Scaling Up Maternal Mental healthcare by Increasing access to Treatment (SUMMIT): A study protocol for perinatal depression and anxiety*

Version 20.0

Funding. This study is funded as a Pragmatic Clinical Study (PCS) by the Patient-Centered Outcomes Research Institute (PCORI), PCS-2018C1-10621.

Trial Sponsor. Sinai Health
Contact Name: **Daisy R. Singla**
600 University Ave., #914A
Toronto, Ontario
M5G 1X5 Canada
Telephone: (416) 562-4438
Email: daisy.singla@utoronto.ca

* This protocol includes details of economic evaluation, which was added for a substudy not funded by PCORI

TABLE OF CONTENTS

Table 1. List of Abbreviations

Abbreviation	Explanation
ADM	Antidepressant Medication
AE	Adverse Event
AUDIT	Alcohol Use and Disorders Identification Test
BA	Behavioural Activation
CAMH	Centre for Addiction and Mental Health
CBT	Cognitive Behavioural Therapy
CEAC	Cost-Effectiveness Acceptability Curve
CSQ-8	Client Satisfaction Questionnaire-8
CTO	Clinical Trials Ontario
DAST	Drug Abuse Screening Test
DSMB	Data and Safety Monitoring Board
EPDS	Edinburgh Postnatal Depression Scale
EQ-5D-5L	EuroQol 5 Dimension 5 Level
GAD-7	Generalized Anxiety Disorder-7 Scale
HAP	Health Activity Program
HIC	High-income countries
HIPAA	Health Insurance Portability and Accountability Act
HTE	Heterogeneity of Treatment Effects
ICER	Incremental Cost-Effectiveness Ratio
ICES	Institute for Clinical Evaluative Sciences
IRB	Institutional Review Board (US)
LMIC	Low and middle-income countries
MINI	Mini International Neuropsychiatric Interview
NSP	Non-specialist provider
OB	Obstetrician
PAI	Personalized Advantage Index
PCL-6	Abbreviated PTSD Checklist
PCORI	Patient-Centered Outcomes Research Institute
PHIPA	Personal Health Information Protection Act
PHQ-9	Patient Health Questionnaire-9
PT	Psychological Treatment
Q-HAP	Quality of Healthy Activity Program
RA	Research Assistant
REB	Research Ethics Board (Canada)
SAE	Serious Adverse Event
SAN	Storage Area Network
SAS	Statistical Analysis System
SD	Standard Deviation
SH	Sinai Health
SHSC	Sunnybrook Health Sciences Centre
SMH	St. Michael's Hospital
TIC	Trial Investigator Committee
TMC	Trial Management Committee
TSC	Trial Steering Committee
TSM	Trial Steering Management

QALY	Quality-Adjusted Life Year
UK	United Kingdom
UNC	University of North Carolina
USA/US	United States of America
USD	U.S. Dollar
UofT	University of Toronto
WAI-SR	Working Alliance Inventory-Short Revise
WCH	Women's College Hospital
WHODAS	World Health Organization Disability Assessment Schedule

Table 2. Investigators, Roles and Responsibilities

Name	Role	Affiliation	Contact Information
Daisy Radha Singla, PhD	Study PI, Site PI of SHS and CAMH	SHS, CAMH, UofT	daisy.singla@utoronto.ca
Samantha Meltzer-Brody, MD, FRCPC	Co-PI, Site PI of UNC	UNC	samantha_meltzer-brody@med.unc.edu
Richard Silver, MD	Co-I, Site PI of NorthShore	NorthShore, UChicago	rsilver@northshore.org
Simone Vigod, MD, FRCPC	Co-I, Site PI of WCH	WCH, UofT	simone.vigod@wchospital.ca
Cindy-Lee Dennis, PhD	Co-I, Site PI of SMH	SMH, UofT	cindylee.dennis@utoronto.ca
UToronto			
Paula Ravitz, MD, FRCPC	Co-I, Clinical Lead	SHS, UofT	paula.ravitz@sinaihealthsystem.ca
Ariel Dalfen, MD, FRCPC	Co-I, Trial Psychiatrist	SHS, UofT	ariel.dalfen@utoronto.ca
Alex Kiss, PhD	Co-I, Biostatistician	SHSC, UofT	alex.kiss@ices.on.ca
Andrea Lawson, PhD	Co-I, Mental Health Researcher	CAMH, UofT	andrea.lawson@camh.ca
Claire de Oliveira, PhD	Health Economist	CAMH, UofT	claire.deoliveira@camh.ca
University of North Carolina			
Crystal Schiller, PhD	Co-I, Clinical Lead	UNC	crystal_schiller@med.unc.edu
Alison Stuebe, MD	Co-I, Maternal-Fetal Medicine	UNC	astuebe@med.unc.edu
Mae Lynn Reyes-Rodriguez, PhD	Co-I, Cultural Competency Lead	UNC	maelynn_reyes@med.unc.edu
Bradley Gaynes, MD, MPH	Co-I, Trial Expert	UNC	bradley_gaynes@med.unc.edu
University of Chicago/NorthShore University HealthSystem			
Laura La Porte, LCSW	Co-I, Clinical Lead	NorthShore, UChicago	LLaPorte@northshore.org
Jo Kim, PhD	Co-I, Clinical Lead	NorthShore, UChicago	jkim@northshore.org
Other			
John Naslund, PhD	Co-I, Digital Expert	Harvard	John_Naslund@hms.harvard.edu
Vikram Patel, PhD, FRCPC	Co-I, Trial and NSP Expert	Harvard	Vikram_Patel@hms.harvard.edu
Steven D. Hollon, PhD	Consultant; Trials & Psychological Treatment Expert	Vanderbilt University	steven.d.hollon@vanderbilt.edu
Sona Dimidjian, PhD	Consultant; BA expert	U Colorado Boulder	sona.dimidjian@colorado.edu

Table 3. Study Management and Committees.

Committee	Role	Members	Frequency
Trial Management Committee (TMC)	To monitor all aspects of the conduct and progress of the trial, ensure that the protocol is adhered to and take appropriate action to safeguard participants and the quality of the trial.	● Principal Investigator from each hub* ● Trial Coordinators ● Project Administrator ● Data Manager	Weekly**
Trial Investigator Committee (TIC)	To monitor all aspects of the conduct and progress of the trial, including site-specific safety protocols within and across sites	● All Investigators ● Research Coordinators from each site ● Data Manager	Biweekly to Monthly
Trial Steering Committee (TSC)	To provide overall supervision of the trial and ensure that it is being conducted in accordance with the protocol and the relevant regulations. The TSC should approve the trial protocol and any protocol amendments and provide advice to the TMC on all aspects of the trial. Decisions about continuation or termination of the trial or substantial amendments to the protocol are finally the responsibility of the TSC.	● Members of the TMC ● All Investigators ● Trial/Content Advisors and Consultants	Six-monthly
Data Safety Monitoring Board (DSMB)	The DSMB will review the accruing trial serious adverse event reports to assess whether there are any safety issues that should be brought to participants' attention or any reasons for the trial not to continue. It is the only body that makes recommendations to unblind data and makes further recommendations to the TMC.	● Sherryl Goodman - clinical psychologist, with experience in BA ● Tim Oberlander - developmental pediatrician ● Catherine Monk - psychologist experience on maternal depression ● Robert Gibbons - biostatistician with experience on mental health trials	Six-monthly

*hub refers to locality (Toronto, Chapel Hill, Chicago) ** weekly meetings also held within sites

Table 4. Protocol Summary Table.

Title	Scaling Up Maternal Mental healthcare by Increasing access to Treatment (SUMMIT)
Sample Size	**N = 1230**
Study Population	Women pregnant up to 36 weeks or 4-30 weeks postpartum, with an EPDS score ≥10 and capable of giving informed consent.
Accrual Period	39 months; Recruitment starts January 2020 and ends Mar 31 2023
Study Design	A multi-center randomized non-inferiority trial to test the choice of delivery mode and provider, implementing a brief evidence-based PT of BA for perinatal depressive and anxiety symptoms.
Study Duration	01 June 2019 – 31 May 2024
Study Intervention	Eligible participants will be randomly assigned to the same BA treatment for perinatal depressive and anxiety symptoms in one of four arms: 1) In-person specialist; 2) Telemedicine specialist; 3) In-person non-specialist; 4) Telemedicine non-specialist.
Primary Objectives	1. To examine if a brief BA psychological treatment delivered by a non-specialist provider is as effective in treating perinatal depressive symptoms as specialist-delivered treatment. 2. To examine if a brief BA psychological treatment delivered through telemedicine is as effective in treating perinatal depressive symptoms as in-person treatment.
Secondary Objectives	1. To examine the aforementioned questions among secondary outcomes including anxiety symptoms (GAD-7). 2. To conduct subgroup analyses including the role of the perinatal period on long-term maternal mental health outcomes at 12-months post-randomization and child development outcomes at 6-24 months post-child birth. 3. To identify relevant and underlying processes related to delivery and scalability of a brief PT for perinatal depressive and anxiety symptoms from a multi-stakeholder perspective including barriers and facilitators. 4. To conduct an economic evaluation to examine the cost-effectiveness of NSP-delivered and telemedicine psychotherapy to specialist-delivered and in-person psychotherapy.

1 INTRODUCTION

1.1. Background

Depression is the leading cause of disability of mothers worldwide[1] with an estimated 10 to 15% of women in the USA and Canada experiencing depression during pregnancy or in the year following childbirth[2,3]. Many of these symptoms begin in the antenatal period[4], and estimates of the lifetime costs of perinatal depression amount to over $45.9 billion dollars (USD) in the USA each year[5]. Similarly, although given less attention than depression, 15 to 20% of mothers experience anxiety symptoms perinatally[6], with up to 10% experiencing both[7]. The negative impact of depression and anxiety on the well-being of the mother and her child[8-10] underscores the importance of addressing perinatal mental health issues.

Psychological treatments (PTs), including cognitive, behavioral and interpersonal therapies, are effective in preventing and treating perinatal depression and anxiety[11,12] and preferred by women over pharmacological treatment[13,14]. Yet, across the US and Canada, as few as 20% of women with perinatal depression are treated with evidence-based PTs[15]. Barriers to accessing PTs in the perinatal period include childcare needs, unpredictable schedules and stigma[13,16]. In addition, a dearth of financial resources and trained professionals may contribute to this burden[16]. These barriers are problematic given the recent US Preventive Task Force Recommendations for universal screening of all pregnant and postpartum women for depression[17] and evidence-based psychological treatments[18]. Thus, there is an urgent need for widely accessible, low-cost, and innovative solutions to improve access to psychological treatments for perinatal depression and anxiety.

1.1.1. Scalable Innovations

1.1.1.2. Non-specialist providers. Task-sharing has been used worldwide to improve access to care, where non-specialist providers (NSPs)—individuals with no formal training in delivering mental health care including nurses, peers, lay counselors, midwives, teachers, and primary care doctors—have been trained to effectively treat perinatal depressive and anxiety symptoms worldwide[19,20]. NSPs are an important human-resource because they are widely available, affordable, considered trustworthy, and have regular and frequent contacts with mothers[19,21,22]. Worldwide, NSPs have successfully implemented behavioural activation (BA) to reduce maternal depressive symptoms[23,24], anxiety symptoms[24], and even intimate partner violence among women of childbearing years[23]. BA can be as effective in targeting depression as traditional and longer courses of CBT[25,26]. However, evidence is lacking as to whether specialists and NSPs are equally effective in targeting perinatal depressive and anxiety symptoms when delivering the *same* treatment.

1.1.1.3. Telemedicine. Telemedicine-based PTs—interventions provided through web-, app- and/or telephone-based platforms—offer a promising, alternative approach for mothers in terms of flexibility[27], efficiency[28], and costs[29]—thus increasing accessibility and scalability of PTs. Recent reviews show that effects on depressive and anxiety symptoms are similar when comparing digital and in-person services[30]. **Access to PTs nonetheless remains limited[15,31].**

1.1.2. Why Behavioural Activation?

BA proposes that the key to feeling less depressed and anxious is to increase enjoyable or fulfilling activities that align with one's values[32,33], targeting the mechanisms of patient activation and avoidant coping. We will be using a brief, 6-8 session version of BA that was shown to be effective in reducing depressive symptoms and enhancing rates of remission when delivered by NSPs, with sustained effects in Goa, India (led by co-I Patel and involved the PI)[23,34]. Among effective treatments for depression, BA is an excellent candidate for this study for three reasons:

1. BA has a **strong evidence base** for effectiveness in the general health care population[35,36]. In a randomized placebo controlled study in the US, BA was as efficacious as, and more enduring than, antidepressant medication (ADM), with fewer patients dropping out of treatment[25,26]; in the UK, BA was as effective as longer courses of CBT[26,37].
2. BA is **effective in treating perinatal depression**. In randomized trials in the US, BA was associated with high satisfaction and treatment engagement[22,38] and, as compared to treatment as usual, offered significant benefit including the reduction of commonly-occurring anxiety symptoms[24,39].
3. BA is a parsimonious approach that is **easy to understand and implement**[32]. This is critical when training NSPs. Trials in Colorado[24], Uganda[40], India[23,34], and the UK[37] have demonstrated that lay counselors, undergraduate psychology students, peers and midwives can all be trained to effectively deliver BA or components to reduce depressive symptoms.

1.2. Rationale

Our overarching goal of SUMMIT (**S**caling **U**p **M**aternal **M**ental healthcare by **I**ncreasing access to **T**reatment) is to examine the scalability and patient-centered provision of brief, evidence-based PTs for perinatal depression and anxiety. Specifically, the current proposal will answer the question of the effectiveness of the integration of telemedicine of PTs within larger health systems, and whether frontline non-mental health workers, with appropriate training, can effectively deliver BA for perinatal depressive and anxiety symptoms. We will identify relevant underlying implementation processes and determine whether, and to what extent, these strategies work differentially for certain women over others. Addressing these aims can support a stepped care model in which we optimize the use of available resources. In doing so, this research has the potential to increase the accessibility, scalability, and cost-effectiveness of evidence-base of PTs for perinatal women worldwide.

1.3 Potential Risk and Benefits

1.3.1. Potential Risks

1.3.1.1. Inconvenience. Participants may feel inconvenienced by the time required to complete the study task. The expected time commitment will be explicitly stated in the consent form.

1.3.1.2. Psychological risks associated with assessment. There is the possibility that participants could be upset or offended by survey questions. In order to minimize psychological

harm, information about potential risks is included in the consent form. Participants also will be instructed that they are free to not answer any questions they choose on the surveys.

1.3.1.3. Psychological risks associated with treatment or training.

1.3.1.3.1. Participants. As in any study related to mental health concerns, discussion of symptoms may cause psychological discomfort and participants may also experience exacerbation of depressive or anxiety symptoms. Established safety protocols will be implemented to offset these risks (*Appendix A – Safety Protocols and Forms*). Furthermore, the recent United States Preventive Task Force guidelines have suggested that there is no-to-minimal harm of evidence-based PTs including cognitive, behavioral and interpersonal psychotherapies[18].

1.3.1.3.2. Providers. It also is possible that a potential treatment provider may be upset if they are withdrawn from the study for not meeting competency standards. We will provide clear orientation as part of the training process, and will make referrals for anyone who is overly bothered by being withdrawn from the study. In addition, NSPs may become stressed while participating as providers. Treatment Providers will receive information and techniques in self-care as part of the BA training curriculum and established safety protocols will be implemented to offset these risks (*see Treatment Provider/Research Assistant Safety, Appendix A – Safety Protocols and Forms,* pg 5).

1.3.1.4. Physical Harm. We do not anticipate any physical risks for participants or treatment providers. Participants' physical care needs will remain in the hands of their primary care physicians or OB provider during the study. We will ask participants to provide contact information and a release of information for their provider in order to communicate any physical health concerns during the study involvement should the need arise. We will also ask participants for secondary personal contacts (*see Secondary Contact Form, Appendix E – Study Documents,* pg 83).

1.3.2 Potential Benefits

1.3.2.1. Participants. This study may provide participants with insight into their mental health states and equip them with strategies to overcome their depressive and anxiety symptoms. Over the course of the treatment, participants may experience an increase in knowledge of mental health symptoms and increase their coping skills over time. In light of COVID-19, this study is considered 'essential research' by its study institutions by contributing a valuable, evidence-based psychotherapy to study participants.

1.3.2.2. Providers. NSPs may receive the benefit of personal satisfaction from making a positive impact on perinatal women struggling with these issues.

1.3.3.3. Society. Preventing and decreasing the severity of perinatal mental health issues offers significant benefits to society, given that depression is a leading cause of disability worldwide with serious consequences for offspring. These benefits would be particularly marked if the BA treatment provided by NSPs and telemedicine are as effective as those delivered by specialists and in-person.

1.4. Objectives.

1.4.1. Primary Objectives

1. To examine if a brief, BA psychological treatment delivered by a non-specialist provider is as effective in treating perinatal depressive symptoms as specialist-delivered treatment, as measured by the Edinburgh Postnatal Depression Scale (EPDS)[41].
2. To examine if a brief BA psychological treatment delivered through telemedicine is as effective in treating perinatal depressive symptoms as in-person treatment, as measured by the EPDS[41].

The assessment period will be extended for those women who encounter a hiatus in treatment due to significant perinatal-related life events (e.g., giving birth, obstetrical complications, COVID-19).

1.4.2. Secondary Objectives

1. To examine the aforementioned questions among anxiety symptoms, as measured by the GAD-7[42,43].
2. To assess moderating effects as to whether clinical severity (mild, moderate and severe) and the timing of the treatment (antenatal vs. postnatal) differentially influences depressive symptoms and whether timing of treatment influences child mental development at 6-24 month post-child birth.
3. To conduct a process evaluation; i.e., identify the underlying processes related to delivery and scalability of a brief PT for perinatal depressive and anxiety symptoms from a multi-stakeholder perspective including relevant barriers and facilitators.
4. To conduct an economic evaluation to examine the cost-effectiveness of NSP-delivered and telemedicine psychotherapy compared to specialist and in-person psychotherapy.

1.5. Hypotheses

1. We hypothesize that among mothers with depressive symptoms, i) telemedicine-PT will be as effective as in-person treatment; and ii) NSP-delivered PT will be as effective as specialist-delivered PT at 3-months, post-randomization.
2. Our secondary hypothesis is identical to those above with anxiety symptoms at 3-months, post-randomization.
3. We hypothesize that non-specialists and telemedicine will be less costly to deliver, equally effective and associated with similar use and cost of other health services, and thus likely cost-effective compared with specialist providers and in-person sessions

1.6. Study Design

This is a multi-center randomized non-inferiority trial to test the choice of delivery mode and provider, implementing a brief evidence-based PT of BA for perinatal depressive and anxiety symptoms, and to determine the underlying processes related to delivery and scalability of the PT from a multi-stakeholder prospective.

2 METHODS – Participants, Interventions and Outcomes

2.1. Participants and Setting.

N=1230 pregnant and postpartum women, aged ≥18, will be recruited from three study centers though their networks of clinics. Participants will be drawn from diverse and representative populations in Toronto, Chapel Hill, and Chicago. The flow chart (Figure 1) shows the process of recruitment and follow-up of participants in each hub. In Toronto, we will recruit from Sinai Health (SH), Women's College Hospital (WCH) and St. Michael's Hospital (SMH)—three major academic hospitals that are all affiliated with the University of Toronto, and our external referral networks. In North Carolina, we will recruit from three clinical sites affiliated with the University of North Carolina (UNC) Women's and Neuroscience Hospitals. In Chicago, we will recruit from fourteen affiliated obstetric and family medicine clinics. Our trained research assistants will approach all attending women or contact via phone for a screening. At one site (WCH), permission will first be required by a member of the participant's circle of care, will introduce the study and ask if she is interested in learning more. If the participant agrees, the trained research assistant will approach the participant in person or via phone for screen. Recruitment personnel will use established safety protocols (*Appendix A - Safety Protocols and Forms*). Inclusion and exclusion criteria are detailed in Table 5. Screening will take place in-person or over the phone, as determined by the referral pathway.

Table 5. Inclusion and Exclusion Criteria

INCLUSION	EXCLUSION
● EPDS≥10 ● ≥18 years ● Pregnant up to 36 weeks or 4 to 30 weeks postpartum ● Speaks English or (US sites) Spanish	● Active suicidal intent (ideation and plan), active symptoms of psychosis or mania ● Psychotropic medication dose or medication change within two weeks of enrollment or beginning treatment ● Ongoing psychotherapy (no more than once every 8 weeks or during the duration of the intervention) ● Active substance abuse or dependence ● Severe fetal anomalies, stillbirth or infant death at time of enrollment for index pregnancy ● Non-English, non-Spanish speakers

Following informed consent, potential participants will be screened for eligibility using a list of previously-used and validated measures. Our screening measures are briefly described below and can be found in *SUMMIT Screening Measures, Appendix C – Study Measures*, pg 4.

- Active suicidal intent will be assessed using the Columbia Suicide Severity Rating Scale[44]. This is the most commonly used scale to assess suicide risk and has been used extensively with adult populations and by our own teams among adult perinatal populations[45].
- Active symptoms of psychosis will be assessed using two screening questions: "Do you hear things that others can't or don't hear?" and "Have you ever felt that someone was playing with your mind?" The measure was validated against clinician ratings of psychosis, and performed with strong psychometric properties including a sensitivity of 53% and specificity of 98% for clinically significant hallucinations or delusions, and sensitivity of 32% and specificity of 99% for identifying people in an early phase of psychosis (clinical high risk or

first episode psychosis). To reduce false positives, these two stem questions have been further adapted for the purposes of this study and perinatal population to be supplemented with probes from the frequently used Mini International Neuropsychiatric Interview (MINI)[46].

- Mania will be assessed using a one-item question that asks "In the last two weeks, have you had periods of feeling so happy or energetic that you did not need to sleep or behaved in ways that were unusual for you and did this worry your friends and family?". This question has been used extensively by UNC Perinatal Mood Disorders Program[47].
- Alcohol use will be assessed using the Alcohol Use and Disorders Identification Test (AUDIT) has been used extensively worldwide to screen for alcohol use and dependence[48]. Its psychometric properties are well-validated[49] and patient-centered in its brevity.
- Three questions, borrowed from the psychometrically-robust DAST Drug Use Questionnaire[50], will assess participants' reported drug use to determine whether participants engage in drug use and the extent to which it influences their lives. We selected three key questions as individuals have found that even single-item questionnaires are valid to assess drug use in primary care [51]. These three questions are:
    a. Do you currently use drugs other than those required for medical reasons? [Probe: [if NO] What about marijuana?]
    b. Does it negatively impact your life?
    c. Does your spouse (or parents) ever complain about your involvement with drugs?
    - Finally, severe fetal anomaly, stillbirth or infant death will be determined by participant self-report and/or verified with a chart review. Language aptitude in English or Spanish will be determined when the potential participant is approached.

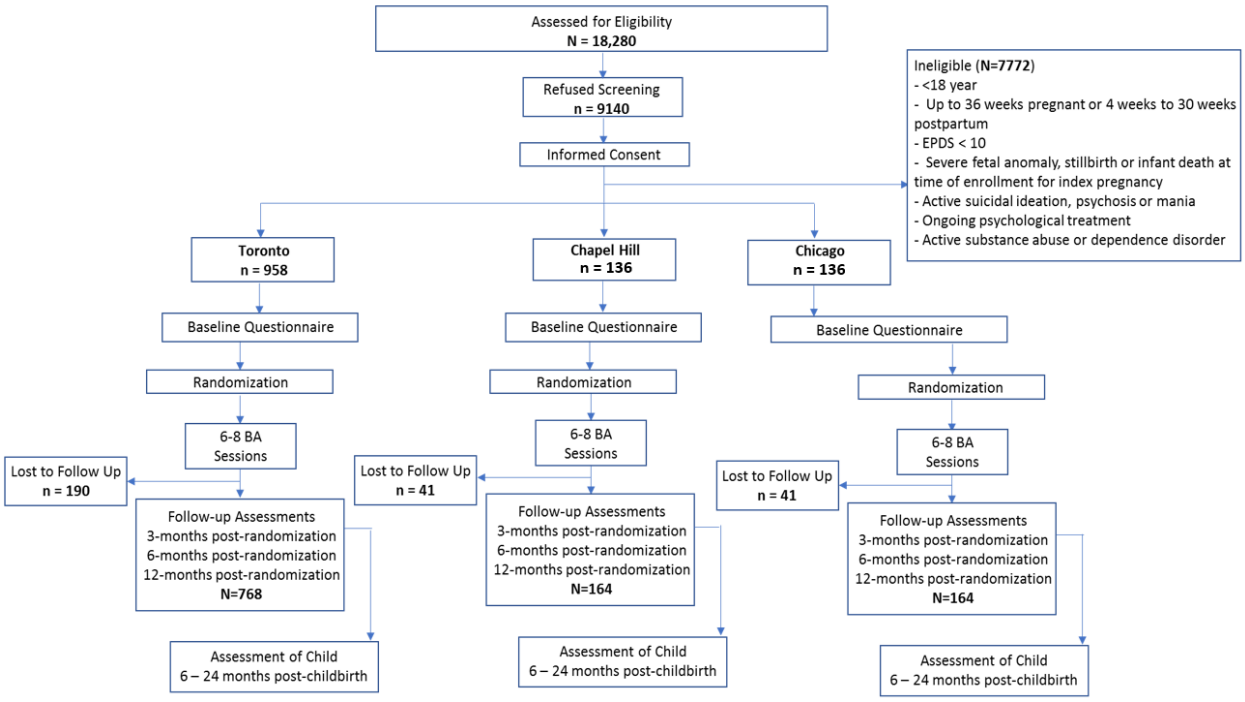

Figure 1: Study Flow Chart

2.2. Treatment and Intervention Arms

2.2.1. Treatment. The intervention will consist of six to eight weekly sessions of BA delivered to each participant individually. The current manual (*Appendices B1 and B2 – Treatment Manual and The PREMIUM Counselling for Relationships Manual, respectively)*) have been adapted from two well-established and tested manuals— the Alma Program for perinatal populations in Colorado and the Healthy Activity Program (HAP) from Goa, India, respectively. BA posits that the key to feeling less depressed is to increase enjoyable or fulfilling activities that align with one's values[32,33] and targeting the mechanisms of patient activation. Unlike other cognitive behavioral interventions for depression, BA explicitly targets avoidant coping, treating it much like any other anxiety disorder. This should facilitate our secondary focus on anxiety. Key strategies include: psychoeducation, behavioral assessment, activity monitoring and structuring, activation of social networks, interpersonal effectiveness, and problem solving.

2.2.2. Treatment indicators. Compliance is defined as attending a minimum of six sessions (reduced if patient has two consistent sessions of EPDS session-wise scores<10) over a period of four months unless a pregnancy- or COVID-related hiatus is experienced. Reasons for dropout will be recorded in the individual patient's treatment log (*see SUMMIT Treatment Log Record Form, Appendix C – Study Measures,* pg 40) using a data capture and storage system called the Research Electronic Data Capture (REDCap™). **Response** and **remission** will be defined as PHQ-9 of less than 10 and less than 5, respectively, at 3-months post-randomization. This period will be extended in the event that there is a hiatus to treatment due to significant perinatal-related life events (e.g., giving birth, obstetrical complications, COVID-19). Another indicator will include the estimation of comorbidity on both the EPDS and GAD-7.

2.2.3. Intervention Arms. During regular operations participants will be randomized to one of four arms (Figure 2):

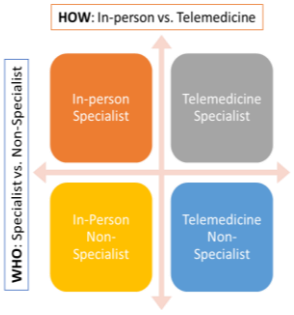

Figure 2. Intervention Arms

2.2.3.1 Telemedicine vs. In-Person. Telemedicine will be implemented via the Ontario Telemedicine Network and Zoom™ in Toronto, via the UNC TelePsychiatry Program in Chapel Hill, and via Zoom™ in Chicago. All offer secure personal videoconferencing service, are free to

registered healthcare providers (including nurses), and PHIPA/HIPAA compliant. They also permit video-visits and scheduling and are accessible on PC, Mac, Android and iOS systems. For participants who do not have access to a phone, tablet, or computer, or whose device malfunctions, study tablets will be available and enabled for use on a temporary basis. Instructions for use will be provided. During telemedicine sessions, established safety protocols will be implemented (see *Appendix A – Safety Protocols and Forms*). Participants assigned to Telemedicine can do their BA sessions in whatever private location they prefer (e.g., home or elsewhere). In-person sessions will be held at participating clinical care sites within UToronto, UNC and NorthShore Chicago.

2.2.3.2 Non-Specialists vs. Specialists. SUMMIT treatment providers will include both non-specialist providers (NSPs) and specialists. NSPs will be healthcare workers with general health care professional skills (as assessed during recruitment) but without formal training in mental health care or experience delivering PTs. Specialists will include individuals with formal training in mental health care delivery (e.g., psychiatrists, psychologists and social workers) with experience in treating perinatal mental illness and a minimum of 5 years of experience delivering psychological treatments. To ensure consistency across hubs, a minimum of two clinical leads (i.e., expert Co-I's designated to oversee the training and clinical implementation of the BA treatment), one from the local hub and one from another hub, will conduct training.

The original study design included a 4-arm, 1:1:1:1 randomization strategy (Phase 1, see Figure 3 below) which was modified to a 2-arm, 1:1 (telemedicine only; Phase 2) strategy at the start of the COVID-19 pandemic. Following vaccine rollout and the subsequent decrease in COVID-19 cases, we transitioned to a weighted 4-arm, 3:1 (in-person:telemedicine; Phase 3) strategy in an effort to equalize the number of randomized in-person to telemedicine cases. However, in light of the frequently changing COVID-19 situation, we will transition to an adaptive study design (Phase 4) in which we move between a balanced 4-arm randomization strategy (1:1:1:1; similar to Phase 1) and a 2-arm (telemedicine only; similar to Phase 2) randomization strategy (1:1) (Figure 3) based on COVID-related institutional restrictions that impact in-person care and recommendations at each site. This protocol modification will allow us to continually adapt to the rapidly-changing COVID-19 situation and reflects the pragmatic and patient-centered nature of this study. Participants will be informed about these ratios in the consent forms (see *Informed Consent Form – 2 Arm, Appendix E – Study Documents*, *pg 4* and *Informed Consent Form – 4 Arm, Appendix E – Study Documents, pg 16)*. The study design and relevant statistics have been reviewed by the study team, biostatistician and independent board of trial advisors.

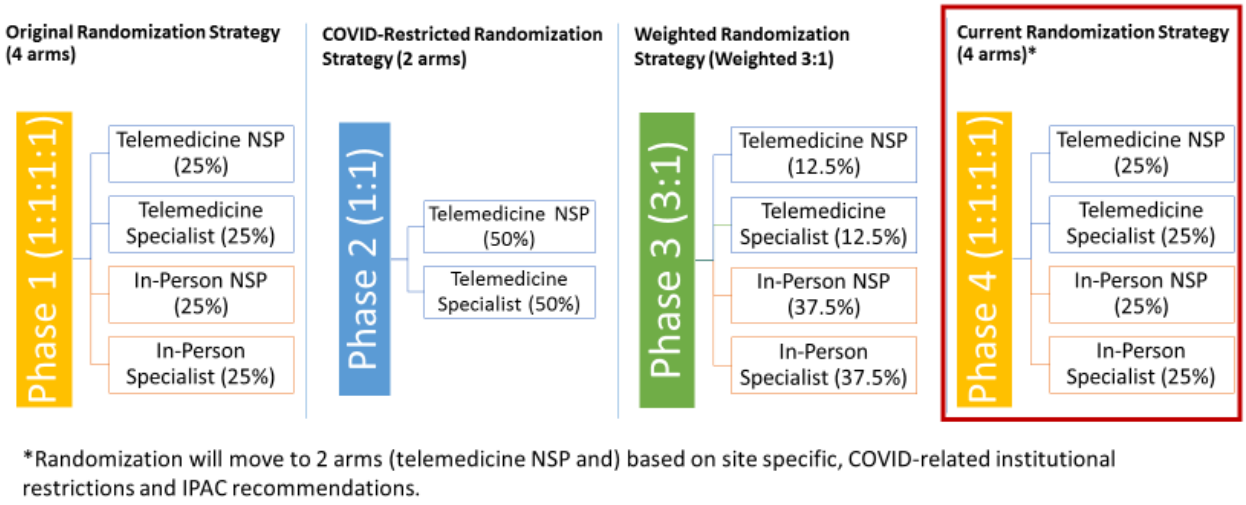

Figure 3. Updated Randomization Strategy

2.3 Non-Specialists: Recruitment, Training, Supervision and Selection

2.3.1 Recruitment. Recruitment of NSPs will be based on the aforementioned inclusion/exclusion criteria as well an expressed desire to help women with depressive and anxiety symptoms. During recruitment, a selection interview will take place, whereby NSPs are assessed based on their common skills—the style with which they interact with a mock patient—using the Therapy Quality Scale[52]. We have used this method previously and found that it correlates with trainee competency. Trainee providers will be recruited by placing advertisements in newspapers and through word of mouth, and selected based on their performance in a structured interview and role-play during an intake interview. The same procedures will be implemented in each hub.

2.3.2. Training & Competency Tests. In Fall 2019, those trainees selected will be invited to participate in a 5-day participatory training workshop covering the manualized BA treatment for perinatal mental health populations (*Appendix B1 – Treatment Manual*) and an adapted version of The PREMIUM Counselling for Relationships Manual[23] (*Appendix B2 – The PREMIUM Counselling for Relationships Manual*). Trainees who meet competency standards (based on a multiple-choice questionnaire used for another BA trial; *see Competency Measure Survey, Appendix C – Study Measures,* pg 49) and in an interaction with a standardized patient (actor) rated by experts using an adapted version of the Q-HAP[52] (*see Quality of Healthy Activity Program (Q-HAP), Appendix C – Study Measures*, pg 45) will be selected for the 8-week internship phase of the trial. During the internship phase, trainees will see up to 2 participants (one via telemedicine and one in-person) to implement the BA treatment. In addition, all trainees will receive training in the safety protocols (*see Appendix A*).

2.3.3. Supervision. The trainees will receive weekly group supervision from the hub intervention lead (see Table 1 above). All intervention leads are co-I's on the current study, with a minimum of 10 years' experience training and supervising lay and specialist therapists in psychological treatments and who have allocated time to focus on providing training and

supervision to NSPs. Therapy quality will be assessed using a reliable and validated tool (an adapted version of the Q-HAP[52]; *see Quality of Healthy Activity Program (Q-HAP), Appendix C – Study Measures*, pg 45). As the trainee counsellors gain experience in delivering the intervention, the supervision format will evolve from expert led (that is, the clinical lead and specialist who is skilled in the delivery of the PT) to peer-led group supervision. All treatment providers will record all BA sessions for training and supervision purposes and for fidelity measures.

2.3.4. Selection. Only trainees who achieve competence, as assessed by standardised role-plays and therapy quality assessments, will be selected to deliver BA during the trial.

2.4 Specialists. Specialists will include a wide range of specialties (e.g., psychiatrists, clinical and counseling psychologists, social workers) with a minimum of 5 years of formal experience in delivering psychological treatments. They will be recruited through word-of-mouth and based on their location, interest and availability, and good standing with their respective colleges. In Fall 2019, specialists will also receive the BA workshop followed by an internship phase where they will see up to 2 participants each with expert-led supervision followed by guided peer-supervision. The primary focus of the workshop for the specialists will be on the contents of the BA manual.

To assess provider proficiency with telemedicine, we will collect information from the trainees about their perceived telemedicine proficiency at the end of training and then after their last session with each participant during the internship phase (*see SUMMIT Perceived Telemedicine Platform Proficiency Self-Assessment Questionnaire (for providers), Appendix C – Study Measures*, pg 105). Participants assigned to the given provider will also be asked to rate their provider's telemedicine proficiency after their last session (*see SUMMIT Perceived Telemedicine Platform Proficiency Self-Assessment Questionnaire (for participants), Appendix C – Study Measures*, pg 105). The questionnaire is a brief, 5-item scale rated from 0 (Strongly Disagree) to 4 (Strongly Agree).

2.5. Outcomes and Timing of Assessments.

Table 6. Outcomes

Study Variable	Instrument	Outcome (Range)
Maternal: Measured at Baseline and 3*-, 6- and 12-months post-randomization		
Maternal Characteristics**	**Trial Baseline Questionnaire**[53,54]	Self-reported age, education level, marital status, immigrant status and ethnicity, clinical history with depression or anxiety (severity, chronicity, number of prior episodes, and age at first episode, occupational status, number of children, pregnancy intention, pregnancy history, delivery and birth, sexual orientation.
Depressive Symptoms	**Edinburgh Postnatal Depression Scale (EPDS)**[41]	Mean continuous score of a 10-item scale (0-30)
Anxiety Symptoms	**Generalized Anxiety Disorder Scale (GAD-7)**[42]	Mean continuous score of a 7-item scale (0-21)
Response & Remission	**Patient Health Questionnaire-9 (PHQ-9)**[55]	Response: PHQ<10 Remission is defined as PHQ<5
Perceived Support	**Multidimensional Scale of Perceived Social Support**[56]	Mean continuous score of a 12-item scale (1-84)
Disability Assessment	**World Health Organization Disability Assessment Schedule (WHODAS)**[57]	Mean continuous score of a 12-item scale (0 – 48)
Quality of Life Assessment	**EQ5D-5 Level (EQ5D-5L)**[58]	Mean continuous score of a 5-item scale (1-25)
Trauma Symptoms	**Abbreviated PTSD Checklist (PCL-6)**[59]	Mean continuous score of a 6-item scale (1-30)
Patient-Reported Activation	**Premium Abbreviated Activation Scale**[34,60]	Mean continuous score of a 5-item scale (0-20)
Patient Satisfaction***	**Client Satisfaction Questionnaire-8 (CSQ-8)**[61]	Mean continuous score of an 8-item scale (1-32)
Therapeutic Alliance***	**Working Alliance Inventory-Short Revise (WAI-SR)**[62]	Mean continuous score of a 12-item scale (1-60)
Health Service Utilization	**Health Service Utilization Form**[63]	3, 6 months - Total score of a 16-item scale (0-32); 12 months – Total score of a 19-item scale (0 – 38)
Treatment Preference**	**Delivery of treatment and treatment provider preference**	Score of 0 or 1
COVID-19 Exposure	**1-item question on COVID-19 exposure**	Self-reported
Health Benefits Access and Use	**2-item question on access and use of health benefits**	Self-reported
Sexual Orientation◊**	**2-item question on sexual orientation**	Self-reported
Home Visit Survey□**	**4-item question on home visit perspective**	Self-reported
Intervention Costs	**A survey of costs associated with the intervention**	Self-reported and health records, related to the economic evaluation portion of the study
Treatment: Measured at every session during treatment, unless otherwise indicated		
Dosage	**Treatment Log**[64]	Frequency of Sessions Attended
Therapy Quality****	**Quality of Healthy Activity Program (Q-HAP)**[52]	Mean continuous score of treatment-specific BA skills (0-4) and general counseling skills (0-4)
Session Depressive / Anxiety	**Session-by-session EPDS**[41] and **GAD-7**[42] **scores**	Mean continuous score of a 10-item scale (0-30) on EPDS and of a 7-item scale (0-21) on GAD-7
Homework Adherence	**Treatment Log**[64]	Mean continuous score of a 1-item question (0-2)
Adverse or Serious AEs	**Anytime an Adverse Event (AE) or Serious AE occurs**	Any event that represents a series threat to the safety of the mother or her child (see *Appendix D*)
Health Service Utilization	**Health Service Utilization Form**[63]	Total score of a 16-item scale (0-32)
List of Medications	**List of Medications**	Self-reported list of medications
COVID-19 Exposure	**1-item question on COVID-19 exposure**	Self-reported
Child: Measured at 6 to 24-months post-child birth unless otherwise indicated		
Birth weight & Length	**Retrieved from hospital chart or self-report†**	Assessed at birth

Breastfeeding	**Whether breastfeeding and if stopped age stopped.**[54]	Total Number of Months (0-12)
Psychosocial Stimulation	**Home Observation Measurement Evaluation**[65]	Total score of a 45-item checklist
Mental Development	**Bayley Mental Development Scales IV**[66]	Mean continuous score of cognitive, receptive and expressive language development

* Assessment period will be extended to account for post-treatment outcomes when there are perinatal-related interruptions to treatment (e.g., giving birth, obstetrical complications, COVID-19); **Only at baseline; ***Measured at 3-months post-randomization only; ****Randomly selected for supervision, rated by self, peers, expert supervisor; †Self-report will be used when hospital charts are outside of the recruiting site; ◊Measured at baseline for participants consented after addition of measure to baseline questionnaire, measured at 12-months post-randomization for those who completed baseline before addition of measure; □Measured once at 3-, 6- or 12-months post-randomization.

All measures proposed in the current study have been previously used and validated in one or more of the investigators' trials targeting perinatal mental health[23,24,34,40,53,54,60,67,68] and selected because of their role in the presumed causal pathway. Our emphasis on mothers' self-reported data adheres to PCORI's methodology standards that emphasize that the patient population is the best source of information.

2.6. Sample Size

The primary outcome measure will be participant's EPDS mean score at 3 months after randomization; the assessment period will be extended in the event that there is a hiatus to treatment due to significant perinatal-related life events (e.g., giving birth, obstetrical complications, COVID-19) or to one week after the last session if participants complete treatment after the 3 month-post randomization period. The study is designed to be a non-inferiority trial with four arms: telemedicine specialist, telemedicine non-specialist, in-person specialist, in-person non-specialist. The two primary hypotheses are that telemedicine will be non-inferior to in-person, and that non-specialist delivery agents will be non-inferior to specialist delivery agents. The sample size calculation is based on an EPDS mean estimate of 7.93 (SD=4.68[53]). Using a non-inferiority margin of 10% (i.e., EPDS score of 0.79) in relation to the mean), and an alpha=0.05 for our first comparison (NSP vs. SP) we require 431 participants in each of the two groups to provide greater than 80% power. To account for 10% drop out, the sample size is inflated to N=958 (479 per group, NSP vs. SP). Using a non-inferiority margin of 13% (i.e., EPDS score of 1.03 in relation to the mean), and an alpha=0.05 for our second comparison (TM vs. IP), we require an additional 268 IP participants. When combined with the 958 TM participants, this yields a total study sample size of N=1,226 which accounts for 10% loss to follow-up. Previous randomized non-inferiority trials[37] have demonstrated that a 14% non-inferiority margin can be considered clinically meaningful.

We are planning a final sample size of N=1,230, four more participants than the power calculation requires. The reason for this n=4 over recruitment, is to provide flexibility due to the fact that we do not receive each site's recruitment numbers until the end of each day. These statistics have been reviewed and approved by our study statistician.

2.7. Screening, Recruitment, and Consent

2.7.1. Quantitative Data. Across the three hubs, we anticipate assessing 18,280 participants for potential eligibility, obtaining informed consent to screen 9,140, and recruit and retain a sample size of N=1,230 women (see Figure 1). We anticipate that 50% of participants will be recruited in Toronto, and 25% each in Chapel Hill and Chicago, respectively. Participants will be recruited from all study sites though their networks of in-patient obstetrical units, family practice, obstetrical and mental health clinics, and community partners (*see Recruitment Poster – 4 Arm Without Tabs, Appendix E – Study Documents, pg 123*). Referrals will also be elicited from clinicians from site hospitals and satellite clinics who will send patient contact information to the research coordinator. Relevant stakeholder groups, such as Life with a Baby, will also be provided with information about the trial to share with their members. Interested individuals can contact the trial for more information.

To help alleviate the mental health burden that COVID-19 imposes on our health care systems, physicians not affiliated with a study site will be able to refer their patients to the

SUMMIT Trial for screening and enrollment, if eligible and willing. This includes referrals from physicians not affiliated with the trial (external referrals) and referrals between recruitment sites (internal referrals). Internal referrals will occur among patients who express interest in participating in the trial at a study site where recruitment is on temporary hold due to treatment provider capacity.

1. **External referrals:** Study Investigators will notify their physician colleagues not affiliated with a study site that the SUMMIT Trial is open to their patients (*see Email for External Physicians, Appendix E – Study Documents, pg 126 and 129*) and provide a description of the study (*Information for External Physicians, Appendix E – Study Documents, pg 125 and 128)*. Physicians will briefly describe the study to their patients, and ask them if they are interested in learning more. If the patient is interested, the physician will then request their patients' permission to send their name and contact information to the study team. Potential participants may also learn of the trial through relevant stakeholder groups in which they are members (for example, Life with a Baby). Interested individuals will be able to contact the trial for more information.
2. **Internal referrals:** Patients who express interest in participating in the trial at a study site may not be able to immediately access treatment if randomization is temporarily paused due to reaching treatment provider capacity. In this event, patients will be given the option of a referral to the trial at a different site that is still able to accept new participants. If they agree, the study team at the referring site will request the patient's permission to send their contact information to the study team at the second site. Contact information will be faxed or encrypted and sent via secure hospital email to a secure hospital email at the second site. No identifying information will be included in the title or body of the email. Participation in the trial at the second site will not affect the patient's care at the referring site. Upon receipt of the contact information, the study team at the second site will contact the patient and follow the consent procedures described in this protocol.

Women will be contacted either by a trained research assistant, a treating physician or another member of the circle of care to ask whether they can be contacted by one of the trained research assistants. They will then receive a brief overview of the study by a research assistant, while an in-patient, visiting an out-patient clinic, or over the phone. Women who agree will then be fully informed of the study procedures by the research assistant, and given an informed consent form (*see Informed Consent Form, Appendix E – Study Documents,* pg 4 and 16) to review. Women will be given as much time as they need to decide whether they wish to participate, and the research assistant will answer all questions they may have. Once they sign the consent form, women will be screened using the Edinburgh Postnatal Scale for Depression (EPDS≥10[41]) and General Anxiety Disorder (GAD)-7[42] for eligibility. If an EPDS administered in the last 7 days is in the patient's clinical chart, we will use and enter that score for screening instead of asking the patient to complete the scale again. In the interest of patient-centered care, women who give their permission may also complete informed consent

over the phone, witnessed by a third-party. These women will be mailed a consent form and be fully informed of the study procedures by the research assistant over the phone. If they decide that they want to participate, they will give verbal consent verified by a third-party (a second study team member), and will then be emailed links to screening measures. These women will be provided with a self-addressed, stamped envelope to mail their signed consent forms back to the site, and will then be sent a consent form with their own signature and the name and signature of the RA who obtained their informed consent for their files.

To reflect the pragmatic and patient-centered nature of the trial and in an effort to uphold recommended social distancing guidelines for both participants and research staff, we will not ask participants to mail back signed consent forms during COVID-19 and will instead rely on verbal consent given over the phone and confirmed by a third party. To ensure a paper trail, participants will also indicate their consent in REDCap™. During verbal consent, participants will be sent a link to a consent form in REDCap™. The consent form will be the same as the form they receive for verbal consent, but will include checkboxes for each permission and a checkbox to indicate that they agree to participate in the trial. The research assistant obtaining informed consent and the third party will add their names and signatures to the form to indicate that they explained the study and confirmed consent. Participants will not enter their names in REDCap™ during this process. Consent forms will only be identified by participant ID and will be linked to participant profiles in REDCap™, respectively.

At the American sites, research assistants will either be bilingual or use a translation service so that they can explain the study to potential participants who only speak Spanish and a consent form will be provided in that language. All study questionnaires and the BA intervention will be conducted in Spanish for Spanish-speaking participants. Recruitment rates will be closely monitored. Self-referrals will also be accepted if recruitment rates are low. Brochures outlining the study's aims, potential impact and inclusion/exclusion criteria will be provided to all recruitment sites (*see Brochure, Appendix E – Study Documents*, pg 115 and 117).Trained research assistants will follow-up with each appropriate referral and obtain informed consent (Figure 3). Recruitment will begin in January 2020 and continue until March 2023 (39 months in total) at all sites.

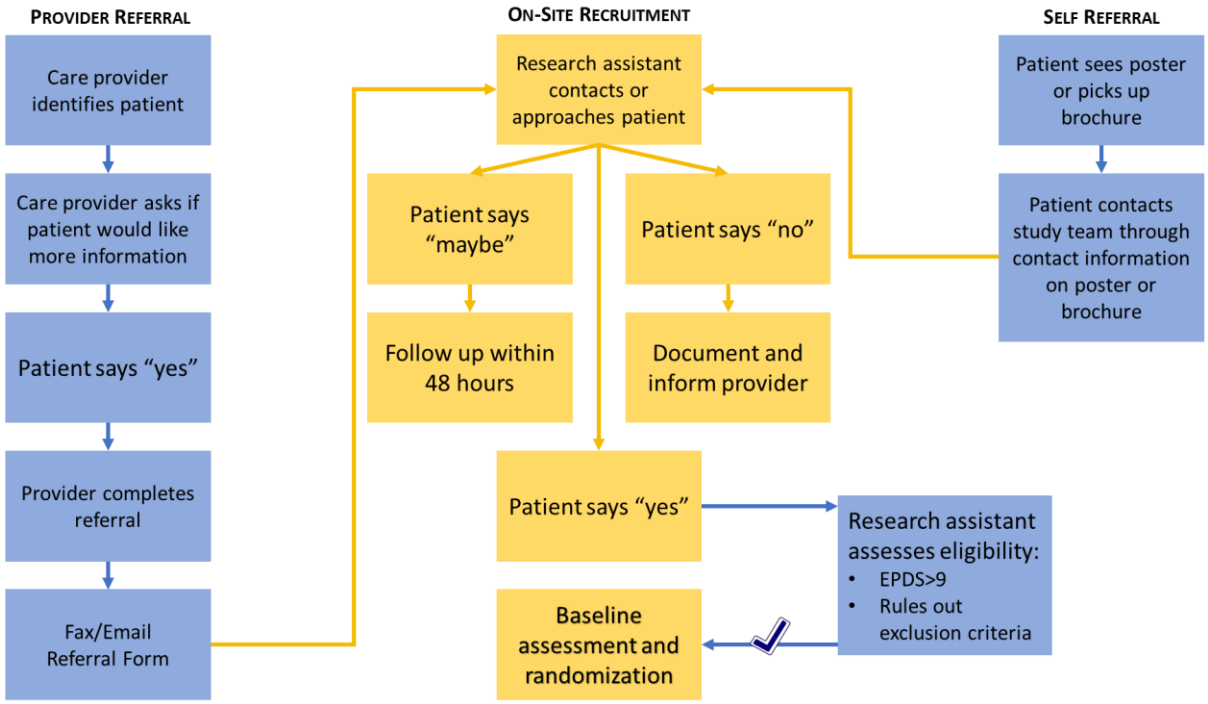

Figure 4: Recruitment Process

2.7.2. Qualitative Interviews. A team of trained interviewers will collect qualitative data through semi-structured, in-depth interviews and focus group discussions. These individuals will be independent from both the recruitment and data teams. Topic guides were developed and reviewed with a range of key stakeholders to ensure questions are relevant and acceptable.

2.7.2.1. Qualitative Phase 1 (Sep 2019 to Apr 2020). All NSPs, specialist providers, and supervisors will take part in qualitative semi-structured interviews to explore their work experience related to training and supervision. All potential treatment providers will be informed that they will be required to complete interviews when they are recruited. All successful applicants who accept the positions as providers or supervisors will understand that in accepting the position, they will participate in semi-structured interviews that will be used in the research study as a part of their job.

2.7.2.2. Qualitative Phase 2 (May 2020 to Mar 2024). A subset of study participants, significant others, treatment providers and supervisors, and a wide range of stakeholders (patient advocates, health professionals, and policy makers) will be invited to participate in qualitative, semi-structured interviews to examine key aspects of the intervention, including content, delivery and modality, enactment of key messages, scalability, and, where applicable, supervision procedures. In the first six months of this process, we will also ask several semi-structured questions to examine whether aspects of the treatment content or delivery were modified in light of COVID-19.

Participants. – We will conduct the qualitative interviews with representative subsets of perinatal participants (n=300) across the four arms, selected to represent different variables such as symptom severity, perinatal period (antenatal and postnatal), sociodemographic

variables, and treatment completion status (completers and non-completers). Potential participants will be identified by the Data Manager and approached by a research assistant if they would consider taking part in the interviews. A research assistant will approach those willing to explain the interview process and secure informed consent (*see Informed Consent Form – Participant Qualitative, Appendix E – Study Documents,* pg 37). The research assistant will answer any questions and the women will be given as much time as they need to make an informed decision. Please note that this research assistant is independent and has not conducted any baseline or follow-up assessments with potential participants.

Significant Others. – Up to 60 significant others will be interviewed as a part of Qualitative Phase 2. Participants will indicate in their informed consent form if their partners can be contacted to participate in interviews about their experiences related to the treatment intervention. If they indicate yes, we will ask the participants to speak with their significant others and let their significant other know that they may be contacted by a research assistant. Spouses or partners are preferred but we will accept any significant other that the participant identifies. If the participant agrees, a research assistant will contact their partner and explain the interview process and secure informed consent (*see Informed Consent Form – Partners Qualitative, Appendix E – Study Documents,* pg 46) to review. The research assistant will answer any questions the partners may have via whatever means of communication they prefer (e.g., phone or email), and the partners will be given as much time as they need to make an informed decision.

Treatment Providers and Clinical Leads. – All NSPs, specialist providers and clinical leads will be expected to participate in the Qualitative Phase 2 interviews as a part of their job description. Treatment providers and clinical leads were previously asked to provide informed consent to participate in up to three interviews and then will be given a consent addendum form to agree to participate in an additional five interviews for a total of eight interviews.

Other Stakeholders. – Other stakeholders include relevant health professionals (e.g., nurses, midwives and physicians from mental health, obstetrics and family medicine), patients and patient advocates, telemedicine and insurance experts, study staff, and policy makers identified by the PIs at each site and invited to engage in the trial as committee members and consultants. Stakeholders are informed that qualitative interviews about their views and experiences are a voluntary part of the study and that they will be sent invitations to participate. The research assistant will send a link to the informed consent in REDCap™ by email (*see Informed Consent Form – Stakeholders Qualitative, Appendix E – Study Documents,* pg 55, *Informed Consent Form – Staff Qualitative, Appendix E – Study Documents,* pg 62 ,*SUMMIT Stakeholder Email Interview Invite – Appendix E – Study Documents,* pg 61 and *SUMMIT Staff Email Interview Invite – Appendix E – Study Documents,* pg 68). The stakeholders will be able to ask the research assistant any questions they may have by phone or email, and the stakeholder will be given as much time as they need to make an informed decision. The RA will follow up with stakeholders who do not respond to the email to check if they have any questions and if they are interested in participating.

2.8. Randomization

All participants will be stratified by perinatal period (antenatal vs. postnatal) and then randomized within site to one of the four arms, unless restricted to two arms due to COVID-related institutional restrictions of in-person care (see section 2.2.3.2) and using a web-based randomization service as part of REDCap™. Randomization will occur only after informed consent has been obtained and all baseline measures have been completed and all inclusion/exclusion criteria verified with the participant. We will enroll equal numbers in each of the four intervention arms. This strategy allows sufficient power to determine whether NSPs are no less efficacious than specialists and telemedicine is no less efficacious than in-person treatment, respectively, in settings that collectively assure ethnic and racial diversity, as well as potential heterogeneity in age, variability in symptom severity as well as access to resources.

2.9. Blinding and Masking

If eligible, baseline assessments will be carried out on site at time of recruitment or at home, as per the participant's preference. The baseline assessment will be accessed by a secure link sent via REDCap™, which participants will be asked to complete before their first treatment session. Once baseline is completed, the participants will be randomized. The outcome assessments at 3, 6, 12-months post-randomization will also be completed by secure links sent via REDCap™. Participants and providers will not be blind to treatment allocation; however, both will be told that we are evaluating the same intervention and there is clinical equipoise about whether one is better than the other. Outcome measures for participants will be administered through REDCap™. Trained research assistants independent of the treatment delivery and blind to allocation status will conduct the participant interviews and child assessments.

2.10. Economic Evaluation

We will evaluate the cost-effectiveness of non-specialist providers and telemedicine, compared to specialist providers and in-person sessions. We will undertake a cost-utility analysis conducted at the 12-month post-randomization follow-up point. The primary outcome will be quality-adjusted life years (QALYs) calculated using the EuroQol 5 Dimension 5 Level (EQ-5D-5L) measure of health-related quality of life, collected at 3, 6 and 12-months.

2.10.1. Economic-related data. The primary outcome related to the economic evaluation is quality adjusted life years (QALYs) calculated using the EuroQol 5 Dimension 5 Level (EQ-5D-5L) measure of health-related quality of life[58]. An advantage of the EQ-5D-5L is that an overall utility score of health-related quality of life can be obtained, which facilitates comparisons with other interventions and health states in other illnesses.

2.10.2. Service use. The use of all health services, other than the intervention (BA), are being recorded using an adapted version of the health services utilization questionnaire (HSUQ)[63]. Data on the use of BA (number and duration of therapy contacts and with whom) will be collected from existing clinical records. This measure (*see Health Service Utilization Questionnaire, Appendix C – Study Measures,* pg. 35) was adapted with reference to other measures used in perinatal depression populations[69] and for use in Canadian and US contexts with input from service users and clinicians. It will be completed at baseline (covering the last 3 months) and each follow-up interview (covering the period since the last interview). The HSUQ records use of health services, specifically: hospital inpatient stays,

outpatient appointments, emergency department visits and ambulance contacts, and community health contacts.

2.10.3. Costs. BA will be directly costed for each comparison taking a standard micro-costing approach[70]. Unit costs for BA will include all employer costs and appropriate overheads (capital, managerial, administrative etc.; *see List of Costs, Appendix C – Study Measures,* pg. 108). The cost of supervision will be included and the time each therapist spends on various direct and indirect participant-related activities (non-participant contact time including, for example, training, administration, meetings with other professionals etc.) will be estimated using a questionnaire. Data on the use of all other health services will be costed for each individual participant over the course of the trial. Unit costs in both Canada and the USA will be obtained from Canadian guidelines on person-level costing[71] and local Medicaid, Medicare, and private insurer fee schedules commonly available in the US[72], respectively.

3 METHODS – Data Collection, Management & Analysis

3.1. Data Collection

Two types of data will be collected**.**

3.1.1. Quantitative Data. All variables will be recorded into the web-based data management software REDCap™. All variables, with the exception of those recorded in the treatment log collected by the delivering provider (*see SUMMIT Treatment Log Record Form, Appendix C – Study Measures,* pg 40) and childbirth outcomes, will be automated for collection via REDCap™. Children will be assessed in-home at 6-24 months post-birth (Figure 1; Table 6) with data recorded in REDCap™.

3.1.2. Qualitative Data. An independent team of trained interviewers will collect qualitative data through audio-recorded, semi-structured, in-depth interviews and focus group discussions. Interviews will examine key aspects of the intervention including intervention content, delivery, and enactment. Interview guides were developed and reviewed with specific advisory committees including mothers and clinicians will be used. A maximum variance, purposive, and snowball sampling approach will be used, when applicable, to adequately capture participants across severity levels, perinatal periods, sites, key sociodemographic factors and stakeholder groups.

3.1.2.1. Qualitative Phase 1. We will explore experiences related to training by examining relevant barriers and facilitators reported by trained NSPs, specialists and supervisors. All providers and supervisors will be interviewed by stakeholder group and within each hub. Individuals will be interviewed at the end of training that will take place in Fall 2019 and before the trial, and interviews with newly on-boarded providers will be conducted on an ongoing basis.

3.1.2.2. Qualitative Phase 2. We will assess the relevant barriers and facilitators of the intervention delivery and implementation with a subset of participants from the existing trial and with multiple stakeholders approached to participate in semi-structured interviews. We will conduct interviews with various subsets of n=300 perinatal participants who have been enrolled in the study, including representation from each of the four arms, treatment completers vs non-completers, and antenatal vs postnatal enrollment. An unblinded data

manager will identify potential participants. In addition, we will interview up to n=6 0 significant others (including spouses or partners of participating perinatal participants), all specialist and non-specialist providers, and n=160 health professionals and relevant stakeholders such as patient advocates. A maximum variance, purposive, or snowball sample will be used, when applicable, to adequately capture participants across severity levels, perinatal periods, sites key sociodemographic factors and stakeholder groups. Each type of group will consent to being interviewed three times, except for treatment providers who may be interviewed up to eight times (*see Treatment Provider Qualitative Addenda, Appendix E – pg. 37 – 40*), and interviews will be conducted during all phases of the trial (beginning, middle, and end) to ensure that we are capturing relevant aspects of the interventions while it is being implemented. In addition, we will interview all treatment providers and clinical leads again to assess their perspectives over a later part of the trial and in the event that there are new providers trained. Finally, we will examine the data to see whether there were specific barriers, facilitators and modifications made to the content and delivery of the BA treatment in light of COVID-19.

3.2. Drop Out and Follow-Up Strategies

We expect 20% of participants to be lost to follow-up. This dropout rate is conservative and has been found in many similar trials using NSPs and specialist delivered PTs for perinatal depression in the US and Canada[24,39,53,54]. Based on a Cochrane review[73], numerous research on study retention (e.g.,[74-76]), and our previous experiences[34,40,53,54], it is realistic to expect a 12-month post-randomization follow-up rate of at least 80% when we incorporate the following strategies:

(1) **Participant Strategies**: We will complete qualitative interviews at the participant's pace, allow flexibility with the interviews, design information material at a low reading level, and provide realistic expectations regarding study involvement. We will send quarterly newsletters to study participants with information that might be of interest or that participants may find useful, and reminders about questionnaire timelines (*see SUMMIT Newsletter Template, Appendix E – pg 103*). A small token of appreciation will be provided to all mothers ($15 USD/$20 CAD gift card) per completed assessment. Upon completion of the 12-month questionnaire, participants will be entered in a raffle to receive a $100 gift card. The raffle will be held for every 100 participants who complete the questionnaire. Upon completion of the child assessment, participants will be provided a $40 token of appreciation. Participants reluctant to have RAs in their homes for the child assessment will have the option to complete the child assessment at the site hospital;

(2) **Tracking Strategies**: REDCap™ will be used as a study-specific computerized tracking system to calculate follow-up dates. We will request permission from each participant for two secondary contacts (e.g., relative or friend) in the event that we lose contact with the participant. In the event that we are not able to reach the participant and we reach out to a secondary contact, no information about the participant will be released to the secondary contact (*see Secondary Contact Form, Appendix E – Study Documents,* pg 83). We will verify and update contact information for the participant at each follow-up session. Email or telephone reminders will be sent to participants prior to each appointment or assessment;

(3) **Personnel Strategies**: All research staff will be trained and possess the following skills: (i) the capacity to communicate enthusiasm about the study; (ii) interpersonal persistence in a manner that is highly respectful; and (iii) a non-judgmental, empathic attitude. Research assistants will work flexible hours In order to accommodate telephone interviews at non-standard times. We used these strategies in our previous prevention (retention = 85% at 24 weeks postpartum) and treatment (retention = 87% at 36-weeks post-randomization) trials.

3.3. Participant Withdrawal

Participants can withdraw consent and end their participation in the study at any time. Non-completers and patients who do not respond to treatment will be referred to a trial psychiatrist or referred pathway (family physician, OBGYN, or external psychiatric services).

3.4. Data Management

3.4.1. Quantitative Data. Quantitative data will be collected in identical REDCap™ databases kept on secure institutional servers within each of the three participating cities. De-identified data from the US hubs will be extracted from the hub's REDCap™ database, encrypted, and transferred to Toronto, where it will be uploaded, entered, and stored in REDCap™ hosted at The HUB: Health Research Solutions in the Li Ka Shing Knowledge Institute. REDCap™ is a secure, web-based application designed to support data capture for research studies, providing 1) an intuitive interface for validated data entry; 2) audit trails for tracking data manipulation and export procedures; 3) automated export procedures for seamless data downloads to common statistical packages; and 4) procedures for importing data from external sources. Data collected via REDCap™ at The HUB, or uploaded to The HUB from US hubs, will be encrypted and stored on the local Storage Area Network (SAN) within St. Michael's Hospital and will be backed up regularly and stored off-site. The data center is designed such that there are daily backups made of all critical data. In addition, the backups are stored both locally, as well as at a remote off-site location, in the case of catastrophic failure at one location. With limited access privileges, 24-hour security, and around-the-clock monitoring, the data center is highly secure. Upon providing informed consent, participants will be assigned a unique study ID. Participants will receive emailed links to complete study assessments linked to their unique study ID profile in REDCap™. The cross walk for the study ID to personal information will be stored at each site in a password protected, encrypted file on a secure server within the institution. Participants will be identified in REDCap™ by study ID only. REDCap™ will store participant emails in a separate server from the study data. The emails will be inaccessible to study staff and protected against data export by being stored in a secure server in the REDCap backend, inaccessible to any study staff. Treatment data will be collected by the provider, child assessments will be collected by trained research assistants, and all entered on the same REDCap™ system.

3.4.2. Qualitative Data. All study interviews and focus groups will be audio-recorded using a digital voice recorder.

3.4.2.1. Qualitative Phase 1. Clinical Leads and Treatment Providers will be consented to the interviews and focus groups over the phone or a secure Zoom™ account from Toronto. All participants will be sent a link to their completed consent form in REDCap™ for their records. A research staff person will arrange a time and call them to review the consent. If they consent to participate, the research staff will ask them if they can record their consent. They will turn on the recorder, summarize the consent discussion, and ask the interviewee if they agree and consent. Once they give recorded informed consent, the interview will begin. Interviews and focus groups with the Clinical Leads and Treatment Providers will be conducted from Toronto, either in-person or via Zoom™—a secure and PHIPA-/HIPAA-compliant software used for remote communication. The recording will be coded with the provider and supervisor ID and the date. No identifying information will be listed on the name of the file. Audio files will be transcribed and identifiers removed during transcription. All study procedures for handling study audio-recorded data will be reviewed with study personnel every quarter to ensure that protocols are implemented with fidelity throughout the life of the study. Only approved research staff with a need to review or analyze data will have access.

3.4.2.2. Qualitative Phase 2.

Participants and significant others will be asked for informed consent, either in-person or over the phone or a secure Zoom™ account, by research staff at the site where they gave their consent to participate in the larger trial. Interviews and focus groups will be scheduled by the site where the participant gave their consent, and will be hosted over Zoom™ and conducted by an interviewer located in Toronto. This will ensure that PHI does not leave each study though the interviewer may be located at a different site. All study procedures for handling study audio-recorded data will be reviewed with study personnel every quarter to ensure that protocols are implemented with fidelity throughout the life of the study. Only approved research staff with a need to review or analyze data will have access.

Stakeholders involved in the SUMMIT Trial (ie. Committee members, attendance at stakeholder meetings, etc.) will be invited via email to participate in a qualitative interview or focus group (*see SUMMIT Stakeholder Email Interview Invite, Appendix E – Study Documents, pg 61*). If they are interested, they will be consented by research staff in the hub nearest and most convenient to them, either in-person or over the phone or a secure Zoom™ account. Interviews will be conducted from Toronto either in-person or via Zoom™—a secure and PHIPA-/HIPAA-compliant software used for remote communication. The recording will be coded with the provider and supervisor ID and the date. No identifying information will be listed on the name of the file. Audio files will be transcribed and identifiers removed during transcription. All study procedures for handling study audio-recorded data will be reviewed with study personnel every quarter to ensure that protocols are implemented with fidelity throughout the life of the study. Only approved research staff with a need to review or analyze data will have access.

3.4.3. Audio-Recordings. Audio-recordings of BA sessions will be stored locally at each site and will adhere to local privacy and data management policies. Audio-recordings may be securely

transferred between sites within a hub to allow the clinical lead and independent consultant's review.

3.5. Statistical Analyses

3.5.1. Primary Analysis. Using SAS 9.4, primary analyses will be performed based on an intent-to-treat basis. Two-sided significance levels of $p<0.05$ will be used for all analyses. Demographic (ethnicity, age, marital status) and other baseline variables (e.g., severity and chronicity) will be compared for potential differences between study groups using descriptive statistics (means, standard deviations, proportions) and the associated statistical tests (t-tests, chi-square tests). Those who drop out of the study will also be compared to those completing the study on baseline indices including maternal education and occupation. The primary outcome of non-inferiority in EPDS scores will be compared between groups (telemedicine vs in-person, and specialist vs NSP) at 3-months using a t-test. The assessment period will be extended in the event that there is a hiatus to treatment due to significant perinatal-related life events (e.g., giving birth, obstetrical complications or COVID-19). One t-test will be run to assess this and compare modes (is telemedicine non-inferior to in-person). Another t-test will be run to assess this and compare agents (is non-specialist non-inferior to specialist). Each t-test will look at the confidence interval around the difference in EPDS scores (say between telemedicine and in-person), and see if the upper bound contains the 10% (provider) or 13% (modality) non-inferiority margin (upper bound for the case in which we take the difference to be telemedicine minus in-person). We will conduct a sensitivity analysis to examine potential differences in baseline and outcome depressive symptoms and anxiety symptom scores to determine whether participants who were recruited and received treatment during the COVID-19 outbreak differed significantly from the larger sample.

In addition, we will run a linear regression model with mode, agent and a mode by agent interaction term to assess whether an interaction between mode and agent exists. In the unlikely event that randomization was not able to reduce bias by balancing out confounders, we will assess for confounding. This will be carried out by assessing each potential confounder individually in relation to the key predictor (mode or agent) and seeing if the potential confounder impacts the parameter estimate of the key predictor using the 10% change in estimate approach. The potential confounder would also need to be significantly related to the EPDS score outcome. Should any confounders exist, they'll will be included in all of the previously mentioned analyses. Rules on the number of variables allowed in a multivariable model will be followed. The rule on number of variables in a linear regression is to take the total number of observations divided by 10. Given our trial sample size, we will not have any concerns with overfitting our linear model for our covariates of interest (education level, marital status, ethnicity, baseline severity and chronicity, timing of treatment, compliance, perceived support and medication).

Our study statistician and two independent methodological consultants have reviewed and approved this plan of analysis.

3.5.2. Secondary Analysis. In our secondary analysis, we are interested in: i) assessing the trend in EPDS scores over time (baseline, 3, 6- and 12-months post-randomization) between

groups (telemedicine vs in-person, and specialist vs NSP). This will involve the use of linear mixed models adjusting for possible covariates including baseline depression symptoms, medication, treatment dosage, and child age when assessing child outcomes. Mothers will be taken as a random effect and the models will include a treatment- by- time interaction term; ii) the same set of analyses will be conducted for anxiety symptoms (GAD-7). All child outcomes, including child mental development and the provision of psychosocial stimulation by the mother, will be compared between two groups (antenatal vs. postnatal) at 6-24 months post-childbirth using a two-sample two-sided t-test. Sensitivity analyses will be carried out excluding those who dropped out after first session to see if the results are comparable to the entire study group. No interim analyses will be carried out.

3.5.3. Subgroup Analyses. Moderating effects will be conducted to determine whether delivery mode (telemedicine vs. in-person) or agent (specialist vs. non-specialist) have differential effects as a function of initial severity, whether mothers who receive antenatal treatment have lower depressive symptoms at 12-months post-enrollment than those who receive postnatal treatment; and whether children whose mothers receive the antenatal treatment have higher child development scores at 6-24 months post child birth than those children whose mothers receive postnatal treatment.

3.5.3.1. Clinical Severity. Tests of moderation will be conducted to determine whether delivery mode (telemedicine vs. in-person) or agent (specialist vs. non-specialist) influence mothers who are severely depressed (EPDS>19)[77]. We expect 10 to 15% of our sample to be severely depressed (EPDS>19). A 2-group comparison with 75 per group and an assumed mean EPDS score of 20.0 (SD=7.0 [77], will provide 80% power, with an alpha of 0.05, to detect a mean change of 2.3 or greater. Our current sample size including those within severity subgroups described above will be adequately powered to detect medium effect size differences of HTE of symptom severity between groups. This indicates that we have adequate power for clinically meaningful tests. The two groups' (specialist vs non-specialist) EPDS scores at 3-months will be compared using a two-sided two-sample t-test. Change in EPDS score over time (baseline and every 3 months, post-randomization) will also be compared using linear mixed models. We will also examine this on anxiety symptoms at 3-months post randomization.

3.5.3.2. Perinatal Period. This model will also test whether expectant mothers who receive antenatal treatment benefit more in terms of reduced depressive symptoms than mothers who receive postnatal treatment at 12-months post-randomization. A 2-group comparison with 200 per group and an assumed mean of EPDS score of 7.93 (SD=4.68) will provide 90% power, with an alpha of 0.05, to detect a mean change of 1.5 points. This allows us to detect a drop corresponding to a small effect size of 0.3 (i.e., a drop to mean 6.43 on the EPDS). The two groups' (antenatal vs. postnatal) EPDS scores at 3-months will be compared using a two-sided two-sample t-test. Change in EPDS scores over time (baseline and every 3 months, post-randomization) also will be compared using linear mixed models including the interaction between group and time. We will examine anxiety symptoms at 12-months post randomization in a linear mixed model controlling for relevant covariates including child age.

We will examine the hypothesis of whether the subset of mothers (up to 75% of the sample) who receive the antenatal treatment will benefit more in terms of improved child

outcomes at 6-24 months than mothers who receive postnatal treatment. A 2-group comparison of 238 per group (antenatal vs. postnatal) and an assumed mean on any Bayley IV raw subscale score of 100, SD=15 to assess child mental development will provide 80% power, with an alpha of 0.05 to detect a mean clinically-significant change of 3.0 units. Raw and standardized Bayley IV scores of the two groups' children (antenatal vs. postnatal) at 12 months will be compared using a two-sided, two-sample t-test.

3.5.3.3. Treatment Providers. Competency measures from the training sessions among both types of treatment providers (specialists and non-specialists) will be compared with participant outcomes to determine if competency scores during training predict participant outcomes. Key variables related to competency will include treatment-specific skills (range mean score of 0-4) and general skills (range mean score of 0-4) as measured by the Q-HAP and a total score (range 0-35) of a 35-item multiple choice exam (range total score of 0-35). All analyses will be conducted using SAS 9.4. Means, SDs and ranges of all competency measures for each individual NSP and across NSPs. This will be followed by estimating the relations between competency measures by calculating the Pearson correlation. Finally, multiple regression analyses will be used to estimate whether competency measures can predict patient outcomes scores of EPDS and GAD-7 scores post-treatment. Covariates including baseline EPDS scores and treatment provider will be utilized to account for potential baseline heterogeneity.

3.5.4. Qualitative Analysis. All qualitative data will be analyzed using NVivo™, a qualitative data analysis software package. We will use content analysis with data analysis (coding) conducted by multiple independent raters, for whom inter-rater reliability will be calculated using Kappa (κ) scores. Coding will be conducted in a step-wise fashion to facilitate iterative revision and then finalization of a coding scheme. Specifically, there will be a process of first independently coding and then discussing a minimum of 3 cases per stakeholder group to achieve a kappa (κ) score of κ=0.75 or higher (defined as substantial to almost perfect agreement). Qualitative data will then be quantified and triangulated across stakeholder groups using our previously established methods[21,78].

3.5.5 Economic Evaluation Analysis. Our economic evaluation will examine the cost-effectiveness of non-specialist vs. specialist provider and telemedicine vs. in-person sessions (in line with the clinical analyses) and will follow CADTH guidelines [79]. As such, we will adopt the health system perspective.

3.5.5.1. Cost-effectiveness analyses. The primary economic evaluation will be a cost-utility analysis (a cost-effectiveness analysis where outcomes are measured using a utility measure) carried out at the 12-month post-randomization follow-up point with outcomes expressed in terms of QALYs based on cases with non-missing data. Costs and outcomes will be compared at 12 months and will be presented as mean values by arm with standard deviations. Mean differences and 95% confidence intervals will be obtained by non-parametric bootstrap regressions to account for the non-normal distribution of the data[80]. A secondary analysis will be a cost-effectiveness analysis using the primary clinical outcome measure, the EPDS[41] at 12-months post-randomization. Cost-effectiveness will be assessed using the net benefit approach[81] with reference to Bosmans' methods for economic evaluations alongside

equivalence or noninferiority trials[82]. Incremental cost-effectiveness ratios (ICERs) will be calculated, defined as the mean difference in cost between two groups divided by the mean difference in effect. Uncertainty will be explored using cost-effectiveness planes and cost-effectiveness acceptability curves (CEACs) based on the net-benefit approach[83]. Cost-effectiveness planes illustrate the uncertainty around the estimates of costs and effects by plotting the bootstrapped cost and effects, with points in each quadrant indicating a different implication for economic evaluation[84]. CEACs are an alternative to confidence intervals around ICERs and show the probability that one intervention is cost-effective compared to another, for a range of values that a decision maker would be willing to pay for a unit improvement in the outcome of interest[83].

3.5.6.2. Sensitivity and subgroup analyses. To explore the potential impact of excluding non-responders, we will examine the sociodemographic and clinical characteristics of those included in the analyses and those in the full sample. Further, we will rerun the primary and secondary cost-effectiveness analyses with missing total costs and outcomes imputed using multiple imputation by chained equations. We will explore the possibility of conducting subgroup analyses related to heterogeneity due to clinical severity and perinatal period (antenatal or postnatal), in line with proposed clinical analyses. We will examine geographic location by postal or zip code to explore the impact of location on various measures.

3.5.6. Missing Data. Strategies to reduce missing data have been discussed above. In addition, study staff will make at least four attempts to contact patients before recording their data as missing. Multiple imputation methods will be used when rates of missing data for a variable exceed 20%. Five imputed datasets will be created and the model results averaged across the five iterations. SAS's Proc MI and Proc MIANALYZE will be used to carry out the procedures. Linear mixed models will be used to assess repeated measures outcomes. These models use maximum likelihood estimation methods that retain patients who do not have complete data across all time points. Reasons for dropout will be ascertained from clinician records or interviewing a subset of the participants who dropped out and coded. Sensitivity analyses will be carried out should missing data lead to the use of multiple imputation methods. These analyses will compare the results of the models on the imputed data to the ones with the actual missing data included. For the economic evaluation, participant health records will be accessed retrospectively to ensure that all health services accessed and costs incurred are captured. In Toronto, we will work with the Institute for Clinical Evaluative Sciences (ICES) to collect these data.

3.6. Data Retention

Study data will be kept for 10 years as per MSH REB recommendations and then destroyed following best practices.

4 MONITORING

4.1. Data and Safety Monitoring Board

The Data and Safety Monitoring Board (DSMB) will consist of four persons, including a psychiatrist with expertise in perinatal depressive or anxiety symptoms, a medical provider with expertise in providing obstetric care for pregnant women, a psychologist with expertise in the design and implementation of pragmatic clinical trials, and a PhD-level statistician. None of these persons will be involved directly in the study. The DSMB will have the following aims:

- To assure the safety, privacy, and confidentiality of human subjects.
- To assure the reliability, validity, completeness, and integrity of the data collection and management process.
- To review implementation of the protection of human subjects protocol, including amendments made in relation to safety concerns.
- To review all serious adverse events, rates of dropout or study withdrawal, and rates of missing data.

4.2. Reporting

The Trial Steering Management (TSM) Subcommittee will present regular reports to the Advisory Committee, and the Data and Safety Monitoring Board. These include:

- **Monthly Recruitment Reports** - reports of the number of women screened and enrolled by month and by clinical center are provided monthly to the TSM Subcommittee and all other members of the Steering Committee.
- **Quarterly Steering Committee Reports** - reports detailing recruitment, baseline patient characteristics, data quality, missing data and protocol adherence by clinical center, are provided quarterly to the TSM Subcommittee and all members of the Steering Committee.
- **Data and Safety Monitoring Board Reports** - a report will be prepared for every DSMB meeting that includes patient recruitment, baseline patient characteristics, center performance with respect to data quality, timeliness of data submission and protocol adherence (in addition to safety and efficacy data). The reports also include serious adverse events, loss to follow-up and all outcome variables as described previously in this protocol.

4.3. Serious Adverse Events

Within 72 hours of a serious adverse event (SAE), the Trial Management Committee will report the event to the DSMB, including the details of the event, the severity of any reactions, the phase of the study, and the procedures for its resolution, and will report to the IRB of record (*Appendix D – Adverse Events*). In addition, the Project Administrator will ensure that the assigned PCORI program officer is informed of any actions taken by the IRB of record as a result of such events. The DSMB meet twice per year for the study duration. Members will also meet via conference calls if SAEs attributable to study procedures are reported.

4.4. Safety Protocols

Safety protocols for both the recruitment research assistant and the treatment provider have been adapted from previous studies[54] (detailed in *Appendix A –Safety Protocols and Forms*). In brief, a multistep process will be followed:

- **Monitoring treatment quality.** BA delivery will be extensively monitored during the study using previously established methods developed and tested by the PI.
  - **Training Phase.** During the training phase, clinical leads will host weekly supervision meetings with **NSPs** and **specialist providers** to review and discuss cases and reinforce BA training. Independent consultants will evaluate treatment among 5% of all audio sessions to assess treatment adherence and fidelity. Treatment providers not reaching the mean cut-off for specific items will receive refresher training by the study clinicians. In addition, quality indicators of NSP and specialist-delivered sessions will be collected including the number, mode of delivery and duration of sessions.
  - **Trial Phase.** During the trial phase, **NSPs** will continue to attend weekly supervision meetings to review and discuss cases. Once per month, clinical leads and NSPs will rate individual audio-recorded sessions for therapy quality—the extent to which a psychological treatment was implemented well enough to achieve its expected effects[85]—of audio-recorded sessions in a monthly structured supervision. Audio recordings will be rated using the 20-item Q-HAP[52] for BA-treatment-specific and common (e.g., collaboration, empathy) skills. The individual NSP who conducted the session (self-rating), two to three peers (peer ratings) and a clinical lead (expert rating) will rate each session. The summed score for each subscale will be used to estimate therapy quality for each NSP and compared to expert ratings. **Specialists** will not attend regular supervision, but supervision will be available to them on an ad hoc basis to reflect real-world conditions.
  - Independent consultants will continue to evaluate treatment among 5% of all audio sessions, by both NSPs and specialists, to assess treatment adherence and fidelity. Treatment providers not reaching the mean cut-off for specific items will receive refresher training by the study clinicians.

- **Notification and Plan for Severe Session-Wise Scores.** Depression scores on the EPDS will be assessed at the beginning of each session. Any mother who has a positive response on the self-harm ideation item will undergo the standard safety protocol and in the case of active suicidal ideation, will be further assessed by a trial perinatal psychiatrist or associated Emergency Department for further assessment as is currently in practice at each of our three hubs. Women with worsening symptom scores on post-baseline EPDS session-by-session assessments during the intervention phase will be reviewed by the study team who will follow up on each case individually to provide recommendations about referral to additional psychiatric care if clinically indicated, including recommendations around a need for emergency assessment. 24-hr emergency psychiatric care is available at each hub. For concerns about maternal self-harm or infant/child harm that arise during the intervention,

the provider will follow an established protocol, used effectively in our previous trials[24,53,54], which will be supervised by the site trial psychiatrist (*Appendix A –Safety Protocols and Forms*).

- **Post-Intervention Protocols.** If a participant has an EPDS >12 at the post-treatment outcome assessment or at any subsequent assessment, the participant's most responsible provider (e.g. midwife, obstetrician, family physician or psychiatrist) will be informed by the research team and, if the participant is not already enrolled in care with a perinatal mental health team, information will be given for how to initiate referral to additional specialty psychiatric care with the participant's consent. For concerns about maternal self-harm or infant/child harm that arise during the follow-up period, we will follow an established protocol, used effectively in our previous trials[24,53,54], which will be supervised by the trial psychiatrist at each site (*Appendix A –Safety Protocols and Forms*).
- **Extensive Training.** All providers, including specialists will have completed an extensive training program focused on the development and evaluation of core competencies in BA and common counseling skills, followed by weekly supervision with a clinical lead.
- **Negative Intervention Effects** will be assessed for through regular evaluations on a monthly basis by our research team including Serious Adverse Events (SAE's).
- **Measuring Child Outcomes.** Trained research assistants will conduct post-child birth assessments (months 6-24) as per hospital protocols for conducting home visits.
- **Informing the Patient's Provider.** With participant consent (*see Participation in SUMMIT Health Care Provider Fax/Email Template, Appendix E – Study Documents,* pg. 101), we will inform the patient's regular provider with the results of screening (i.e., EPDS score and eligibility into SUMMIT) via electronic message (using the electronic health record when available) or via fax or mailed letter. We also will communicate with the patient's provider at the end of the study for each participant with a letter summarizing all follow-up and management strategies implemented and final symptomatology results. All provider information will be kept on site in an encrypted file on a secure server and linked to the participant by Study ID. We have completed these interdisciplinary communication strategies before with no difficulties in previous maternal depression trials[24,53,86,87].
- **Incidental Findings.** If any unexpected, clinically important information about a participant's or their child's health is suspected, the study team may contact them to make them aware of that information. The Site PI will review the information, consult with specialists if needed while protecting the participant's identity, and determine whether the participant should be contacted and informed.

5 ETHICS AND DISSEMINATION

5.1. Research Ethics Approval

In Toronto, we will seek ethical approval from Clinical Trials Ontario (CTO) Stream —a system that allows one CTO-Qualified Research Ethics Board (REB) to oversee the ethical approval and conduct of research occurring at multiple sites in Ontario, streamlining the REB application process (SHS, SMH, and WCH). In North Carolina, we will seek ethical approval from the UNC Biomedical Institution Review Board. In Chicago, we will seek ethical approval from the

Institutional Review Board of NorthShore University HealthSystems, which provides approval across all NorthShore-affiliated study sites. We will submit for approval after receiving input on the current protocol following our first Stakeholder Advisory Committee meeting in May 2019.

5.2. Protocol Amendments

Any modifications to the protocol that may impact study conduct, potentially benefit patients or affect their safety, including changes of study objectives, study design, patient population, sample sizes, study procedures, or significant administrative aspects will require a formal amendment to the protocol. Such amendment will be agreed upon by the TSM and approved by the REB/IRB (Institutional Review Board) prior to implementation and to the health authorities notified in accordance with local regulations. Amendments in relation to safety concerns will also be reviewed by the DSMB prior to implementation. Administrative changes are minor corrections or clarifications that have no effect on the way the study is to be conducted. These administrative changes will be agreed upon by TSM, and will be documented in a memorandum. The REB/IRB will be notified of administrative changes as necessary.

5.3. Confidentiality

To ensure confidentiality, data dispersed to relevant team members (including the PI and study biostatistician) will be blinded of any identifying participant information and treatment allocation. Only the site recruitment research assistants, research assistants responsible for scheduling appointments, and treatment providers will have access to the identifying participant information. All information will be kept in password-protected, encrypted files on encrypted computers and secure servers. Unique Study IDs will be used to identify the participants with Key Files kept at the site from which the participant was recruited. The Key File will be encrypted and stored on an encrypted computer or a secure server within each trial site institution.

5.4. Declaration of Interests

There are no conflicts of interest.

5.5. Data Management

The Data Manager will oversee the intra-study data sharing process, with input from the Trial Steering Committee. The PI and study biostatistician will be given access to the cleaned data sets. Quantitative data will be collected via REDCap™ at each hub separately, exported, encrypted and securely transferred to Toronto, where it will be housed in encrypted files on the local Storage Area Network (SAN) within St. Michael's Hospital, and will be backed up regularly and stored off-site. The data center is designed such that there are daily backups made of all critical data. In addition, the backups are stored both locally, as well as at a remote off-site location, in the case of catastrophic failure at one location. With limited access privileges, 24-hour security, and around-the-clock monitoring, the data center is highly secure. Qualitative data, as well as audio-recorded sessions, will be saved as de-identified transcribed interviews and stored in encrypted files on an encrypted computer within each hub. The site PIs will have direct access to their own site's data and may have access to other sites' data by request.

5.6. Testimonials

As a part of the dissemination phase of the trial, we will ask some participants, significant others, and treatment providers who have consented to be contacted to participate in qualitative interviews, offered to provide testimonies based on previous visits, and/or to be contacted for future research, if they would like to provide testimonials about their experiences with the SUMMIT Trial. These testimonials will be used for SUMMIT dissemination and knowledge mobilization purposes, and may include text quotes and/or or audio-visual materials such as photos, audio or voice. Those participants, significant others, and treatment providers who would like to provide testimonials will be provided with the opportunity to consent to, or not consent to, each specific type of testimonial (i.e., text, photos, audio, video). They will be made aware that publication of these testimonials may appear in media available to the public, for example in newsletters, brochures, websites and videos, and will also be informed of the potential risks involved (*see ICF Testimonial Addenda*, *Appendix E – Study Documents, pg. 41 – 56*).

6 REFERENCES

1. Ferrari AJ, Charlson FJ, Norman RE, et al. Burden of depressive disorders by country, sex, age, and year: findings from the global burden of disease study 2010. *PLoS Med*. 2013;10(11):e1001547.
2. O'hara MW, Swain AM. Rates and risk of postpartum depression—a meta-analysis. *International review of psychiatry*. 1996;8(1):37-54.
3. Wisner KL, Sit DKY, McShea MC, et al. Onset Timing, Thoughts of Self-harm, and Diagnoses in Postpartum Women With Screen-Positive Depression Findings. *JAMA psychiatry*. 2013;70(5):490-498. doi:10.1001/jamapsychiatry.2013.87
4. Robertson E, Grace S, Wallington T, Stewart DE. Antenatal risk factors for postpartum depression: a synthesis of recent literature. *General Hospital Psychiatry*. 7// 2004;26(4):289-295. doi:https://doi.org/10.1016/j.genhosppsych.2004.02.006
5. Bauer A, Knapp M, Parsonage M. Lifetime costs of perinatal anxiety and depression. *Journal of affective disorders*. 2016;192:83-90.
6. Dennis C-L, Falah-Hassani K, Shiri R. Prevalence of antenatal and postnatal anxiety: systematic review and meta-analysis. *The British Journal of Psychiatry*. 2017:bjp. bp. 116.187179.
7. Falah-Hassani K, Shiri R, Dennis C-L. Prevalence and risk factors for comorbid postpartum depressive symptomatology and anxiety. *Journal of affective disorders*. 2016;198:142-147.
8. Stein A, Pearson RM, Goodman SH, et al. Effects of perinatal mental disorders on the fetus and child. *The Lancet*. 2014;384(9956):1800-1819.
9. Field T. Prenatal depression effects on early development: a review. *Infant Behav Dev*. Feb 2011;34(1):1-14. doi:10.1016/j.infbeh.2010.09.008
10. Meltzer-Brody S, Howard LM, Bergink V, et al. Postpartum psychiatric disorders. *Nat Rev Dis Primers*. Apr 26 2018;4:18022. doi:10.1038/nrdp.2018.22
11. Sockol LE, Epperson CN, Barber JP. A meta-analysis of treatments for perinatal depression. *Clinical psychology review*. 2011;31(5):839-849.
12. O'hara MW, McCabe JE. Postpartum depression: current status and future directions. *Annual review of clinical psychology*. 2013;9:379-407.
13. Goodman JH. Women's attitudes, preferences, and perceived barriers to treatment for perinatal depression. *Birth*. 2009;36(1):60-69.
14. O'Mahen HA, Flynn HA. Preferences and perceived barriers to treatment for depression during the perinatal period. *Journal of women's health*. 2008;17(8):1301-1309.
15. Byatt N, Xiao RS, Dinh KH, Waring ME. Mental health care use in relation to depressive symptoms among pregnant women in the USA. *Archives of women's mental health*. 2016;19(1):187-191.
16. Dennis CL, Chung-Lee L. Postpartum depression help-seeking barriers and maternal treatment preferences: A qualitative systematic review. *Birth*. 2006;33(4):323-331.
17. O'Connor E, Rossom RC, Henninger M, Groom HC, Burda BU. Primary Care Screening for and Treatment of Depression in Pregnant and Postpartum Women: Evidence Report and

Systematic Review for the US Preventive Services Task Force. *Jama*. Jan 26 2016;315(4):388-406. doi:10.1001/jama.2015.18948
18. Curry SJ, Krist AH, Owens DK, et al. Interventions to Prevent Perinatal Depression: US Preventive Services Task Force Recommendation Statement. *Jama*. 2019;321(6):580-587.
19. Singla DR, Kohrt B, Murray LK, Anand A, Chorpita BF, Patel V. Psychological treatments for the world: lessons from low-and middle-income countries. *Annual Review of Clinical Psychology*. 2017;13(1)
20. Hoeft TJ, Fortney JC, Patel V, Unützer J. Task-Sharing Approaches to Improve Mental Health Care in Rural and Other Low-Resource Settings: A Systematic Review. *The Journal of rural health*. 2016;
21. Singla D, Lazarus A, Atif N, et al. "Someone like us": Delivering maternal mental health through peers in two South Asian contexts. *Journal of affective disorders*. 2014;168:452-458.
22. O'Mahen HA, Grieve H, Jones J, McGinley J, Woodford J, Wilkinson EL. Women's experiences of factors affecting treatment engagement and adherence in internet delivered Behavioural Activation for Postnatal Depression. *Internet Interventions*. 2015;2(1):84-90.
23. Patel V, Weobong B, Weiss HA, et al. The Healthy Activity Program (HAP), a lay counsellor-delivered brief psychological treatment for severe depression, in primary care in India: a randomised controlled trial. *The Lancet*. 2017;389(10065):176-185.
24. Dimidjian S, Goodman SH, Sherwood NE, et al. A pragmatic randomized clinical trial of behavioral activation for depressed pregnant women. *Journal of Consulting and Clinical Psychology*. 2017;85(1):26.
25. Dimidjian S, Hollon SD, Dobson KS, et al. Randomized trial of behavioral activation, cognitive therapy, and antidepressant medication in the acute treatment of adults with major depression. *Journal of consulting and clinical psychology*. 2006;74(4):658.
26. Dobson KS, Hollon SD, Dimidjian S, et al. Randomized trial of behavioral activation, cognitive therapy, and antidepressant medication in the prevention of relapse and recurrence in major depression. *Journal of consulting and clinical psychology*. 2008;76(3):468.
27. Lee EW, Denison FC, Hor K, Reynolds RM. Web-based interventions for prevention and treatment of perinatal mood disorders: a systematic review. *BMC pregnancy and childbirth*. 2016;16(1):38.
28. Andrews G. Utility of computerised cognitive–behavioural therapy for depression. *The British Journal of Psychiatry*. 2010;196(4):257-258.
29. Lokkerbol J, Adema D, Cuijpers P, et al. Improving the cost-effectiveness of a healthcare system for depressive disorders by implementing telemedicine: a health economic modeling study. *The American Journal of Geriatric Psychiatry*. 2014;22(3):253-262.
30. Andrews G, Cuijpers P, Craske MG, McEvoy P, Titov N. Computer therapy for the anxiety and depressive disorders is effective, acceptable and practical health care: a meta-analysis. *PloS one*. 2010;5(10):e13196.
31. Singla DR, Raviola G, Patel V. Scaling up psychological treatments for common mental disorders: a call to action. *World Psychiatry*. in press;
32. Dimidjian S, Barrera Jr M, Martell C, Muñoz RF, Lewinsohn PM. The origins and current status of behavioral activation treatments for depression. *Annual Review of Clinical Psychology*. 2011;7:1-38.

33. Jacobson NS, Martell CR, Dimidjian S. Behavioral activation treatment for depression: Returning to contextual roots. *Clinical Psychology: science and practice*. 2001;8(3):255-270.
34. Weobong B, Weiss HA, McDaid D, et al. Sustained effectiveness and cost-effectiveness of the Healthy Activity Program, a brief psychological treatment for depression delivered by lay counsellors in primary care: twelve-month follow-up of a randomised controlled trial. *PLoS Medicine*. 2017;17(9)
35. Cuijpers P, Van Straten A, Warmerdam L. Behavioral activation treatments of depression: A meta-analysis. *Clinical psychology review*. 2007;27(3):318-326.
36. Mazzucchelli T, Kane R, Rees C. Behavioral activation treatments for depression in adults: a meta-analysis and review. *Clinical Psychology: Science and Practice*. 2009;16(4):383-411.
37. Richards DA, Ekers D, McMillan D, et al. Behavioural Therapy for Depression (COBRA): a randomised, controlled, non-inferiority trial. 2016;
38. Dimidjian S, Goodman SH. Preferences and attitudes toward approaches to depression relapse/recurrence prevention among pregnant women. *Behaviour research and therapy*. 2014;54:7-11.
39. O'mahen H, Richards D, Woodford J, et al. Netmums: a phase II randomized controlled trial of a guided Internet behavioural activation treatment for postpartum depression. *Psychological medicine*. 2014;44(08):1675-1689.
40. Singla DR, Kumbakumba E, Aboud FE. Effects of a parenting intervention to address maternal psychological wellbeing and child development and growth in rural Uganda: a community-based, cluster-randomised trial. *The Lancet Global Health*. 2015;3(8):e458-e469.
41. Cox JL, Holden JM, Sagovsky R. Detection of postnatal depression. Development of the 10-item Edinburgh Postnatal Depression Scale. *The British journal of psychiatry*. 1987;150(6):782-786.
42. Spitzer RL, Kroenke K, Williams JB, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Archives of internal medicine*. 2006;166(10):1092-1097.
43. O'Mahen HA, Richards DA, Woodford J, et al. Netmums: a phase II randomized controlled trial of a guided Internet behavioural activation treatment for postpartum depression. *Psychol Med*. Jun 2014;44(8):1675-89. doi:10.1017/S0033291713002092
44. Posner K, Brent D, Lucas C, et al. Columbia-suicide severity rating scale (C-SSRS). *New York, NY: Columbia University Medical Center*. 2008;
45. Kim JJ, La Porte LM, Saleh MP, et al. Suicide risk among perinatal women who report thoughts of self-harm on depression screens. *Obstetrics & Gynecology*. 2015;125(4):885-893.
46. Sheehan DV, Lecrubier Y, Sheehan KH, et al. The Mini-International Neuropsychiatric Interview (MINI): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *The Journal of clinical psychiatry*. 1998;
47. Kimmel MC, Lara-Cinisomo S, Melvin K, Di Florio A, Brandon A, Meltzer-Brody S. Treatment of severe perinatal mood disorders on a specialized perinatal psychiatry inpatient unit. *Archives of women's mental health*. 2016;19(4):645-653.
48. Babor TF, Higgins-Biddle JC, Saunders JB, Monteiro MG. The alcohol use disorders identification test. *Guidelines for use in primary health care Geneva: World Health Organization*. 1992;

49. Maisto SA, Carey MP, Carey KB, Gordon CM, Gleason JR. Use of the AUDIT and the DAST-10 to identify alcohol and drug use disorders among adults with a severe and persistent mental illness. *Psychological assessment*. 2000;12(2):186.
50. Staley D, El-Guebaly N. Psychometric properties of the Drug Abuse Screening Test in a psychiatric patient population. *Addictive Behaviors*. 1990;15(3):257-264.
51. Smith PC, Schmidt SM, Allensworth-Davies D, Saitz R. A single-question screening test for drug use in primary care. *Archives of internal medicine*. 2010;170(13):1155-1160.
52. Singla DR, Weobong B, Nadkarni A, et al. Improving the scalability of psychological treatments in developing countries: an evaluation of peer-led therapy quality assessment in Goa, India. *Behaviour research and therapy*. 2014;60:53-59.
53. Dennis C-L, Hodnett E, Kenton L, et al. Effect of peer support on prevention of postnatal depression among high risk women: multisite randomised controlled trial. *Bmj*. 2009;338:a3064.
54. Dennis C-L, Ravitz P, Grigoriadis S, et al. The effect of telephone-based interpersonal psychotherapy for the treatment of postpartum depression: study protocol for a randomized controlled trial. *Trials*. 2012;13(1):38.
55. Spitzer RL, Kroenke K, Williams JB, Group PHQPCS. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. *Jama*. 1999;282(18):1737-1744.
56. Zimet GD, Dahlem NW, Zimet SG, Farley GK. The multidimensional scale of perceived social support. *Journal of personality assessment*. 1988;52(1):30-41.
57. Üstün TB, Chatterji S, Kostanjsek N, et al. Developing the World Health Organization disability assessment schedule 2.0. *Bulletin of the World Health Organization*. 2010;88:815-823.
58. Herdman M, Gudex C, Lloyd A, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Quality of Life Research*. 2011-12-01 2011;20(10):1727-1736. doi:10.1007/s11136-011-9903-x
59. Lang AJ, Stein MB. An abbreviated PTSD checklist for use as a screening instrument in primary care. *Behaviour research and therapy*. 2005;43(5):585-594.
60. Singla DR, MacKinnon DP, Fuhr DC, Sikander S, Rahman A, Patel V. Multiple mediation analysis of the peer-delivered Thinking Healthy Programme for perinatal depression: findings from two parallel, randomised controlled trials. *The British Journal of Psychiatry*. 2019:1-8.
61. Attkisson CC, Zwick R. The Client Satisfaction Questionnaire: Psychometric properties and correlations with service utilization and psychotherapy outcome. *Evaluation and program planning*. 1982;5(3):233-237.
62. Munder T, Wilmers F, Leonhart R, Linster HW, Barth J. Working Alliance Inventory-Short Revised (WAI-SR): psychometric properties in outpatients and inpatients. *Clin Psychol Psychother*. May-Jun 2010;17(3):231-9. doi:10.1002/cpp.658
63. Brown G. Health service utilization and cost of care questionnaire. *Health and Social Service Utilization Research Unit Hamilton, ON: McMaster University*. 2001;
64. Fuhr DC, Calvert C, Ronsmans C, et al. Contribution of suicide and injuries to pregnancy-related mortality in low-income and middle-income countries: a systematic review and meta-analysis. *Lancet Psychiatry*. Aug 2014;1(3):213-25. doi:10.1016/s2215-0366(14)70282-2
65. Bradley RH, Caldwell BM. The HOME Inventory and family demographics. *Developmental Psychology*. 1984;20(2):315.

66. Bayley N, Aylward G. *Bayley scales of infant and toddler development*. 4 edn ed. MN: NCS Pearson, Inc; 2019.
67. Aboud FE, Singla DR. Challenges to changing health behaviours in developing countries: a critical overview. *Social science & medicine*. 2012;75(4):589-594.
68. Singla DR, Hollon SD, Velleman R, et al. Temporal pathways of change in two randomized controlled trials for depression and harmful drinking in Goa, India. *Psychological Medicine*. 2019:1-9. doi:10.1017/S0033291718003963
69. Trevillion K, Ryan E, Pickles A, et al. An exploratory parallel-group randomised controlled trial of antenatal Guided Self-Help (plus usual care) versus usual care alone for pregnant women with depression: DAWN trial. *Journal of affective disorders*. 2020;261:187-197.
70. Drummond MF, Sculpher MJ, Claxton K, Stoddart GL, Torrance GW. *Methods for the economic evaluation of health care programmes*. Oxford university press; 2015.
71. Wodchis WP, Bushmeneva K, Nikitovic M, McKillop I. Guidelines on person-level costing using administrative databases in Ontario. 2013;
72. Pollack LM, Chen J, Cox S, et al. Healthcare Utilization and Costs Associated With Perinatal Depression Among Medicaid Enrollees. *American Journal of Preventive Medicine*. 2022/02/25/ 2022;doi:https://doi.org/10.1016/j.amepre.2021.12.008
73. Edwards P, Roberts I, Clarke M, et al. Methods to influence response to postal questionnaires. *The Cochrane Library*. 2001;
74. Seed M, Juarez M, Alnatour R. Improving recruitment and retention rates in preventive longitudinal research with adolescent mothers. *Journal of Child and Adolescent Psychiatric Nursing*. 2009;22(3):150-153.
75. Hough RL, Tarke H, Renker V, Shields P, Glatstein J. Recruitment and retention of homeless mentally ill participants in research. *Journal of consulting and clinical psychology*. 1996;64(5):881.
76. Katz KS, El-Mohandes A, Johnson DM, Jarrett M, Rose A, Cober M. Retention of low income mothers in a parenting intervention study. *Journal of Community Health*. 2001;26(3):203-218.
77. McCabe-Beane JE, Segre LS, Perkhounkova Y, Stuart S, O'Hara MW. The identification of severity ranges for the Edinburgh Postnatal Depression Scale. *Journal of Reproductive and Infant Psychology*. 2016;34(3):293-303.
78. Singla DR, Kumbakumba E. The development and implementation of a theory-informed, integrated mother-child intervention in rural Uganda. *Social Science & Medicine*. 2015;147:242-251.
79. Drugs CAf, Health Ti. Guidelines for the economic evaluation of health technologies: Canada. Ottawa: Canadian Agency for Drugs and Technologies in Health; 2006.
80. Barber JA, Thompson SG. Analysis of cost data in randomized trials: an application of the non-parametric bootstrap. *Statistics in medicine*. 2000;19(23):3219-3236.
81. Stinnett AA, Mullahy J. Net health benefits: a new framework for the analysis of uncertainty in cost-effectiveness analysis. *Medical decision making*. 1998;18(2_suppl):S68-S80.
82. Bosmans JE, De Bruijne MC, Van Hout HP, Hermens ML, Adèr HJ, Van Tulder MW. Practical guidelines for economic evaluations alongside equivalence trials. *Value in Health*. 2008;11(2):251-258.

83. Briggs AH. A Bayesian approach to stochastic cost-effectiveness analysis. *Health economics*. 1999;8(3):257-261.
84. Fenwick E, Byford S. A guide to cost-effectiveness acceptability curves. *The British Journal of Psychiatry*. 2005;187(2):106-108.
85. Fairburn CG, Cooper Z. Therapist competence, therapy quality, and therapist training. *Behaviour Research and Therapy*. 2011;49(6):373-378.
86. Sikander S, Lazarus A, Bangash O, et al. The effectiveness and cost-effectiveness of the peer-delivered Thinking Healthy Programme for perinatal depression in Pakistan and India: the SHARE study protocol for randomised controlled trials. *Trials*. 2015;16(1):534.
87. Fuhr DC, Weobong B, Lazarus A, et al. Delivering the Thinking Healthy Programme for perinatal depression through peers: an individually- randomised controlled trial in India. *The Lancet Psychiatry*. 2019;6(2):115-127.