MInD Study Statistical analysis plan

IRB approval date 7-9-2025

Participants

Recruitment occurred at seven integrated care sites identified through listservs of integrated mental health programs. Participants were consented and enrolled across two sites. Inclusion criteria included pregnant under 38 weeks gestation, receiving prenatal care and integrated mental health care at the study site, English speaking, age 18 years or older, PHQ-9 score ≥5, and ability to send and receive text messages. Exclusion criteria included self-reported multiple gestation, mental health treatment from psychiatric prescriber (but not a primary care provider or obstetrician), and severe mental illness such as bipolar disorder, schizophrenia-spectrum disorders, and severe substance use disorders.

Procedures

Research coordinators masked to randomization status collected outcome measures by online video conferencing. Participants were randomized into the two study arms using a 1:1 ratio blocked randomization with randomly selected block sizes. Research assessments were administered at baseline (enrollment), an additional prenatal timepoint, and at 6 weeks, 3 months, and 6 months postpartum. Dyads were videotaped by the participant with assistance from the research coordinator during virtual study visits to assess mother-infant dyadic interactions. Participants received a $35 gift card after completion of each assessment.

Measures

Primary outcome measures included the Patient Health Questionnaire-9 (PHQ-9), a 9-item measure assessing depression symptom severity. Total scores range from 0 to 27, with higher scores indicating greater depression symptom severity. Functioning was assessed with the 12-item World Health Organization Disability Assessment Scale (WHODAS 2.0). WHODAS 2.0 raw total scores range from 0-48 and were converted into a standardized score ranging from 0 to 100, with higher scores indicating greater disability. The Barkin Index of Maternal Functioning (BIMF)[27] is a 20-item measure of maternal functioning. BIMF total scores range from 0 to 120, with higher scores indicating higher functioning. The Self-Efficacy in a Nurturing Role (SENR prenatal and postnatal versions) is a 16-item measure of maternal self-efficacy with higher scores (range, 16-122) indicating higher self-efficacy in the parenting role.

Secondary outcome measures included the Generalized Anxiety Disorder-7 (GAD-7) to measure anxiety symptom severity (range, 0-21; higher scores indicate greater severity) and the Posttraumatic Stress Disorder (PSTD) Checklist for DSM-5 (PCL-5) to measure PTSD symptom severity (range, 0-80; higher scores indicate greater severity). We also included an assessment of parent sensitivity, evaluated using the Nursing Child Assessment Satellite Training Scale (NCAST) Parent-Child Interaction Teaching Scale (PCI-T). We used scores from the first 50 items comprising the caregiver subscale, with higher total scores (range, 0-50) indicating better dyadic interactions. The 17-item Hamilton Depression Rating Scale (HAM-D), administered by a trained research coordinator is a measure of depression symptom severity with higher scores indicating greater severity (range, 0-52), and was included for sensitivity analyses.

Statistical Analysis

Participants' demographic and clinical characteristics were descriptively analyzed. We also descriptively report on number of sessions completed with the care manager for all participants, and number of PFR-B sessions completed for participants in the intervention arm. Change scores were computed from baseline to 3 months and 6 months postpartum for the PHQ-9, WHODAS 2.0, SENR, GAD-7, and PCL-5. Change scores for the BIMF and NCAST PCI-T caregiver subscale were computed from 6 weeks postpartum to 3 and 6 months postpartum. Scores were prorated when data for at least 60% of the items were available. Due to small sample sizes ($n \leq 6$ per condition), Mann-Whitney U tests were conducted to compare PHQ-9 change scores between IC and IC+PFR-B. We conducted a sensitivity analysis using a Mann-Whitney U test comparing HAM-D change scores from baseline to 6 months postpartum between conditions. Between-subject effect sizes and 95% CIs are reported for the remaining measures and were calculated using the control (IC) group mean pre-post change minus the intervention (IC+PFR-B) group pre-post change divided by the pooled pre-test standard deviation. A correction factor was applied to adjust for small sample bias. Effect sizes should be interpreted with caution due to small sample sizes.

We explored the association between parenting self-efficacy and parenting sensitivity with depression symptom severity to inform a future mediational hypothesis. We used median cut points on SENR and NCAST caregiver subscale scores at 3 and 6 months postpartum within condition at each timepoint to categorize participants in low and high self-efficacy and sensitivity groups, respectively, and descriptively analyzed depression outcomes. Missing data were handled via listwise deletion, such that Mann-Whitney U tests were conducted among participants with data available at both timepoints of interest. Analyses were conducted in R Version 4.4.3. We also report on the feasibility of recruitment and success of randomization.