THE LANCET
Obstetrics, Gynaecology, & Women's Health

Supplementary appendix

This appendix formed part of the original submission and has been peer reviewed. We post it as supplied by the authors.

Supplement to: Wolitzky-Taylor K, Richards MC, Metts A, et al. An integrated system of digital therapy and clinician care versus perinatal psychiatric care for women with perinatal depression and anxiety in the USA: a single-centre, randomised controlled trial. *Lancet Obstet Gynaecol Womens Health* 2026; published online June 4. https://doi.org/10.1016/S3050-5038(26)00096-8.

Supplement to

Integrated system of digital therapy and clinician care for perinatal depression and anxiety: a randomized controlled trial

Kate Wolitzky-Taylor, Misty C. Richards, Allison Metts, Amelia Welborn, Vanessa McDonald, Inna Arnaudova, Scott Fears, Heather O'Mahen, Jill M. Newby, Michael R Millard, Alan Stein, Nelson Freimer, and Michelle G. Craske

Table of Contents

Inclusion/Exclusion Criteria 2
STAND System of Care Condition 3
Perinatal Psychiatric Care Condition 8
Measures and Statistical Analysis Plan 9
Supplementary Results 13
Primary and Secondary Outcome Variable Descriptives and Available Data Table.....15
STAND Model of Care Figure 16
References 17

Details on Participant Inclusion and Exclusion Criteria (text below derived and adapted from Wolitzky-Taylor et al., 2024)(1).

Inclusion required women to be between 18 and 65 years old, fluent in English, sum score ≥ 11 on the Edinburgh Postnatal Depression Scale (EPDS) at screening, not currently in individual treatment for a behavioral or emotional problem (e.g., anxiety, depression), between week 28 of a pregnancy to 6 months postpartum, willing to follow study procedures, willing to participate in treatment through the study and follow all study procedures (including provide HIPAA Authorization for research), and have access to the internet via mobile or desktop device. Psychometrics and cut-offs for severity of depression do not differ between the full, 10-item version and the 9-item version of the EPDS from which the item that measures thoughts about self-harm is removed(2). A range of elevated EPDS scores has been shown to identify depression (e.g., 10-14(2), though a cut-off score of 11 or higher was found to be optimal to meet both DSM-5 and ICD-10 criteria for depression(3). Therefore, EPDS scores of 11 or greater were required to be eligible for the current study.

Exclusion criteria included current treatment by a therapist or a psychiatrist, unstable suicidality (e.g., two or more suicide attempts or self-injurious behaviors resulting in hospitalization in the last 6 months, combined with high ratings on self-reported negative urgency, as assessed via self-report instrument at baseline), current substance use disorder that interfered with treatment (e.g., unable to attend session not under the influence of that substance), current use of cocaine or non-prescribed opioids, principal diagnosis of psychosis unrelated to unipolar or bipolar depression, neurological conditions, severe uncontrolled medical conditions (e.g., anorexia nervosa, cardiac conditions requiring continuous monitoring), and cognitive impairment (e.g., developmental disability, dementia) as identified upon clinical assessment.

Details on the STAND system of care and Perinatal Psychiatric Care Conditions (text below derived and adapted from Wolitzky-Taylor et al., 2024)(1).

A. Screening and Treatment of Anxiety and Depression (STAND)

STAND for PND is a version of STAND (shown in Figure S1) tailored specifically to PND. STAND for PND assigns women to either (1) Digital therapy with coaching or (2) Clinical care, dependent upon baseline CAT-MH scores. Those with CAT-MH depression severity scores ≤75 and no current suicidality were allocated to perinatal depression digital cognitive behavioral therapy with coaching by bachelors' level associates and clinical psychology doctoral students. Those with depression scores in the severe range on the CAT-MH (≥75.1) or who endorsed items on the CAT-MH indicating risk of suicide were assigned to clinical care.

A.1 Ongoing Assessment/Measurement-based Care. Participants in STAND completed CAT-MH assessments weekly throughout the 6 months of study participation. CAT-MH scores during treatment and following treatment informed potential adaptations to care and were used to identify and respond to suicide risk (see below). Research staff monitored the completion of CAT-MH by STAND participants and prompted participants if >3 consecutive CAT-MH measures were missed. Participants were encouraged to respond to these measures by their providers as well as research staff.

A.2 Adaptation of Care. CAT-MH scores throughout the entire six months of participation informed tier switching. Moderately depressed participants assigned originally to digital therapy with coaching could be switched to clinical care at any time, during the period of time they were completing the digital therapy with coaching (~10 weeks) or thereafter, should their depression worsen to the severe level or in the case of suicidality. While in active treatment, two consecutive weeks of a severe depression score on the CAT-MH triggered a symptom worsening

alert for digital therapy participants, indicating they should be moved to clinical care. During remote monitoring (following the completion of acute treatment), worsening depression was defined as a CAT-MH depression score that increased 30% increase from prior score at treatment completion and initiated outreach to re-engage in care if that score exceeded the mild range.

Those who completed their clinical care prior to the final six-month assessment continued to complete weekly CAT-MH assessments of depression, anxiety and suicidality, at which times signs of symptom worsening or suicidality could activate re-initiation of either digital therapy with coaching or clinical care, depending on symptom severity at that point. Those who required additional care beyond the six months were referred to appropriate mental health resources that included therapy and psychiatry in the community.

A.3 Suicide Risk. If participant responses to their weekly assessments (CAT-MH with suicide risk items) indicated a risk for suicide, a computer-automated alert was sent to the coach/clinician (if the patient is actively in care), the research team, and a third-party contracted service. The third party contracted service made up to three outreach attempts to the patient within 24 hours to assess for, and address safety concerns. If the third-party service could not reach the patient or determined that additional follow-up was needed, the clinical team made efforts alongside the research team to reach the patient and to address safety needs as clinically indicated. This approach to risk management was implemented for both levels of care within STAND (digital therapy with coaching, and clinical care). After triggering a positive suicide risk alert, participants from digital therapy with coaching were moved up to clinical care.

A.4 Digital therapy with Coaching. Participants with CAT-MH depression severity scores $\leq$75 and no current suicidality at baseline were initially allocated to a perinatal depression digital cognitive-behavioral therapy course. After enrolling 94 participants, a new digital therapy program was introduced due to a mid-stream change in access to digital programs. For the first 94

participants, those who were assigned to digital therapy + coaching within the STAND condition received a digital therapy program developed by THIS WAY UP (see https://thiswayup.org.au/), , a non-profit entity that provides digital CBT for anxiety and depression. This program included two parts: 1) a 3-lesson CBT course specifically tailored to perinatal and postnatal anxiety and depression(4-6), shown to be effective in randomized controlled trials, and 2) six optional CBT lessons from TWU for mixed anxiety and depression(7), also shown to be effective in randomized controlled trials(5). Courses in TWU include illustrated lessons that follow a fictional character who experiences mental health difficulties. Throughout the courses, the character learns about the symptoms they experience and how to apply CBT or mindfulness skills. Each individual lesson ends with a summary and homework exercises to be completed prior to the next lesson. Participants assigned to TWU were provided with a link to a separate platform/software took to complete lessons that was distinct from the STAND assessment system. An initial research agreement with TWU was for a pre-determined specified period of time; after that agreement ended, we switched to another digital therapy for PND created specifically for this study. The two digital therapies comprised very similar cognitive and behavioral therapy principles, skills, and exercises applied to perinatal depression.

Starting with the 95th participant, those who were assigned to digital therapy + coaching within STAND received a new digital therapy program for perinatal depression. This second digital therapy was part of the UCLA DGC suite of digital CBT programs for anxiety and depression. The specific DGC PND program included (and extended from) many of the principles used in Netmum(8), a digital program shown to be effective for PND. The DGC PND program includes 10 lessons, with five core modules followed by five additional, optional modules. Lessons include skills of behavioral activation scheduling and monitoring, avoidance traps, challenging negative thoughts, self-compassion, imaginal recounting of positive autobiographical experiences,

sleep hygiene, present-moment focus with kindness and gratitude intentions and actions, if-then planning to break cycles of excessive worry, building support/communication skills, and concrete problem solving to manage stressful situations. Each lesson provides examples relevant to PND and includes a homework exercise to reinforce the content of the lesson. Participants are encouraged to practice their lesson homework over the course of the week before the next session. Lessons are completed sequentially and accessed online. As the DGC PND program was developed in-house for the STAND program, the software tool was embedded into the STAND digital platform.

In both digital therapies (i.e., TWU and DGC PND), participants were offered a review of material with a STAND coach who answered questions about the digital content, reviewed portions of the digital material where appropriate, addressed barriers to home practice completion, and made home practice recommendations. Coaching sessions were delivered via zoom and designed to support the lesson after it is reviewed, but lessons could be reviewed during the coaching session if the participant had not had a chance to review it prior to the coaching session. The number of coaching sessions (typically weekly, 45 min sessions) offered matched the number of lessons; 9 for TWU program vs 10 for the DGC PND program. All participants were given 10 weeks to complete their lessons and coaching sessions.

Bachelor's level staff associates and clinical psychology doctoral students served as the coaches and received training in CBT for anxiety and depression, and didactics relevant to the perinatal patient population. They also received training in how to "coach" (vs. provide therapy) and in the digital therapy, including digital platform navigation and CBT for PND. Coaches received didactic training and experiential role play practice for each module of the digital therapy program, and weekly supervision by licensed clinicians (one PhD in clinical psychology and one licensed social worker).

A.5 Clinical care. Participants with severe depressive symptoms (CAT-MH ≥ 75.1) or significant suicidality on the baseline CAT-MH were allocated to clinical care, which entailed weekly psychotherapy sessions by PhD students in clinical psychology, with the goal of 12-16 sessions but more as needed (i.e., if CAT-MH scores remained in the severe range). CAT-MH scores were monitored weekly by clinicians through the STAND dashboard. Psychotherapy was based on a modular treatment approach that was personalized to the patient's needs. A functional assessment was completed at intake sessions during which clinicians assessed the client's principal problem (most commonly depression given the inclusion criteria, but may be an anxiety or trauma disorder). The functional assessment guided selection of an evidence-based intervention suited to the principal problem. Examples of evidence-based interventions included behavioral activation for depressed mood (adapted for perinatal populations in accordance with Stein et al., 2018(9)), cognitive restructuring for excessive worry, exposure for panic and social anxiety, and distress tolerance skills for affective instability/self-harm/chronic suicidality. If the first-line intervention did not result in meaningful clinical change after six sessions per CAT-MH scores (i.e., scores remained in severe range), clinicians revise the functional assessment and shifted to a second-line treatment for the principal problem, after consultation with their licensed supervisor. Interventions to address patient non-adherence, acute suicidality, and major life stressors were incorporated as needed. See Wolitzky-Taylor et al. (2023)(10) for a more detailed description of the modular, process-based approach to therapy in STAND and a full list of intervention modalities.

Psychotropic medications were administered by psychiatry residents and offered as an adjunct treatment when deemed necessary or clinically appropriate during case conference consultation (see below). The schedule of psychiatric care was determined by the provider for up to a maximum of six months. Medication prescriptions followed best practices for perinatal

populations, typically SSRIs shown to be safe during pregnancy and breastfeeding (e.g., sertraline(11, 12)).

A.6 Clinician Training. Clinical fellows or psychiatry residents conducted all treatment appointments and received weekly supervision by licensed clinicians (psychotherapy services) and attending physicians (psychiatric services). Training on psychotherapy interventions involved a once per week (3-hour seminar) training series over the course of approximately three months. Seminars were led by experts on each psychological intervention option provided in STAND clinical care. Seminars involved didactics on the intervention content and role-plays to ensure understanding. Weekly case conferences were attended by clinical fellows, psychiatry residents, licensed clinicians, and attending physicians during which clinical care and patient needs were discussed, including recommendations for addition of psychotropic medications to the psychological intervention.

B. Perinatal Psychiatric Care (PPC)

The PPC protocol consisted of a comprehensive, diagnostic psychiatric assessment that included a history and mental status examination plus up to three follow-up visits as deemed appropriate for up to six months; the follow-up visits were standardized to include supportive therapy using a standardized approach(13) and shown to be effective for adult depression(14), and pharmacotherapy (using a medication algorithm based on Kimmel et al., 2018(12)) where considered appropriate. Medication prescriptions followed best practices for perinatal populations, typically in the form of SSRIs shown to be safe during pregnancy and breastfeeding (e.g., sertraline(11)).

The rationale for a brief intervention with referral to a longer-term provider was to mimic real-world experiences of patients seeking specialty care for PND and is in line with the practices

of the UCLA reproductive psychiatry clinic model. Efforts to begin transferring care were encouraged early on in the protocol. Community referrals for longer-term care were provided at the final visit to a psychiatrist, therapist, or primary care provider. These referrals included information about crisis resources, psychiatric, therapeutic, and primary care providers, community support groups for parenting and postpartum support, as well as sleep consultants and financial assistance resources. Throughout treatment, participants completed CAT-MH assessments every other week (without suicidality assessment), but these data were not shared with the treating psychiatrist and did not inform care

B.1 Psychiatry Training. Psychiatry residents conducted all treatment appointments and received supervision by attending psychiatrists, in the context of a training rotation in reproductive psychiatry. Residents were trained by the same reproductive psychiatrist in how to evaluate, diagnose, and treat MDD with peripartum onset, including medication management and provision of supportive psychotherapy(13).

Details on Measures and Statistical Analysis Plan (see primary manuscript text for overview)

Details about CAT-MH development. The CAT-MH was developed by obtaining data on large item banks (e.g., depression over 400 items) which were then administered to large groups of psychiatric outpatients and healthy controls based on a balanced incomplete block design so that each subject received approximately 200 items, maximizing the pairing of each item with every other item in the bank. The data were then calibrated using a multidimensional item response theory model (the bifactor model(15)) and those items with primary dimension loadings in excess of 0.3 were retained. Based on the calibrated item parameters adaptive testing algorithms were developed and then validated against structured clinical interviews (e.g., the SCID for DSM-5). A

unique feature of the CAT-MH is that it can be used to adaptively measure psychopathology (e.g., depression and anxiety) and provide an estimate of the precision or uncertainty in that estimate, which traditional psychiatric rating scales cannot. For example, the CAT-MH Depression Inventory measures depression on a 100-point scale with 5 points of precision. As such, change less than 5 points is not meaningful, representing only noise in the data, whereas change of more than 5 points is meaningful and change of more than 10 points statistically significant(15, 16).

Notably, high correlations have been observed between EPDS scores and CAT-MH severity scores for depression and anxiety (i.e., r=.82 and r=.78, N=358; Kim et al., 2016(17)) and between PHQ-9 scores and CAT-MH depression severity (r=.73, N=373; Wenzel et al., 2021(18)) in perinatal samples. Concordance between the EPDS suicidality item and the CAT-MH suicide severity score was very high (91%; Kim et al., 2016(17)).

Power Analysis. Power was originally tested for purposes of superiority. Power for the comparison between PPC and STAND was based on published effect sizes for the effect of CBT on perinatal depression symptoms. The power analysis considered ranges of effect sizes of STAND compared to PPC centered on .65 (i.e., moderate effect) and included the lower and upper limits of the 95% CI of the effect size estimate from Sockol et al. (2015)(19) (.54 – .76). Based on these estimates, we expected to have sufficient power to detect depressive symptom differences previously reported in the literature in a sample of 120 women.

After enrolling 94 participants, a new digital CBT program was introduced into the STAND condition. We thus generated a secondary aim of evaluating equivalence between the two digital CBT programs within STAND (TWU and DGC PND). A power analysis using an effect size estimated from the participants who were assigned to TWU digital CBT in the first set of participants (n= 30) was conducted to determine the sample size needed for the second set of participants. The power analysis tested bounds of (-1,1) to (-10,10) to indicate the sample sizes

needed to show equivalence to a given point amount on CAT-MH depression severity (the primary outcome). Analyses demonstrated that we could show equivalence for bound values above (-5,5) for sample sizes > 20 and all standard errors explored. A value of 5 points on the CAT-MH depression severity measure is considered to be meaningful within an individual person given that the CAT-MH is designed to have a 5-point uncertainty margin on the 100-point scale(15, 16).

Hence, to be sufficiently powered to compare the two versions of digital CBT, we added enough participants to have at least 20 new participants who were randomized to STAND and assigned the digital DGC PND therapy. Specifically, 72 new randomizations were made, with 36 randomized to PPC and 36 randomized to STAND. Based on the first set of participants, we expected that 75% (n=27) of those randomized to STAND would be allocated to receive the digital therapy with coaching. With an expected 25% attrition, our final sample size receiving DGC PND digital therapy was expected to be 21 women, which would provide sufficient power to assess equivalence between those assigned to TWU digital therapy versus the DGC PND digital therapy. For the tests of equivalence, bounds were then determined based on a difference between the two versions of digital therapy that would be considered clinically equivalent. These were (-1,1) to (-5,5) for CAT-MH depression and CAT-MH anxiety, since, as described above, a value of 5 points is considered to be meaningful within an individual. In addition, we tested bounds of (-1,1) to (-4,4) for EPDS(20) and for SDS(21) since change scores of 4 or more are considered to represent clinically significant change for each of those scales.

With the addition of 72 participants to the initial cohort of 94 participants, our final sample size for the primary analysis comparing STAND to PPC was N=166.

As a secondary analysis, we also conducted tests of equivalence for STAND vs PPC, using the same equivalence bounds: (-1,1) to (-5,5) for CAT-MH outcomes and (-1,1) to (-4,4) for EPDS(20) and SDS(21). We also completed a power analysis to demonstrate that our study

maintained adequate statistical power to detect equivalence between treatment conditions. While equivalence testing (like TOST) does not require a separate power analysis in the same way as traditional hypothesis testing—since achieving statistical significance in the TOST test itself demonstrates equivalence—we conducted this analysis to provide additional transparency about our study's statistical capabilities. We used two-one-sided tests (TOST) to assess our power to demonstrate equivalence in treatment effects, using clinically meaningful equivalence bounds based on prior literature (CAT-MH outcomes: [−1, 1] to [−5, 5]; EPDS and SDS outcomes: [−1, 1] to [−4, 4]). Across all tested equivalence bounds, our study achieved very high statistical power (approaching 1.0) to detect equivalence between treatment conditions.

Clinically Reliable Change Index Calculation. Reliable change on the CAT-MH depression scale involved two steps. In Step 1 for CAT-MH depression, a metric ("SE") was calculated by multiplying the standard deviation of a normalized score at baseline (*SD* = 0.182) by the square root of 1 minus the reliability of the CAT-MH Depression from our data (baseline to week 2) (test-retest reliability = .69). In Step 2, the reliable change metric was calculated by taking the square root of 2 times the square of SE. From these calculations, reliable change in the CAT-MH Depression was determined to be 15.74. Clinically significant change was operationalized as reliable change, meaning a reduction by at least 15.74 points, and a final CAT-MH Depression of 65 or less, which represents mild or less depression severity. Doing this same process with CAT-MH anxiety (baseline to week 2) (test-retest reliability = .63). Clinically significant change was operationalized as reliable change, meaning a reduction by at least 17.10 points and a final CAT-MH Anxiety of 65 or less.

For the EPDS, we calculated the percent who achieved week 26 remission according to EPDS < 10. These indices were calculated for participants with data available at Week 26 (*n* =

87 for CAT-MH, $n = 91$ for EPDS). χ^2 and risk ratio tests were performed to examine potential meaningful difference in outcome likelihood between groups.

Missing Data and Sample Sizes. Participants were included if they had at least one observed outcome measure. Longitudinal mixed-effects models were estimated using maximum likelihood, which incorporates all available observations without imputing missing values; therefore, participants with incomplete data across timepoints were retained. Although we applied intent-to-treat principles by retaining the full randomized sample (N = 166) in our models, estimation was based only on participants with available outcome and covariate data (n = 145) (15 participants did not complete any assessments, and 6 additional participants were missing covariates). Tables S1 shows the outcome variables counts, descriptives and available data points.

Supplementary Results

Protocol Deviation Analyses

The pattern of results was consistent when STAND participants were compared to PPC participants with ≤ 4 psychiatry visits for all outcomes (i.e., excluding protocol deviators).

Two versions of digital CBT

Within STAND, 30 received TWU and 25 received DGC PND digital therapy. Tests of equivalence showed no meaningful differences (according to the effects of time) between the two sets of participants for any measure (all *p*s < .0001): CAT-MH Depression (DGC PND: β = -0.04, *SE* = 0.005, TWU: β = -0.04, *SE* = 0.004); EPDS (DGC PND: β = -0.05, *SE* = 0.009, TWU: β = -0.04, *SE* = 0.008); CAT-MH Anxiety (DGC PND: β = -0.03, *SE* = 0.006, TWU: β = -0.04, *SE* = 0.005); SDS (DGC PND: β = -0.03, *SE* = 0.008; TWU: β = -0.04, *SE* = 0.008).

Treatment Credibility and Expectancies

CEQ scores at 4 weeks averaged 30.98 (*SD* = 10.36, median = 32) for PPC and 32.74 (*SD* = 8.55, median = 33) for STAND, indicative of moderately high credibility and expectancy with no differences between conditions, $t(104) = -0.920$, $p = .360$. At post-treatment, CEQ scores were 39.37 (*SD* = 14.40, median = 44) for PPC and 47.28 (*SD* = 17.41, median = 44.50) for STAND, with no significant differences between conditions, $t(35) = -1.51$, $p = .140$, Cohen's $d = 0.50$.

Table S1

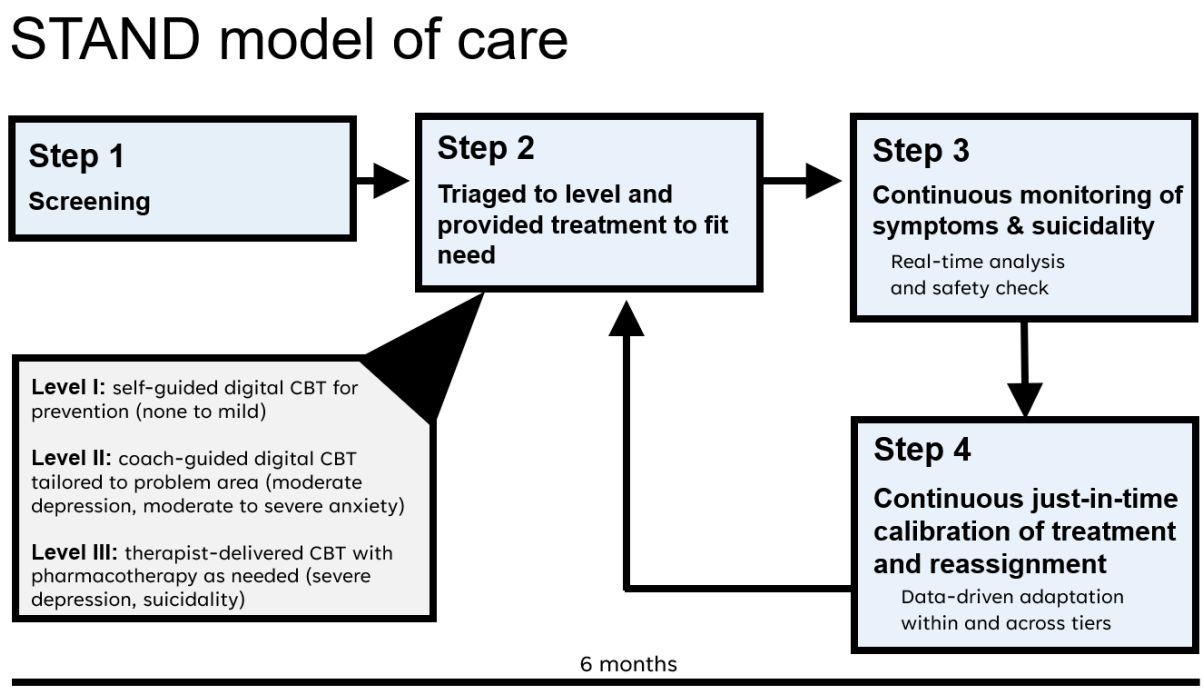
Primary and Secondary Outcome Variable Descriptives and Available Data (N)

	CAT-MH Depression Primary		CAT-MH Anxiety Secondary		EPDS Secondary		SDS Secondary	
Time	**M (SD)**	**N**	**M (SD)**	**N**	**M (SD)**	**N**	**M (SD)**	**N**
BL	60.06 (14.47)	151	54.50 (17.48)	151	16.17 (4.17)	151	16.73 (7.31)	151
Wk 1	56.41 (15.57)	41	48.70 (16.96)	41				
Wk 2	54.64 (15.46)	94	48.84 (19.20)	94				
Wk 3	53.45 (17.24)	39	47.07 (19.39)	39				
Wk 4	52.85 (15.27)	88	44.84 (18.70)	88				
Wk 5	48.29 (17.39)	30	42.03 (20.98)	30				
Wk 6	46.62 (15.80)	84	40.89 (17.78)	84				
Wk 7	46.88 (17.86)	31	41.09 (19.36)	31				
Wk 8	43.87 (18.01)	76	39.74 (18.77)	76				
Wk 9	49.04 (19.09)	28	40.83 (21.82)	28				
Wk 10	43.79 (15.69)	75	37.46 (16.83)	75				
Wk 11	43.64 (20.88)	25	39.86 (19.11)	25				
Wk 12	42.04 (19.57)	68	35.42 (21.28)	68				
Wk 13	41.69 (17.86)	88	33.22 (19.95)	88	10.48 (5.09)	98	10.40 (8.09)	98
Wk 14	38.46 (17.27)	58	30.81 (19.04)	58				
Wk 15	40.70 (17.80)	27	38.38 (19.57)	27				
Wk 16	37.60 (17.80)	69	31.40 (18.26)	69				
Wk 17	34.27 (16.53)	19	33.83 (20.74)	19				
Wk 18	36.69 (17.34)	56	30.66 (18.42)	56				
Wk 19	35.10 (21.97)	26	33.26 (20.95)	26				
Wk 20	35.08 (20.16)	66	29.00 (19.90)	66				
Wk 21	32.97 (15.41)	21	30.62 (15.64)	21				
Wk 22	36.27 (17.89)	64	28.73 (21.15)	64				
Wk 23	32.60 (14.78)	17	33.38 (18.33)	17				
Wk 24	38.31 (22.19)	59	32.77 (24.56)	59				
Wk 25	33.09 (20.44)	20	31.12 (15.35)	20				
Wk 26	31.97 (19.68)	87	24.21 (17.93)	87	7.85 (4.88)	91	7.65 (7.40)	91

Note. CAT-MH Depression = computerized adaptive testing of depression severity; CAT-MH Anxiety = computerized adaptive testing of anxiety severity; EPDS = Edinburgh Postnatal Depression Severity; SDS = Sheehan Disability Scale. Wk = Week of assessment. N=number of data points available at each assessment occasion

Figure S1

STAND (Screening and Treatment for Anxiety and Depression) model of care



References

1. Wolitzky-Taylor K, Richards MC, Welborn A, et al. Study protocol: perinatal mood treatment study. Trials. 2024;25(1):525.
2. Qiu X, Wu Y, Sun Y, et al. Individual participant data meta-analysis to compare EPDS accuracy to detect major depression with and without the self-harm item. Scientific Reports. 2023;13(1):4026.
3. Smith-Nielsen J, Matthey S, Lange T, Væver MS. Validation of the Edinburgh Postnatal Depression Scale against both DSM-5 and ICD-10 diagnostic criteria for depression. BMC Psychiatry. 2018;18(1):393.
4. Mahoney A, Shiner CT, Grierson AB, et al. Online cognitive behaviour therapy for maternal antenatal and postnatal anxiety and depression in routine care. J Affect Disord. 2023;338:121-8.
5. Loughnan SA, Sie A, Hobbs MJ, et al. A randomized controlled trial of 'MUMentum Pregnancy': Internet-delivered cognitive behavioral therapy program for antenatal anxiety and depression. J Affect Disord. 2019;243:381-90.
6. Loughnan SA, Butler C, Sie AA, et al. A randomised controlled trial of 'MUMentum postnatal': Internet-delivered cognitive behavioural therapy for anxiety and depression in postpartum women. Behav Res Ther. 2019;116:94-103.
7. Newby JM, Mewton L, Williams AD, Andrews G. Effectiveness of transdiagnostic Internet cognitive behavioural treatment for mixed anxiety and depression in primary care. J Affect Disord. 2014;165:45-52.
8. O'Mahen HA, Woodford J, McGinley J, et al. Internet-based behavioral activation--treatment for postnatal depression (Netmums): a randomized controlled trial. J Affect Disord. 2013;150(3):814-22.
9. Stein A, Netsi E, Lawrence PJ, et al. Mitigating the effect of persistent postnatal depression on child outcomes through an intervention to treat depression and improve parenting: a randomised controlled trial. Lancet Psychiatry. 2018;5(2):134-44.
10. Wolitzky-Taylor K, LeBeau R, Arnaudova I, et al. A Novel and Integrated Digitally Supported System of Care for Depression and Anxiety: Findings From an Open Trial. JMIR Ment Health. 2023;10:e46200.
11. Kulin NA, Pastuszak A, Sage SR, et al. Pregnancy outcome following maternal use of the new selective serotonin reuptake inhibitors: a prospective controlled multicenter study. Jama. 1998;279(8):609-10.
12. Kimmel MC, Cox E, Schiller C, Gettes E, Meltzer-Brody S. Pharmacologic Treatment of Perinatal Depression. Obstet Gynecol Clin North Am. 2018;45(3):419-40.
13. Markowitz JC, Sacks, M., & Frances, A. . Manual for brief supportive psychotherapy for chronic depression. 2002.
14. Cuijpers P, Driessen E, Hollon SD, van Oppen P, Barth J, Andersson G. The efficacy of non-directive supportive therapy for adult depression: a meta-analysis. Clin Psychol Rev. 2012;32(4):280-91.
15. Gibbons RD, Weiss DJ, Pilkonis PA, et al. Development of a computerized adaptive test for depression. Archives of General Psychiatry. 2012;69(11):1104-12.
16. Gibbons RD, Weiss DJ, Frank E, Kupfer D. Computerized Adaptive Diagnosis and Testing of Mental Health Disorders. Annual Reviews of Clinical Psychology. 2016;12:83-104.
17. Kim JJ, Silver RK, Elue R, et al. The experience of depression, anxiety, and mania among perinatal women. Archives of Women's Mental Health. 2016;19(5):883-90.
18. Wenzel E, Penalver Bernabe B, Dowty S, et al. Using computerised adaptive tests to screen for perinatal depression in underserved women of colour. Evidence-Based Mental Health. 2022;25(1):23-8.

19. Sockol LE. A systematic review of the efficacy of cognitive behavioral therapy for treating and preventing perinatal depression. J Affect Disord. 2015;177:7-21.
20. Matthey S. Calculating clinically significant change in postnatal depression studies using the Edinburgh Postnatal Depression Scale. J Affect Disord. 2004;78(3):269-72.
21. Turkoz I, Alphs L, Singh J, et al. Clinically meaningful changes on depressive symptom measures and patient-reported outcomes in patients with treatment-resistant depression. Acta Psychiatr Scand. 2021;143(3):253-63.