; 2012.

21. Robin Cook Kopelman, Moel J, Mertens C, Stuart S, Arndt S, O'Hara MW. Barriers to Care for Antenatal Depression. *Psychiatr Serv*. 2008;59(4):429-432. doi:10.1176/appi.ps.59.4.429.Barriers.

22. Bakker D, Hons BP, Kazantzis N, et al. Mental Health Smartphone Apps : Review and Evidence-Based Recommendations for Future Developments. *JMIR*. 2016;3(1):e7:1-31. doi:10.2196/mental.4984.

23. FDA Digital Health. https://www.fda.gov/MedicalDevices/DigitalHealth/default.htm. Accessed March 11, 2017.

24. Batra S, Baker R, Wang T, Forma F, DiBiasi F, Peters-Strickland T. Digital health technology for use in patients with serious mental illness: a systematic review of the literature. *Med Devices Evid Res*. 2017;Volume 10:237-251. doi:10.2147/MDER.S144158.

25. Donker T, Petrie K, Proudfoot J, Clarke J, Birch M-R, Christensen H. Smartphones for smarter delivery of mental health programs: a systematic review. *J Med Internet Res*. 2013;15(11):e247. doi:10.2196/jmir.2791.

26. Depressive D, Episodes M, Overview S. Smartphones in Mental Health : 2015;14(3):10-13.

27. Nicholas J, Fogarty AS, Boydell K, Christensen H. The Reviews Are in: A Qualitative Content Analysis of Consumer Perspectives on Apps for Bipolar Disorder. *J Med Internet Res*. 2017;19(4):e105. doi:10.2196/jmir.7273.

28. Editor M, Reviewed GE, Atienza A, et al. Cognitive Behavior Therapy for Anxious and Depressed Youth : Improving Homework Adherence Through Mobile Technology.

2016;5(4). doi:10.2196/resprot.5841.

29. Arean PA, Hallgren KA, Jordan JT, et al. The use and effectiveness of mobile apps for depression: Results from a fully remote clinical trial. *J Med Internet Res*. 2016;18(12):1-13. doi:10.2196/jmir.6482.

30. Mohr DC, Tomasino KN, Lattie EG, et al. IntelliCare : An Eclectic , Skills-Based App Suite for the Treatment of Depression and Anxiety Corresponding Author : 19(1):1-14. doi:10.2196/jmir.6645.

31. Fitzpatrick KK, Darcy A, Vierhile M. 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 Ment Heal*. 2017;4(2):e19. doi:10.2196/mental.7785.

32. Marko KI, Krapf JM, Meltzer AC, et al. Testing the Feasibility of Remote Patient Monitoring in Prenatal Care Using a Mobile App and Connected Devices: A Prospective Observational Trial. *JMIR Res Protoc*. 2016;5(4):e200. doi:10.2196/resprot.6167.

33. Hantsoo L, Criniti S, Khan A, et al. A Mobile Application for Monitoring and Management of Depressed Mood in a Vulnerable Pregnant Population. *Psychiatr Serv*. 2017;(September):appi.ps.2016005. doi:10.1176/appi.ps.201600582.

34. Goetz M, Müller M, Matthies LM, et al. Perceptions of Patient Engagement Applications During Pregnancy: A Qualitative Assessment of the Patient's Perspective. *JMIR mHealth uHealth*. 2017;5(5):e73. doi:10.2196/mhealth.7040.

35. Bass PF, Wilson JF, Griffith CH. A Shortened Instrument for Literacy Screening. *J Gen Intern Med*. 2003;18(12):1036-1038. doi:10.1111/j.1525-1497.2003.10651.x.

36. Kupfer DJ. Long-term treatment of depression. *J Clin Psychiatry*. 1991;52 Suppl:28-34.

37. Ferster CB. Schedules of Reinforcement with Skinner. *J Exp Anal Behav*. 2002;77(3):303-311. doi:10.1901/jeab.2002.77-303.

38. Gibbons RD, Weiss DJ, Pilkonis PA, et al. Development of a computerized adaptive test for depression. *Arch Gen Psychiatry*. 2012;69(11):1104-1112. doi:10.1001/archgenpsychiatry.2012.14.

39. Kim JJ, Silver RK, Elue R, et al. The experience of depression, anxiety, and mania among perinatal women. *Arch Womens Ment Health*. 2016;19(5):883-890. doi:10.1007/s00737-016-0632-6.

40. Cox JL, Holden JM, Sagovsky R. Detection of postnatal depression . Development of the 10-item Edinburgh Postnatal Depression Scale Detection of Postnatal Depression Development of the 10-item Edinburgh Postnatal Depression Scale. 2010:782-786. doi:10.1192/bjp.150.6.782.

41. Wisniewski SR, Rush AJ, Balasubramani GK, Trivedi MH, Nierenberg AA, STARD Investigators. Self-Rated Global Measure of the Frequency, Intensity, and Burden of Side Effects. *J Psychiatr Pract*. 2006;12(2):71-79. doi:10.1097/00131746-200603000-00002.

42. Manos RC, Kanter JW, Luo W. The Behavioral Activation for Depression Scale-Short Form: Development and Validation. *Behav Ther*. 2011;42(4):726-739. doi:10.1016/j.beth.2011.04.004.

43. Barkin JL, Wisner KL, Bromberger JT, et al. Development of the Barkin Index of Maternal Functioning. *J Womens Health (Larchmt)*. 2010;19(12):2239-2246. doi:10.1089/jwh.2009.1893.

44. Ninan PT, Poole RM, Stiles GL. Selective publication of antidepressant trials. *N Engl J Med*. 2008;358(20):2181; author reply 2181-2182. doi:10.1056/NEJMc080313.