Woebot for Postpartum Mood and Anxiety: A Randomized Controlled Trial Evaluating Feasibility, Acceptability, and Preliminary Efficacy of a Mobile CBT Intervention

Toluwalase Ajayi, Jacqueline Kueper, Jill Waalen, Lauren Ariniello, Edward Ramos, Katie Baca-Motes, Andre Williams, Lauren Bullard,, Athena Robinson, Athena Robinson

Submitted to: Journal of Medical Internet Research
on: March 17, 2025

Table of Contents

Woebot for Postpartum Mood and Anxiety: A Randomized Controlled Trial Evaluating Feasibility, Acceptability, and Preliminary Efficacy of a Mobile CBT Intervention

Toluwalase Ajayi[1, 2] MD; Jacqueline Kueper[2] PhD; Jill Waalen[1] PhD; Lauren Ariniello[2] MBA; Edward Ramos[2, 3] PhD; Katie Baca-Motes[2, 3] MBA; Andre Williams[4] PhD; Lauren Bullard,[4] PhD; Athena Robinson[4]; Athena Robinson[4] PhD

[1] University of California, San Diego La Jolla US
[2] Scripps Research Institute La Jolla US
[3] 625 N Main Street Care Evolution Ann Arbor US
[4] 535 Mission Street Woebot Health San Francisco US

Corresponding Author:
Toluwalase Ajayi MD

University of California, San Diego
University of California San Diego 9300 Campus Point Drive
MC 7196
La Jolla
US

Abstract

Background: Postpartum psychological distress, ranging from transient mood and anxiety disturbances to full-syndrome postpartum depression (PPD), is prevalent. Many postpartum individuals lack access to evidence-based interventions due to stigma and insufficient provider availability. The treatment gap is particularly pronounced among historically marginalized groups, including Black, Hispanic/Latina, and low-income mothers, who face higher PPD prevalence and systemic barriers to care. Digital health interventions offer scalable, accessible, and culturally informed emotional support to address these disparities.

Objective: To evaluate the feasibility, acceptability, and preliminary efficacy of a smartphone application-delivered intervention for managing stress, anxiety, and mood in a diverse postpartum population.

Methods: This randomized controlled trial recruited participants from the PowerMom study, a digital platform for maternal health research. Eligible individuals (?16 years, <3 months postpartum) were randomized to Woebot for Postpartum Mood and Anxiety (W-PPMA) or a waitlist control condition. W-PPMA, an investigational digital mental health intervention, features a relational agent delivering cognitive behavioral therapy (CBT)-based psychoeducation via text-based conversations.
Primary outcomes included feasibility, acceptability, and satisfaction at 8-week end-of-intervention (EOI). The secondary outcome was change in self-reported depressive symptoms (PHQ-8) at 8-week EOI among participants with elevated baseline symptoms. Exploratory outcomes included anxiety (GAD-7), perinatal depression (EPDS), stress (PSS), mother-infant bond (MIB), and therapeutic alliance (WAI-SR), assessed at baseline, mid-treatment (4 weeks), EOI, and follow-ups at 12 and 16 weeks. The study followed CONSORT guidelines and received IRB approval.

Results: Participants (N=267; W-PPMA=144, Waitlist=123) represented diverse sociodemographic, mental health, and pregnancy backgrounds. W-PPMA users engaged with the app a median (Q1, Q3) of 9.0 (5.0, 23.8) days over 4.0 (2.0, 7.0) active weeks and reported high feasibility, acceptance, and satisfaction (URPI-F= 31 (28, 34), URPI-A= 30 (28, 34), CSQ-8= 26 (24, 29)).
The secondary outcome indicated a small but favorable effect of W-PPMA on depressive symptoms (Cohen’s d=-0.16). Exploratory analyses showed positive trends in perceived stress (PSS) and perinatal depression (EPDS) at EOI. Therapeutic alliance (WAI-SR) was highest among Black participants and those from socioeconomically disadvantaged neighborhoods (ADI ?75) at Baseline, and at EOI, among those with military healthcare insurance and socioeconomically disadvantaged neighborhoods (ADI ?75). Satisfaction (CSQ-8) was highest among those with a high school or GED education, highlighting

accessibility.

Conclusions: Among a diverse postpartum cohort, W-PPMA demonstrated feasibility, acceptability, and modest preliminary efficacy in reducing depressive symptoms. Exploratory findings suggest broader benefits for stress and mood management. High engagement and satisfaction highlight W-PPMA's potential as a scalable, accessible, and culturally informed digital mental health tool. These findings underscore its potential to bridge gaps in postpartum mental health care, particularly for marginalized populations. Further research is warranted. Clinical Trial: Clinicaltrials.gov NCT05662605

(JMIR Preprints 17/03/2025:74083)
DOI: https://doi.org/10.2196/preprints.74083

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

✓ **Please make my preprint PDF available to anyone at any time (recommended).**
Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.
Only make the preprint title and abstract visible.
No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**
Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain v
Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <a href="http
No. Please do not make my accepted manuscript PDF available to anyone.

Original Manuscript

Woebot for Postpartum Mood and Anxiety: A Randomized Controlled Trial Evaluating Feasibility, Acceptability, and Preliminary Efficacy of a Mobile CBT Intervention

Toluwalase Ajayi [1,2], Jacqueline K. Kueper[1], Jill Waalen[1,2], Lauren Ariniello[1], Edward Ramos[1,4], Katie Baca Motes[1,4], Andre Williams[3], [,3], Lauren Bullard[,3], Athena Robinson[,3]

[1] Scripps Research Translational Institute, 3344 N Torrey Pines Ct Plaza Level, La Jolla, California, 92037 USA

[2] University of California San Diego, La Jolla, CA, 92037 USA

[3] Woebot Health, 535 Mission Street, 14th Floor

San Francisco CA 94105

[4]CareEvolution, Ann Arbor, MI 48104

Correspondence to:

Tolúwalàṣẹ Àjàyí, Scripps Research Translational Institute, 3344 N Torrey Pines Ct Plaza Level, La Jolla CA 92037 USA; tajayi@health.ucsd.edu

Funding

This work received partial support from National Center for Advancing Translational Sciences (UM1TR004407) and from The Patrick J. McGovern Foundation

Data Availability

The data sets generated and analyzed during the PowerMom study are not publicly available due to privacy and confidentiality restrictions. The study collected sensitive health and demographic data from participants, which are de-identified but remain protected under institutional review board (IRB) protocols. Researchers who wish to access the data for legitimate academic purposes can submit a formal request to PowerMom PI, Toluwalase Ajayi. Requests will be evaluated on a case-by-case basis to ensure compliance with ethical guidelines and data use agreements. Code used for data analysis is available upon request to the corresponding author at tajayi@health.ucsd.edu.

Abstract

Background

Postpartum psychological distress, ranging from transient mood and anxiety disturbances to full-syndrome postpartum depression (PPD), is prevalent. Many postpartum individuals lack access to evidence-based interventions due to stigma and insufficient provider availability. The treatment gap is particularly pronounced among historically marginalized groups, including Black, Hispanic/Latina, and low-income mothers, who face higher PPD prevalence and systemic barriers to care. Digital health interventions offer scalable, accessible, and culturally informed emotional support to address these disparities.

Objective

To evaluate the feasibility, acceptability, and preliminary efficacy of a smartphone application-delivered intervention for managing stress, anxiety, and mood in a diverse postpartum population.

Methods

This randomized controlled trial recruited participants from the PowerMom study, a digital platform for maternal health research. Eligible individuals (≥16 years, <3 months postpartum) were randomized to Woebot for Postpartum Mood and Anxiety (W-PPMA) or a waitlist control condition. W-PPMA, an investigational digital mental health intervention, features a relational agent delivering cognitive behavioral therapy (CBT)-based psychoeducation via text-based conversations.

Primary outcomes included feasibility, acceptability, and satisfaction at 8-week end-of-intervention (EOI). The secondary outcome was change in self-reported depressive symptoms (PHQ-8) at 8-week EOI among participants with elevated baseline symptoms. Exploratory outcomes included anxiety (GAD-7), perinatal depression (EPDS), stress (PSS), mother-infant bond (MIB), and therapeutic alliance (WAI-SR), assessed at baseline, mid-treatment (4 weeks), EOI, and follow-ups at 12 and 16 weeks. The study followed CONSORT guidelines and received IRB approval.

Results

Participants (N=267; W-PPMA=144, Waitlist=123) represented diverse sociodemographic, mental health, and pregnancy backgrounds. W-PPMA users engaged with the app a median (Q1, Q3) of 9.0 (5.0, 23.8) days over 4.0 (2.0, 7.0) active weeks and reported high feasibility, acceptance, and

satisfaction (URPI-F= 31 (28, 34), URPI-A= 30 (28, 34), CSQ-8= 26 (24, 29)).

The secondary outcome indicated a small but favorable effect of W-PPMA on depressive symptoms (Cohen's d=-0.16). Exploratory analyses showed positive trends in perceived stress (PSS) and perinatal depression (EPDS) at EOI. Therapeutic alliance (WAI-SR) was highest among Black participants and those from socioeconomically disadvantaged neighborhoods (ADI ≥75) at Baseline, and at EOI, among those with military healthcare insurance and socioeconomically disadvantaged neighborhoods (ADI ≥75). Satisfaction (CSQ-8) was highest among those with a high school or GED education, highlighting accessibility.

Conclusions

Among a diverse postpartum cohort, W-PPMA demonstrated feasibility, acceptability, and modest preliminary efficacy in reducing depressive symptoms. Exploratory findings suggest broader benefits for stress and mood management. High engagement and satisfaction highlight W-PPMA's potential as a scalable, accessible, and culturally informed digital mental health tool. These findings underscore its potential to bridge gaps in postpartum mental health care, particularly for marginalized populations. Further research is warranted.

Trial Registration

Clinicaltrials.gov NCT05662605

Keywords

postpartum mental health; postpartum depression; health equity; digital health; mobile apps; relational agent; chatbot; cognitive behavioral therapy

Main Manuscript Word Limit-3000 and 6000 words

Introduction

Prevalence and Burden of Postpartum Blues and Postpartum Depression (PPD)

The postpartum period is a time of significant lifestyle, relational, and physical health changes, often accompanied by psychological distress and profound impacts on maternal mood. Understandably, many mothers experience various mood and anxiety symptoms and labile emotions after birth, such as mood swings, crying spells, anxiety, irritability, sadness, and feelings of being overwhelmed [1-3]. These symptoms typically resolve within two weeks without intervention. However, recognition and management of persistent postpartum mood and anxiety symptoms is critical, as they are not only distressing to mothers and their families but also represent a risk factor for the development of full-syndrome postpartum depression (PPD) and anxiety disorders [3, 4].

As many as 75% of women report some degree of mood and or anxiety symptoms after birth [5]. Full-syndrome PPD affects approximately 13-20% of postpartum individuals in the United States, with prevalence varying significantly by state and demographic group [6, 7]. PPD is associated with adverse outcomes, including increased maternal morbidity and mortality [8], impaired maternal-child interactions [9], diminished parenting behaviors, and cognitive, emotional, and behavioral challenges in child development [10].

Current Treatment

Treatment of peripartum mood and anxiety symptoms typically involves psychotherapy and/or antidepressant therapy. Cognitive behavioral therapy (CBT) and interpersonal psychotherapy (IPT) are recommended as first-line treatments for mild to moderate postpartum mood and anxiety concerns, while antidepressant pharmacotherapy is reserved for more severe cases of PPD [11]. CBT aims to address the reciprocal interactions between thoughts, feelings, and behaviors by challenging negative thought patterns and promoting adaptive behaviors [12]. IPT focuses on alleviating symptoms by improving interpersonal functioning, communication, and role clarification in response to relational challenges [13]. Techniques derived from these psychotherapies are beneficial not only for mothers diagnosed with PPD but may also provide effective mental health and emotional support for non-clinical populations.

Barriers to Mental Health Care Support

Despite the existence of evidence-based treatments for peripartum mood and anxiety, many postpartum individuals do not seek or engage in care, even when recommended by healthcare

providers [14, 15]. Stigma surrounding mental health is a significant barrier to help-seeking and treatment utilization [16]. Additional barriers include a shortage of qualified therapists, inflexible scheduling, long wait times, childcare challenges, transportation difficulties, and affordability concerns [14]. For pharmacotherapy, concerns about antidepressant transmission through breast milk further limit its use [17]. These barriers are especially pronounced among Black, Hispanic/Latina, and low-income mothers, who face higher risks of PPD and significant systemic inequities in accessing care [18].

Low representation of historically marginalized groups in health research exacerbates disparities, reducing the generalizability of findings and hindering the development of effective interventions [19, 20]. Ensuring diverse inclusion in research is critical to addressing health inequities and improving outcomes.

Digital Interventions

Digital mental health interventions (DMHI) offer promise for postpartum mental health support. DMHI, including those delivered via smartphone applications, can offer techniques derived from evidence-based interventions, augment traditional care, and reduce barriers to access such as time, location, and provider availability [21]. Such DMHI provide private, stigma-free support, which is particularly relevant for postpartum mothers navigating the unique demands of early parenthood [22]. Moreover, their scalability enables broader population reach, including individuals who might otherwise receive no care.

To address these challenges, this study leveraged the PowerMom mobile research platform [23], chosen for its ability to facilitate decentralized, bilingual, and scalable data collection for maternal health. PowerMom's capacity to enable remote participation and support for both English and Spanish speakers enhance inclusivity and generalizability, making it particularly valuable for engaging underrepresented populations in research.

The intervention evaluated in this study is *Woebot for Postpartum Mood and Anxiety (W-PPMA)*, an investigational DMHI that offers guided self-help program that delivers cognitive behavioral therapy (CBT) psychoeducation and self-management techniques through brief text-based conversations with a relational agent named Woebot. Woebot is intended to present to users a friendly, helpful, self-help ally that is explicitly not a human or a therapist. Neither W-PPMA nor Woebot have been evaluated, cleared, or approved by the FDA and are not available for general use. W-PPMA integrates evidence-supported techniques from CBT and IPT, specifically tailored to the postpartum experience. Earlier

prototypes of Woebot have demonstrated feasibility, acceptability, and preliminary efficacy in improving anxiety and depression symptoms among young adults [24], postpartum women [25, 26], and adults addressing substance use [27, 28]. Building on these foundations, W-PPMA aims to offer a specialized approach to address the unique mental health needs of postpartum mothers.

This study has several aims. The primary aim is to evaluate the feasibility and acceptability of W-PPMA among a diverse group of postpartum women. The secondary aim is to assess the preliminary efficacy of W-PPMA relative to a waitlist control condition in managing symptoms of depression at 8-weeks end-of-intervention (EOI) among participants with elevated baseline depressive symptoms. Exploratory aims include: (1) examining potential between-group differences in perceived mother-to-infant bonding, mood and anxiety symptoms, and perceived stress; (2) evaluating therapeutic bond with Woebot at 1- and 8-weeks; (3) assessing feasibility, acceptability, perceptions of stigma as a barrier to mental health care, and mood outcomes across key sociodemographic variables (e.g., race and ethnicity, age, education, gender, marital status, health insurance, and employment status) at baseline and 8-weeks.

Despite the significant prevalence and burden of postpartum psychological distress, existing interventions remain underutilized, particularly among historically marginalized groups who face systemic barriers to care. While evidence-based psychotherapies such as CBT and IPT are effective in treating postpartum mood disorders, there is a critical need to explore how these interventions can be feasibly and equitably delivered to diverse populations in non-clinical settings. Digital health tools, such as mobile applications, hold immense promise for addressing these gaps by overcoming logistical and stigma-related barriers. However, few digital interventions have been rigorously studied in postpartum populations. By evaluating the feasibility, acceptability, and preliminary efficacy of W-PPMA in a diverse and bilingual cohort, this study aims to advance understanding of scalable, accessible mental health solutions. Ultimately, this research provides critical insights into closing the postpartum mental health treatment gap with an emphasis on equity and inclusion.

Methods

Ethics

This study was approved by the Institutional Review Board (IRB) of the Scripps Research Translational Institute.

Study Design

This randomized waitlist-controlled trial evaluated the *Woebot for Postpartum Mood and Anxiety (W-PPMA)* digital health tool. The trial consisted of three phases: Screening/Baseline, Treatment, and Follow-up. Participants were assessed at Screening/Baseline, Day 3, Week 4, Week 8 (End of Intervention [EOI]), Week 12, and Week 16 (End of Study [EOS]). All primary endpoints were evaluated at Week 8 (EOI).

Recruitment Strategy

Participants were recruited through the PowerMom baseline research study (IRB-21-7738), a mobile app-based research initiative hosted on the HIPAA-compliant MyDataHelps™ clinical research platform. PowerMom enables expectant mothers to access prenatal health resources while contributing pregnancy-related data to research. Recruitment was supported by a bilingual, multi-channel campaign (e.g., email, social media, in-app nudges) involving PowerMom Consortium [29] partners from health and technology organizations, maternal advocacy groups, and community health centers. Key partners included Microsoft, Google/Fitbit, Philips, March of Dimes, and others, including community-based organizations serving Black and Hispanic/Latina populations. Recruitment campaigns outside of consortium partners also included Ovia Health.

English-speaking PowerMom participants received information about the Woebot sub-study through email and in-app notifications. Interested participants were directed to the eligibility screening, which included a disclaimer about the sensitive nature of some questions. Eligible participants provided informed consent for the sub-study, and participation in the Woebot sub-study did not affect their involvement in the PowerMom baseline study.

The recruitment period for the Woebot sub-study was from February 2023 and ended January 2024. initial goal was to recruit 225 participants in each arm of the study. To promote inclusivity, recruitment aimed to enroll at least 30% of participants from underrepresented racial and ethnic groups (URGs) and 50% from groups underrepresented in biomedical research (UBRs), as defined by the NIH *All of Us* Research Program [30]. UBR criteria included racial/ethnic minorities, low SES, rural residency, and other categories. However, the sponsor nominated to end enrollment early (after N=296 consented) given that the a-priori goal of sample diversity (50% UBR) was successfully achieved early and also to advance time to results given the importance and relevance of the hypotheses tested herein.

Inclusion and Exclusion Criteria

Eligible participants were PowerMom enrollees aged 16 years or older, less than three months postpartum, English-literate, smartphone users, and willing to engage with the program and assessments for 16 weeks.

Exclusion criteria included a lifetime diagnosis of psychotic or bipolar disorders, suicidal ideation or attempts within the past 12 months, substance abuse within the past 12 months, pregnancy loss in the past 18 months, or prior use of Woebot programs.

Procedures

Participation spanned 16 weeks, including an 8-week intervention or waitlist control period, followed by exploratory assessments at 12- and 16-weeks. After completing baseline assessments, participants were randomized 1:1 into the W-PPMA group or the waitlist control group. Neither participants nor staff were blinded to group assignments.

Participants randomized to W-PPMA accessed the app from randomization through 16-week (EOS) and were encouraged to use it daily, with a recommended minimum of five minutes per day during the treatment phase. Waitlist participants accessed W-PPMA starting at 8-weeks (EOI). Assessments were conducted via the PowerMom app at Baseline, Day 3, 4-week, 8-week (EOI), 12-week, and 16-week (EOS) timepoints. Each assessment took 10–15 minutes, and participants received reminders via email and PowerMom-app push notifications.

Participants were eligible to receive up to $175 in Amazon gift cards as compensation for completing study assessments. However, during the enrollment period, fraudulent activity was detected. The platform's automated system was exploited by 580 fraudulent accounts created to illegitimately claim gift cards. This led to an unusual and rapid spike in enrollment across both the baseline study and the Woebot sub-study. To ensure data integrity, fraudulent enrollments were excluded from randomization, and participants who failed to complete baseline assessments were classified as screen failures and not included in the study.

Intervention Groups

Woebot for Postpartum Mood and Anxiety (W-PPMA)

W-PPMA is an investigational DMHI that offers an 8-week guided self-help program that combines CBT- and IPT-based techniques tailored to provide emotional support to postpartum women. Through a text-based interface, a relational agent named Woebot engages users in conversations and provides users with self-guided psychoeducation, mood and anxiety management strategies. The program includes daily push notifications to re-engage users and encourage consistent participation. Neither W-PPMA nor Woebot have been evaluated, cleared, or approved by the FDA and are not available for general use.

Waitlist Control

Waitlist control participants had access to W-PPMA from the 8-week (EOI) through to 16-week (EOS) period with similar usage recommendations.

Safety Considerations

W-PPMA followed safety recommendations from the American Psychiatric Association [31] and American Medical Association [32]. W-PPMA has a Language Detection Protocol (LDP) that is based on a natural language processing algorithm designed to detect if a user inputs, via free-response text entry, concerning phrases or words that match an a priori identified and thorough list that may be used in crisis situations. Upon detection of any concerning language, the LDP reminds the participant of the application's limitations of services – which are stated in the study's informed consent, the app's privacy policy and terms of service, as well as the intervention's onboarding screens – and offer a resource list of readily accessible support channels, including emergency contact phone numbers, and suicide ideation and domestic violence hotline information.

Additionally, in the event that a participant endorses having current suicidal ideation or a plan when answering the screening questions, an automatic email would be sent to the study team. Study personnel would then call 911, notify the operator that it is a psychiatric emergency, and provide a trained emergency response person with the participant's phone number to follow up with them. All participants were informed of these procedures in the study consent form, which explicitly outlined the protocol for responding to safety concerns, including emergency escalation and study personnel involvement in crisis situations.

Assessments

Participant Demographics, Pregnancy History and Psychiatric History

Baseline assessments collected demographic data, including age, race, ethnicity, gender identity, marital status, education, employment, and insurance status. Pregnancy history included the number of prior pregnancies, outcomes, and any fetal demise within the exclusionary period. Psychiatric history assessed diagnosis of psychotic disorders, mood disorders and suicidality within the past 12 months, and substance use during the same period. Current or prior use of psychiatric medications and therapy was also documented.

Primary Outcomes: Feasibility and Acceptability Measures

Usage Rating Profile–Intervention Revised (URP-IR): [33] : The 6-item Feasibility and 6-item Acceptability subscales were used to measure feasibility (e.g., “The total time required to do the treatment procedures was manageable”) and acceptability (e.g., “I liked the procedures used in this treatment”). Responses ranged from 1 (“strongly disagree”) to 6 (“strongly agree”), with total scores ranging from 6 to 36 with higher average scores indicating greater feasibility or acceptability. The URP-IR demonstrates strong reliability, internal consistency, and discriminant validity [33].

Client Satisfaction Questionnaire (CSQ-8): [34]: This widely used 8-item measure evaluated participants' satisfaction with W-PPMA on a 4-point scale (e.g., “How would you rate the quality of service you received?”). Total scores ranged from 8 to 32, with higher scores indicating greater satisfaction.

App Engagement: Engagement with the W-PPMA app was assessed using quantitative metrics, including the number of active days, messages sent per week, and the number of content modules completed.

Secondary Outcome: Symptoms of Depressions

Patient Health Questionnaire (PHQ-8): The PHQ-8 is an 8-item measure evaluating depressive symptom severity over the past two weeks [35]. Scores range from 0 to 24, with cut points of 5, 10, 15, and 20 corresponding to mild, moderate, moderately severe, and severe depression, respectively [36, 37]. The PHQ-8 is a reliable and valid tool for screening depressive symptoms [38].

Exploratory Outcomes

Edinburgh Prenatal Depression Scale (EPDS): This 10-item self-report tool screens for perinatal depression [39]. Scores range from 0 to 30, with higher scores indicating greater depressive symptom severity. The EPDS omits the suicidality item for this study and is widely validated for postpartum populations [40].

Generalized Anxiety Disorder (GAD-7): The GAD-7 is a 7-item measure assessing anxiety symptoms over the past two weeks [41]. Scores range from 0 to 21, with cut points of 5, 10, and 15 indicating mild, moderate, and severe anxiety, respectively. This measure has demonstrated high reliability and validity [42].

Perceived Stress Scale (PSS-10): This 10-item scale measures perceived stress in the past month [43]. Responses range from 0 (“never”) to 4 (“very often”) and scores range between 0 and 40, with higher scores indicating greater perceived stress. The PSS-10 demonstrates strong reliability and validity [44].

Mother Infant Bonding Scale (MIBS) The MIBS includes eight one-word emotional descriptors (e.g., "loving," "resentful") rated on a 4-point scale from 0 (“very much”) to 3 (“not at all”). Total scores range from 0 to 24, with lower scores indicating stronger mother-infant bonding [45, 46].

Working Alliance Inventory-Short Revised (WAI-SR): The 12-item WAI-SR measures therapeutic alliance using three subscales: task agreement, goal agreement, and affective bond [47]. Items were adapted to refer to Woebot as the therapist. Responses range from 1 (“never”) to 5 (“always”), with higher scores indicating a stronger therapeutic alliance.

*Perceived Barriers to Psychological Treatment (PBPT)**:*** The PBPT [48] identifies potential barriers to engaging in mental health treatment. This study used the 7-item Stigma subscale (e.g., “My concern about being judged makes it ____ for me to attend counseling”), with responses ranging from 1 (“not difficult at all”) to 5 (“impossible”).

Figure 1 presents the schedule of assessment by study group.

Statistical Analysis

Primary analyses were conducted using an *Intention To Treat (ITT)* strategy whereby participants were analyzed in the treatment group that they were assigned to. *Per Protocol (PP) sensitivity analyses* were performed for the CSQ-8, URPI-A, URPI-F, PHQ-8, GAD-7, EPDS, PSS, MIBS, and

WAI-SR outcome measures. These sensitivity analyses excluded participants who deviated from the protocol: relevant measure(s) completed in full within the preplanned timing window, did not withdraw from the study, and if in the W-PPMA group, opened the app at least once per week for at least half of the intervention weeks. The latter criteria was based on previous research into app usage patterns.

Sample Characteristics including participant self-reported sociodemographic, psychiatric and pregnancy history, and prior and concomitant medications and therapy characteristics were described overall and by ITT study group.

The Primary Outcomes of feasibility and acceptability were described for the W-PPMA study group at the 8-week timepoint (EOI).

The Secondary Outcome to assess the preliminary efficacy of W-PPMA relative to a waitlist control condition to manage symptoms of depression was analyzed among the subgroup of participants with PHQ-8 greater than or equal to 5 at baseline (minimal to several symptoms of depression). Mean eight-week change scores were calculated and compared using a two-sided independent t-test. Cohen's d effect size was also calculated. Preplanned sensitivity analyses for this outcome included 1) per protocol sensitivity analysis as described above, 2) comparing 8-week scores instead of change scores, 3) multiple linear regression to adjust for potential residual confounding by baseline covariates (age, race, psychotherapy at baseline, mental health medication at baseline, previous postpartum depression diagnosis, and previous depression diagnosis), and 4) describing scores by clinical categories. Bonferroni adjustment of P-values was performed to account for the four statistical tests (primary and first three sensitivity) run on the PHQ-8 outcome.

The Exploratory Outcomes for preliminary assessment of mood and efficacy (PSS, EPDS, GAD-7, MIBS) were summarized at each time point with Cohen's d to assess between group differences. Change scores from baseline to 8-weeks were additionally compared and summarized using the mean difference with 95% confidence interval. Exploratory outcomes of therapeutic bond were summarized at baseline and 8-weeks.

Exploratory stratified analyses were performed to describe outcomes and the measure of perceptions of stigma by strata of age, Hispanic/Latina, racial background, marital status, sexual orientation, employment status, education, health insurance, and neighborhood atlas. To maintain high enough sample size for meaningful grouping, within each stratification some response options were grouped together or removed. While grouping removes some important between-group differences, we opted

to do this to be able to share some information rather than excluding minority groups altogether. Due to variability in circumstances related to being a student, we do not report student strata for Employment and Education analyses.

Missing Data. For the primary between-group comparison of treatment effect size (secondary outcome PHQ-8), we used multiple imputation to complete any missing PHQ-8 scores to minimize potential bias from missing data. Multiple imputation by chained equations was performed using the miceforest library [49] with the default random forests predictive model. Covariates were the same as for the multiple regression analysis, with missing data in these additionally imputed as needed: study group, age, race, psychotherapy at baseline, mental health medication at baseline, previous postpartum depression diagnosis, and previous depression diagnosis. For single group and exploratory outcome analyses, where the analytical goals were more descriptive or preliminary, participants that did not complete a measure in full at any given time point were excluded from pertinent analyses. We report the number of participants included at each time point for each measure.

For descriptive analyses, measures were summarized by mean and standard deviation or by median and interquartile range, whether normally distributed or not, respectively. Analyses were completed in Python v3.8.

Results

Study Sample

Of 296 participants that consented, n=15 (5.1%) were screen fails, and n=14 (4.7%; W-PPMA: n=12, Waitlist: n=2) withdrew from the study, leaving 267 participants (W-PPMA: n=144, Waitlist: n=123) for analyses. Table 1 shows participant self-reported sociodemographic characteristics overall and by treatment group. Overall, the mean (standard deviation (SD)) participant age was 32.2 (5.1) years (range: 16.3, 48.7), all participants identified as female gender, n=34 (12.7%) as Hispanic/Latina, and n=183 (68.5%) as White racial background. After randomization, the W-PPMA and Waitlist groups were similar on most characteristics, with notable exceptions that a greater proportion of the W-PPMA group was Hispanic/Latina (16.0% vs 8.9%) and identified with multiple races (16.7% vs 8.9%) whereas the Waitlist group had a higher proportion of White participants (73.2% vs 63.6%), participants with a College degree (37.4% vs 27.1%), and participants with ADI_NATRANK of at least 75 (11.4% vs 4.9%).

Table 2 shows psychiatric and pregnancy history. Overall n=31 (11.6%) participants had a previous postpartum depression diagnosis and n=113 (42.3%) a previous depression diagnosis. At study baseline, n=41 (15.4%) of participants reported engaging in psychotherapy and n=67 (25.1%) taking a mental health medication. The number of past pregnancies ranged from 0 to over 5, and regarding participants' most recent pregnancy, n=38 (14.2%) had their child spend time in the NICU and n=22 (8.2%) reported receiving a medical and/or genetic disorder diagnosis.

Primary Outcomes: Feasibility and Acceptability (URPI, CSQ)

Figure 2 shows app engagement metrics. Across the 8-week intervention, the median (quartile 1 (Q1), quartile 3 (Q3)) number of days that participants used W-PPMA was 9.0 (5.0, 23.8) and 4.0 (2.0, 7.0) active weeks. Participants sent a median 217 (89, 688) in-app messages and completed 6.0 (2.3, 18.8) stories and 1.5 (1.0, 3.0) tools.

Of the ITT W-PPMA group, n=122 (84.7%) completed the 8-week (EOI) measures for acceptability, feasibility, and satisfaction. The median (Q1, Q3) score for URPI-feasibility was 31.0 (28.0, 34.0), for URPI - Acceptability was 30.0 (28.0, 34.0), and for CSQ-8 was 26.0 (24.0, 29.0). Score distributions are presented in Supplementary Figure 1. Slightly higher scores were seen in the per protocol sensitivity analysis that included 76 (52.8%) of the W-PPMA group, with median (Q1, Q3) scores of 32 (29.0, 36.0) for URPI-feasibility, 31.5 (28.0, 35.2) for URPI-acceptability, and 27 (24.0, 29.2) for CSQ-8.

Secondary Outcome: Symptoms of Depression (PHQ-8)

There were 97 participants (n=48 W-PPMA, n=49 Waitlist) with a baseline PHQ-8 score of at least 5, the minimal total score of the PHQ-8 indicating the presence of at least mild symptoms of depression. Multiple imputation was used to impute missing 8-week PHQ-8 scores for n=14 participants (n=5 W-PPMA, n=9 Waitlist). Overall mean (SD) change in PHQ-8 from baseline to EOI was -0.3 (4.2) whereby W-PPMA averaged slight improvement (-0.7 (3.4)) and the Waitlist group averaged no change (0.0 (4.8)). The between-group effect size was small (Cohen's d: -0.16) and not statistically significant after adjustment for multiple testing (P-value=1). The per protocol sensitivity analysis showed a greater separation of groups (W-PPMA: -1.4 (3.5), Waitlist: 0.7 (3.9),

adjusted P-value=0.13) with medium effect size (Cohen's d: -0.57). Figure 3 shows the distribution of change scores for the ITT and per protocol subgroups.

Additional planned sensitivity analyses found similar trends. Comparison of the 8-week EOI scores alone showed lower average symptoms in the W-PPMA (mean (SD): 6.9 (3.4)) than Waitlist (8.4 (5.1)) group (Cohen's d -0.48, adjusted P-value=0.38). Multiple regression analyses to adjust for potential residual confounding by baseline characteristics found access to W-PPMA was associated with a one-point reduction in PHQ-8 change score as compared to the Waitlist control group (95% CI: -2.69, 0.69, adjusted P-value=0.96). All regression coefficients are presented in Supplementary Table 1. Examining PHQ-8 scores by clinical category found similar distributions at baseline and a trend towards lower severity among the W-PPMA than Waitlist group (Supplementary Table 2). ITT results indicated that among those with elevated symptoms at baseline, n=9 (18.8%) W-PPMA and n=9 (18.4%) Waitlist participants improved by at least four points. Per protocol subgroup results indicated that among those with elevated symptoms at baseline n=7 (30.4%) of W-PPMA and n=5 (12.5%) of the Waitlist group improved by at least 4 points.

Exploratory Outcomes: Postnatal Anxiety (GAD-7), Depression (EPDS), Stress (PSS), Mother-to-Infant Bond (MIB), and Therapeutic Bond (WAI-SR)

Preliminary mood and efficacy measures. Table 3 displays the median (Q1, Q3) scores with Cohen's d between-group effect size for each exploratory outcome at each time point. The PSS and EPDS showed small between-group effect sizes in favor of W-PPMA at 4-, 8-, and 12-week points. By the 16-week end of study point differences had mostly disappeared, driven by a reduction in Waitlist scores while the W-PPMA scores stayed relatively stable after 8-weeks; the Waitlist group had access to the app from 8-weeks to 16-weeks while the W-PPMA group had access for the full 16-weeks. There were very small or no differences between W-PPMA and Waitlist groups across all time points for the GAD-7 and MIBS. The number of participants that completed the measure and included in these calculations was similar for each measure and ranged from 80.1% to 98.5% (complete breakdown of data completion is in Supplementary Table 3). In the per protocol sensitivity analyses, differences were not statistically significant but tended to be larger than in the primary ITT analysis, trending in favor of the W-PPMA group, with the greatest difference on the PSS at EOI (Cohen's d = -0.47) (Supplementary Table 4). Figure 4 shows the change scores for each measure across the study period. The mean difference from baseline to EOI was -1.59 (95% CI: -2.44, -0.73) for PSS, -0.51

(95% CI: -1.36, 0.34) for EPDS, -0.05 (95% CI: -0.91, 0.80) for GAD-7, and -0.43 (95% CI: -1.28, 0.42) for MIB. For all measures except MIB the difference was larger in favor of W-PPMA in the per protocol sensitivity analyses (Supplementary Table 4 legend)

Therapeutic bond. The WAI-SR median (Q1, Q3) score was 42.0 (36.0, 51.0) at Baseline among the n=94 (65.3%) W-PPMA participants who completed it, and increased to 48.0 (41.0, 53.0) at EOI among the 113 (78.5%) who completed it at the second time measure. Per protocol sensitivity analysis scores were similar (Baseline: 43.0 (37.0, 51.0) with n=54, Week 8: 51.0 (45.0, 54.8) with n = 70).

Exploratory Analyses: Stratified Analyses

Perceptions of stigma. The overall baseline PBC - Stigma Subscale score median (Q1, Q3) was 12.0 (9.0, 15.0), similar in W-PPMA (12.0 (9.0,14.0)) and Waitlist (11.0 (10.0, 15.0)) groups at Baseline. Variation across strata was within two points. The highest point estimate was among the racial background strata of Hispanic/Latina (14.0 (13.8, 15.2)) and those who did not have health insurance (14.0 (11.5,17.5)). The lowest point estimates were among those employed part time (10.0 (8.0, 13.5)) and with military health insurance (10.0 (9.0, 11.5) (Supplementary Table 5).

Usability and Acceptability Measures. The median URPI-F score was similar across strata with the lowest median scores being 2 points below the overall median among those not employed and with ADI ranking 75 and above, and 5 points lower than the overall median among those without health insurance. Two overall strata had median scores 2 points higher, including military healthcare insurance and those with high school or GED education. The URPI-A had a wider range of scores across strata with a median score 6 points above the overall median among those single, separated, or divorced and 5 points higher among ADI ranking 75 and above; and median 6 points below the overall median amount those without health insurance. The CSQ-8 had the highest median score among those with a high school or GED education and lowest among those without health insurance. All stratified usability and acceptability measure scores, including by study subgroups, are in Supplementary Table 6.

Preliminary mood and efficacy measures. The PSS Baseline scores were highest among those who

reported being Black, not employed, or not having health insurance. The EOI scores were highest among those who reported being Asian, Hispanic/Latina, or not employed (Supplementary Table 7). EPDS baseline scores were highest among those who reported being Black and without health insurance, with Week 8 scores additionally highest among those who reported being Black or not employed (Supplementary Table 8). GAD-7 and MIB median scores among the entire cohort were within 2 points across all strata (Supplementary Tables 9 and 10, respectively). WAI-SR scores were highest among Black, and ADI ranking of at least 75 strata at baseline and among those with military health insurance or ADI ranking of at least 75 at Week 8 (Supplementary Table 11).

Discussion

Principal Results

In this study, we assessed the feasibility, acceptability, and preliminary efficacy of the Woebot for Postpartum Mood and Anxiety (W-PPMA) intervention to provide emotional support among a diverse and bilingual cohort. Results demonstrated high feasibility and acceptability among participants, as evidenced by strong scores on the Usage Rating Profile–Intervention Revised (URP-IR) and Client Satisfaction Questionnaire (CSQ-8). The intervention achieved notable engagement, with participants using W-PPMA for a median of nine days and sending an average of 217 messages across the 8-week intervention period.

Across all analyses comparing the W-PPMA to the Waitlist control group, the W-PPMA showed similar or more positive trends than the Waitlist group, suggesting access to W-PPMA has the potential to provide meaningful support during the postpartum period. Although the primary efficacy outcome (PHQ-8) revealed a small effect size that was not statistically significant in the ITT analysis, secondary and exploratory analyses provided promising trends. Participants in the W-PPMA group experienced improvements in perceived stress (PSS) and depressive symptoms (EPDS) compared to the waitlist control group at 8-weeks EOI. Per protocol analyses indicated moderate effect sizes, particularly for perceived stress, highlighting the potential of W-PPMA in reducing postpartum psychological distress. Therapeutic bonding with Woebot was also positively rated, suggesting the relational agent's ability to foster a supportive user experience.

Despite challenges in achieving the full recruitment target, this study successfully enrolled a diverse cohort of postpartum individuals, ensuring representation of underrepresented racial, ethnic, and socioeconomic groups. Notably, W-PPMA demonstrated strong

engagement across these diverse populations, reinforcing its promise as a scalable and inclusive intervention. The findings further support W-PPMA's potential to foster positive therapeutic engagement among groups historically underrepresented in digital health research. For example, working alliance (WAI-SR) scores were highest among Black participants and those residing in the most socioeconomically disadvantaged neighborhoods (ADI ≥75) at baseline. By Week 8, Hispanic/Latina participants, those of multiple racial backgrounds, and those in areas with an ADI ranking of at least 75 demonstrated the highest WAI-SR scores, suggesting that W-PPMA effectively engaged these populations over time.

Additionally, participant satisfaction with W-PPMA was notably high among those with a high school or GED education, as indicated by the CSQ-8 scores. This underscores W-PPMA's accessibility and acceptability among individuals who may have limited access to traditional mental health services. The intervention's ability to resonate across educational backgrounds, racial and ethnic identities, and socioeconomic strata highlights its potential as an equitable and scalable digital mental health tool.

The small trends and non-statistically significant differences should also be interpreted in the context that the study did not reach target recruitment efforts, as described in the methods section, and thus statistical tests were underpowered. However, the overall findings suggest that digital interventions like W-PPMA hold promise for reducing postpartum psychological distress, particularly in populations with historically limited access to mental health care.

Limitations

First, while the sample included a more diverse population than the general U.S. postpartum population, it was still predominantly White and English-speaking, which may limit the generalizability of findings to broader postpartum populations. Despite efforts to recruit underrepresented racial and ethnic groups (URGs), challenges in reaching some populations persisted. Future studies should consider additional strategies to ensure even greater representation from URGs, such as leveraging community partnerships and multilingual recruitment materials.

Second, while the study demonstrated high engagement, the intervention's duration (8 weeks) may have been insufficient to capture the full impact of W-PPMA on all outcomes, particularly for participants with more severe baseline symptoms of depression. Additionally, the reliance on self-reported measures introduces the potential for response bias, and app engagement metrics may not fully capture the depth and/or nuance of participants' interactions with the intervention.

Third, an unexpected challenge was the fraudulent redemption of gift cards offered as participation incentives. The platform's automated system was exploited by 580 fraudulent enrollments, leading to a rapid spike in participation in both the baseline study and the Woebot Substudy. To counter this, we implemented anomaly detection measures, including monitoring the time taken to complete e-consent and flagging participants with irregular patterns of completion. These measures successfully curtailed fraudulent activity. Future studies should incorporate automated fraud detection tools from the outset to prevent disruptions and ensure data integrity.

Fourth, the lack of blinding and the use of a waitlist control group, rather than an active control, may have influenced participants' responses. While ethical considerations guided this choice, future research should explore the efficacy of W-PPMA against active comparators to better contextualize its effectiveness.

Finally, attrition and missing data were challenges, particularly among participants with higher baseline depressive symptoms. Although imputation methods and sensitivity analyses were used to address this, the results should be interpreted with caution.

Conclusions

This research showcases W-PPMA's feasibility, acceptability, and potential as a scalable emotional support intervention in the postpartum period among a diverse group of mothers. W-PPMA demonstrated high user engagement and satisfaction, and exploratory findings suggest its capacity to reduce depressive symptoms and perceived stress. Importantly, the intervention's scalability and accessibility address critical barriers to mental health care, such as stigma, cost, and provider shortages, offering an equitable solution for postpartum individuals across diverse demographic groups.

While the intervention showed promise, additional research is needed to evaluate its long-term efficacy and implementation in real-world settings. Studies involving larger, more diverse cohorts and active comparators are recommended to strengthen the evidence base. Furthermore, addressing systemic barriers to digital health adoption, including technological literacy and access to devices, will be essential for achieving equitable mental health outcomes.

Acknowledgements

This study was made possible through the support of the PowerMom research platform and its

Consortium partners. We extend our gratitude to the community-based organizations that assisted in recruitment efforts, particularly those serving Black and Hispanic/Latina populations. We also acknowledge the contributions of the study participants, whose engagement and feedback have provided invaluable insights.

From the Scripps Research team, we would like to thank Shiri Warshawsky for her instrumental role in study operations and monitoring, and Felipe Delgado for his contributions to custom development within MyDataHelps and preliminary data exports, Giorgio Quer and his data science team members for setting up the initial coding necessary for this project.

From the Woebot team, we would like to thank Stephanie Rappoport and Emily Durden for thier early contributions to drafting the study protocol, Megan Flom for quality control of the statistical analysis plan, and Stephanie Eaneff for developing the code and pulling app utilization data.

Lastly, we thank the Institutional Review Board of the Scripps Research Translational Institute for their oversight and guidance in conducting this research.

Conflicts of Interest

A.Robinson, A. Williams, and S. Rappoport are employed by Woebot Health. L.Bullard is a former employee of Woebot Health.

References

1. Henshaw C. Mood disturbance in the early puerperium: a review. Arch Womens Ment Health. 2003;6 Suppl 2:S33-42. Epub 20030809. doi: 10.1007/s00737-003-0004-x. PubMed PMID: 14615921.
2. Howard LM, Molyneaux E, Dennis CL, Rochat T, Stein A, Milgrom J. Non-psychotic mental disorders in the perinatal period. Lancet. 2014;384(9956):1775-88. Epub 20141114. doi: 10.1016/S0140-6736(14)61276-9. PubMed PMID: 25455248.
3. Balaram K, Marwaha R. Postpartum Blues(Archived). StatPearls. Treasure Island (FL)2025.
4. Reck C, Stehle E, Reinig K, Mundt C. Maternity blues as a predictor of DSM-IV depression and anxiety disorders in the first three months postpartum. J Affect Disord. 2009;113(1-2):77-87. Epub 20080624. doi: 10.1016/j.jad.2008.05.003. PubMed PMID: 18573539.
5. Tosto V, Ceccobelli M, Lucarini E, Tortorella A, Gerli S, Parazzini F, Favilli A. Maternity Blues: A Narrative Review. J Pers Med. 2023;13(1). Epub 20230113. doi: 10.3390/jpm13010154. PubMed PMID: 36675815; PMCID: PMC9863514.
6. Bauman BL, Ko JY, Cox S, D'Angelo Mph DV, Warner L, Folger S, Tevendale HD, Coy KC,

Harrison L, Barfield WD. Vital Signs: Postpartum Depressive Symptoms and Provider Discussions About Perinatal Depression - United States, 2018. MMWR Morb Mortal Wkly Rep. 2020;69(19):575-81. Epub 20200515. doi: 10.15585/mmwr.mm6919a2. PubMed PMID: 32407302; PMCID: PMC7238954.
7. Prevention CfDCa. Symptoms of Depression Among Women May 15, 2024 [March 12, 2025]. Available from: https://www.cdc.gov/reproductivehealth/depression/index.htm.
8. Gavin NI, Gaynes BN, Lohr KN, Meltzer-Brody S, Gartlehner G, Swinson T. Perinatal depression: a systematic review of prevalence and incidence. Obstet Gynecol. 2005;106(5 Pt 1):1071-83. doi: 10.1097/01.AOG.0000183597.31630.db. PubMed PMID: 16260528.
9. Grisbrook MA, Letourneau N. Improving maternal postpartum mental health screening guidelines requires assessment of post-traumatic stress disorder. Can J Public Health. 2021;112(2):240-3. Epub 20200811. doi: 10.17269/s41997-020-00373-8. PubMed PMID: 32780347; PMCID: PMC7910351.
10. Brummelte S, Galea LA. Postpartum depression: Etiology, treatment and consequences for maternal care. Horm Behav. 2016;77:153-66. Epub 20150828. doi: 10.1016/j.yhbeh.2015.08.008. PubMed PMID: 26319224.
11. Bobo WV, Yawn BP. Concise review for physicians and other clinicians: postpartum depression. Mayo Clin Proc. 2014;89(6):835-44. doi: 10.1016/j.mayocp.2014.01.027. PubMed PMID: 24943697; PMCID: PMC4113321.
12. Li X, Laplante DP, Paquin V, Lafortune S, Elgbeili G, King S. Effectiveness of cognitive behavioral therapy for perinatal maternal depression, anxiety and stress: A systematic review and meta-analysis of randomized controlled trials. Clin Psychol Rev. 2022;92:102129. Epub 20220125. doi: 10.1016/j.cpr.2022.102129. PubMed PMID: 35123346.
13. Miniati M, Callari A, Calugi S, Rucci P, Savino M, Mauri M, Dell'Osso L. Interpersonal psychotherapy for postpartum depression: a systematic review. Arch Womens Ment Health. 2014;17(4):257-68. Epub 20140624. doi: 10.1007/s00737-014-0442-7. PubMed PMID: 24957781.
14. Prevatt BS, Desmarais SL. Facilitators and Barriers to Disclosure of Postpartum Mood Disorder Symptoms to a Healthcare Provider. Matern Child Health J. 2018;22(1):120-9. doi: 10.1007/s10995-017-2361-5. PubMed PMID: 28766092.
15. Fonseca A, Gorayeb R, Canavarro MC. Women´s help-seeking behaviours for depressive symptoms during the perinatal period: Socio-demographic and clinical correlates and perceived barriers to seeking professional help. Midwifery. 2015;31(12):1177-85. Epub 20150921. doi: 10.1016/j.midw.2015.09.002. PubMed PMID: 26433622.
16. Grissette BG, Spratling R, Aycock DM. Barriers to Help-Seeking Behavior Among Women With Postpartum Depression. J Obstet Gynecol Neonatal Nurs. 2018;47(6):812-9. Epub 20181006. doi: 10.1016/j.jogn.2018.09.006. PubMed PMID: 30296405.
17. Dennis CL, Chung-Lee L. Postpartum depression help-seeking barriers and maternal treatment preferences: a qualitative systematic review. Birth. 2006;33(4):323-31. doi: 10.1111/j.1523-536X.2006.00130.x. PubMed PMID: 17150072.
18. Sherman LJ, Ali MM. Diagnosis of Postpartum Depression and Timing and Types of Treatment Received Differ for Women with Private and Medicaid Coverage. Womens Health Issues. 2018;28(6):524-9. Epub 20181009. doi: 10.1016/j.whi.2018.08.007. PubMed PMID: 30314906.
19. Cook BL, Trinh NH, Li Z, Hou SS, Progovac AM. Trends in Racial-Ethnic Disparities in Access to Mental Health Care, 2004-2012. Psychiatr Serv. 2017;68(1):9-16. Epub 20160801. doi: 10.1176/appi.ps.201500453. PubMed PMID: 27476805; PMCID: PMC5895177.
20. Summers-Gabr NM. Rural-urban mental health disparities in the United States during COVID-19. Psychol Trauma. 2020;12(S1):S222-S4. Epub 20200601. doi: 10.1037/tra0000871. PubMed PMID: 32478541.
21. Feldman N, Perret S. Digital mental health for postpartum women: perils, pitfalls, and promise. NPJ Digit Med. 2023;6(1):11. Epub 20230126. doi: 10.1038/s41746-023-00756-4. PubMed

PMID: 36702915; PMCID: PMC9879244.
22. Gega L, Marks I, Mataix-Cols D. Computer-aided CBT self-help for anxiety and depressive disorders: experience of a London clinic and future directions. J Clin Psychol. 2004;60(2):147-57. doi: 10.1002/jclp.10241. PubMed PMID: 14724922.
23. Ajayi T KJ, Ariniello L, Ho D, Delgado F, Beal M, Waalen J, Baca Motes K, Ramos E. Leveraging a Digital Health Platform for Observational Maternal Health Research: Design, Recruitment Strategies, and Lessons Learned from the PowerMom Study. JMIR Formative Research. 25/02/2025:70149. doi: 10.2196/70149.
24. Fitzpatrick KK, Darcy A, Vierhile M. Delivering Cognitive Behavior Therapy to Young Adults With Symptoms of Depression and Anxiety Using a Fully Automated Conversational Agent (Woebot): A Randomized Controlled Trial. JMIR Ment Health. 2017;4(2):e19. Epub 20170606. doi: 10.2196/mental.7785. PubMed PMID: 28588005; PMCID: PMC5478797.
25. Ramachandran M, Suharwardy S, Leonard SA, Gunaseelan A, Robinson A, Darcy A, Lyell DJ, Judy A. 74: Acceptability of postnatal mood management through a smartphone-based automated conversational agent. American Journal of Obstetrics & Gynecology. 2020;222(1):S62. doi: 10.1016/j.ajog.2019.11.090.
26. Suharwardy S, Ramachandran M, Leonard SA, Gunaseelan A, Robinson A, Darcy A, Lyell DJ, Judy A. 116: Effect of an automated conversational agent on postpartum mental health: A randomized, controlled trial. American Journal of Obstetrics & Gynecology. 2020;222(1):S91. doi: 10.1016/j.ajog.2019.11.132.
27. Prochaska JJ, Vogel EA, Chieng A, Kendra M, Baiocchi M, Pajarito S, Robinson A. A Therapeutic Relational Agent for Reducing Problematic Substance Use (Woebot): Development and Usability Study. J Med Internet Res. 2021;23(3):e24850. Epub 20210323. doi: 10.2196/24850. PubMed PMID: 33755028; PMCID: PMC8074987.
28. Prochaska JJ, Vogel EA, Chieng A, Baiocchi M, Maglalang DD, Pajarito S, Weingardt KR, Darcy A, Robinson A. A randomized controlled trial of a therapeutic relational agent for reducing substance misuse during the COVID-19 pandemic. Drug Alcohol Depend. 2021;227:108986. Epub 20210827. doi: 10.1016/j.drugalcdep.2021.108986. PubMed PMID: 34507061; PMCID: PMC8423936.
29. Institute TSR. PowerMom: A consortium advancing research into maternal health: Digital Trials Center; 2024 [March 12, 2025]. Available from: https://digitaltrials.scripps.edu/powermom/.
30. All of Us Research Program I, Denny JC, Rutter JL, Goldstein DB, Philippakis A, Smoller JW, Jenkins G, Dishman E. The "All of Us" Research Program. N Engl J Med. 2019;381(7):668-76. doi: 10.1056/NEJMsr1809937. PubMed PMID: 31412182; PMCID: PMC8291101.
31. Association AP. ASSESSING AND TREATING SUICIDAL BEHAVIORS : A Quick Reference Guide November, 2003 [March 12, 2025]. Available from: https://psychiatryonline.org/pb/assets/raw/sitewide/practice_guidelines/guidelines/suicide-guide-1410457253647.pdf.
32. Association AM. Suicide Prevention: How-To Guide May 8, 2024 [March 12, 2025]. Available from: https://www.ama-assn.org/system/files/bhi-suicide-prevention-how-to-guide.pdf.
33. Briesch AM, Chafouleas SM, Neugebauer SR, Riley-Tillman TC. Assessing influences on intervention implementation: revision of the usage rating profile-intervention. J Sch Psychol. 2013;51(1):81-96. Epub 20120920. doi: 10.1016/j.jsp.2012.08.006. PubMed PMID: 23375174.
34. Larsen DL, Attkisson CC, Hargreaves WA, Nguyen TD. Assessment of client/patient satisfaction: development of a general scale. Eval Program Plann. 1979;2(3):197-207. doi: 10.1016/0149-7189(79)90094-6. PubMed PMID: 10245370.
35. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med. 2001;16(9):606-13. doi: 10.1046/j.1525-1497.2001.016009606.x. PubMed PMID: 11556941; PMCID: PMC1495268.
36. Dhingra SS, Kroenke K, Zack MM, Strine TW, Balluz LS. PHQ-8 Days: a measurement option for DSM-5 Major Depressive Disorder (MDD) severity. Popul Health Metr. 2011;9:11. Epub

20110428. doi: 10.1186/1478-7954-9-11. PubMed PMID: 21527015; PMCID: PMC3108293.
37. Pavlov C, Egan K, Limbers C. Reliability and validity of the PHQ-8 in first-time mothers who used assisted reproductive technology. Human Reproduction Open. 2022;2022(2). doi: 10.1093/hropen/hoac019.
38. Kroenke K, Spitzer RL. The PHQ-9: A new depression diagnostic and severity measure. Psychiatric Annals. 2002;32(9):509-15. doi: 10.3928/0048-5713-20020901-06.
39. Cox JL, Holden JM, Sagovsky R. Detection of postnatal depression. Development of the 10-item Edinburgh Postnatal Depression Scale. Br J Psychiatry. 1987;150:782-6. doi: 10.1192/bjp.150.6.782. PubMed PMID: 3651732.
40. Yawn BP, Pace W, Wollan PC, Bertram S, Kurland M, Graham D, Dietrich A. Concordance of Edinburgh Postnatal Depression Scale (EPDS) and Patient Health Questionnaire (PHQ-9) to assess increased risk of depression among postpartum women. J Am Board Fam Med. 2009;22(5):483-91. doi: 10.3122/jabfm.2009.05.080155. PubMed PMID: 19734393.
41. Spitzer RL, Kroenke K, Williams JB, Lowe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med. 2006;166(10):1092-7. doi: 10.1001/archinte.166.10.1092. PubMed PMID: 16717171.
42. Dear BF, Titov N, Sunderland M, McMillan D, Anderson T, Lorian C, Robinson E. Psychometric comparison of the generalized anxiety disorder scale-7 and the Penn State Worry Questionnaire for measuring response during treatment of generalised anxiety disorder. Cogn Behav Ther. 2011;40(3):216-27. Epub 20110720. doi: 10.1080/16506073.2011.582138. PubMed PMID: 21770844.
43. Cohen S, Kamarck T, Mermelstein R. A Global Measure of Perceived Stress. Journal of Health and Social Behavior. 1983;24(4):385-96. doi: 10.2307/2136404.
44. Lee EH. Review of the psychometric evidence of the perceived stress scale. Asian Nurs Res (Korean Soc Nurs Sci). 2012;6(4):121-7. Epub 20120918. doi: 10.1016/j.anr.2012.08.004. PubMed PMID: 25031113.
45. Taylor A, Atkins R, Kumar R, Adams D, Glover V. A new Mother-to-Infant Bonding Scale: links with early maternal mood. Arch Womens Ment Health. 2005;8(1):45-51. Epub 20050504. doi: 10.1007/s00737-005-0074-z. PubMed PMID: 15868385.
46. Wittkowski A, Wieck A, Mann S. An evaluation of two bonding questionnaires: a comparison of the Mother-to-Infant Bonding Scale with the Postpartum Bonding Questionnaire in a sample of primiparous mothers. Arch Womens Ment Health. 2007;10(4):171-5. Epub 20070704. doi: 10.1007/s00737-007-0191-y. PubMed PMID: 17607505.
47. Hatcher RL, Gillaspy JA. Development and validation of a revised short version of the Working Alliance Inventory. Psychotherapy Research. 2006;16(1):12-25. doi: 10.1080/10503300500352500.
48. Mohr DC, Ho J, Duffecy J, Baron KG, Lehman KA, Jin L, Reifler D. Perceived barriers to psychological treatments and their relationship to depression. J Clin Psychol. 2010;66(4):394-409. doi: 10.1002/jclp.20659. PubMed PMID: 20127795; PMCID: PMC2907887.
49. 6.0.3 m. miceforest: Fast, Memory Efficient Imputation with LightGBM Aug 1, 2024 [March 12, 2025]. Available from: https://pypi.org/project/miceforest/.

Tables

	Overall n (%)	W-PPMA n (%)	Waitlist n (%)
Number of Participants	267	144	123

Age category			
<35	190 (71.2)	103 (71.5)	87 (70.7)
>=35	76 (28.5)	40 (27.8)	36 (29.3)
Missing	1 (0.4)	1 (0.7)	0 (0.0)
Gender			
Woman	267 (100.0)	144 (100.0)	123 (100.0)
Sex assigned at birth			
Female	266 (99.6)	143 (99.3)	123 (100.0)
Missing	1 (0.4)	1 (0.7)	0 (0.0)
Hispanic/Latina			
Yes	34 (12.7)	23 (16.0)	11 (8.9)
No	231 (86.5)	120 (83.3)	111 (90.2)
Missing	2 (0.7)	1 (0.7)	1 (0.8)
Racial background			
Asian	8 (3.0)	5 (3.5)	3 (2.4)
Black/African American/African	23 (8.6)	11 (7.6)	12 (9.8)
Hispanic/Latina/Spanish*	15 (5.6)	11 (7.6)	4 (3.3)
White	183 (68.5)	93 (64.6)	90 (73.2)
Multiple	35 (13.1)	24 (16.7)	11 (8.9)
Missing	3 (1.1)	0 (0.0)	3 (2.4)
Marital status			
Married/Partnered/Cohabitating	240 (89.9)	127 (88.2)	113 (91.9)
Divorced/Separated	4 (1.5)	3 (2.1)	1 (0.8)
Single	22 (8.2)	13 (9.0)	9 (7.3)
Missing	1 (0.4)	1 (0.7)	0 (0.0)
Education			
Graduate or postgraduate degree	107 (40.1)	60 (41.7)	47 (38.2)
College degree	85 (31.8)	39 (27.1)	46 (37.4)
Some college or technical school	50 (18.7)	27 (18.8)	23 (18.7)
High school graduate or GED (grade 12)	18 (6.7)	12 (8.3)	6 (4.9)
Some high school (grades 9-11)	6 (2.2)	5 (3.5)	1 (0.8)
Student	4 (1.5)	2 (1.4)	2 (1.6)
Missing	1 (0.4)	1 (0.7)	0 (0.0)
Sexual Orientation			

Straight or heterosexual	230 (86.1)	126 (87.5)	104 (84.6)
Bisexual	23 (8.6)	10 (6.9)	13 (10.6)
Lesbian or gay	3 (1.1)	2 (1.4)	1 (0.8)
Asexual	2 (0.7)	1 (0.7)	1 (0.8)
Pansexual	4 (1.5)	3 (2.1)	1 (0.8)
Demisexual	1 (0.4)	0 (0.0)	1 (0.8)
Queer	2 (0.7)	0 (0.0)	2 (1.6)
Don't know	1 (0.4)	1 (0.7)	0 (0.0)
Missing	1 (0.4)	1 (0.7)	0 (0.0)
Employment status			
Employed, full time	161 (60.3)	83 (57.6)	78 (63.4)
Employed, part time	23 (8.6)	12 (8.3)	11 (8.9)
Primary caretaker at home, not employed outside of the home	57 (21.3)	33 (22.9)	24 (19.5)
Not employed, looking for work	10 (3.7)	6 (4.2)	4 (3.3)
Not employed, not looking for work	9 (3.4)	6 (4.2)	3 (2.4)
On disability, not able to work	2 (0.7)	1 (0.7)	1 (0.8)
Student	4 (1.5)	2 (1.4)	2 (1.6)
Missing	1 (0.4)	1 (0.7)	0 (0.0)
Health insurance			
Private insurance	188 (70.4)	101 (70.1)	87 (70.7)
Medicaid, Medical Assistance, or any kind of government assistance plan for those with low incomes or a disability	56 (21.0)	28 (19.4)	28 (22.8)
TRICARE or other military health care	12 (4.5)	7 (4.9)	5 (4.1)
VA (including those who have ever used or enrolled for VA health care)	1 (0.4)	1 (0.7)	0 (0.0)
Medicare for people 65 and older, or people with certain disabilities	2 (0.7)	2 (1.4)	0 (0.0)
Do not have health insurance	5 (1.9)	2 (1.4)	3 (2.4)
Don't know	1 (0.4)	1 (0.7)	0 (0.0)
Missing	2 (0.7)	2 (1.4)	0 (0.0)
Level of access to maternity care			
Moderate Access to Care / Access to Maternity	245 (91.8)	133 (92.4)	112 (91.1)

Care			
Maternity Care Desert / Low Access to Care	17 (6.4)	8 (5.6)	9 (7.3)
Missing	5 (1.9)	3 (2.1)	2 (1.6)
Neighborhood atlas			
ADI_NATRANK >= 75	21 (7.9)	7 (4.9)	14 (11.4)
50 <= ADI_NATRANK < 75	82 (30.7)	42 (29.2)	40 (32.5)
25 <= ADI_NATRANK < 50	95 (35.6)	55 (38.2)	40 (32.5)
ADI_NATRANK < 25	66 (24.7)	38 (26.4)	28 (22.8)
Missing	3 (1.1)	2 (1.4)	1 (0.8)

Table 1 Sociodemographic characteristics of participants overall and by assigned treatment group. *These self-reported survey data were collected at study baseline. Note the one participant missing age value answered the screening questionnaire to confirm they were aged 16 years or older but did not provide their exact date of birth.* ***Legend:*** *W-PPMA: Woebot-Postpartum Mood and Anxiety , n = Number of Participants, GED = General Education Degree, VA = Veterans Affairs, ADI_Natrank = Area Deprivation Index - National Rank.* ******* *indicates participants who self identified their race as Hispanic/Latina in addition to their ethnicity has Hispanic/Latina*

	Overall	**W-PPMA**	**Waitlist**
Number of participants	267	144	123
Psychiatric History			
Previous postpartum depression diagnosis			
Yes	31 (11.6)	18 (12.5)	13 (10.6)
No	234 (87.6)	125 (86.8)	109 (88.6)
Prefer not to answer	1 (0.4)	0 (0.0)	1 (0.8)
Missing	1 (0.4)	1 (0.7)	0 (0.0)
Previous depression diagnosis			
Yes	113 (42.3)	58 (40.3)	55 (44.7)
No	152 (56.9)	84 (58.3)	68 (55.3)

Prefer not to answer	1 (0.4)	1 (0.7)	0 (0.0)
Missing	1 (0.4)	1 (0.7)	0 (0.0)
Prior and Concomitant Medications and Therapy at Baseline			
Any treatment at Connect study baseline	86 (32.2)	44 (30.6)	42 (34.1)
Psychotherapy			
Yes	41 (15.4)	24 (16.7)	17 (13.8)
No	221 (82.8)	118 (81.9)	103 (83.7)
Not sure	4 (1.5)	1 (0.7)	3 (2.4)
Missing	1 (0.4)	1 (0.7)	0 (0.0)
Mental Health Medication			
Yes	67 (25.1)	33 (22.9)	34 (27.6)
No	198 (74.2)	109 (75.7)	89 (72.4)
Missing	2 (0.7)	2 (1.4)	0 (0.0)
Connected to or attending activities at Connect study baseline	94 (35.2)	51 (35.4)	43 (35.0)
Online support groups for Moms	55 (20.6)	31 (21.5)	24 (19.5)
Mommy and me or similar gatherings	26 (9.7)	14 (9.7)	12 (9.8)
Online neighborhood groups (i.e. Nextdoor)	15 (5.6)	7 (4.9)	8 (6.5)
Other activity	19 (7.1)	9 (6.2)	10 (8.1)
Missing	1 (0.4)	1 (0.7)	0 (0.0)
Follow-up visit(s) scheduled during the next 8 weeks	251 (94.0)	130 (90.3)	121 (98.4)
OB/Gyn	206 (77.2)	105 (72.9)	101 (82.1)
Pediatrician	179 (67.0)	89 (61.8)	90 (73.2)
Primary care doctor	40 (15.0)	22 (15.3)	18 (14.6)
Psychiatrist, therapist, psychologist, or clinical social worker	53 (19.9)	31 (21.5)	22 (17.9)
Other provider	22 (8.2)	9 (6.2)	13 (10.6)
Missing	2 (0.7)	2 (1.4)	0 (0.0)
Pregnancy History			
Number of previous pregnancies			
0	7 (2.6)	4 (2.8)	3 (2.4)
1	116 (43.4)	66 (45.8)	50 (40.7)

2	70 (26.2)	32 (22.2)	38 (30.9)
3	38 (14.2)	21 (14.6)	17 (13.8)
4	17 (6.4)	8 (5.6)	9 (7.3)
5+	17 (6.4)	11 (7.6)	6 (4.9)
Missing	2 (0.7)	2 (1.4)	0 (0.0)
Number of term births			
0	5 (1.9)	3 (2.1)	2 (1.6)
1	142 (53.2)	74 (51.4)	68 (55.3)
2	72 (27.0)	39 (27.1)	33 (26.8)
3	25 (9.4)	14 (9.7)	11 (8.9)
4	9 (3.4)	5 (3.5)	4 (3.3)
5+	5 (1.9)	3 (2.1)	2 (1.6)
Not applicable	7 (2.6)	4 (2.8)	3 (2.4)
Missing	2 (0.7)	2 (1.4)	0 (0.0)
Number of premature births			
0	224 (83.9)	120 (83.3)	104 (84.6)
1	24 (9.0)	12 (8.3)	12 (9.8)
2	9 (3.4)	6 (4.2)	3 (2.4)
3	1 (0.4)	0 (0.0)	1 (0.8)
Not applicable	7 (2.6)	4 (2.8)	3 (2.4)
Missing	2 (0.7)	2 (1.4)	0 (0.0)
Number of miscarriages			
0	208 (77.9)	117 (81.2)	93 (75.6)
1	40 (15.0)	16 (11.1)	24 (19.5)
2	4 (1.5)	2 (1.4)	2 (1.6)
3	4 (1.5)	3 (2.1)	1 (0.8)
5+	2 (0.7)	2 (1.4)	0 (0.0)
Not applicable	7 (2.6)	4 (2.8)	3 (2.4)
Missing	2 (0.7)	2 (1.4)	0 (0.0)
Number of stillbirths			
0	257 (96.2)	137 (95.1)	120 (97.6)
1	1 (0.4)	1 (0.7)	0 (0.0)
Not applicable	7 (2.6)	4 (2.8)	3 (2.4)
Missing	2 (0.7)	2 (1.4)	0 (0.0)

Did your most recently born child spend time in the NICU?			
Yes	38 (14.2)	20 (13.9)	18 (14.6)
No	222 (83.1)	119 (82.6)	103 (83.7)
Prefer not to answer	3 (1.1)	1 (0.7)	2 (1.6)
Missing	4 (1.5)	4 (2.8)	0 (0.0)
Has your most recently born child been diagnosed with any of the following?			
No	241 (90.3)	132 (91.7)	109 (88.6)
Medical condition (e.g., loss, heart condition, other birth defect)	8 (3.0)	5 (3.5)	3 (2.4)
Genetic disorder (e.g., Down syndrome)	1 (0.4)	0 (0.0)	1 (0.8)
Medical and Genetic conditions	1 (0.4)	0 (0.0)	1 (0.8)
Other	12 (4.5)	3 (2.1)	9 (7.3)
Missing	4 (1.5)	4 (2.8)	0 (0.0)

Table 2 Participant psychiatric and pregnancy history overall and by study group. *These self-reported survey data were collected at the study baseline.* ***Legend:*** *W-PPMA: Woebot-Postpartum Mood and Anxiety, n = Number of Participants.*

	Baseline	Week 4	Week 8	Week 12	Week 16
Perceived Stress Scale (PSS) [Range 0,40]					
Overall	13.0 (8.8, 20.0)	15.0 (9.0, 21.0)	15.0 (8.0, 20.0)	14.0 (8.5, 20.5)	13.5 (8.0, 20.8)
W-PPMA	13.0 (8.0, 19.0)	15.0 (8.0, 20.0)	13.0 (8.0, 19.5)	13.0 (7.0, 20.0)	13.0 (8.0, 20.0)
Waitlist	14.0 (9.0, 20.0)	16.0 (10.2, 22.8)	16.0 (10.0, 22.0)	15.0 (10.0, 22.0)	15.0 (8.0, 21.0)
Cohen's d	-0.16	-0.25	-0.33	-0.23	-0.13
The Edinburgh Postnatal Depression Scale (EPDS) [Range 0, 27]					
Overall	5.0 (2.0, 9.0)	5.0 (2.0, 10.0)	5.0 (2.0, 9.0)	5.0 (2.0, 10.0)	4.0 (2.0, 8.0)
W-PPMA	5.0 (2.0, 8.0)	4.0 (2.0, 9.0)	4.0 (2.0, 8.0)	4.0 (1.0, 9.0)	4.0 (1.0, 8.0)
Waitlist	5.5 (2.0, 10.0)	7.0 (3.0, 11.0)	6.0 (2.2, 10.0)	5.0 (2.0, 11.0)	5.0 (2.0, 8.0)

Cohen's d	-0.16	-0.26	-0.22	-0.23	-0.10
The Generalized Anxiety Disorder Scale (GAD-7) [Range 0,21]					
Overall	4.0 (2.0, 7.0)	5.0 (2.0, 8.0)	5.0 (2.0, 8.0)	5.0 (1.0, 8.0)	4.0 (1.0, 7.0)
W-PPMA	4.0 (2.0, 7.0)	5.0 (2.0, 8.0)	4.0 (1.0, 7.0)	4.0 (1.0, 7.8)	4.0 (1.0, 7.0)
	5.0 (3.0, 7.0)	5.0 (3.0, 8.0)	5.0 (2.0, 8.0)	5.0 (2.0, 8.0)	4.0 (1.0, 7.0)
Waitlist					
Cohen's d	-0.18	-0.11	-0.14	-0.10	0.06
The Mother-to-Infant Bonding Scale (MIBS) [Range 0, 24]					
Overall	3.0 (1.0, 5.0)	3.0 (1.0, 6.0)	3.0 (1.0, 5.0)	2.0 (0.0, 6.0)	2.0 (1.0, 5.0)
W-PPMA	3.0 (1.0, 5.0)	3.0 (1.0, 5.0)	2.5 (1.0, 4.0)	2.0 (0.0, 5.2)	2.0 (1.0, 5.0)
	3.0 (1.0, 5.0)	3.0 (1.0, 6.0)	3.0 (1.0, 6.0)	2.0 (0.0, 6.0)	2.0 (0.0, 5.0)
Waitlist					
Cohen's d	0.04	-0.14	-0.12	-0.08	0.08

Table 3 Exploratory measures of mood and efficacy at each study time point. *Median (Quartile 1, Quartile 3) scores with Cohen's d measure of effect size between W-PPMA and Waitlist Groups for exploratory measures of mood and preliminary efficacy at each time point. Negative Cohen's d represents lower scores in the W-PPMA group. The number of participants who completed each measure and are represented in the calculations are reported in Supplementary Table 3.*

Figures

	Treatment Period				Follow-up	
W-PPMA Access:	W-PPMA Group				All Participants	
	Baseline	***Day 3***	***Week 4***	***Week 8 (EOI)***	***Week 12***	***Week 16***
PSS	All		All	All	All	All
PHQ-8	All		All	All	All	All
EPDS	All		All	All	All	All
GAD-7	All		All	All	All	All
MIB	All		All	All	All	All
WAI-SR	W-PPMA only			W-PPMA only		
PBC, Stigma		All				
CSQ-8				W-PPMA only		
UPRI				W-PPMA only		

Figure 1 Schedule of assessment. Time point where each outcome measure was administered and whether this was for all participants or only those in the W-PPMA group. Week 8 represents End of Intervention (EOI) after which the Waitlist control group additionally gained access to the Woebot app. *Legend*: W-PPMA = Woebot for Postpartum Mood and Anxiety; group, PSS = Perceived Stress Scale, PHQ-8 = Patient Health Questionnaire - 8, EPDS = Edinburgh Postnatal Depression Scale, GAD-7 = The Generalized Anxiety Disorder Scale, MIB = Mother-to-Infant Bond Scale, WAI-SR = Working Alliance Inventory - Short Revised, PBC = Perceptions of Stigma, Stigma Subscale, CSQ - 8 = Client Satisfaction Questionnaire - 8 items, UPRI = Usage Rating Profile Inventory - Feasibility and Acceptability subscales.

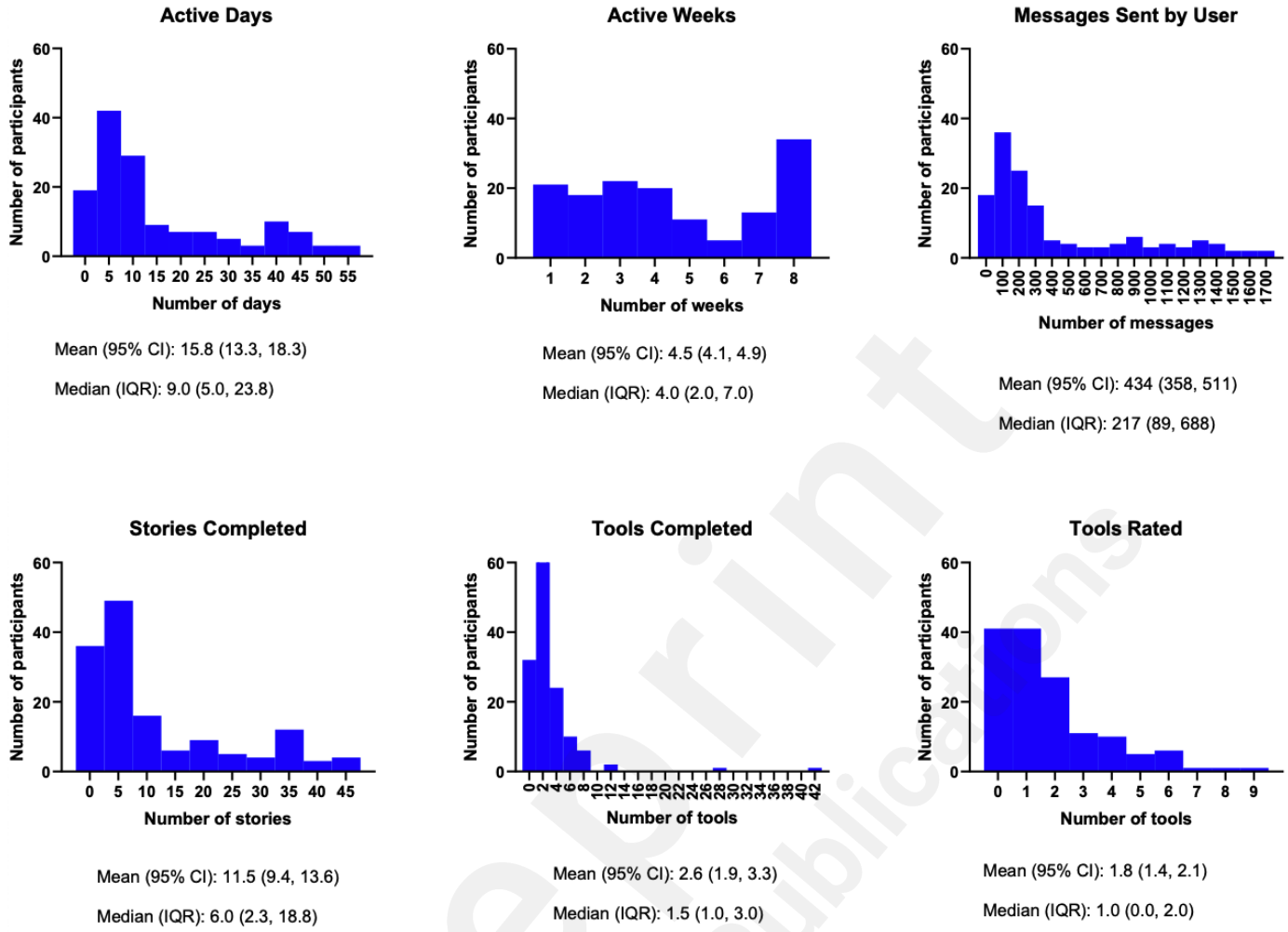


Figure 2. Woebot App engagement metrics. Figures represent primary outcome measures of app engagement among the W-PPMA group for the initial eight week treatment period. Because the W-PPMA user experience is tailored entirely to the user's input, there is no a-priori set number of stories and or tools a user must complete; rather each conversation with Woebot is adapted to the users in the moment need with offering of stories and or tools.

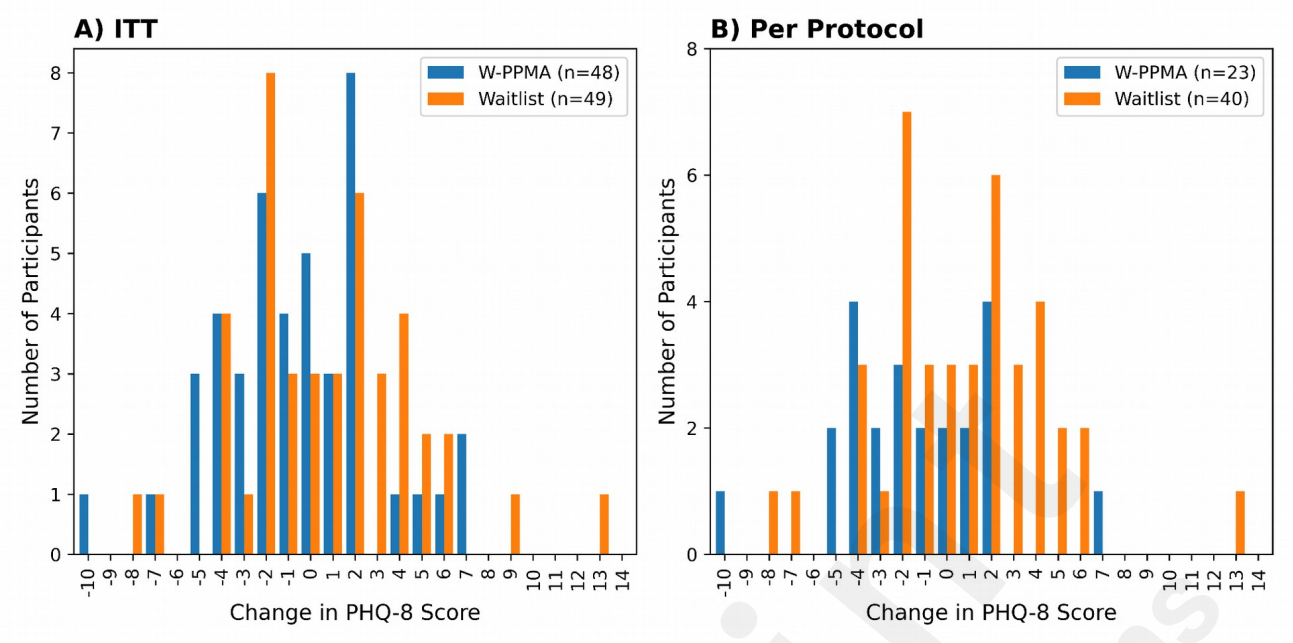


Figure 3 Secondary outcome PHQ-8 baseline to 8-week (end of intervention)) change scores for the Intention to Treat (ITT) primary and Per Protocol sensitivity analyses. Negative change scores represent a lowering of PHQ-8 scores (symptom reduction) over time; positive scores represent worsening scores. *Legend:* W-PPMA = Woebot for Postpartum Mood and Anxiety, PHQ-8 = Patient Health Questionnaire - 8 items, n = Number of Participants.

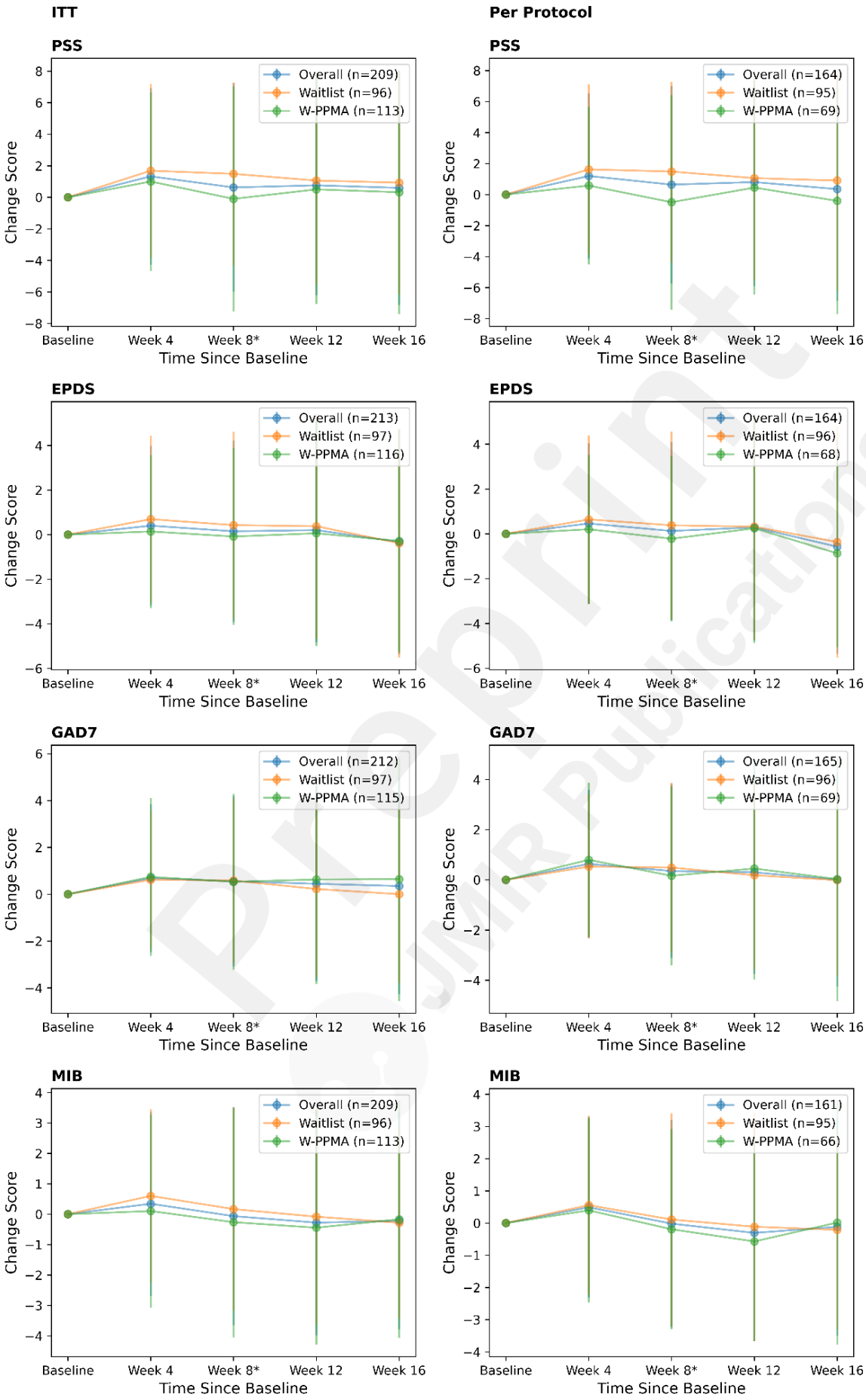


Figure 4 Change scores for exploratory mood and preliminary efficacy measures. Mean and standard deviation of change scores for each time point relative to the baseline overall and by treatment group. * represents the eight week end of intervention point where the waitlist group additionally gained access to the intervention. W-PPMA= Woebot for Postpartum Mood and Anxiety.

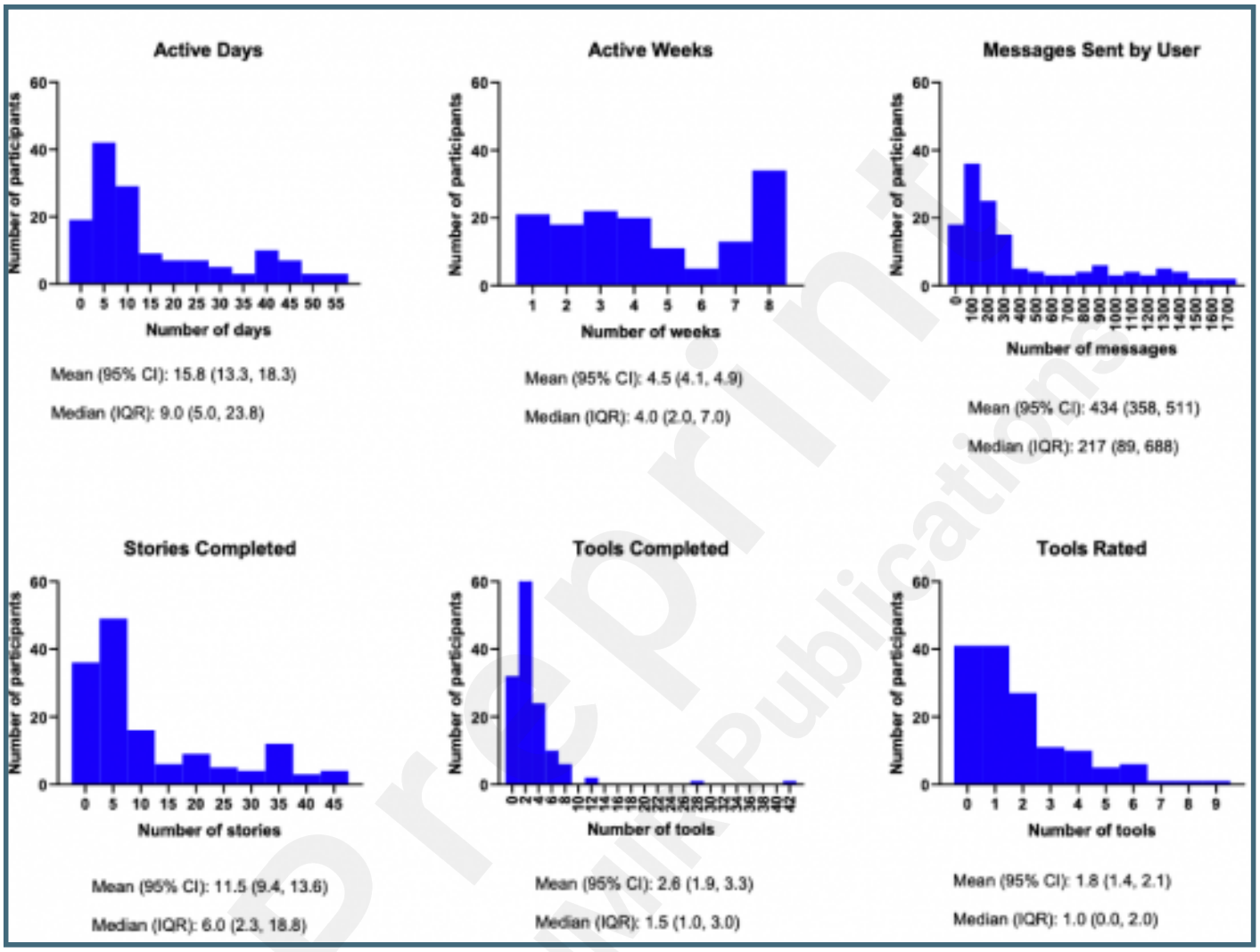
Supplementary Files

Figures

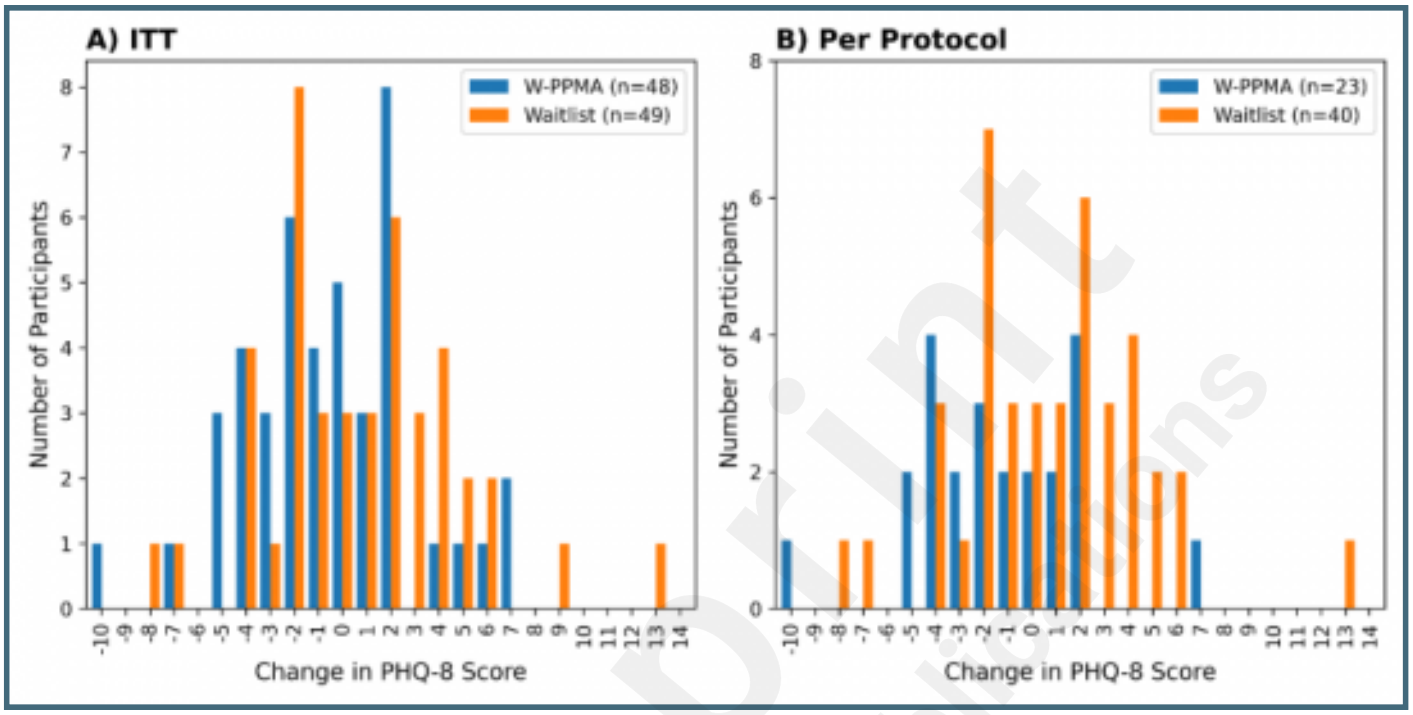
Time point where each outcome measure was administered and whether this was for all participants or only those in the W-PPMA group. Week 8 represents End of Intervention (EOI) after which the Waitlist control group additionally gained access to the Woebot app. Legend: W-PPMA = Woebot for Postpartum Mood and Anxiety; group, PSS = Perceived Stress Scale, PHQ-8 = Patient Health Questionnaire - 8, EPDS = Edinburgh Postnatal Depression Scale, GAD-7 = The Generalized Anxiety Disorder Scale, MIB = Mother-to-Infant Bond Scale, WAI-SR = Working Alliance Inventory - Short Revised, PBC = Perceptions of Stigma, Stigma Subscale, CSQ - 8 = Client Satisfaction Questionnaire - 8 items, UPRI = Usage Rating Profile Inventory - Feasibility and Acceptability subscales.

	Treatment Period				**Follow-up**	
W-PPMA Access:	W-PPMA Group				All Participants	
	Baseline	***Day 3***	***Week 4***	***Week 8 (EOI)***	***Week 12***	***Week 16***
PSS	All		All	All	All	All
PHQ-8	All		All	All	All	All
EPDS	All		All	All	All	All
GAD-7	All		All	All	All	All
MIB	All		All	All	All	All
WAI-SR	W-PPMA only			W-PPMA only		
PBC, Stigma		All				
CSQ-8				W-PPMA only		
UPRI				W-PPMA only		

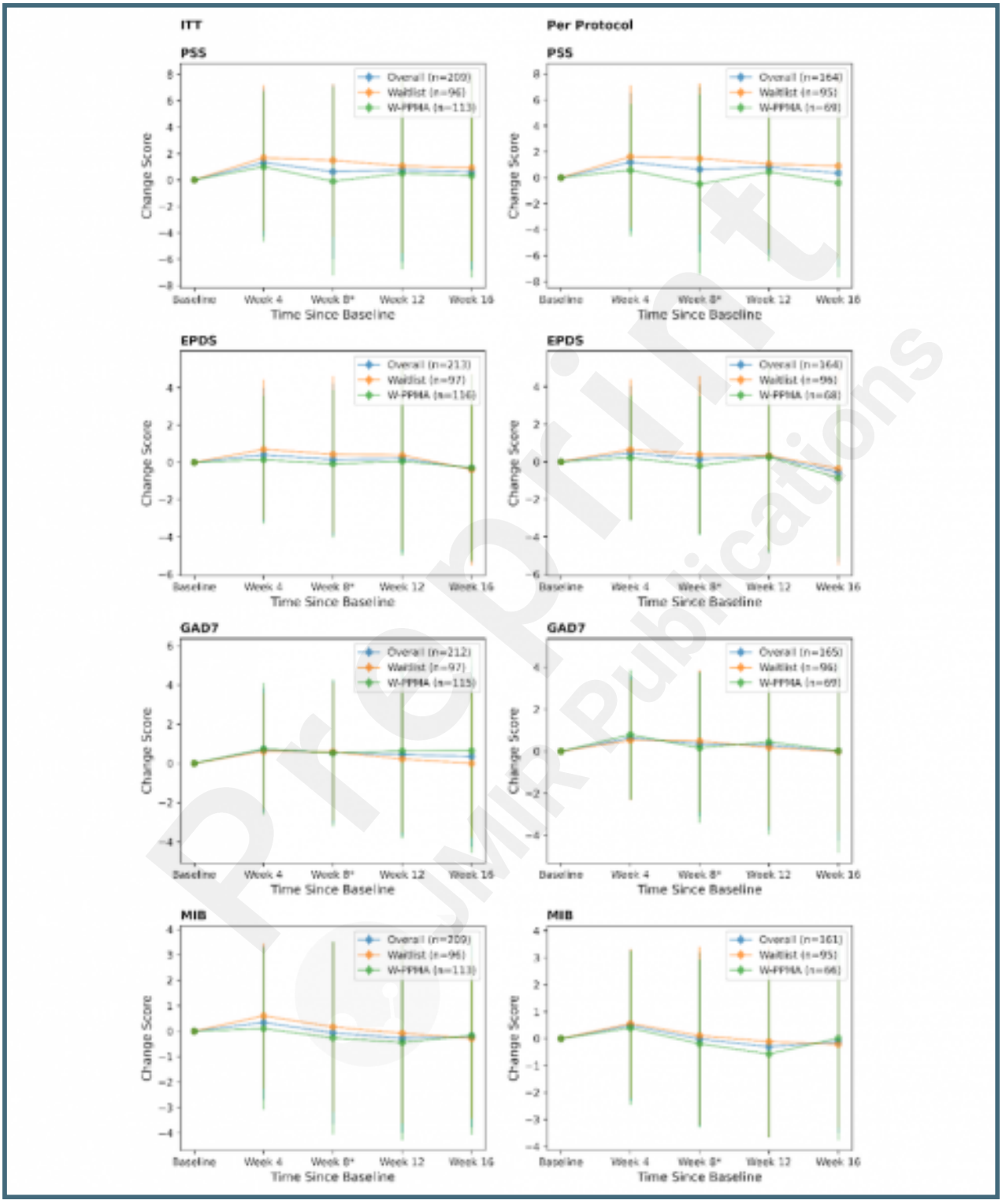
This represent primary outcome measures of app engagement among the W-PPMA group for the initial eight week treatment period. Because the W-PPMA user experience is tailored entirely to the user's input, there is no a-priori set number of stories and or tools a user must complete; rather each conversation with Woebot is adapted to the users in the moment need with offering of stories and or tools.



Secondary outcome PHQ-8 baseline to 8-week (end of intervention)) change scores for the Intention to Treat (ITT) primary and Per Protocol sensitivity analyses. Negative change scores represent a lowering of PHQ-8 scores (symptom reduction) over time; positive scores represent worsening scores. Legend: W-PPMA = Woebot for Postpartum Mood and Anxiety, PHQ-8 = Patient Health Questionnaire - 8 items, n = Number of Participants.



Mean and standard deviation of change scores for each time point relative to the baseline overall and by treatment group. * represents the eight week end of intervention point where the waitlist group additionally gained access to the intervention. W-PPMA= Woebot for Postpartum Mood and Anxiety.



CONSORT (or other) checklists

Completed checklist, doesn't look like my answers were saved, but the form was submitted online using the docs for.
URL: http://asset.jmir.pub/assets/fcef3a52bab3402b1b211871ccf4f4c4.pdf