Effectiveness of a peer-delivered, psychosocial intervention on maternal depression and child development at 3 years postnatal: a cluster randomised trial in Pakistan

Effectiveness of a peer-delivered, psychosocial intervention on maternal depression and child development at 3 years of age: a cluster randomized trial in Pakistan

Joanna Maselko ScD[1] and Siham Sikander PhD[2]*, Elizabeth L. Turner PhD[3], Lisa M. Bates ScD[4], Ikhlaq Ahmad MPhil[5], Najia Atif PhD[5], Victoria Baranov PhD[6], Sonia Bhalotra PhD (full professor)[7], Amina Bibi MPH[5], Tayyaba Bibi MS[5], Samina Bilal MS[5], Pietro Biroli PhD[8], Esther Chung MS[1], John A. Gallis MS[3], Ashley Hagaman PhD[9], Anam Jamil BA[5], Katherine LeMasters MS[1], Karen O'Donnell PhD[10], Elissa Scherer BS[1,11], Maria Sharif MSPH[5], Ahmed Waqas MBBS[5], Ahmed Zaidi MS[5], Shaffaq Zulfiqar MS[5], & Atif Rahman PhD (full professor)[12]

1. Department of Epidemiology, Gillings School of Global Public Health, and Carolina Population Center, University of North Carolina, Chapel Hill, NC, USA
2. Human Development Research Foundation (HDRF) and Health Services Academy, Islamabad, Pakistan
3. Department of Biostatistics and Bioinformatics, and Duke Global Health Institute, Duke University, Durham, NC USA
4. Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, USA
5. Human Development Research Foundation (HDRF)
6. Department of Economics, University of Melbourne, Melbourne, Australia
7. Department of Economics, Univesity of Essex, Colchester, UK
8. Department of Economics, University of Zurich, Zurich, Switzerland
9. Department of Social and Behavioral Sciences, Yale School of Public Health, Yale University, New Haven, CT, USA
10. Center for Child and Family Health, Durham, NC USA
11. Research Triangle International, Research Triangle Park, USA
12. Institute of Psychology, Health and Society, University of Liverpool, Liverpool UK

Corresponding Authors
Joanna Maselko, ScD
jmaselko@unc.edu
Dept. of Epidemiology
Carolina Population Center Fellow
2105e McGavran-Greenberg Hall
Campus Box 7435
Gillings School of Global Public Health, UNC
Chapel Hill, NC 27599-7435 USA

Abstract

Background

Maternal depression has a recurring course that can influence offspring outcomes. There is limited evidence about how to treat maternal depression to improve longer term maternal outcomes and reduce intergenerational transmission of psychopathology using task-shifted, low-intensity scalable psychosocial interventions. We sought to fill this gap, evaluating the effects of a peer-delivered psychosocial depression intervention on maternal depression and child development at 3 years of age.

Methods

Forty village clusters in Pakistan were randomly allocated to treatment or enhanced usual care (EUC). Pregnant women aged 18 years or over screening positive for moderate or severe depression symptoms (Patient Health Questionnaire (PHQ-9) score 10+) were recruited into the trial (n=570) and a non-depressed cohort was also enrolled (n=584). Primary outcomes were maternal depression symptoms and remission (PHQ-9<10) and child socioemotional skills (Strengths and Difficulties Questionnaire- Total Difficulties (SDQ-TD) at 36-monrha postnatal. Analyses were intention-to-treat. The trial was registered with ClinicalTrials.gov, number NCT02658994.

Findings

At 36-months postnatal, complete data were available from 889 mother-child dyads: 206 treatment (72.5%), 216 EUC (75.3%), and 467 prenatally non-depressed (80.0%). We did not observe significant outcome differences between treatment and EUC arms of the trial (PHQ-9 total score: Standardized Mean Difference = -0.13, 95% CI -0.33 to 0.07; PHQ-9 remission: RR= 1.08, 95% CI 0.88 to 1.33; SDQ-TD treatment estimate: -0·10; 95%CI -1·39, 1·19;). Approximately 40% of women did not complete their treatment sessions, and competence levels of peers dropped over time.

Interpretation

Reduced symptom severity and high remission rates were seen across both arms, possibly masking any treatment effects. A multi-year, psychosocial intervention can be task-shifted via peers but are susceptible to reductions in fidelity and dosage over time. Early intervention efforts might need to rely on multiple models (e.g. collaborative care), be of greater intensity, and potentially targeted toward higher risk mothers to reduce the intergenerational transmission of psychopathology from mothers to children.

This study was funded by a grant from the NIH R01HD075875, NIMH U19MH95687, and NICHD T32HD007168.

Research in context

Evidence before this study.

Recent systematic reviews of psychotherapy interventions for depression have highlighted the limited evidence on long-term effects of psychotherapy on either maternal mental health or child outcomes. We conducted a search to identify studies designed to evaluate interventions for perinatal depression, whose intervention either lasted beyond 12 months postnatal (e.g. booster sessions), or whose follow-up was more than 12 after the completion of the intervention, in years 2002-April 2020. We limited our search to randomized clinical trials or meta-analyses. We did not place restrictions on language or country. We used Pubmed and Web of Science, with the following search terms: ((maternal depression) or (perinatal depression) or (postpartum depression) or (postpartum depression)) AND ((treatment) or (therapy) or (intervention) or (psychotherapy) or (cognitive behavioral therapy)) AND ((longer-term) or (longer) or (booster)). We identified six RCTs specific to perinatal depression with the longer follow-up period, ranging from 1.5 to 7 years. None utilized an extended duration design (booster sessions) that continued past 12 months postnatal; two studies included a non-depressed comparison group. Most common intervention models were Cognitive Behavioral Therapy (CBT) and Interpersonal Therapy (IPT). Evidence generally showed that interventions improved outcomes which then weakened over time so that, overall, there is limited evidence of intervention effects of perinatal depression interventions that persist beyond the perinatal period. Of the six studies, two reported some lasting impact. One study of 884 mother-child dyads assessed maternal and child outcomes 7 years after the end of a CBT intervention and found a 5 point percent lower rates of depression among mothers who received the intervention, but no significant effects on child outcomes. With this one exception, sample sizes were small, with studies having fewer than 60 participants per group available at follow-up.

Added value of this study.

Given the chronic and recurring nature of depression, longer lasting interventions may be necessary to effectively reduce disease burden and potentially reduce the intergenerational transmission of psychopathology. This study, in rural Pakistan, is the first large multi-year randomized controlled trial, focusing on both maternal and child outcomes, where individuals with depression received psychotherapy beginning prenatally. The extended duration psychosocial intervention evaluated in the current study did not show evidence of meaningfully reducing depression symptom levels, nor improving child outcomes, at the 3 years postnatal mark. These findings highlight the challenges of implementing a peer-delivered psychosocial intervention over a longer period in a low resource community setting.

Implications of all the available evidence. These findings point to several implementation lessons for such task-shifted, low-intensity interventions when delivered at scale alongside existing health systems. These include importance of ensuring high levels of fidelity of the intervention, potentially through use of technology platforms. It is also important that any intervention be situated within a collaborative care model that can help detect and respond to women in need of other services to help social determinants like poverty and domestic violence or pharmacological interventions.

Introduction

Global prevalence estimates of depression in the perinatal period range from 4% to over 50%, with the highest burden in low-resource settings, making depression a public health priority.[1] In addition to the effect of maternal depression on the woman's functioning, physical health, and risk of suicide, observational evidence suggests that maternal depression is associated with higher risk of multiple negative child outcomes, including stunting, socioemotional difficulties, problems with school readiness and performance, and depressive symptoms over their lifecourse.[1,2] Women experiencing perinatal depression are at much higher risk of subsequent or recurrent episodes of depression and this chronic or episodic depression is most deleterious for numerous maternal and child outcomes.[3,4] This risk of intergenerational transmission of psychopathology is most heavily borne by poorer families and those in low resource settings with limited access to quality healthcare, thus exacerbating economic and social inequality.[3]

Because of the lack of specialists in many LMIC settings, task shifting for maternal depression is necessitated to bridge the treatment gap. Evidence-based, task-shifted, targeted maternal depression or universal psychosocial interventions can be delivered through community health workers as well as lay peers.[5,6] However, most of these interventions are delivered either during pregnancy or in the early postnatal months, focusing on the acute phase of maternal depression, without tackling issues of recurrence and chronicity. To our knowledge no depression interventions that begin prenatally are designed specifically to prevent recurrence after the perinatal period. Hence, the extent to which such interventions can break the cycle of recurrence of depression beyond the first postnatal year remains unknown.

While interventions have demonstrated efficacious reductions in shorter term (i.e. 12 months or less) maternal depression and improved maternal behaviours,[7] we do not know whether such interventions can reduce intergenerational transmission to children. Many depression interventions in the perinatal period include a child development component, opening the possibility that such depression interventions, including the one studied here, may affect child outcomes through pathways that are independent of changes in depression symptoms themselves.[5] While maternal depression interventions have been shown to improve key parenting practices,[8] evidence of long-term effects on child socioemotional development is scarce.[9] Studies showing improved child outcomes have short post-intervention follow-up periods, typically less than 12 months,[10-12] leaving uncertainty about longer lasting program impacts. The few studies with follow-up longer than one year have reported mixed or even incongruent effects.[4,6,13,14] For example, analysis of the subset of women who were depressed when beginning the Philani+ program in South Africa, which broadly focused on improving child outcomes and lasted through 6 months postnatally, showed improved child physical and cognitive outcomes at 18 months but higher levels of aggression at 5 years of age.[6,15] The challenges of differential attrition in longer-term follow-ups, diminishing sample sizes, and heterogeneous responses among particular sub-groups (such as those exposed to poverty or intimate partner violence) make clear conclusions difficult.[2,14]

The Thinking Healthy Program, Peer-delivered (THPP), delivered individual and group sessions from pregnancy to 6 months postnatal and has been evaluated through two randomized controlled trials, one in Pakistan and one in India.[16-18] Although the country specific findings were weak, the pooled analyses of these trials showed greater recovery from perinatal depression among the intervention group at 6 months postnatal. It also showed that delivering this psychosocial intervention through peers was a cost-effective, feasible and acceptable approach.[16]

We evaluate a 3-year, task-shifted psychosocial peer-delivered intervention for maternal depression, Thinking Healthy Program, Peer-delivered Plus (THPP+),[19] that followed up on the THPP. The project is located in rural Pakistan, a low resource context characterized by high levels of maternal depression and limited access to clinical mental healthcare.[20]

Although our hypothesis was that the children in the intervention arm will be less high risk (as compared to those in the control arm), the full impact of the intervention can only be discerned if we know the level of *excess* risk remaining- that is, the difference between the reduced level of risk among children (of prenatally depressed mothers) in the intervention arm and the risk among children whose mothers were not depressed. If outcomes of these two groups are comparable, we can infer that the intervention is capable of preventing the intergenerational transmission of risk. To achieve this we gathered data on women who were not depressed in pregnancy.The resulting pregnancy-birth cohort of both prenatally depressed (trial participants) and non-depressed women is referred to as the Bachpan Cohort (*Bachpan* means Childhood in the local Urdu language). Finally, we examined whether intervention effects differ by key social contextual factors such as socioeconomic status and intimate partner violence.

Methods

Study design and participants

We conducted a stratified cluster-randomized controlled trial in Kallar Syedan, a rural subdistrict of Rawalpindi, Pakistan. The sub-district is a socioeconomically deprived area with poverty rates around 50%, female literacy of 40-45%, and a high fertility rate (3·8 births per woman).[21] It is primarily agrarian with close knit communities co-residing in large households (average 6·2 people per household). The sub-district has 11 Union Councils (UC), the smallest administrative unit, each with a population of about 22,000-25,000. Each UC is serviced by a Primary Health Care Facility which houses a physician, midwife, vaccinator, dispenser, and village-based Lady Health Workers (LHWs).

This trial maintained the original cluster criterion, randomization sequence and procedures under the previous trial.[16,22] Pregnant women in the 3rd trimester, aged 18+ years and registered with their LHWs, were eligible. Approximately 95% of the women in the study area were registered with LHWs. All pregnant women were approached by trained research staff either within the pregnant woman's residence or that of their LHW and, if they consented, were assessed. Women who needed immediate medical or psychiatric inpatient care were excluded from the study. All eligible women were invited to be screened for depression using the Urdu version of the Patient Health Questionnaire 9 (PHQ-9), which has been used extensively as a screening tool in the study setting and has an acceptable criterion validity and reliability for this population.[23] Women screening positive (PHQ-9 score $\geq$ 10) were eligible for enrolment into the trial and follow-up as part of the Bachpan Cohort.[16] One out of every three women who screened with a < 10 score on PHQ were enrolled to participate in the Bachpan Cohort only, resulting in a roughly equal size of prenatally non-depressed and depressed women at the beginning the cohort.[24]

Randomization and masking

The trial was conducted at a sub-district level. Forty village-clusters (population of 2,400 to 3,600) were the unit of randomization and were geographically separate to minimize contamination risk. The sub-district is administratively subdivided into 11 Union Councils (as explained above and, within each of

these 11 union councils, we identified an even number of village clusters to ensure that equal numbers of clusters are randomized into intervention or control condition (ie 1:1 ratio) by an independent statistician using a computerized randomization sequence. Research teams responsible for identifying, obtaining consent and recruiting trial participants were blind to the allocation status. The trial PI, site PIs/coordinators, trial statisticians, and members of the Trial Steering Committee were blinded to the allocation status until the analysis of the six-month data for the initial THPP trial.[25]

The Thinking Healthy Program, Peer-delivered Plus (THPP+) Intervention

The intervention arm received the longer-duration peer-delivered psychosocial intervention (THPP+). It consisted of 18 group-based "booster" sessions (from 7th to 36th month postnatal). Of these, the first 6 sessions were delivered monthly, then bi-monthly until 36 months. These sessions built on the shorter duration intervention and were delivered by the same peers. The peers were lay married women who lived in the same community as that of the depressed women and volunteered their time.

The key features of this psychosocial intervention, delivered by non-specialists, were peer-support, behavioural activation, and problem-solving in a culturally sensitive, non-medicalized format, and developmental activities for children up to the 36th month (See Appendix p.1-3 for the overall structure of the intervention and peer characteristics.[19] A cascaded model of training and supervision was used which included frequent competence assessments.[26] A competency checklist was used for supervision, including direct observation of the peers, and to provide feedback. Each peer received this assessment six times over the course of the program. During the feedback meetings, the supervisor discussed checklist information along with session logs maintained by the peers. Competence was assessed through observations and the checklist captured whether content was delivered as intended. Refresher trainings were done 6 and 18 months after the initial training. The two-day training included re-orientation to the intervention and its principles, as well as training on materials and use of job aids.

The intervention group sessions provided a safe environment for women to voice their problems, share experiences of childcare, and provide support to women. Peer volunteers were trained to use culturally grounded vignettes that served as tools to deliver health and well-being messages. The sessions aimed for maternal well-being but also child-care and development by encouraging mother-infant interaction and play. The intervention provided examples of age-appropriate activities, derived from the UNICEF/WHO's Care for Development Package and encouraged demonstration of these activities during the sessions. While these 'booster' group sessions did not focus on specific strategies to address depression, the peer could still draw on her prior knowledge and skills of specific psychotherapeutic elements such as behavioural activation when required. Additional details of the intervention are reported elsewhere.[19,26] We defined overall treatment completion as attending 10 (out of 14) sessions from pregnancy through 6 months postnatal (Phase 1: THPP, individual and group sessions) and 12 (out of 18) sessions from 7 to 36 months postnatal (Phase 2: THPP+, booster group sessions).

Both intervention and control arms received Enhanced Usual Care (EUC). No treatment was offered to the prenatally non-depressed women in either of the arms. EUC consisted of informing participants about their depression status and ways to seek help for it, informing their respective LHWs about women's depression status at enrollment, training all the 11 primary care facility-based physicians in the subdistrict on the mental health Gap Action Programme (mhGAP) treatment guidelines for maternal

depression,[27] and providing depressed participants with a leaflet on how and where to seek appropriate health care during pregnancy and beyond.

Procedures

Assessments were conducted 6 times over the course of the study (in pregnancy, and 3, 6, 12, 24, and 36 months postnatal). As originally specified, the current analysis utilizes mother and child outcome data at 36 months.[22]All measures were extensively piloted. Assessments were done at the community level within households of women by trained interviewers blind to the allocation status, all questions were interviewer administered. Assessors inter-rater reliability was ensured through classroom-based training which included role plays, followed by field practice sessions. During these sessions each pair of assessors assessed up to 10 participants jointly and discussed their coding on each item to establish inter-rater reliability prior to start of actual data collection.

The project received approval for the IRBs of Human Development Research Foundation (HDRF), Duke University, and University of North Carolina. The study protocol for the effectiveness trial of THPP+ and inclusion of prenatally non-depressed pregnant women in the Bachpan Cohort study has been published previously.[16,22]

Outcomes

The primary maternal outcome was depression symptoms assessed using the Patient Health Questionnaire-9 (PHQ-9) and analyzed as symptom severity (total score) and remission (score < 10). The secondary maternal outcomes were disability, assessed using WHO's Disability Assessment Schedule, WHO-DAS[28], and current major depressive episode based on the *Structured Clinical Interview for DSM-IV (SCID) Disor*ders.[29] Since it provides a clinically salient diagnostic outcome, the SCID was also included to increase measurement robustness. Process data regarding THPP+ sessions attended, competence scores of peers, duration of sessions, and peers' supervision attendance are described elsewhere[30] and in the appendix (p.4-8).

The primary child outcome was child socioemotional development measured using the Strengths and Difficulties Questionnaire, Total Difficulties (SDQ-TD) score. The SDQ is a parent-reported measure of 25 child attributes with five subscales: emotional symptoms, conduct problems, hyperactivity, peer problems, and prosocial behavior.[31] The Total Difficulties (TD) score is calculated based on four subscales (omitting prosocial behaviour) with a score range of 0-40 points. The SDQ is widely used in low- and middle-income countries and has been translated into Urdu[32]; in our sample, internal consistency measured by Cronbach's alpha was 0.78.

The secondary child outcomes were two developmental domains. Given language differences, two subscales from the Bayley Scales of Infant and Toddler Development, 3rd edition (BSITD III) were selected to assess achievement of developmental milestones, the Receptive Language and Fine Motor subscales.[33] The BSITD was administered in the family's home; scaled scores were calculated using the child's chronological age. The BSITD has been widely used and validated internationally.[34]

An additional outcome of interest, child growth, was analysed using weight- and length-for-age Z-scores. Additional variables included demographic and psychosocial factors hypothesized to moderate the effect of the intervention. These included household assets as an indicator of socioeconomic status,[35]

maternal education (coded as none vs. any), household composition (nuclear family status), intimate partner violence (IPV) in the previous 12 months, child gender, maternal age (18-24 vs. 25+), number of siblings (0 vs. 1+), treatment expectations (very/moderately useful vs. somewhat/not useful), depression chronicity (<12 weeks vs. ≥12 weeks), and depression severity (PHQ-9 10-14 vs. 15+).[23,36] We collected information on a number of domains that may have been differentially distributed across the treatment and EUC arms or correlated with loss to follow-up (table 1 and Appendix p.9-14).

Statistical analyses

For mother outcomes, anticipating a sample of 480 prenatally depressed women, we were powered at 90% to detect a remission rate of 65% in the prenatally depressed-intervention versus 45% in prenatally depressed-control at the two-tailed 5% significance level and assuming a conservative ICC of 0.07 in intervention arm and 0.05 in control.[22] For child outcomes, with this sample size, we were powered at 90% to detect a difference between treatment arms of 3 points on the SDQ-TD at the two-tailed 5% significance level assuming a standard deviation of 5.2 points and ICC of 0.04-0.08.[22]

Statistical analyses were done according to the CONSORT guidelines in Stata software version 16.1 (StataCorp, College Station, TX) and SPSS. All analyses compare the three groups (prenatally depressed in intervention, prenatally depressed in control, and prenatally non-depressed) across the two arms, using the 36-month outcomes. We had pre-specified a comparison of outcomes between intervention and control arms within the prenatally non-depressed women and, in the absence of such an effect, present results for the overall prenatally non-depressed cohort.

The primary analyses were designed as intention-to-treat. Data from prenatally depressed and non-depressed participants were analyzed jointly using linear mixed effects models so that all comparisons of interest could be estimated from the same model. The identity link was used for continuous outcomes to estimate differences in mean outcomes. Standardized mean differences (SMDs) and their 95% confidence intervals (CIs) were obtained using the method of Hedges.[37] In the primary model, we included a random intercept for cluster and fixed effects for treatment arm (depressed intervention, depressed control, non-depressed), union council (11 levels; the stratification variable), and variables found to be imbalanced by loss to follow-up or at baseline (determined using $p<0.10$; see Appendix p.9-14). Mixed models assume missing at random conditional on the covariates included in the statistical model. Therefore, we include variables predictive of loss to follow-up to account for missing data. We used restricted maximum likelihood (REML). The between-within method was used to apply small-sample bias corrections to the intervention effect standard errors in the mixed effects framework.[38] These models also generated the intra-cluster correlation (ICC) values.

All binary outcomes were analyzed using generalized estimating equations (GEE). As with the continuous outcomes, we include as fixed effects treatment arm and union council, as well as fixed effects for any variables found to be imbalanced by loss to follow-up or at baseline. In the GEE framework, we took into account clustering using an exchangeable working correlation matrix. We used a modified Poisson approach[39] and Kauermann-Carroll bias-corrected standard errors to account for the relatively small number of clusters (i.e. 40).[40] When analyzing SCID major depressive episode, we included SCID at all time points as the outcome in a GEE model with exchangeable working correlation for village cluster. Additional analyses focus on *a priori* identified potential moderators of any main associations with the primary outcomes (described above). We tested for moderation of the intervention effect by including these potential moderators in the model as individual interaction terms.

ClinicalTrials.gov Identifier: NCT02658994 (registered on 21 January 2016).

Role of Funding Source

The Funding Source (NIH) had no influence in study design, data collection, data analysis, data interpretation, or writing of the report.The corresponding authors had full access to all of the data and the final responsibility to submit for publication.

Results

From Oct 15, 2014 to Feb 25, 2016, we identified and randomly selected 40 village clusters out of 46 and randomly assigned 20 village clusters each to intervention (THPP+) and control (EUC) arms. In all we approached 1910 pregnant women; 287 prenatally depressed women in the control arm and 283 in the intervention arm completed the baseline questionnaire. Of the prenatally non-depressed women approached, 584 were enrolled, yielding comparable numbers of prenatally depressed and prenatally non-depressed women in each of the 40 village clusters.
Of the 1,154 participants enrolled at baseline, 889 (77.0%) were successfully interviewed at 36 months: 206 (72.5%) intervention, 216 (75.3%) EUC, and 467 (80.0%) prenatally non-depressed (Figure 1). There was no differential loss to follow up by treatment arm and no differences in adverse events (Appendix p.9-14).

At baseline, the mean age of the women in the sample was 26.7 (SD: 4.5) years, with 30.2% of women being in their first pregnancy. The mean PHQ-9 scores across the treatment arms were similar (14.5 control and 14.9 intervention) with a mean score of 2.8 among prenatally non-depressed women. Further baseline characteristics are summarized in Table 1. At 36 months, 49.6% of the infants were girls.

Baseline variables that were imbalanced between intervention and control groups include the life events checklist score with a higher mean score in the intervention arm. Other variables were found to be imbalanced but were not included because of collinearity or conceptual overlap (e.g. subjective religiosity). Additional baseline demographic variables that were associated with loss to follow-up at 36 months and adjusted for include number of people per room, child's grandmother living with him/her, nuclear family status, number of living children and the asset score.

Just over two-thirds of prenatally depressed women in the intervention arm completed the THPP+ intervention (Appendix p.5). Only 63% of women completed treatment during Phase 2 compared to nearly 80% treatment completion in Phase 1. The competence levels of the peers declined over the implementation period, particularly in the period after time point 3 (i.e. at 12 months postnatal) where we introduced new content for the booster sessions and the frequency of peer supervisions dropped from monthly to every two months (Appendix p.6-8) show high and low scoring peers and their competence levels over time

For all prenatally depressed women, depression scores dropped meaningfully, regardless of their arm allocation (Figure 2). There were no significant differences in depression outcomes between arms (THPP+ vs EUC) at 36 months postnatal (Table 2). The adjusted standardized mean difference (SMD) in depressive symptom severity (PHQ-9) between arms was -0.13 (95% CI -0.33 to 0.07) and the risk ratio (RR) for depression remission (PHQ-9<10) was 1.00 (95% CI 1.13 to 0.97). Turning to the secondary outcomes, we observed a relatively larger difference between arms on the secondary outcomes of SCID-based major depressive episode at 36 months (22% control vs 16.5% intervention) (RR=0.67, 95% CI 0.43

to 1.05). The intervention effects on disability (SMD=-0.12, 95% CI -0.33 to 0.09) were not statistically significant.

The prevalence of major depressive episodes (SCID-based) among the three groups of women (the prenatally depressed intervention arm, prenatally depressed control arm, and prenatally non-depressed women) became increasingly similar in the proportion depressed by 36 months postnatal (Figure 2). The THPP+ intervention arm showed higher convergence with the prenatally non-depressed women at 36 months compared to the control arm, so much so that there was not a statisticaly significant difference in the probability of being depressed (using SCID) between the intervention arm women (at 16.5%) and the prenatally non-depressed (9.0% of whom were depressed at 36 months, RR 0·74 (0·46 to 1·18).

We also examined intervention arm differences in depression severity (PHQ-9 scores) at 36 months postnatal by potential moderators assessed at baseline. There was no strong evidence of meaningful moderation of the intervention effect by any of these factors (Appendix p.18).

For the child primary outcome of SDQ-TD, the mean adjusted difference between intervention and control arms was -0·10; 95% CI -1·39 to 1·19 (Table 2). There were also no meaningful differences between the two arms in the secondary outcomes of Receptive Language and Fine Motor scores (from BSITD) (Table 2),

Similar to maternal depression results, we do not find strong support for the hypothesis that baseline characteristics moderated treatment effect on the SDQ-TD scores (Appendix p.19).

Children of prenatally non-depressed mothers had somewhat better SDQ-TD scores than the intervention or control arm children (e.g. SDQ-TD=13.7 among the children of the prenatally non-depressed vs. 14.7 in the intervention arm, p-value=0.07, Table 2). In other words, the prenatal depression episode predicted slightly worse SDQ-TD scores at 36 months of age, independent of the intervention. Exploratory analyses of the five sub-scales of the SDQ separately showed that this overall difference was driven by the hyperactivity and conduct problems sub-scales, with negligible differences by prenatal depression status for the peer problems, emotional problems, and the pro-social scales (Appendix p.19). As an example, the adjusted mean difference in the hyperactivity scores between the non-depressed and the control arm was 0·31 (95% CI -0·62, 0·01,). The receptive language, fine motor and physical growth indicators did not meaningfully differ between the children of prenatally non-depressed and depressed mothers, regardless of intervention arm.

Discussion

Our study showed that a peer-delivered intervention beginning in pregnancy with booster sessions through 36-months postnatal did not measurably affect a range of maternal depression symptom and child developmental outcomes. Though women in theintervention arm did show greater convergence in depression symptoms with the prenatally non-depressed women at 36 months, relative to women in the control arm, evidence of a meaningful intervention effect is lacking. We also find only weak evidence that the prenatal depression episode was itself associated with child socioemotional outcomes and no

evidence of associations with other developmental outcomes; for the most part, children of prenatally depressed and prenatally non-depressed mothers had similar outcomes at 36-months of age.

The overall small effect sizes and lack of statistical significance on maternal outcomes might be attributable to several factors. First, the intervention was a non-specialist, lay peer delivered psychosocial intervention. The lay peers were housewives from rural villages, without prior training or work experience. They were trained and supervised by non-specialists using a cascaded model (with no direct specialist contact).[26] Second, this lay peer delivered approach was used to inform scaling-up of maternal mental health services through existing health systems and community resources. It is possible that this non-specialist, low-intensity design, coupled with longer-duration implementation, led to what Chambers and colleagues refer to as 'voltage drop' (the intervention loses some degree of its potency (or fidelity) when moving from efficacy to effectiveness in the real world) and 'program drift' (the intervention deviates from its manualized or implementation protocols).[41] We saw a substantial drop in women attending the maintenance group sessions delivered every two months beyond the 6th month postnatal period. This implementation challenge of attendance is reported in other community based interventions targeting maternal outcomes.[42] Women reported that they lost interest in attending these group sessions or found it demanding on their time. This challenge of poor attendance and how to best address it came up regularly in peer supervision meetings. In addition to attempts to add more interesting content to the sessions, we made sure that community health workers reminded women about the sessions and followed up with those who missed a session. Addressing sustaining participant interest and limited time is seminal in community-based program success and has been highlighted as an important challenge in other low resource settings.[43] Finally, it is possible that, since the booster group sessions did not focus on specific strategies to address depression as mentioned earlier, the intervention arm was not so different from the EUC arm. There were no detectable treatment effects at the 6 month postnatal wave.[16]The boosters introduced after the 6 month mark did not change this - we continue to see no intervention effect at 36 months.

Maintaining the competence levels and motivation of the peers over the multi-year long implementation period was challenging. Perhaps the cascaded model of supervision via non-specialists led to dropped potency (or fidelity) of sessions. We explored the perceptions of the peers about this intervention in a nested qualitative study at six months postnatal.[26] We did not find any negative perceptions towards the intervention which might have led to change in motivation or competence levels.. The reduced frequency of supervision sessions of the peers from every month to every two months seems to have contributed to the drop in competence levels. This drop in supervisory intensity has been shown to reduce effectiveness of known approaches.[44,45] Another contributing factor to the drop in competence levels was the addition of new content, beyond the 6th month postnatal period: both the high and low scoring peers experienced a drop in competence levels when more content was added. Finding the right balance of content vs capacity of peers, and maintaining fidelity, is an implementation challenge reported in other programs.[46,47]

Finally, we saw a substantial improvement in the control arm (lowered rates and higher recovery from depression). The control arm recovery rates were higher compared to our previous studies from 2008 and 2020.[8,48] This could be attributable to "regression to the mean" or to spillover of the intervention through the LHWs who regularly interact with LHWs responsible for women residing in control arm sites. The active control arm received enhanced usual care and some evidence suggests that informing people

about their illness status improves outcomes,[49] which raises important methodological (study design) issues for trials that are embedded within community settings. Perhaps future trials, using similar active control arms ought to consider equivalence or non-inferiority trial designs to avoid a nonsignificant superiority trial being wrongly interpreted, as proof of no difference between the two active comparisons.[50]

We found no clear indications of sub-group differences. This points to the challenge of intervention targeting especially given that approximately one-fifth of women did not respond to this treatment, indicating a need for different interventions. A collaborative care model where non-responders can be detected earlier and connected to more specialized care may be needed.

The findings suggesting an association (albeit weak) between prenatal depression and child socioemotional outcomes at age 3, which is not mitigated by the intervention, mirrors results from our previous trial with a different sample, designed to examine a similar intervention's effect on children at age seven.[4] Children's socioemotional development may be less likely to 'catch up' during the resolution phase of depression within the postnatal period. If confirmed in other studies, differences in socioemotional, but not other developmental outcomes, may point to specific mechanisms in maternal-child intergenerational transmission of risk.[11]

While chronic depression likely has the greatest effect on child development,[51] the majority of children in our study were exposed to varying maternal depression levels, including periods of low or no depression symptoms over the study period. An association between the prenatal episode and child socioemotional outcomes at 36 months would be consistent with 'foetal programming,'[52] and is supported by evidence of linking prenatal depression to a number of child outcomes regardless of postnatal depression.[53] However, inconsistent with foetal programming, another study concluded that a reduction in maternal depression levels postnatally lead to 'near normal levels' of child behavioural problems.[54]

A possible explanation for the lack of intervention effects on child outcomes is that, unlike the trained specialist delivery in Stein et al's study, this intervention was delivered by lay peers in a low-income country.[54] As mentioned above, we experienced a 'voltage drop' over time in terms of reductions in both session frequency and attendance, posing significant challenges to sustained implementation.[41] Another potential explanation of no group differences in socioemotional outcomes is that the variability in maternal depression symptoms in the first 3 years could itself be a risk factor. Given that depression is a chronic, recurring disorder, mothers who were prenatally depressed likely recovered and had a recurrence. Variation in symptoms over time would be larger among women who entered the study depressed compared to those non-depressed. It is possible that these children were exposed to inconsistent or unpredictable parenting, which has been linked with negative behavioral and socioemotional indicators.[55] This hypothesis is supported by results from prior work, showing a tendency toward worse child anxiety symptoms among children whose mothers relapsed when compared with those who had chronic depression.[4]

Our findings that receptive language, fine motor, and length- and weight-for-age Z scores were not a function of maternal prenatal depression status or whether she was treated with the intervention are consistent with our previous study and also with Tomlinson et al who did not observe intervention effects when looking at the continuous versions of the Bayley Scales and growth outcomes, although they did report a difference when outcomes were dichotomized.[4,6] These results suggest that a different set of pathways operates for these outcomes relative to socioemotional development. It is possible that the child's experience of maternal depression did not, on average, reach severe or chronic enough levels to affect language and fine motor development and growth. The possibility is consistent with prior literature that suggests that the most deleterious effects on child outcomes are from severe and persistent depression levels.[51] The presence of other family members, such as the grandmother, might also buffer any negative impacts of depression on the child. A complementary possibility is that the effects of maternal depression on these child outcomes are a function of baseline levels of other risk factors, such as illness, low maternal literacy, intimate partner violence, and others. In samples such as ours where a large fraction of women carry these risk factors, the effects of depression *per se* on child development may be overshadowed.

The study has limitations. The indicators of child socioemotional outcomes, although measured with validated and extensively used instruments, were mother-reported and, as such, susceptible to bias.We addressed this by utilizing only validated instruments used extensively globally. Additionally, if depression symptoms biased reporting, we would expect this to affect the overall socioemotional domain. However, in exploratory analyses we found that only specific socioemotional domains were predicted by prenatal depression; we have no reason to believe that reports of hyperactivity or conduct problems would be more influenced by depression than reporting of peer or emotional problems.

Overall, our results on the lack of intervention effectiveness for maternal depression and child development at 36 months suggest three potential scenarios. First, specific to the socioemotional domains, the intervention may have been too "light touch" to reverse the effect of prenatal depression exposure. The peer delivered version of this psychosocial intervention had a weaker effect on maternal depression than found in a previous study where community health workers were used to deliver the intervention instead of peer volunteers.[16,17] It is likely that women at risk of depression (and their children) need more than bimonthly group sessions delivered by peers for sustained changes to occur that will reduce depression's effect on children. A different model of delivery such as collaborative, or stepped care, merit consideration, with a number of interventions simultaneously targeting specific population needs, e.g. domestic violence, poverty alleviation, social services, social health protection Second, single screening in pregnancy for elevated depression symptoms may not be sufficient to identify the highest risk women, those who will go on to have the most chronic depression trajectories, with the worst effects on their children. Targeting women with a history of depression in combination with other risk factors, such as poverty or IPV, and tailoring the intervention for them, may yield stronger results. Finally, the program was delivered in a high poverty context. It thus remains possible that if the entire socio-political environment were to prioritize women and children's wellbeing and health, interventions such as ours would have more power to make a difference.

In sum, our findings suggest that prenatal depression may have persistent effects on the child's socioemotional skills that are not easily reversed by a psychosocial intervention. Future preventive and early intervention efforts might benefit from being higher intensity and target the highest risk mothers. Importantly, interventions need to be attuned to the social context and ideally implemented as part of a suite of health promoting policies that address social determinants of maternal and child health.

Data Sharing Statement

Per NIH guidelines, data from the project supported by the NIH will be made available 2 years after the end of the project, which will be May 2022.

Authors' contributions

JM and SS are joint first authors* and drafted the manuscript, to which all authors contributed extensively. All authors reviewed and approved the final version. JM, SS, LB, ET, AR, KO designed the study. IA, NA, SS, KO led the implementation of the fieldwork and instrument development/adaptation.

ET, JG designed the analytical database, planned and led the analysis. AZ designed and managed the data collection databases throughout the data collection process.VB, SoBh, PB, EC, AH, KL, and ES contributed to the analysis of the data. JM, SS, ET, LB, IA, AW, NA, AB, SZ, JG, VB, SB, PB, TB, SB, AH, AJ, MS, AZ, EC, KL, ES, KO, AR had full access to all the data in the study and had final responsibility for the decision to submit for publication

Conflict of Interest

We declare we have no conflict of interest.

References

1. Gelaye B, Rondon MB, Araya R, Williams MA. Epidemiology of maternal depression, risk factors, and child outcomes in low-income and middle-income countries. *The Lancet Psychiatry* 2016; **3**(10): 973-82.10.1016/S2215-0366(16)30284-X
2. Goodman SH, Rouse MH, Connell AM, Broth MR, Hall CM, Heyward D. Maternal Depression and Child Psychopathology: A Meta-Analytic Review. *Clin Child Fam Psychol Rev* 2011; **14**(1): 1-27.10.1007/s10567-010-0080-1
3. Patel V, Saxena S, Lund C, Thornicroft G, Baingana F, Bolton P, Chisholm D, Collins PY, Cooper JL, Eaton J, Herrman H, Herzallah MM, Huang Y, Jordans MJD, Kleinman A, Medina-Mora ME, Morgan E, Niaz U, Omigbodun O, Prince M, Rahman A, Saraceno B, Sarkar BK, De Silva M, Singh I, Stein DJ, Sunkel C, UnÜtzer J. The Lancet Commission on global mental health and sustainable development. *The Lancet* 2018; **392**(10157): 1553-98.10.1016/S0140-6736(18)31612-X
4. Maselko J, Sikander S, Bhalotra S, Bangash O, Ganga N, Mukherjee S, Egger H, Franz L, Bibi A, Liaqat R, Kanwal M, Abbasi T, Noor M, Ameen N, Rahman A. Effect of an early perinatal depression intervention on long-term child development outcomes: follow-up of the Thinking Healthy Programme randomised controlled trial. *Lancet Psychiatry* 2015; **2**(7): 609-17.10.1016/s2215-0366(15)00109-1
5. Rahman A, Fisher J, Bower P, Luchters S, Tran T, Yasamy MT, Saxena S, Waheed W. Interventions for common perinatal mental disorders in women in low- and middle-income countries: a systematic review and meta-analysis. *Bulletin of the World Health Organization* 2013; **91**(8): 593-601I.10.2471/BLT.12.109819
6. Tomlinson M, Rotheram-Borus MJ, Scheffler A, le Roux I. Antenatal depressed mood and child cognitive and physical growth at 18-months in South Africa: a cluster randomised controlled trial of home visiting by community health workers. *Epidemiol Psychiatr Sci* 2018; **27**(6): 601-10.10.1017/s2045796017000257
7. Goodman SH, Garber J. Evidence-Based Interventions for Depressed Mothers and Their Young Children. *Child Dev* 2017; **88**(2): 368-77.10.1111/cdev.12732
8. Baranov V, Bhalotra S, Biroli P, Maselko J. Maternal Depression, Women's Empowerment, and Parental Investment: Evidence from a Randomized Control Trial. *American Economic Review* 2020; **110**(3): 824-59
9. Gunlicks ML, Weissman MM. Change in child psychopathology with improvement in parental depression: A systematic review. *Journal of the American Academy of Child and Adolescent Psychiatry* 2008; **47**(4): 379-89.10.1097/CHI.0b013e3181640805
10. Cuijpers P, Weitz E, Karyotaki E, Garber J, Andersson G. The effects of psychological treatment of maternal depression on children and parental functioning: a meta-analysis. *Eur Child Adolesc Psych* 2015; **24**(2): 237-45.10.1007/s00787-014-0660-6
11. Goodman SH, Cullum KA, Dimidjian S, River LM, Kim CY. Opening windows of opportunities: Evidence for interventions to prevent or treat depression in pregnant women being associated with changes in offspring's developmental trajectories of psychopathology risk. *Development and Psychopathology* 2018; **30**(3): 1179-96.10.1017/S0954579418000536
12. Loechner J, Starman K, Galuschka K, Tamm J, Schulte-Körne G, Rubel J, Platt B. Preventing depression in the offspring of parents with depression: A systematic review and meta-analysis of randomized controlled trials. *Clinical Psychology Review* 2018; **60**: 1-14.https://doi.org/10.1016/j.cpr.2017.11.009
13. Milgrom J, Holt CJ, Bleker LS, Holt C, Ross J, Ericksen J, Glover V, O'Donnell KJ, de Rooij SR, Gemmill AW. Maternal antenatal mood and child development: an exploratory study of treatment effects on child outcomes up to 5 years. *J Dev Orig Health Dis* 2019; **10**(2): 221-31.10.1017/s2040174418000739

14. Kersten-Alvarez LE, Hosman CMH, Riksen-Walraven JM, van Doesum KTM, Hoefnagels C. Long-term effects of a home-visiting intervention for depressed mothers and their infants. *Journal of Child Psychology and Psychiatry* 2010; **51**(10): 1160-70.10.1111/j.1469-7610.2010.02268.x
15. Rotheram-Borus MJ, Arfer KB, Christodoulou J, Comulada WS, Stewart J, Tubert JE, Tomlinson M. The association of maternal alcohol use and paraprofessional home visiting with children's health: A randomized controlled trial. *J Consult Clin Psychol* 2019; **87**(6): 551-62.10.1037/ccp0000408PMC6775769
16. Sikander S, Ahmad I, Atif N, Zaidi A, Vanobberghen F, Weiss HA, Nisar A, Tabana H, Ain QU, Bibi A, Bilal S, Bibi T, Liaqat R, Sharif M, Zulfiqar S, Fuhr DC, Price LN, Patel V, Rahman A. Delivering the Thinking Healthy Programme for perinatal depression through volunteer peers: a cluster randomised controlled trial in Pakistan. *Lancet Psychiatry* 2019; **6**(2): 128-39.10.1016/s2215-0366(18)30467-x
17. Vanobberghen F, Weiss HA, Fuhr DC, Sikander S, Afonso E, Ahmad I, Atif N, Bibi A, Bibi T, Bilal S, De Sa A, D'Souza E, Joshi A, Korgaonkar P, Krishna R, Lazarus A, Liaqat R, Sharif M, Weobong B, Zaidi A, Zuliqar S, Patel V, Rahman A. Effectiveness of the Thinking Healthy Programme for perinatal depression delivered through peers: Pooled analysis of two randomized controlled trials in India and Pakistan. *J Affect Disord* 2019.https://doi.org/10.1016/j.jad.2019.11.110
18. Fuhr DC, Weobong B, Lazarus A, Vanobberghen F, Weiss HA, Singla DR, Tabana H, Afonso E, De Sa A, D'Souza E, Joshi A, Korgaonkar P, Krishna R, Price LN, Rahman A, Patel V. Delivering the Thinking Healthy Programme for perinatal depression through peers: an individually randomised controlled trial in India. *Lancet Psychiatry* 2019; **6**(2): 115-27.10.1016/s2215-0366(18)30466-8
19. Atif N, Bibi A, Nisar A, Zulfiqar S, Ahmed I, LeMasters K, Hagaman A, Sikander S, Maselko J, Rahman A. Delivering maternal mental health through peer volunteers: a 5-year report. *International Journal of Mental Health Systems* 2019; **13**(1): 62.10.1186/s13033-019-0318-3
20. National Institute of Population Studies - NIPS/Pakistan & ICF Intternational. Pakistan Demographic and Health Survey 2012-2013. Islamabad: NIPS/Pakistan and ICF International, 2012-3.
21. National Institute of Population Studies Islamabad. Pakistan Demographic and Health Survey. Islamabad, Pakistan: National Institute of Population Studies, 2013.
22. Turner EL, Sikander S, Bangash O, Zaidi A, Bates L, Gallis J, Ganga N, O'Donnell K, Rahman A, Maselko J. The effectiveness of the peer-delivered Thinking Healthy PLUS (THPP+) Program for maternal depression and child socioemotional development in Pakistan: study protocol for a randomized controlled trial. *Trials* 2016; **17**(1): 442.10.1186/s13063-016-1530-yPMC5017048
23. Gallis J, Maselko J, O'Donnell K, Song KE, Saqib K, Turner EL, Sikander S. Criterion-related validity and reliability of the Urdu version of the patient health questionnaire in community-based women in Pakistan. *PeerJ* 2018; **6**(e5185).https://doi.org/10.7717/peerj.5185
24. Sikander S, Ahmad I, Bates LM, Gallis J, Hagaman A, O'Donnell K, Turner EL, Zaidi A, Rahman A, Maselko J. Cohort Profile: Perinatal depression and child socioemotional development ; the Bachpan cohort study from rural Pakistan. *BMJ Open* 2019; **9**(5): e025644.10.1136/bmjopen-2018-025644
25. Sikander S, Lazarus A, Bangash O, Fuhr DC, Weobong B, Krishna RN, Ahmad I, Weiss HA, Price L, Rahman A, Patel V. The effectiveness and cost-effectiveness of the peer-delivered Thinking Healthy Programme for perinatal depression in Pakistan and India: the SHARE study protocol for randomised controlled trials. *Trials* 2015; **16**: 14.10.1186/s13063-015-1063-9
26. Atif N, Nisar A, Bibi A, Khan S, Zulfiqar S, Ahmad I, Sikander S, Rahman A. Scaling-up psychological interventions in resource-poor settings: training and supervising peer volunteers to deliver the 'Thinking Healthy Programme' for perinatal depression in rural Pakistan. *Global mental health (Cambridge, England)* 2019; **6**: e4.10.1017/gmh.2019.4PMC6521132
27. Organization. WH. WHO Guidelines Approved by the Guidelines Review Committee. mhGAP Intervention Guide for Mental, Neurological and Substance Use Disorders in Non-Specialized Health

659 Settings: Mental Health Gap Action Programme (mhGAP): Version 20. Geneva: World Health
660 Organization

661 Copyright (c) World Health Organization 2016.; 2016.
662 28. Organization" WH. WHO Disability Assessment Schedule II (WHO DAS II).
663 http://www.who.int/classifications/icf/whodasii/en/index.html (accessed May 7, 2010.
664 29. Gorman LL, O'Hara MW, Figueiredo B, Hayes S, Jacquemain F, Kammerer MH, Klier CM, Rosi S,
665 Seneviratne G, Sutter-Dallay AL. Adaptation of the structured clinical interview for DSM-IV disorders for
666 assessing depression in women during pregnancy and post-partum across countries and cultures. *Br J*
667 *Psychiatry Suppl* 2004; **46**(23): s17-23
668 30. Ahmad I, Suleman N, Waqas A, Atif N, Malik AA, Bibi A, Zulfiqar S, Nisar A, Javed H, Zaidi A, Khan
669 ZS, Sikander S. Measuring the implementation strength of a perinatal mental health intervention
670 delivered by peer volunteers in rural Pakistan. *Behaviour Research and Therapy* 2020:
671 103559.https://doi.org/10.1016/j.brat.2020.103559
672 31. Goodman R. The strengths and difficulties questionnaire: A research note. *Journal of Child*
673 *Psychology and Psychiatry and Allied Disciplines* 1997; **38**(5): 581-6
674 32. Samad L, Hollis C, Prince M, Goodman R. Child and adolescent psychopathology in a developing
675 country: testing the validity of the Strengths and Difficulties Questionnaire (Urdu version). *International*
676 *Journal of Methods in Psychiatric Research* 2005; **14**(3): 158-66.10.1002/mpr.3
677 33. Bayley N, Reuner G. Bayley scales of infant and toddler development: Bayley-III: Harcourt
678 Assessment, Psych. Corporation; 2006.
679 34. O'Donnell K, Murphy R, Ostermann J, Masnick M, Whetten RA, Madden E, Thielman NM,
680 Whetten K. A brief assessment of learning for orphaned and abandoned children in low and middle
681 income countries. *AIDS Behav* 2012; **16**(2): 480-90.10.1007/s10461-011-9940-zPMC3817622
682 35. Maselko J, Bates L, Bhalotra S, Gallis JA, O'Donnell K, Sikander S, Turner EL. Socioeconomic
683 status indicators and common mental disorders: Evidence from a study of prenatal depression in
684 Pakistan. *SSM - Population Health* 2018; **4**: 1-9.https://doi.org/10.1016/j.ssmph.2017.10.004
685 36. Kroenke K, Spitzer RL, Williams JBW. The PHQ-9 - Validity of a brief depression severity measure.
686 *Journal of General Internal Medicine* 2001; **16**(9): 606-13.10.1046/j.1525-1497.2001.016009606.x
687 37. Hedges LV. Effect Sizes in Cluster-Randomized Designs. *JEBS* 2007; **32**(4): 341-
688 70.10.3102/1076998606298043
689 38. Li P, Redden DT. Comparing denominator degrees of freedom approximations for the
690 generalized linear mixed model in analyzing binary outcome in small sample cluster-randomized trials.
691 *BMC Medical Research Methodology* 2015; **15**(1): 38.10.1186/s12874-015-0026-x
692 39. Gallis J, Turner EL. Relative measures of association for binary outcomes: Challenges and
693 recommendations for the global health researchers. *Ann Glob Health* 2019; **85**(1): 137
694 40. Gallis JA, Li F, Turner EL. xtgeebcv: A command for bias-corrected sandwich variance estimation
695 for GEE analyses of cluster randomized trials. *Stata Journal* (in press):
696 41. Chambers DA, Glasgow RE, Stange KC. The dynamic sustainability framework: addressing the
697 paradox of sustainment amid ongoing change. *Implementation science : IS* 2013; **8**: 117.10.1186/1748-
698 5908-8-117PMC3852739
699 42. Tandon SD, Ward EA, Hamil JL, Jimenez C, Carter M. Perinatal depression prevention through
700 home visitation: a cluster randomized trial of mothers and babies 1-on-1. *J Behav Med* 2018; **41**(5): 641-
701 52.10.1007/s10865-018-9934-7
702 43. Sondaal AEC, Tumbahangphe KM, Neupane R, Manandhar DS, Costello A, Morrison J.
703 Sustainability of community-based women's groups: reflections from a participatory intervention for
704 newborn and maternal health in Nepal. *Community development journal* 2019; **54**(4): 731-
705 49.10.1093/cdj/bsy017PMC6924535

44. Baqui AH, El-Arifeen S, Darmstadt GL, Ahmed S, Williams EK, Seraji HR, Mannan I, Rahman SM, Shah R, Saha SK, Syed U, Winch PJ, Lefevre A, Santosham M, Black RE. Effect of community-based newborn-care intervention package implemented through two service-delivery strategies in Sylhet district, Bangladesh: a cluster-randomised controlled trial. *The Lancet* 2008; **371**(9628): 1936-44.10.1016/S0140-6736(08)60835-1
45. Tripathy P, Nair N, Barnett S, Mahapatra R, Borghi J, Rath S, Rath S, Gope R, Mahto D, Sinha R, Lakshminarayana R, Patel V, Pagel C, Prost A, Costello A. Effect of a participatory intervention with women's groups on birth outcomes and maternal depression in Jharkhand and Orissa, India: a cluster-randomised controlled trial. *The Lancet* 2010; **375**(9721): 1182-92.10.1016/S0140-6736(09)62042-0
46. Yousafzai AK, Rasheed MA, Rizvi A, Armstrong R, Bhutta ZA. Parenting Skills and Emotional Availability: An RCT. *Pediatrics* 2015; **135**(5): e1247-e57.10.1542/peds.2014-2335
47. Hafeez A, Mohamud B, Shiekh M, Shah I, Jooma R. Lady health workers programme in Pakistan: challenges, achievements and the way forward. *JPMA* 2011; **61**(210):
48. Rahman A, Malik A, Sikander S, Roberts C, Creed F. Cognitive behaviour therapy-based intervention by community health workers for mothers with depression and their infants in rural Pakistan: a cluster-randomised controlled trial. *Lancet* 2008; **372**(9642): 902-9
49. Paterick TE, Patel N, Tajik AJ, Chandrasekaran K. Improving health outcomes through patient education and partnerships with patients. *Proc (Bayl Univ Med Cent)* 2017; **30**(1): 112-3.10.1080/08998280.2017.11929552
50. Lesaffre E. Superiority, equivalence, and non-inferiority trials. *Bulletin of the NYU hospital for joint diseases* 2008; **66**(2): 150-4
51. Netsi E, Pearson RM, Murray L, Cooper P, Craske MG, Stein A. Association of Persistent and Severe Postnatal Depression With Child OutcomesAssociation of Persistent and Severe Postnatal Depression With Child OutcomesAssociation of Persistent and Severe Postnatal Depression With Child Outcomes. *JAMA psychiatry* 2018; **75**(3): 247-53.10.1001/jamapsychiatry.2017.4363
52. Robinson R, Lahti-Pulkkinen M, Heinonen K, Reynolds RM, Räikkönen K. Fetal programming of neuropsychiatric disorders by maternal pregnancy depression: a systematic mini review. *Pediatric Research* 2019; **85**(2): 134-45.10.1038/s41390-018-0173-y
53. O'Donnell KJ, Glover V, Barker ED, O'Connor TG. The persisting effect of maternal mood in pregnancy on childhood psychopathology. *Development and Psychopathology* 2014; **26**(2): 393-403.10.1017/S0954579414000029
54. Stein A, Netsi E, Lawrence PJ, Granger C, Kempton C, Craske MG, Nickless A, Mollison J, Stewart DA, Rapa E, West V, Scerif G, Cooper PJ, Murray L. 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): 134-44.10.1016/s2215-0366(18)30006-3
55. Frankenhuis WE, Del Giudice M. When Do Adaptive Developmental Mechanisms Yield Maladaptive Outcomes? *Developmental Psychology* 2012; **48**(3): 628-42.10.1037/a0025629

Effectiveness of a peer-delivered, psychosocial intervention on maternal depression and child development at 3 years of age: a cluster randomized trial in Pakistan

Joanna Maselko ScD[1] and Siham Sikander PhD[2]*, Elizabeth L. Turner PhD[3], Lisa M. Bates ScD[4], Ikhlaq Ahmad MPhil[5], Najia Atif PhD[5], Victoria Baranov PhD[6], Sonia Bhalotra PhD (full professor)[7], Amina Bibi MPH[5], Tayyaba Bibi MS[5], Samina Bilal MS[5], Pietro Biroli PhD[8], Esther Chung MS[1], John A. Gallis MS[3], Ashley Hagaman PhD[9], Anam Jamil BA[5], Katherine LeMasters MS[1], Karen O'Donnell PhD[10], Elissa Scherer BS[1,11], Maria Sharif MSPH[5], Ahmed Waqas MBBS[5], Ahmed Zaidi MS[5], Shaffaq Zulfiqar MS[5], & Atif Rahman PhD (full professor)[12]

1. Department of Epidemiology, Gillings School of Global Public Health, and Carolina Population Center, University of North Carolina, Chapel Hill, NC, USA
2. Human Development Research Foundation (HDRF) and Health Services Academy, Islamabad, Pakistan
3. Department of Biostatistics and Bioinformatics, and Duke Global Health Institute, Duke University, Durham, NC USA
4. Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, USA
5. Human Development Research Foundation (HDRF)
6. Department of Economics, University of Melbourne, Melbourne, Australia
7. Department of Economics, Univesity of Essex, Colchester, UK
8. Department of Economics, University of Zurich, Zurich, Switzerland
9. Department of Social and Behavioral Sciences, Yale School of Public Health, Yale University, New Haven, CT, USA
10. Center for Child and Family Health, Durham, NC USA
11. Research Triangle International, Research Triangle Park, USA
12. Institute of Psychology, Health and Society, University of Liverpool, Liverpool UK

Corresponding Authors

Joanna Maselko, ScD
jmaselko@unc.edu
Dept. of Epidemiology
Carolina Population Center Fellow
2105e McGavran-Greenberg Hall
Campus Box 7435
Gillings School of Global Public Health, UNC
Chapel Hill, NC 27599-7435 USA

~~Siham Sikander, PhD~~
~~siham.sikander@hdrfoundation.org, sihamsikander@gmail.com~~
~~Health Services Academy~~
~~H.No 55, Street 55, F-7/4~~
~~Islamabad Pakistan~~

Abstract

Background

Maternal depression has a recurring course that can influence offspring outcomes. There is limited evidence about how to treat maternal depression to improve longer term maternal outcomes and reduce intergenerational transmission of psychopathology using task-shifted, low-intensity scalable psychosocial interventions. We sought to fill this gap, evaluating the effects of a peer-delivered psychosocial depression intervention on maternal depression and child development at 3 years of age.

Methods

Forty village clusters in Pakistan were randomly allocated to treatment or enhanced usual care (EUC). Pregnant women aged 18 years or over screening positive for moderate or severe depression symptoms (Patient Health Questionnaire (PHQ-9) score 10+) were recruited into the trial (n=570) and a non-depressed cohort was also enrolled (n=584). Primary outcomes were maternal depression symptoms and remission (PHQ-9<10) and child socioemotional skills (Strengths and Difficulties Questionnaire- Total Difficulties (SDQ-TD) at 36-monrha postnatal~~/~~. Analyses were intention-to-treat. The trial was registered with ClinicalTrials.gov, number NCT02658994.

Findings

At 36-months postnatal, complete data were available from 889 mother-child dyads: 206 treatment (72.5%), 216 EUC (75.3%), and 467 prenatally non-depressed (80.0%). We did not observe significant outcome differences between treatment and EUC arms of the trial (PHQ-9 total score: Standardized Mean~~d~~ Difference = -0.13, 95% CI -0.33 to 0.07; PHQ-9 remission: RR= 1.08, 95% CI 0.88 to 1.33; SDQ-TD treatment estimate: -0·10; 95%CI -1·39, 1·19;). Approximately 40% of women did not complete their treatment sessions, and competence levels of peers dropped over time.

Interpretation

Reduced symptom severity and high remission rates were seen across both arms, possibly masking any treatment effects. A multi-year, psychosocial intervention~~s~~ can be task-shifted via peers but are susceptible to reductions in fidelity and dosage over time. Early intervention efforts might need to rely on multiple models (e.g. collaborative care), be of greater intensity, and potentially targeted toward higher risk mothers to reduce the intergenerational transmission of psychopathology from mothers to children.

This study was funded by a grant from the NIH R01HD075875, NIMH U19MH95687, and NICHD T32HD007168.

Research in context

Evidence before this study.

Recent systematic reviews of psychotherapy interventions for depression have highlighted the limited evidence on long-term effects of psychotherapy on either maternal mental health or child outcomes. We conducted a search to identify studies designed to evaluate interventions for perinatal depression, whose intervention either lasted beyond 12 months postnatal (e.g. booster sessions), or whose follow-up was more than 12 after the completion of the intervention, in years 2002-April 2020. We limited our search to randomized clinical trials or meta-analyses. We did not place restrictions on language or country. We used Pubmed and Web of Science, with the following search terms: ((maternal depression) or (perinatal depression) or (postpartum depression) or (postpartum depression)) AND ((treatment) or (therapy) or (intervention) or (psychotherapy) or (cognitive behavioral therapy)) AND ((longer-term) or (longer) or (booster)). We identified six RCTs specific to perinatal depression with the longer follow-up period, ranging from 1.5 to 7 years. None utilized an extended duration design (booster sessions) that continued past 12 months postnatal; two studies included a non-depressed comparison group. Most common intervention models were Cognitive Behavioral Therapy (CBT) and Interpersonal Therapy (IPT). Evidence generally showed that interventions improved outcomes which then weakened over time so that, overall, there is limited evidence of intervention effects of perinatal depression interventions that persist beyond the perinatal period. Of the six studies, two reported some lasting impact. One study of 884 mother-child dyads assessed maternal and child outcomes 7 years after the end of a CBT intervention and found a 5 point percent lower rates of depression among mothers who received the intervention, but no significant effects on child outcomes. With this one exception, sample sizes were small, with studies having fewer than 60 participants per group available at follow-up.

Added value of this study.

Given the chronic and recurring nature of depression, longer lasting interventions may be necessary to effectively reduce disease burden and potentially reduce the intergenerational transmission of psychopathology. This study, in rural Pakistan, is the first large multi-year randomized controlled trial, focusing on both maternal and child outcomes, where individuals with depression received psychotherapy beginning prenatally. The extended duration psychosocial intervention evaluated in the current study did not show evidence of meaningfully reducing depression symptom levels, nor improving child outcomes, at the 3 years postnatal mark. These findings highlight the challenges of implementing a peer-delivered psychosocial intervention over a longer period in a low resource community setting.

Implications of all the available evidence. These findings point to several implementation lessons for such task-shifted, low-intensity interventions when delivered at scale alongside existing health systems. These include importance of ensuring high levels of fidelity of the intervention, potentially through use of technology platforms. It is also important that any intervention be situated within a collaborative care model that can help detect and respond to women in need of other services to help social determinants like poverty and domestic violence or pharmacological interventions.

Introduction

Global prevalence estimates of depression in the perinatal period range from 4% to over 50%, with the highest burden in low-resource settings, making depression a public health priority.[1] In addition to the effect of maternal depression on the woman's functioning, physical health, and risk of suicide, observational evidence suggests that maternal depression is associated with higher risk of multiple negative child outcomes, including stunting, socioemotional difficulties, problems with school readiness and performance, and depressive symptoms over their lifecourse.[1,2] Women experiencing perinatal depression are at much higher risk of subsequent or recurrent episodes of depression and this chronic or episodic depression is most deleterious for numerous maternal and child outcomes.[3,4] This risk of intergenerational transmission of psychopathology is most heavily borne by poorer families and those in low resource settings with limited access to quality healthcare, thus exacerbating economic and social inequality.[3]

Because of the lack of specialists in many LMIC settings, task shifting for maternal depression is necessitated to bridge the treatment gap. Evidence-based, task-shifted, targeted maternal depression or universal psychosocial interventions can be delivered through community health workers as well as lay peers.[5,6] However, most of these interventions are delivered either during pregnancy or in the early postnatal months, focusing on the acute phase of maternal depression, without tackling issues of recurrence and chronicity. To our knowledge no depression interventions that begin prenatally are designed specifically to prevent recurrence after the perinatal period. Hence, the extent to which such interventions can break the cycle of recurrence of depression beyond the first postnatal year remains unknown.

While interventions have demonstrated efficacious reductions in shorter term (i.e. 12 months or less) maternal depression and improved maternal behaviours,[7] we do not know whether such interventions can reduce intergenerational transmission to children. Many depression interventions in the perinatal period include a child development component, opening the possibility that such depression interventions, including the one studied here, may affect child outcomes through pathways that are independent of changes in depression symptoms themselves.[5] While maternal depression interventions have been shown to improve key parenting practices,[8] evidence of long-term effects on child socioemotional development is scarce.[9] Studies showing improved child outcomes have short post-intervention follow-up periods, typically less than 12 months,[10-12] leaving uncertainty about longer lasting program impacts. The few studies with follow-up longer than one year have reported mixed or even incongruent effects.[4,6,13,14] For example, analysis of the subset of women who were depressed when beginning the Philani+ program in South Africa, which broadly focused on improving child outcomes and lasted through 6 months postnatally, showed improved child physical and cognitive outcomes at 18 months but higher levels of aggression at 5 years of age.[6,15] The challenges of differential attrition in longer-term follow-ups, diminishing sample sizes, and heterogeneous responses among particular sub-groups (such as those exposed to poverty or intimate partner violence) make clear conclusions difficult.[2,14]

The Thinking Healthy Program, Peer-delivered (THPP), delivered individual and group sessions from pregnancy to 6 months postnatal and has been evaluated through two randomized controlled trials, one in Pakistan and one in India.[16-18] Although the country specific findings were weak, the pooled analyses of these trials showed greater recovery from perinatal depression among the intervention group at 6 months postnatal. It also showed that delivering this psychosocial intervention through peers was a cost-effective, feasible and acceptable approach.[16]

We evaluate a 3-year, task-shifted psychosocial peer-delivered intervention for maternal depression, Thinking Healthy Program, Peer-delivered Plus (THPP+),[19] that followed up on the THPP. The project is located ~~,~~ in rural Pakistan, a low resource context characterized by high levels of maternal depression and limited access to clinical mental healthcare.[20]

Although our hypothesis was that the children in the intervention arm will be less high risk (as compared to those in the control~~s~~ arm), the full impact of the intervention can only be discerned if we know the level of *excess* risk remaining- that is, the difference between the reduced level of risk among children (of prenatally depressed mothers) in the intervention arm and the risk among children whose mothers were not depressed. If outcomes of these two groups are comparable, we can infer that the intervention is capable of preventing the intergenerational transmission of risk. To achieve this we gathered data on women who were not depressed in pregnancy.The resulting pregnancy-birth cohort of both prenatally depressed (trial participants) and non-depressed women is referred to as the Bachpan Cohort (*Bachpan* means Childhood in the local Urdu language). Finally, we examined whether intervention effects differ by key social contextual factors such as socioeconomic status and intimate partner violence.

Methods

Study design and participants

We conducted a stratified cluster-randomized controlled trial in Kallar Syedan, a rural subdistrict of Rawalpindi, Pakistan. The sub-district is a socioeconomically deprived area with poverty rates around 50%, female literacy of 40-45%, and a high fertility rate (3·8 births per woman).[21] It is primarily agrarian with close knit communities co-residing in large households (average 6·2 people per household). The sub-district has 11 Union Councils (UC), the smallest administrative unit, each with a population of about 22,000-25,000. Each UC is serviced by a Primary Health Care Facility which houses a physician, midwife, vaccinator, dispenser, and village-based Lady Health Workers (LHWs).

This trial maintained the original cluster criterion, randomization sequence and procedures under the previous trial.[16,22] Pregnant women in the 3rd trimester, aged 18+ years and registered with their LHWs, were eligible. Approximately 95% of the women in the study area were registered with LHWs. All pregnant women were approached by trained research staff either within the pregnant woman's residence or that of their LHW and, if they consented, were assessed. Women who needed immediate medical or psychiatric inpatient care were excluded from the study. All eligible women were invited to be screened for depression using the Urdu version of the Patient Health Questionnaire 9 (PHQ-9), which has been used extensively as a screening tool in the study setting and has an acceptable criterion validity and reliability for this population.[23] Women screening positive (PHQ-9 score $\geq$ 10) were eligible for enrolment into the trial and follow-up as part of the Bachpan Cohort.[16] One out of every three women who screened with a < 10 score on PHQ were enrolled to participate in the Bachpan Cohort only, resulting in a roughly equal size of prenatally non-depressed and depressed women at the beginning the cohort.[24]

Randomization and masking

The trial was conducted at a sub-district level. Forty village-clusters (population of 2,400 to 3,600) were the unit of randomization and were geographically separate to minimize contamination risk. The sub-district is administratively subdivided into 11 Union Councils (as explained above and, within each of these 11 union councils, we identified an even number of village clusters to ensure that equal numbers of clusters are randomized into intervention or control condition (ie 1:1 ratio) by an independent statistician using a computerized randomization sequence. Research teams responsible for identifying, obtaining consent and recruiting trial participants were blind to the allocation status. The trial PI, site PIs/coordinators, trial statisticians, and members of the Trial Steering Committee were blinded to the allocation status until the analysis of the six-month data for the initial THPP trial.[25]

The Thinking Healthy Program, Peer-delivered Plus (THPP+) Intervention

The intervention arm received the longer-duration peer-delivered psychosocial intervention (THPP+). It consisted of 18 group-based "booster" sessions (from 7th to 36th month postnatal). Of these, the first 6 sessions were delivered monthly, then bi-monthly until 36 months. These sessions built on the shorter duration intervention and were delivered by the same peers. The peers were lay married women who lived in the same community as that of the depressed women and volunteered their time.

The key features of this psychosocial intervention, delivered by non-specialists, were peer-support, behavioural activation, and problem-solving in a culturally sensitive, non-medicalized format, and developmental activities for children up to the 36th month (See Appendix p.1-3 for the overall structure of the intervention and peer characteristics.[19] A cascaded model of training and supervision was used which included frequent competence assessments.[26] A competency checklist was used for supervision, including direct observation of the peers, and to provide feedback. Each peer received this assessment six times over the course of the program. During the feedback meetings, the supervisor discussed checklist information along with session logs maintained by the peers. Competence was assessed through observations and the checklist captured whether content was delivered as intended. Refresher trainings were done 6 and 18 months after the initial training. The two-day training included re-orientation to the intervention and its principles, as well as training on materials and use of job aids.

The intervention group sessions provided a safe environment for women to voice their problems, share experiences of childcare, and provide support to women. Peer volunteers were trained to use culturally grounded vignettes that served as tools to deliver health and well-being messages. The sessions aimed for maternal well-being but also child-care and development by encouraging mother-infant interaction and play. The intervention provided examples of age-appropriate activities, derived from the UNICEF/WHO's Care for Development Package and encouraged demonstration of these activities during the sessions. While these 'booster' group sessions did not focus on specific strategies to address depression, the peer could still draw on her prior knowledge and skills of specific psychotherapeutic elements such as behavioural activation when required. Additional details of the intervention are reported elsewhere.[19,26] We defined overall treatment completion as attending 10 (out of 14) sessions from pregnancy through 6 months postnatal (Phase 1: THPP, individual and group sessions) and 12 (out of 18) sessions from 7 to 36 months postnatal (Phase 2: THPP+, booster group sessions).

Both intervention and control arms received Enhanced Usual Care (EUC). No treatment was offered to the prenatally non-depressed women in either of the arms. EUC consisted of informing participants about their depression status and ways to seek help for it, informing their respective LHWs about

women's depression status at enrollment, training all the 11 primary care facility-based physicians in the subdistrict on the mental health Gap Action Programme (mhGAP) treatment guidelines for maternal depression,[27] and providing depressed participants with a leaflet on how and where to seek appropriate health care during pregnancy and beyond.

Procedures

Assessments were conducted 6 times over the course of the study (in pregnancy, and 3, 6, 12, 24, and 36 months postnatal). As originally specified, the current analysis utilizes mother and child outcome data at 36 months.[22]All measures were extensively piloted. Assessments were done at the community level within households of women by trained interviewers blind to the allocation status, all questions were interviewer administered. Assessors inter-rater reliability was ensured through classroom-based training which included role plays, followed by field practice sessions. During these sessions each pair of assessors assessed up to 10 participants jointly and discussed their coding on each item to establish inter-rater reliability prior to start of actual data collection.

The project received approval for the IRBs of Human Development Research Foundation (HDRF), Duke University, and University of North Carolina. The study protocol for the effectiveness trial of THPP+ and inclusion of prenatally non-depressed pregnant women in the Bachpan Cohort study has been published previously.[16,22]

Outcomes

The primary maternal outcome was depression symptoms assessed using the Patient Health Questionnaire-9 (PHQ-9) and analyzed as symptom severity (total score) and remission (score < ~~5~~10). The secondary maternal outcomes were disability, assessed using WHO's Disability Assessment Schedule, WHO-DAS[28], and current major depressive episode based on the *Structured Clinical Interview for DSM-IV (SCID) Disor*ders.[29] Since it provides a clinically salient diagnostic outcome, the SCID was also included to increase measurement robustness. Process data regarding THPP+ sessions attended, competence scores of peers, duration of sessions, and peers' supervision attendance are described elsewhere[30] and in the appendix (p.4-8).

The primary child outcome was child socioemotional development measured using the Strengths and Difficulties Questionnaire, Total Difficulties (SDQ-TD) score. The SDQ is a parent-reported measure of 25 child attributes with five subscales: emotional symptoms, conduct problems, hyperactivity, peer problems, and prosocial behavior.[31] The Total Difficulties (TD) score is calculated based on four subscales (omitting prosocial behaviour) with a score range of 0-40 points. The SDQ is widely used in low- and middle-income countries and has been translated into Urdu[32]; in our sample, internal consistency measured by Cronbach's alpha was 0.78.

The secondary child outcomes were two developmental domains. Given language differences, two subscales from the Bayley Scales of Infant and Toddler Development, 3rd edition (BSITD III) were selected to assess achievement of developmental milestones, the Receptive Language and Fine Motor subscales.[33] The BSITD was administered in the family's home; scaled scores were calculated using the child's chronological age. The BSITD has been widely used and validated internationally.[34]

An additional outcome of interest, child growth, was analysed using weight- and length-for-age Z-scores.

Additional variables included demographic and psychosocial factors hypothesized to moderate the effect of the intervention. These included household assets as an indicator of socioeconomic status,[35] maternal education (coded as none vs. any), household composition (nuclear family status), intimate partner violence (IPV) in the previous 12 months, child gender, maternal age (18-24 vs. 25+), number of siblings (0 vs. 1+), treatment expectations (very/moderately useful vs. somewhat/not useful), depression chronicity (<12 weeks vs. ≥12 weeks), and depression severity (PHQ-9 10-14 vs. 15+).[23,36] We collected information on a number of domains that may have been differentially distributed across the treatment and EUC arms or correlated with loss to follow-up (table 1 and Appendix p.9-14).

Statistical analyses

For mother outcomes, anticipating a sample of 480 prenatally depressed women, we were powered at 90% to detect a remission rate of 65% in the prenatally depressed-intervention versus 45% in prenatally depressed-control at the two-tailed 5% significance level and assuming a conservative ICC of 0.07 in intervention arm and 0.05 in control.[22] For child outcomes, with this ~~e~~ sample ~~sample~~ size, we were powered at 90% to detect a difference between treatment arms of 3 points on the SDQ-TD at the two-tailed 5% significance level assuming a standard deviation of 5.2 points and ICC of 0.04-0.08.[22] ~~We were also well-powered to test for equivalence of SDQ-TD score between children of prenatally depressed mothers in the treatment arm and children of prenatally non-depressed mothers in the EUC arm, with equivalence defined as the 95% confidence interval of the mean difference being between -2 and 2 units.~~

Statistical analyses were done according to the CONSORT guidelines in Stata software version 16.1 (StataCorp, College Station, TX) and SPSS. All analyses compare the three groups (prenatally depressed in intervention, prenatally depressed in control, and prenatally non-depressed) across the two arms, using the 36-month outcomes. We had pre-specified a comparison of outcomes between intervention and control arms within the prenatally non-depressed women and, in the absence of such an effect, present results for the overall prenatally non-depressed cohort.

The primary analyses were designed as intention-to-treat. Data from prenatally depressed and non-depressed participants were analyzed jointly using linear mixed effects models so that all comparisons of interest could be estimated from the same model. The identity link was used for continuous outcomes to estimate differences in mean outcomes. Standardized mean differences (SMDs) and their 95% confidence intervals (CIs) were obtained using the method of Hedges.[37] In the primary model, we included a random intercept for cluster and fixed effects for treatment arm (depressed intervention, depressed control, non-depressed), union council (11 levels; the stratification variable), and variables found to be imbalanced by loss to follow-up or at baseline (determined using $p<0.10$; see Appendix p.9-14). Mixed models assume missing at random conditional on the covariates included in the statistical model. Therefore, ~~W~~we include variables predictive of loss to follow-up to account for missing data. ~~Mixed models assume missing at random, since we include baseline data in the model and adjusting for variables lost to follow-up will help account for missing data.~~ We used restricted maximum likelihood (REML). The between-within method was used to apply small-sample bias corrections to the intervention effect standard errors in the mixed effects framework.[38] These models also generated the intra-cluster correlation (ICC) values.

All binary outcomes were analyzed using generalized estimating equations (GEE). As with the continuous outcomes, we include as fixed effects treatment arm and union council, as well as fixed effects for any variables found to be imbalanced by loss to follow-up or at baseline. In the GEE framework, we took into account clustering using an exchangeable working correlation matrix. We used a modified Poisson approach[39] and Kauermann-Carroll bias-corrected standard errors to account for the

relatively small number of clusters (i.e. 40).[40] When analyzing SCID major depressive episode, we included SCID at all time points as the outcome in a GEE model with exchangeable working correlation for village cluster. Additional analyses focus on *a priori* identified potential moderators of any main associations with the primary outcomes (described above). We tested for moderation of the intervention effect by including these potential moderators in the model as individual interaction terms.

ClinicalTrials.gov Identifier: NCT02658994 (registered on 21 January 2016).

Role of Funding Source

The Funding Source (NIH) had no influence in study design, data collection, data analysis, data interpretation, or writing of the report.The corresponding authors had full access to all of the data and the final responsibility to submit for publication.

Results

From Oct 15, 2014 to Feb 25, 2016, we identified and randomly selected 40 village clusters out of 46 and randomly assigned 20 village clusters each to intervention (THPP+) and control (EUC) arms. In all we approached 1910 pregnant women; 287 prenatally depressed women in the control arm and 283 in the intervention arm completed the baseline questionnaire. Of the prenatally non-depressed women approached, 584 were enrolled, yielding comparable numbers of prenatally depressed and prenatally non-depressed women in each of the 40 village clusters.
Of the 1,154 participants enrolled at baseline, 889 (77.0%) were successfully interviewed at 36 months: 206 (72.5%) intervention, 216 (75.3%) EUC, and 467 (80.0%) prenatally non-depressed (Figure 1). There was no differential loss to follow up by treatment arm and no differences in adverse events (Appendix p.9-14).

At baseline, the mean age of the women in the sample was 26.7 (SD: 4.5) years, with 30.2% of women being in their first pregnancy. The mean PHQ-9 scores across the treatment arms were similar (14.5 control and 14.9 intervention) with a mean score of 2.8 among prenatally non-depressed women. Further baseline characteristics are summarized in Table 1. At 36 months, 49.6% of the infants were girls.

Baseline variables that were imbalanced between intervention and control groups include the life events checklist score with a higher mean score in the intervention arm. Other variables were found to be imbalanced but were not included because of collinearity or conceptual overlap (e.g. subjective religiosity). Additional baseline demographic variables that were associated with loss to follow-up at 36 months and adjusted for include number of people per room, child's grandmother living with him/her, nuclear family status, number of living children and the asset score.

Just over two-thirds of prenatally depressed women in the intervention arm completed the THPP+ intervention (Appendix p.5). Only 63% of women completed treatment during Phase 2 compared to nearly 80% treatment completion in Phase 1. The competence levels of the peers declined over the implementation period, particularly in the period after time point 3 (i.e. at 12 months postnatal) where we introduced new content for the booster sessions and the frequency of peer supervisions dropped from monthly to every two months (Aappendix p.6-8) show high and low scoring peers and their competence levels over time

For all prenatally depressed women, depression scores dropped meaningfully, regardless of their arm allocation (Figure 2). There were no significant differences in depression outcomes between arms (THPP+ vs EUC) at 36 months postnatal (Table 2). The adjusted standardized mean difference (SMD) in depressive symptom severity (PHQ-9) between arms was -0.13 (95% CI -0.33 to 0.07) and the risk ratio (RR) for depression remission (PHQ-9<10) was 1.00 (95% CI 1.13 to 0.97). Turning to the secondary outcomes, we observed a relatively larger difference between arms on the secondary outcomes of SCID-based major depressive episode at 36 months (22% control vs 16.5% intervention) (RR=0.67, 95% CI 0.43 to 1.05). The intervention effects on disability (SMD=-0.12, 95% CI -0.33 to 0.09) were not statistically significant.

The prevalence of ~~SCID-based~~ major depressive episodes (SCID-based) among the three groups of women (~~:~~ the prenatally depressed intervention arm, prenatally depressed control arm, and prenatally non-depressed women) became increasingly similar in the proportion depressed by 36 months postnatal (F~~f~~igure 2). The THPP+ intervention arm showed higher convergence with the prenatally non-depressed women at 36 months compared to the control arm, so much so that there was not a statisticaly significant difference in the probability of being depressed (using SCID) between the intervention arm women (at 16.5%) and the prenatally non-depressed (9.0% of whom were depressed at 36 months, RR 0·74 (0·46 to 1·18).

We also examined intervention arm differences in depression severity (PHQ-9 scores) at 36 months postnatal by potential moderators assessed at baseline. There was no strong evidence of meaningful moderation of the intervention effect by any of these factors (Appendix p.18).

For the child primary outcome of SDQ-TD, the mean adjusted difference between intervention and control arms was -0·10; 95% CI -1·39 to 1·19 (Table 2). There were also no meaningful differences between the two arms in the secondary outcomes of Receptive Language and Fine Motor scores (from BSITD) (Table 2),

Similar to maternal depression results, we do not find strong support for the hypothesis that baseline characteristics moderated treatment effect on the SDQ-TD scores (Appendix p.19).

Children of prenatally non-depressed mothers had somewhat better SDQ-TD scores than the intervention or control arm children (e.g. SDQ-TD=13.7 among the children of the prenatally non-depressed vs. 14.7 in the intervention arm, p-value=0.07, Table 2). In other words, the prenatal depression episode predicted slightly worse SDQ-TD scores at 36 months of age, independent of the intervention. Exploratory analyses of the five sub-scales of the SDQ separately showed that this overall difference was driven by the hyperactivity and conduct problems sub-scales, with negligible differences by prenatal depression status for the peer problems, emotional problems, and the pro-social scales (Appendix p.19). As an example, the adjusted mean difference in the hyperactivity scores between the non-depressed and the control arm was 0·31 (95% CI -0·62, 0·01,). The receptive language, fine motor and physical growth indicators did not meaningfully differ between the children of prenatally non-depressed and depressed mothers, regardless of intervention arm.

Discussion

Our study showed that a peer-delivered intervention beginning in pregnancy with booster sessions through 36-months postnatal did not measurably affect a range of maternal depression symptom and child developmental outcomes. Though women in the ~~treatment~~ intervention arm did show greater convergence in depression symptoms with the prenatally non-depressed women at 36 months, relative to women in the control arm, evidence of a meaningful intervention effect is lacking. We also find only weak evidence that the prenatal depression episode was itself associated with child socioemotional outcomes and no evidence of associations with other developmental outcomes; for the most part, children of prenatally depressed and prenatally non-depressed mothers had similar outcomes at 36-months of age.

The overall small effect sizes and lack of statistical significance on maternal outcomes might be attributable to several factors. First, the intervention was a non-specialist, lay peer delivered psychosocial intervention. The lay peers were housewives from rural villages, without prior training or work experience. They were trained and supervised by non-specialists using a cascaded model (with no direct specialist contact).~~-~~[26] Second, this lay peer delivered approach was used to inform scaling-up of maternal mental health services through existing health systems and community resources. It is possible that this non-specialist, low-intensity design, coupled with longer-duration implementation, led to what Chambers and colleagues refer to as 'voltage drop' (the intervention loses some degree of its potency (or fidelity) when moving from efficacy to effectiveness in the real world) and 'program drift' (the intervention deviates from its manualized or implementation protocols).[41] We saw a substantial drop in women attending the maintenance group sessions delivered every two months beyond the 6th month postnatal period. This implementation challenge of attendance is reported in other community based interventions targeting maternal outcomes.[42] Women reported that the ~~women~~y lost interest in attending these group sessions or found it demanding on their time. This challenge of poor attendance and how to best address it came up regularly in peer supervision meetings. In addition to attempts to add more interesting content to the sessions, we made sure that community health workers reminded women about the sessions and followed up with those who missed a session. Addressing sustaining participant interest and limited time is seminal in community-based program success and has been highlighted as an important challenge in other low resource settings.[43] Finally, it is possible that, since the booster group sessions did not focus on specific strategies to address depression as mentioned earlier, the intervention arm was not so different from the EUC arm. There were no detectable treatment effects at the 6 month postnatal wave.[16]The boosters introduced after the 6 month~~s~~ mark did not change this - we continue to see no intervention effect at 36 months.

Maintaining the competence levels and motivation of the peers over the multi-year long implementation period was challenging. Perhaps the cascaded model of supervision via non-specialists led to dropped potency (or fidelity) of sessions. We explored the perceptions of the peers about this intervention in a nested qualitative study at six months postnatal.[26] We did not find any negative perceptions towards the intervention which might have led to change in motivation or competence levels.. The reduced frequency of supervision sessions of the peers from every month to every two months seems to have contributed to the drop in competence levels. This drop in supervisory intensity has been shown to reduce effectiveness of known approaches.[44,45] Another contributing factor to the drop in competence levels was the addition of new content, beyond the 6th month postnatal period: both the high and low scoring peers experienced a drop in competence levels when more content was

added. Finding the right balance of content vs capacity of peers, and maintaining fidelity, is an implementation challenge reported in other programs.[46,47]

Finally, we saw a substantial improvement in the control arm (lowered rates and higher recovery from depression). The control arm recovery rates were higher compared to our previous studies from 2008 and 2020.[8,48] This could be attributable to "regression to the mean" or to spillover of the intervention through the LHWs who regularly interact with LHWs responsible for women residing in control arm sites. The active control arm received enhanced usual care and some evidence suggests that informing people about their illness status improves outcomes,[49] which raises important methodological (study design) issues for trials that are embedded within community settings. Perhaps future trials, using similar active control arms ought to consider equivalence or non-inferiority trial designs to avoid a nonsignificant superiority trial being wrongly interpreted, as proof of no difference between the two active comparisons.[50]

We found no clear indications of sub-group differences. This points to the challenge of intervention targeting especially given that approximately one-fifth of women did not respond to this treatment, indicating a need for different interventions. A collaborative care model where non-responders can be detected earlier and connected to more specialized care may be needed.

The findings suggesting an association (albeit weak) between prenatal depression and child socioemotional outcomes at age 3, which is not mitigated by the intervention, mirrors results from our previous trial with a different sample, designed to examine a similar intervention's effect on children at age seven.[4] Children's socioemotional development may be less likely to 'catch up' during the resolution phase of depression within the postnatal period. If confirmed in other studies, differences in socioemotional, but not other developmental outcomes, may point to specific mechanisms in maternal-child intergenerational transmission of risk.[11]

While chronic depression likely has the greatest effect on child development,[51] the majority of children in our study were exposed to varying maternal depression levels, including periods of low or no depression symptoms over the study period. An association between the prenatal episode and child socioemotional outcomes at 36 months would be consistent with 'foetal programming,'[52] and is supported by evidence of linking prenatal depression to a number of child outcomes regardless of postnatal depression.[53] However, inconsistent with foetal programming, another study concluded that a reduction in maternal depression levels postnatally lead to 'near normal levels' of child behavioural problems.[54]

A possible explanation for the lack of intervention effects on child outcomes is that, unlike the trained specialist delivery in Stein et al's study, this intervention was delivered by lay peers in a low-income country.[54] As mentioned above, we experienced a 'voltage drop' over time in terms of reductions in both session frequency and attendance, posing significant challenges to sustained implementation.[41] Another potential explanation of no group differences in socioemotional outcomes is that the variability in maternal depression symptoms in the first 3 years could itself be a risk factor. Given that depression is

a chronic, recurring disorder, mothers who were prenatally depressed likely recovered and had a recurrence. Variation in symptoms over time would be larger among women who entered the study depressed compared to those non-depressed. It is possible that these children were exposed to inconsistent or unpredictable parenting, which has been linked with negative behavioral and socioemotional indicators.[55] This hypothesis is supported by results from prior work, showing a tendency toward worse child anxiety symptoms among children whose mothers relapsed when compared with those who had chronic depression.[4]

Our findings that receptive language, fine motor, and length- and weight-for-age Z scores were not a function of maternal prenatal depression status or whether she was treated with the intervention are consistent with our previous study and also with Tomlinson et al who did not observe intervention effects when looking at the continuous versions of the Bayley Scales and growth outcomes, although they did report a difference when outcomes were dichotomized.[4,6] These results suggest that a different set of pathways operates for these outcomes relative to socioemotional development. It is possible that the child's experience of maternal depression did not, on average, reach severe or chronic enough levels to affect language and fine motor development and growth. The possibility is consistent with prior literature that suggests that the most deleterious effects on child outcomes are from severe and persistent depression levels.[51] The presence of other family members, such as the grandmother, might also buffer any negative impacts of depression on the child. A complementary possibility is that the effects of maternal depression on these child outcomes are a function of baseline levels of other risk factors, such as illness, low maternal literacy, intimate partner violence, and others. In samples such as ours where a large fraction of women carry these risk factors, the effects of depression *per se* on child development may be overshadowed.

The study has limitations. The indicators of child socioemotional outcomes, although measured with validated and extensively used instruments, were mother-reported and, as such, susceptible to bias.We addressed this by utilizing only validated instruments used extensively globally. Additionally, if depression symptoms biased reporting, we would expect this to affect the overall socioemotional domain. However, in exploratory analyses we found that only specific socioemotional domains were predicted by prenatal depression; we have no reason to believe that reports of hyperactivity or conduct problems would be more influenced by depression than reporting of peer or emotional problems.

Overall, our results on the lack of intervention effectiveness for maternal depression and child development at 36 months suggest three potential scenarios. First, specific to the socioemotional domains, the intervention may have been too "light touch" to reverse the effect of prenatal depression exposure. The peer delivered version of this psychosocial intervention had a weaker effect on maternal depression than found in a previous study where community health workers were used to deliver the intervention instead of peer volunteers.[16,17] It is likely that women at risk of depression (and their children) need more than bimonthly group sessions delivered by peers for sustained changes to occur that will reduce depression's effect on children. A different model of delivery such as collaborative, or stepped care, merit consideration, with a number of interventions simultaneously targeting specific population needs, e.g. domestic violence, poverty alleviation, social services, social health protection Second, single screening in pregnancy for elevated depression symptoms may not be sufficient to identify the highest risk women, those who will go on to have the most chronic depression trajectories, with the worst effects on their children. Targeting women with a history of depression in combination with other risk factors, such as poverty or IPV, and tailoring the intervention for them, may yield stronger results. Finally, the program was delivered in a high poverty context. It thus remains possible that if the entire socio-political environment were to prioritize women and children's wellbeing and health, interventions such as ours would have more power to make a difference.

In sum, our findings suggest that prenatal depression may have persistent effects on the child's socioemotional skills that are not easily reversed by a psychosocial intervention. Future preventive and early intervention efforts might benefit from being higher intensity and target the highest risk mothers. Importantly, interventions need to be attuned to the social context and ideally implemented as part of a suite of health promoting policies that address social determinants of maternal and child health.

Data Sharing Statement

Per NIH guidelines, data from the project supported by the NIH will be made available 2 years after the end of the project, which will be May 2022.

Authors' contributions

JM and SS are joint first authors* and drafted the manuscript, to which all authors contributed extensively. All authors reviewed and approved the final version. JM, SS, LB, ET, AR, KO designed the study. IA, NA, SS, KO led the implementation of the fieldwork and instrument development/adaptation.

ET, JG designed the analytical database, planned and led the analysis. AZ designed and managed the data collection databases throughout the data collection process.VB, SoBh, PB, EC, AH, KL, and ES contributed to the analysis of the data. JM, SS, ET, LB, IA, AW, NA, AB, SZ, JG, VB, SB, PB, TB, SB, AH, AJ, MS, AZ, EC, KL, ES, KO, AR had full access to all the data in the study and had final responsibility for the decision to submit for publication

Conflict of Interest

We declare we have no conflict of interest.

References

1. Gelaye B, Rondon MB, Araya R, Williams MA. Epidemiology of maternal depression, risk factors, and child outcomes in low-income and middle-income countries. *The Lancet Psychiatry* 2016; **3**(10): 973-82.10.1016/S2215-0366(16)30284-X
2. Goodman SH, Rouse MH, Connell AM, Broth MR, Hall CM, Heyward D. Maternal Depression and Child Psychopathology: A Meta-Analytic Review. *Clin Child Fam Psychol Rev* 2011; **14**(1): 1-27.10.1007/s10567-010-0080-1
3. Patel V, Saxena S, Lund C, Thornicroft G, Baingana F, Bolton P, Chisholm D, Collins PY, Cooper JL, Eaton J, Herrman H, Herzallah MM, Huang Y, Jordans MJD, Kleinman A, Medina-Mora ME, Morgan E, Niaz U, Omigbodun O, Prince M, Rahman A, Saraceno B, Sarkar BK, De Silva M, Singh I, Stein DJ, Sunkel C, UnÜtzer J. The Lancet Commission on global mental health and sustainable development. *The Lancet* 2018; **392**(10157): 1553-98.10.1016/S0140-6736(18)31612-X
4. Maselko J, Sikander S, Bhalotra S, Bangash O, Ganga N, Mukherjee S, Egger H, Franz L, Bibi A, Liaqat R, Kanwal M, Abbasi T, Noor M, Ameen N, Rahman A. Effect of an early perinatal depression intervention on long-term child development outcomes: follow-up of the Thinking Healthy Programme randomised controlled trial. *Lancet Psychiatry* 2015; **2**(7): 609-17.10.1016/s2215-0366(15)00109-1
5. Rahman A, Fisher J, Bower P, Luchters S, Tran T, Yasamy MT, Saxena S, Waheed W. Interventions for common perinatal mental disorders in women in low- and middle-income countries: a systematic review and meta-analysis. *Bulletin of the World Health Organization* 2013; **91**(8): 593-601I.10.2471/BLT.12.109819
6. Tomlinson M, Rotheram-Borus MJ, Scheffler A, le Roux I. Antenatal depressed mood and child cognitive and physical growth at 18-months in South Africa: a cluster randomised controlled trial of home visiting by community health workers. *Epidemiol Psychiatr Sci* 2018; **27**(6): 601-10.10.1017/s2045796017000257
7. Goodman SH, Garber J. Evidence-Based Interventions for Depressed Mothers and Their Young Children. *Child Dev* 2017; **88**(2): 368-77.10.1111/cdev.12732
8. Baranov V, Bhalotra S, Biroli P, Maselko J. Maternal Depression, Women's Empowerment, and Parental Investment: Evidence from a Randomized Control Trial. *American Economic Review* 2020; **110**(3): 824-59
9. Gunlicks ML, Weissman MM. Change in child psychopathology with improvement in parental depression: A systematic review. *Journal of the American Academy of Child and Adolescent Psychiatry* 2008; **47**(4): 379-89.10.1097/CHI.0b013e3181640805
10. Cuijpers P, Weitz E, Karyotaki E, Garber J, Andersson G. The effects of psychological treatment of maternal depression on children and parental functioning: a meta-analysis. *Eur Child Adolesc Psych* 2015; **24**(2): 237-45.10.1007/s00787-014-0660-6
11. Goodman SH, Cullum KA, Dimidjian S, River LM, Kim CY. Opening windows of opportunities: Evidence for interventions to prevent or treat depression in pregnant women being associated with changes in offspring's developmental trajectories of psychopathology risk. *Development and Psychopathology* 2018; **30**(3): 1179-96.10.1017/S0954579418000536
12. Loechner J, Starman K, Galuschka K, Tamm J, Schulte-Körne G, Rubel J, Platt B. Preventing depression in the offspring of parents with depression: A systematic review and meta-analysis of randomized controlled trials. *Clinical Psychology Review* 2018; **60**: 1-14.https://doi.org/10.1016/j.cpr.2017.11.009
13. Milgrom J, Holt CJ, Bleker LS, Holt C, Ross J, Ericksen J, Glover V, O'Donnell KJ, de Rooij SR, Gemmill AW. Maternal antenatal mood and child development: an exploratory study of treatment effects on child outcomes up to 5 years. *J Dev Orig Health Dis* 2019; **10**(2): 221-31.10.1017/s2040174418000739

14. Kersten-Alvarez LE, Hosman CMH, Riksen-Walraven JM, van Doesum KTM, Hoefnagels C. Long-term effects of a home-visiting intervention for depressed mothers and their infants. *Journal of Child Psychology and Psychiatry* 2010; **51**(10): 1160-70.10.1111/j.1469-7610.2010.02268.x
15. Rotheram-Borus MJ, Arfer KB, Christodoulou J, Comulada WS, Stewart J, Tubert JE, Tomlinson M. The association of maternal alcohol use and paraprofessional home visiting with children's health: A randomized controlled trial. *J Consult Clin Psychol* 2019; **87**(6): 551-62.10.1037/ccp0000408PMC6775769
16. Sikander S, Ahmad I, Atif N, Zaidi A, Vanobberghen F, Weiss HA, Nisar A, Tabana H, Ain QU, Bibi A, Bilal S, Bibi T, Liaqat R, Sharif M, Zulfiqar S, Fuhr DC, Price LN, Patel V, Rahman A. Delivering the Thinking Healthy Programme for perinatal depression through volunteer peers: a cluster randomised controlled trial in Pakistan. *Lancet Psychiatry* 2019; **6**(2): 128-39.10.1016/s2215-0366(18)30467-x
17. Vanobberghen F, Weiss HA, Fuhr DC, Sikander S, Afonso E, Ahmad I, Atif N, Bibi A, Bibi T, Bilal S, De Sa A, D'Souza E, Joshi A, Korgaonkar P, Krishna R, Lazarus A, Liaqat R, Sharif M, Weobong B, Zaidi A, Zuliqar S, Patel V, Rahman A. Effectiveness of the Thinking Healthy Programme for perinatal depression delivered through peers: Pooled analysis of two randomized controlled trials in India and Pakistan. *J Affect Disord* 2019.https://doi.org/10.1016/j.jad.2019.11.110
18. Fuhr DC, Weobong B, Lazarus A, Vanobberghen F, Weiss HA, Singla DR, Tabana H, Afonso E, De Sa A, D'Souza E, Joshi A, Korgaonkar P, Krishna R, Price LN, Rahman A, Patel V. Delivering the Thinking Healthy Programme for perinatal depression through peers: an individually randomised controlled trial in India. *Lancet Psychiatry* 2019; **6**(2): 115-27.10.1016/s2215-0366(18)30466-8
19. Atif N, Bibi A, Nisar A, Zulfiqar S, Ahmed I, LeMasters K, Hagaman A, Sikander S, Maselko J, Rahman A. Delivering maternal mental health through peer volunteers: a 5-year report. *International Journal of Mental Health Systems* 2019; **13**(1): 62.10.1186/s13033-019-0318-3
20. National Institute of Population Studies - NIPS/Pakistan & ICF Intternational. Pakistan Demographic and Health Survey 2012-2013. Islamabad: NIPS/Pakistan and ICF International, 2012-3.
21. National Institute of Population Studies Islamabad. Pakistan Demographic and Health Survey. Islamabad, Pakistan: National Institute of Population Studies, 2013.
22. Turner EL, Sikander S, Bangash O, Zaidi A, Bates L, Gallis J, Ganga N, O'Donnell K, Rahman A, Maselko J. The effectiveness of the peer-delivered Thinking Healthy PLUS (THPP+) Program for maternal depression and child socioemotional development in Pakistan: study protocol for a randomized controlled trial. *Trials* 2016; **17**(1): 442.10.1186/s13063-016-1530-yPMC5017048
23. Gallis J, Maselko J, O'Donnell K, Song KE, Saqib K, Turner EL, Sikander S. Criterion-related validity and reliability of the Urdu version of the patient health questionnaire in community-based women in Pakistan. *PeerJ* 2018; **6**(e5185).https://doi.org/10.7717/peerj.5185
24. Sikander S, Ahmad I, Bates LM, Gallis J, Hagaman A, O'Donnell K, Turner EL, Zaidi A, Rahman A, Maselko J. Cohort Profile: Perinatal depression and child socioemotional development ; the Bachpan cohort study from rural Pakistan. *BMJ Open* 2019; **9**(5): e025644.10.1136/bmjopen-2018-025644
25. Sikander S, Lazarus A, Bangash O, Fuhr DC, Weobong B, Krishna RN, Ahmad I, Weiss HA, Price L, Rahman A, Patel V. The effectiveness and cost-effectiveness of the peer-delivered Thinking Healthy Programme for perinatal depression in Pakistan and India: the SHARE study protocol for randomised controlled trials. *Trials* 2015; **16**: 14.10.1186/s13063-015-1063-9
26. Atif N, Nisar A, Bibi A, Khan S, Zulfiqar S, Ahmad I, Sikander S, Rahman A. Scaling-up psychological interventions in resource-poor settings: training and supervising peer volunteers to deliver the 'Thinking Healthy Programme' for perinatal depression in rural Pakistan. *Global mental health (Cambridge, England)* 2019; **6**: e4.10.1017/gmh.2019.4PMC6521132
27. Organization. WH. WHO Guidelines Approved by the Guidelines Review Committee. mhGAP Intervention Guide for Mental, Neurological and Substance Use Disorders in Non-Specialized Health

672 Settings: Mental Health Gap Action Programme (mhGAP): Version 20. Geneva: World Health
673 Organization

674 Copyright (c) World Health Organization 2016.; 2016.
675 28. Organization" WH. WHO Disability Assessment Schedule II (WHO DAS II).
676 http://www.who.int/classifications/icf/whodasii/en/index.html (accessed May 7, 2010.
677 29. Gorman LL, O'Hara MW, Figueiredo B, Hayes S, Jacquemain F, Kammerer MH, Klier CM, Rosi S,
678 Seneviratne G, Sutter-Dallay AL. Adaptation of the structured clinical interview for DSM-IV disorders for
679 assessing depression in women during pregnancy and post-partum across countries and cultures. *Br J*
680 *Psychiatry Suppl* 2004; **46**(23): s17-23
681 30. Ahmad I, Suleman N, Waqas A, Atif N, Malik AA, Bibi A, Zulfiqar S, Nisar A, Javed H, Zaidi A, Khan
682 ZS, Sikander S. Measuring the implementation strength of a perinatal mental health intervention
683 delivered by peer volunteers in rural Pakistan. *Behaviour Research and Therapy* 2020:
684 103559.https://doi.org/10.1016/j.brat.2020.103559
685 31. Goodman R. The strengths and difficulties questionnaire: A research note. *Journal of Child*
686 *Psychology and Psychiatry and Allied Disciplines* 1997; **38**(5): 581-6
687 32. Samad L, Hollis C, Prince M, Goodman R. Child and adolescent psychopathology in a developing
688 country: testing the validity of the Strengths and Difficulties Questionnaire (Urdu version). *International*
689 *Journal of Methods in Psychiatric Research* 2005; **14**(3): 158-66.10.1002/mpr.3
690 33. Bayley N, Reuner G. Bayley scales of infant and toddler development: Bayley-III: Harcourt
691 Assessment, Psych. Corporation; 2006.
692 34. O'Donnell K, Murphy R, Ostermann J, Masnick M, Whetten RA, Madden E, Thielman NM,
693 Whetten K. A brief assessment of learning for orphaned and abandoned children in low and middle
694 income countries. *AIDS Behav* 2012; **16**(2): 480-90.10.1007/s10461-011-9940-zPMC3817622
695 35. Maselko J, Bates L, Bhalotra S, Gallis JA, O'Donnell K, Sikander S, Turner EL. Socioeconomic
696 status indicators and common mental disorders: Evidence from a study of prenatal depression in
697 Pakistan. *SSM - Population Health* 2018; **4**: 1-9.https://doi.org/10.1016/j.ssmph.2017.10.004
698 36. Kroenke K, Spitzer RL, Williams JBW. The PHQ-9 - Validity of a brief depression severity measure.
699 *Journal of General Internal Medicine* 2001; **16**(9): 606-13.10.1046/j.1525-1497.2001.016009606.x
700 37. Hedges LV. Effect Sizes in Cluster-Randomized Designs. *JEBS* 2007; **32**(4): 341-
701 70.10.3102/1076998606298043
702 38. Li P, Redden DT. Comparing denominator degrees of freedom approximations for the
703 generalized linear mixed model in analyzing binary outcome in small sample cluster-randomized trials.
704 *BMC Medical Research Methodology* 2015; **15**(1): 38.10.1186/s12874-015-0026-x
705 39. Gallis J, Turner EL. Relative measures of association for binary outcomes: Challenges and
706 recommendations for the global health researchers. *Ann Glob Health* 2019; **85**(1): 137
707 40. Gallis JA, Li F, Turner EL. xtgeebcv: A command for bias-corrected sandwich variance estimation
708 for GEE analyses of cluster randomized trials. *Stata Journal* (in press):
709 41. Chambers DA, Glasgow RE, Stange KC. The dynamic sustainability framework: addressing the
710 paradox of sustainment amid ongoing change. *Implementation science : IS* 2013; **8**: 117.10.1186/1748-
711 5908-8-117PMC3852739
712 42. Tandon SD, Ward EA, Hamil JL, Jimenez C, Carter M. Perinatal depression prevention through
713 home visitation: a cluster randomized trial of mothers and babies 1-on-1. *J Behav Med* 2018; **41**(5): 641-
714 52.10.1007/s10865-018-9934-7
715 43. Sondaal AEC, Tumbahangphe KM, Neupane R, Manandhar DS, Costello A, Morrison J.
716 Sustainability of community-based women's groups: reflections from a participatory intervention for
717 newborn and maternal health in Nepal. *Community development journal* 2019; **54**(4): 731-
718 49.10.1093/cdj/bsy017PMC6924535

44. Baqui AH, El-Arifeen S, Darmstadt GL, Ahmed S, Williams EK, Seraji HR, Mannan I, Rahman SM, Shah R, Saha SK, Syed U, Winch PJ, Lefevre A, Santosham M, Black RE. Effect of community-based newborn-care intervention package implemented through two service-delivery strategies in Sylhet district, Bangladesh: a cluster-randomised controlled trial. *The Lancet* 2008; **371**(9628): 1936-44.10.1016/S0140-6736(08)60835-1
45. Tripathy P, Nair N, Barnett S, Mahapatra R, Borghi J, Rath S, Rath S, Gope R, Mahto D, Sinha R, Lakshminarayana R, Patel V, Pagel C, Prost A, Costello A. Effect of a participatory intervention with women's groups on birth outcomes and maternal depression in Jharkhand and Orissa, India: a cluster-randomised controlled trial. *The Lancet* 2010; **375**(9721): 1182-92.10.1016/S0140-6736(09)62042-0
46. Yousafzai AK, Rasheed MA, Rizvi A, Armstrong R, Bhutta ZA. Parenting Skills and Emotional Availability: An RCT. *Pediatrics* 2015; **135**(5): e1247-e57.10.1542/peds.2014-2335
47. Hafeez A, Mohamud B, Shiekh M, Shah I, Jooma R. Lady health workers programme in Pakistan: challenges, achievements and the way forward. *JPMA* 2011; **61**(210):
48. Rahman A, Malik A, Sikander S, Roberts C, Creed F. Cognitive behaviour therapy-based intervention by community health workers for mothers with depression and their infants in rural Pakistan: a cluster-randomised controlled trial. *Lancet* 2008; **372**(9642): 902-9
49. Paterick TE, Patel N, Tajik AJ, Chandrasekaran K. Improving health outcomes through patient education and partnerships with patients. *Proc (Bayl Univ Med Cent)* 2017; **30**(1): 112-3.10.1080/08998280.2017.11929552
50. Lesaffre E. Superiority, equivalence, and non-inferiority trials. *Bulletin of the NYU hospital for joint diseases* 2008; **66**(2): 150-4
51. Netsi E, Pearson RM, Murray L, Cooper P, Craske MG, Stein A. Association of Persistent and Severe Postnatal Depression With Child OutcomesAssociation of Persistent and Severe Postnatal Depression With Child OutcomesAssociation of Persistent and Severe Postnatal Depression With Child Outcomes. *JAMA psychiatry* 2018; **75**(3): 247-53.10.1001/jamapsychiatry.2017.4363
52. Robinson R, Lahti-Pulkkinen M, Heinonen K, Reynolds RM, Räikkönen K. Fetal programming of neuropsychiatric disorders by maternal pregnancy depression: a systematic mini review. *Pediatric Research* 2019; **85**(2): 134-45.10.1038/s41390-018-0173-y
53. O'Donnell KJ, Glover V, Barker ED, O'Connor TG. The persisting effect of maternal mood in pregnancy on childhood psychopathology. *Development and Psychopathology* 2014; **26**(2): 393-403.10.1017/S0954579414000029
54. Stein A, Netsi E, Lawrence PJ, Granger C, Kempton C, Craske MG, Nickless A, Mollison J, Stewart DA, Rapa E, West V, Scerif G, Cooper PJ, Murray L. 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): 134-44.10.1016/s2215-0366(18)30006-3
55. Frankenhuis WE, Del Giudice M. When Do Adaptive Developmental Mechanisms Yield Maladaptive Outcomes? *Developmental Psychology* 2012; **48**(3): 628-42.10.1037/a0025629

Table

Table 1: Baseline characteristics of women in study population (n=1154)

Characteristics	Control Arm (EUC) (N = 287) N (%) or mean (SD)	Intervention Arm (THPP/THPP+) (N = 283) N (%) or mean (SD)	Prenatally Non-depressed (N = 584) N (%) or mean (SD)
Mother's age	27·29 (4·97)	26·80 (4·60)	26·37 (4·26)
18-24	83 (28·9%)	88 (31·1%)	203 (34·8%)
25+	204 (71·1%)	195 (68·9%)	381 (65·2%)
Mother's Education (in years)			
None (0)	55 (19·2%)	52 (18·4%)	63 (10·8%)
Primary (1-5)	71 (24·7%)	68 (24·0%)	87 (14·9%)
Middle/Secondary (6-12)	134 (46·7%)	145 (51·2%)	338 (57·9%)
Tertiary (13+)	27 (9·4%)	18 (6·4%)	96 (16·4%)
Mother's Education (in years)			
None (0)	55 (19·2%)	52 (18·4%)	63 (10·8%)
Any (1+)	232 (80·8%)	231 (81·6%)	521 (89·2%)
Nuclear Family	49 (17·1%)	50 (17·7%)	59 (10·1%)
Total Number of Children in the Household	2·97 (2·67)	3·09 (2·71)	2·51 (2·63)
Number of Children			
First pregnancy	72 (25·1%)	65 (23·0%)	212 (36·3%)
1 to 3	183 (63·8%)	180 (63·6%)	336 (57·5%)
4+	32 (11·1%)	38 (13·4%)	36 (6·2%)
SCID (MDE)	210 (73·2%)	218 (77·0%)	14 (2·4%)
PHQ-9 Total Score	14·48 (3·58)	14·89 (3·72)	2·80 (2·46)
Severity (PHQ-9 score)			
10-14	167 (58·2%)	145 (51·2%)	584 (100·0%)
≥15	120 (41·8%)	138 (48·8%)	0 (0·0%)
MSPSS Total Score	3·95 (1·33)	3·92 (1·41)	4·97 (1·01)
WHO-DAS Total Score	16·11 (9·12)	16·71 (8·52)	5·61 (6·46)
Duration of Depression (Chronicity)			
<12 weeks	38 (13·2%)	35 (12·4%)	
≥12 weeks	155 (54·0%)	171 (60·4%)	N/A
missing	94 (32·8%)	77 (27·2%)	
SES (Assets)			
Lowest Quintile	74 (25·8%)	85 (30·0%)	71 (12·2%)
Lower Middle Quintile	67 (23·3%)	71 (25·1%)	93 (15·9%)
Middle Quintile	55 (19·2%)	50 (17·7%)	126 (21·6%)
Upper Middle Quintile	47 (16·4%)	39 (13·8%)	145 (24·8%)
Upper Quintile	44 (15·3%)	38 (13·4%)	149 (25·5%)
SES (Assets index)			
Bottom 1/3rd	109 (38·0%)	119 (42·0%)	113 (19·3%)
Top 2/3rds	178 (62·0%)	164 (58·0%)	471 (80·7%)
Number of people per room	2·47 (1·87)	2·79 (2·03)	2·22 (1·79)
Life Events Checklist Score	4·10 (2·33)	4·70 (2·44)	2·90 (2·16)
Any IPV (last 12 months)	165 (59·1%)	178 (65·4%)	179 (33·0%)

Treatment Expectations			
None/somewhat	~~76 (26·5%)~~	~~70 (24·7%)~~	N/A
Moderate/very	~~211 (73·5%)~~	~~213 (75·3%)~~	N/A

Table 2. Primary and Secondary maternal and child trial outcomes at 36 months postnatal

	Control Arm (EUC) (N=216) mean (SD) or N (%)	Intervention Arm (N=206) mean (SD) or N (%)	Non-depressed (N=467) mean (SD) or N (%)	Adjusted Standardized Mean Difference (SMD) or adjusted RR (95% CI) for Intervention vs Control arm	ICC*
PRIMARY OUTCOMES					
Maternal PHQ-9 score	6·48 (6·25)	5·84 (5·80)	3.44 (4.53)	SMD= -0·13 (-0·33, 0·07)	0·009
Remission: PHQ-9 < 10 score	54 (75.0%)	51 (75.2%)	61 (86.9%)	1.00 (0.88, 1.13)	<0.001
Child SDQ-Total Difficulties	14·72 (6·13)	14·73 (6·04)	13·69 (6·34)	-0·10 (-1·39, 1·19)	0.020
SECONDARY OUTCOMES					
Maternal Major Depressive episode (SCID)	48 (22·2%)	34 (16·5%)	42 (9.0%)	RR= 0·67 (0·43, 1·05)	<0·001
Maternal Disability (WHO-DAS)	6·78 (9·44)	5·87 (8·06)	3.34 (6.19)	SMD= -0·12 (-0·33,0·09)	0·018
Child Bayley Scaled Receptive Score	9·98 (2·60)	10·42 (2·81)	10·41 (2·79)	0·38 (-0·19, 0·96)	0.027
Child Bayley Scaled Fine Motor Score	11·38 (4·12)	11·42 (4·05)	11·31 (3·99)	0·03 (-0·83, 0·90)	0.039

PHQ-9: Patient Healthy Questionnaire; SDQ: Strength and Difficulties Questionnaire; SCID: Structured Clinical Interview for DSM-IV; WHO-DAS: World Health Organization – Disability Assessment Schedule; SD: Standard Deviation; RR: Risk Ratio; ICC: Intra-cluster correlation. * ICC comes from the mixed models; for the binary outcomes the ICCs are based on mixed-effects logit models, although GEE was used to estimate the effects).

Figure 1. CONSORT sample flow chart

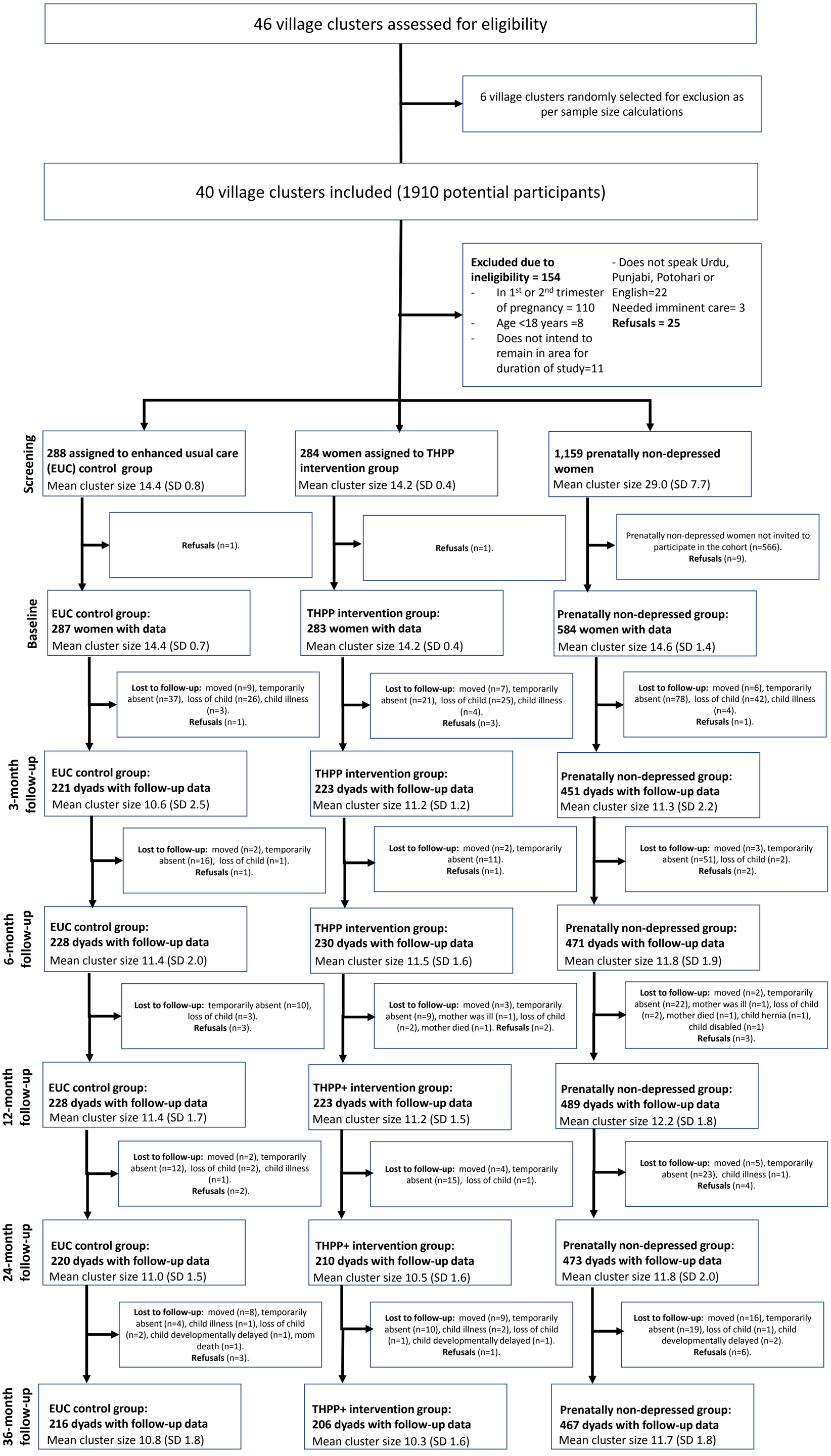

Figure 2

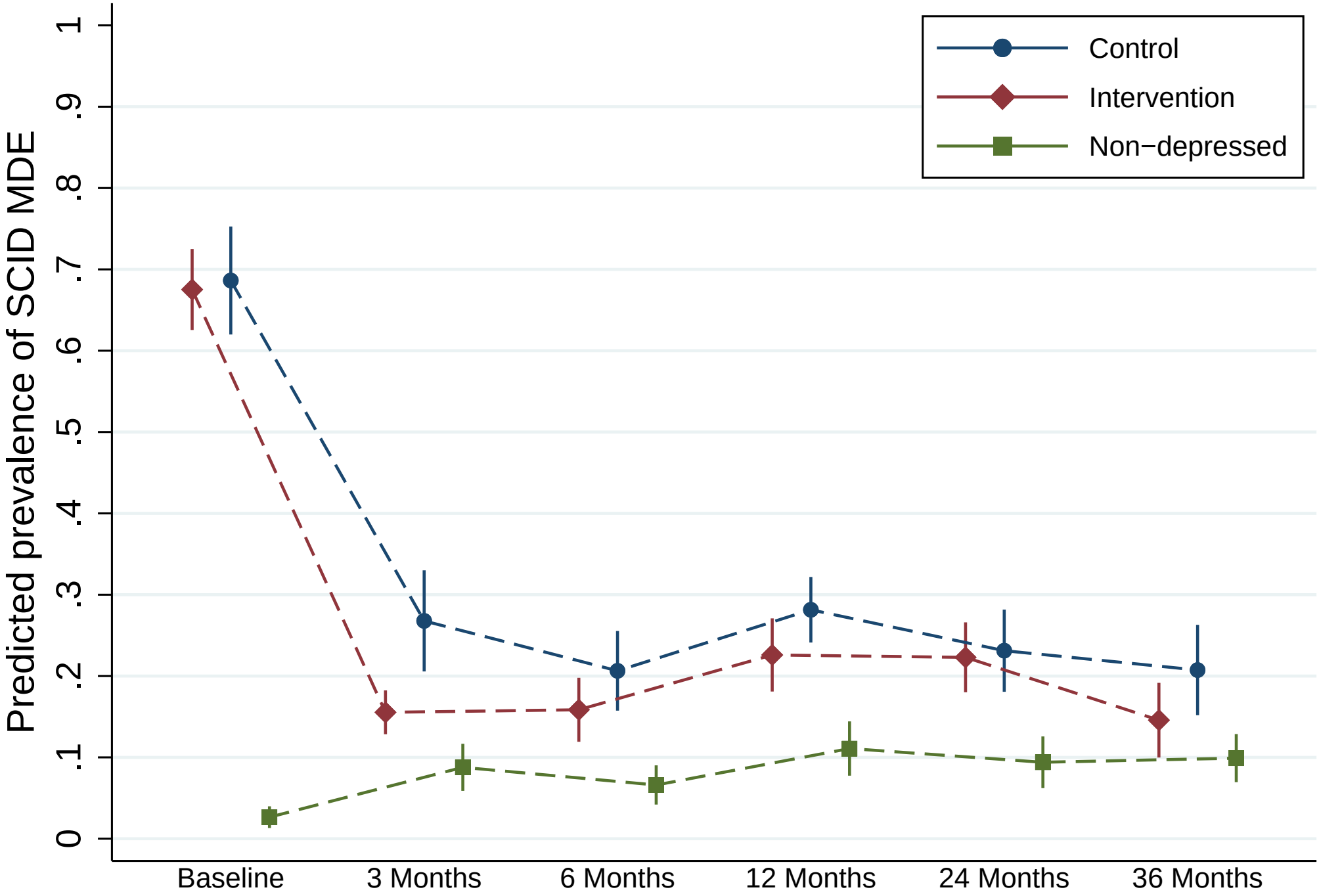
Predicted prevalence of SCID MDE
0
.1
.2
.3
.4
.5
.6
.7
.8
.9
1
Baseline
3 Months
6 Months
12 Months
24 Months
36 Months
Control
Intervention
Non−depressed

Figure 2 title

Figure 2: Depressive episodes (recurrence) across time points and comparison groups

Footnote:

Note:

*Abbreviations: SCID – Structured Clinical Interview for Depression; MDE – Major Depressive Episode.

**Predicted prevalences come from a longitudinal GEE model with SCID at all time points as the outcome. Time point, intervention arm (depressed in intervention, depressed in control, and non-depressed), and the interaction between these two variables are included in the model. In addition, the model is adjusted for the variables imbalanced at baseline or differential by missingness at any time point (see footnote to Table 3). An exchangeable working correlation matrix is used to take into account clustering by village cluster. Marginal predicted prevalences are computed from this model—at the average of continuous adjustors and the population percentages of categorical adjustors—and these are graphed by intervention arm.

Turner et al. Trials (2016) 17:442
DOI 10.1186/s13063-016-1530-y

Trials

STUDY PROTOCOL Open Access

The effectiveness of the peer delivered Thinking Healthy Plus (THPP+) Programme for maternal depression and child socio-emotional development in Pakistan: study protocol for a three-year cluster randomized controlled trial

Elizabeth L. Turner[1,2], Siham Sikander[3], Omer Bangash[3], Ahmed Zaidi[3], Lisa Bates[4], John Gallis[1,2], Nima Ganga[1], Karen O'Donnell[1], Atif Rahman[5*] and Joanna Maselko[6*]

Abstract

Background: The negative effects of perinatal depression on the mother and child start early and persist throughout the lifecourse (Lancet 369(9556):145–57, 2007; Am J Psychiatry 159(1):43-7, 2002; Arch Dis Child 77(2): 99–101, 1997; J Pak Med Assoc 60(4):329; J Psychosoma Res 49(3):207–16, 2000; Clin Child Fam Psychol Rev 14(1): 1–27, 2011). Given that 10–35% of children worldwide are exposed to perinatal depression in their first year of life (Int Rev Psychiatry 8(1):37–54, 1996), mitigating this intergenerational risk is a global public health priority (Perspect Public Health 129(5):221–7, 2009; Trop Med Int Health 13(4):579–83, 2008; Br Med Bull 101(1):57–79, 2012). However, it is not clear whether intervention with depressed women can have long-term benefits for the mother and/or her child. We describe a study of the effectiveness of a peer-delivered depression intervention delivered through 36 postnatal months, the Thinking Healthy Program Peer-delivered PLUS (THPP+) for women and their children in rural Pakistan.

Methods/design: The THPP+ study aims are: (1) to evaluate the effects of an extended 36-month perinatal depression intervention on maternal and index child outcomes using a cluster randomized controlled trial (c-RCT) and (2) to determine whether outcomes among index children of perinatally depressed women in the intervention arm converge with those of index children born to perinatally nondepressed women. The trial is designed to recruit 560 pregnant women who screened positive for perinatal depression (PHQ-9 score ≥10) from 40 village clusters, of which 20 receive the THPP+ intervention. An additional reference group consists of 560 perinatally nondepressed women from the same 40 clusters as the THPP+ trial. The women in the nondepressed group are not targeted to receive the THPP+ intervention; but, by recruiting pregnant women from both intervention and control clusters, we are able to evaluate any carryover effects of the THPP+ intervention on the women and their children. Perinatally depressed women in the THPP+ intervention arm receive bimonthly group-based sessions. Primary outcomes are 3-year maternal depression and 3-year child development indicators. Analyses are intention-to-treat and account for the clustered design.
(Continued on next page)

* Correspondence: atif.rahman@liverpool.ac.uk; jmaselko@unc.edu
[5]Institute of Psychology, Health and Society, University of Liverpool, Liverpool, UK
[6]Department of Epidemiology, Gillings School of Public Health, University of North Carolina Chapel Hill, Chapel Hill, USA
Full list of author information is available at the end of the article

(Continued from previous page)

Discussion: This trial, together with the reference group, has the potential to further our understanding of the early developmental lifecourse of children of both perinatally depressed and perinatally nondepressed women in rural Pakistan and to determine whether intervening with women's depression in the perinatal period can mitigate the negative effects of maternal depression on 36-month child development.

Trial registration: THPP-P ClinicalTrials.gov Identifier: NCT02111915 (registered on 9 April 2014).
THPP+ ClinicalTrials.gov Identifier: NCT02658994 (registered on 21 January 2016).
Sponsor: Human Development Research Foundation (HDRF).

Keywords: Thinking healthy program, Psychological treatment, Peer volunteers, Nonmental health professionals, Perinatal depression, Maternal depression, Task-shifting, Randomized trials, Low- and middle-income countries, Child development

Background

Perinatal maternal depression, defined by at least one depressive episode during pregnancy and/or the first postnatal year, has been shown to have negative health effects for both the mother and the child. Negative effects on the mother include reductions in daily functioning as well as early mortality. Negative effects on the child, including illness and poor growth, start early and persist throughout the child's life [1–6]. Given that 10–35% of children worldwide are exposed to perinatal depression in their first year [7], mitigating this intergenerational risk is a global public health priority [8–10]. Pakistan has one of the highest rates of maternal depression globally, and one of the only studies examining potential long-term benefits of maternal depression interventions on child outcomes found no significant effects [11].

The Thinking Health Program (THP), a community health worker (CHW)-delivered intervention developed and evaluated in Pakistan, was shown to have beneficial effects on both perinatal maternal depression and short-term child outcomes including reductions in diarrheal episodes and increased vaccination rates [12]. In 2015, the THP was formally designated by the World Health Organization (WHO) as an evidence-based intervention that could be implemented in a variety of global settings using an established CHW healthcare delivery system [13]. Unfortunately, many CHW systems, such as Pakistan's, are underfunded and stretched to capacity; and alternative delivery methods are required. In response to this need, the Thinking Health Program Peer-delivered (THPP) was developed by adapting the THP to be delivered primarily by peers who operate within the existing CHW system. An ongoing study, the THPP-Pakistan trial [14], seeks to evaluate THPP for 6 postnatal months.

Although effective in reducing maternal perinatal depression, our recent work failed to show that the 6-month CHW-led THP led to improved longer-term child outcomes [11]. At age 7 years, children of perinatally depressed mothers who received the intervention did not show better outcomes than children of control group mothers. To improve the longer-term outcomes of both perinatally depressed mothers and their children, we have developed the Thinking Health Program Peer-delivered PLUS (THPP+), an extension of the 6-month THPP intervention delivered at a lower intensity for an additional 30 postnatal months to the same women who have been receiving the THPP. The THPP+ is an extension and a continuation of the THPP intervention for mothers until the child is 3 years old.

The aim of this manuscript is to describe the protocol for the THPP+ study in Pakistan. The THPP+ study is a cluster randomized controlled trial (c-RCT), which compares outcomes among three groups of mother-child dyads: (1) those receiving the intervention, (2) those receiving Enhanced Usual Care (EUC) in the control clusters, and (3) a reference group of mother-child dyads in which the woman was not depressed in pregnancy and resides in the same intervention and control clusters where the trial is being implemented. Focusing on outcomes at 36 postnatal months, the goal of this c-RCT is to evaluate the cumulative effectiveness of the combined THPP and THPP+ interventions on mothers and their children. The goal of the embedded reference group of perinatally nondepressed women and their children is two-fold: (1) to evaluate whether the intervention is able to meaningfully reduce the gap in child outcomes that is traditionally observed when comparing children of depressed and nondepressed mothers; and (2) to determine whether there are any beneficial carryover effects of the intervention on this nondepressed group.

This manuscript complements and extends the THPP trial protocol [14]. To ensure that the current protocol is able to stand alone, we present the necessary key features of the THPP design and the ways in which the THPP+ trial builds on, and is different from, the ongoing THPP trial in Pakistan.

Objectives and hypotheses

The primary objective of the study is to evaluate the impact of a 36-month perinatal peer-delivered community-based perinatal depression intervention on (1) maternal depression and (2) child development. Our primary hypothesis for the perinatally depressed mothers is that the intervention will result in lower prevalence of depression at 3 years postnatal. Our primary hypothesis for the children is that the perinatal depression intervention will lead to improved developmental outcomes (see "Measures and constructs" in Table 1) at 3 years of age. Additional child hypotheses address proposed mediators and moderators of the effects of the perinatal depression intervention on child outcomes.

The second objective is to determine whether outcomes of perinatally depressed mothers and children in the intervention arm will converge to those in the reference group of perinatally nondepressed mothers and children as well as, secondarily, to determine whether there are any carryover effects of the intervention to benefit perinatally nondepressed mothers and children.

Methods/design

Trial settings

The study will be conducted in rural Pakistan in the rural Sub-District of Kallar Syedan, Rawalpindi, Pakistan.

Design

The THPP+ trial is a stratified cluster randomized controlled trial (c-RCT) of 40 village clusters allocated in a 1:1 ratio to receive intervention or EUC within 11 strata defined by Union Councils (sub-district units), each with an even number of village clusters [15]. Cluster randomization is used to avoid contamination between women since the THPP intervention is delivered at the community level through CHWs and peer women in the community. Stratification is used to minimize imbalance in baseline covariates.

THPP+ is conducted in the same 40 village clusters as the THPP trial. The same study population of perinatally depressed women is invited to consent to participate in THPP+. An equal number of perinatally nondepressed women are also recruited from each village cluster. The latter forms the reference group that enables us to evaluate whether convergence of maternal and child outcomes occurs during the 3-year postnatal period. In summary, all depressed women enrolled in the THPP+ trial were enrolled in THPP, while all nondepressed women are only recruited to the THPP+ study. See Fig. 1 for details of the distinction.

In brief, the ongoing THPP trial focuses on the effects of the THPP intervention on maternal outcomes at 6 postnatal months, with a limited number of child outcomes measured. The THPP+ protocol is designed to recruit the same 560 pregnant women who screen positive for perinatal depression from the 40 village clusters described above for the THPP trial, of which 20 clusters receive the THPP intervention delivered by trained lay peer volunteers.

Participants and procedures

Figures 1 and 2 show recruitment and flow of both the perinatally depressed and perinatally nondepressed mother-child dyads through the study. After collecting prebirth baseline information, we assess each mother and her index child born during the study at 3, 6, 12, 24 and 36 postnatal months. The 3- and 6-month assessments will coincide with those of the THPP trial. The 12-, 24- and 36-month assessments are unique to THPP+ (further details in Additional file 1).

Recruitment of study participants: inclusion and exclusion criteria

The current THPP+ study consists of perinatally depressed women and the index children who are participating in the THPP study and an additional sample of nonperinatally depressed mothers and of the index child

Table 1 Primary outcome measures for women and children in the Thinking Health Program Peer-delivered PLUS (THPP+)

	Source of data	Postnatal months				
Outcomes	Measure	3	6	12	24	36
Mother: depression	Patient Health Questionnaire (PHQ-9)			✓	✓	✓
	WHO Disability Assessment Schedule (WHO-DAS)			✓	✓	✓
Child: socioemotional	Total Difficulties score from the Strengths and Difficulties Questionnaire (SDQ-TD)					✓
	Ages and Stages Questionnaire (ASQ)		✓	✓	✓	✓
Child: developmental milestones	Bayley Scales of Infant and Toddler Development III (BSITD-III)			✓	✓	✓
Child: physical	Length, weight (WHO weight-for-length *z*-scores)	✓	✓	✓	✓	✓
	Head circumference	✓	✓	✓	✓	✓
	Diarrhea/ARI		✓	✓	✓	✓

ARI acute respiratory infection, *WHO* World Health Organization

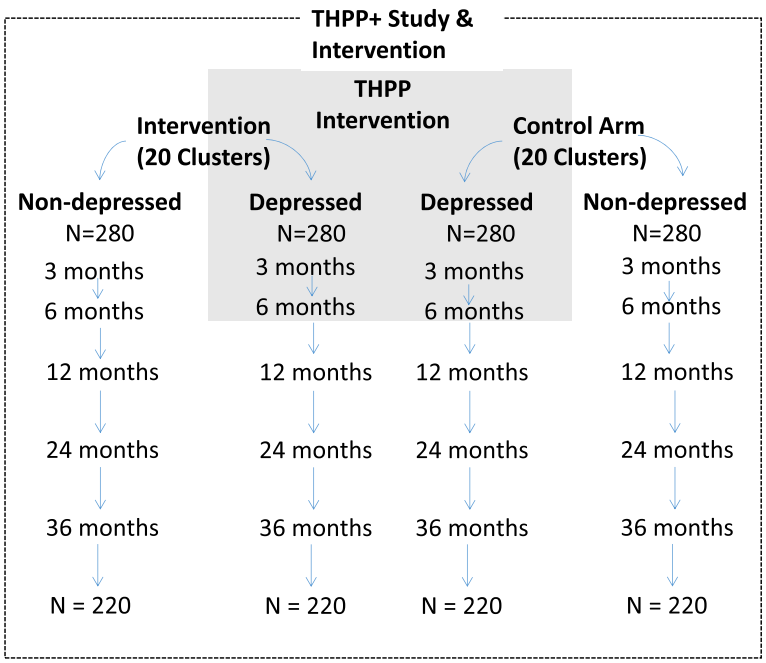

Fig. 1 THPP+ study (both the trial of perinatally depressed mothers and the reference group of perinatally nondepressed mothers) in relation to the THPP intervention trial showing anticipated loss to follow-up. *N* = number of mother-child dyads. Unshaded area is unique to the THPP+ study

of each mother. For THPP, pregnant women registered with the CHW (called Lady Health Workers) were approached. The study team has been engaged with the Lady Health Workers and the community in the past and enrollment rates have been consistently high. All eligible women in their third trimester of pregnancy were assessed for depression using the Patient Health Questionnaire (PHQ-9) and those scoring above the 10-point cutoff were invited to participate in the trial. For THPP, a random sample of approximately a third of women scoring less than 10 (i.e., screening negative on PHQ-9) are asked to serve as an additional reference group of equal size as the number of perinatally depressed women. In order to be eligible to participate, women need to be married, to reside in the study area, to understand one of the study languages (Urdu, Punjabi or Potohari), and to not require immediate medical attention. Following a live birth, the mother-infant dyads remain eligible to continue in the study unless the woman develops a psychotic or manic episode, or the dyad is broken through death, disability or relocation of the woman or child. Any participant who develops severe symptoms over the course of the study will be immediately referred for additional treatment.

Informed consent

Women are informed about the study goals and study design in the third trimester of pregnancy by trained research staff. Those who agree to participate consent to be followed up for 3 years postnatally and to participate in an intervention if they screen positive for perinatal depression. This consent covers the THPP+ period. The additional THPP+ sessions are seamlessly added to the existing intervention content for depressed women in the intervention clusters.

Randomization

The current THPP+ study is designed to maintain the randomization that was performed at the start of the THPP. According to the randomization procedure, 11 UC strata were selected with an even number of village clusters identified in each. Within each UC, village clusters were then randomized in a 1:1 ratio. In total, there are 20 intervention and 20 control arm clusters.

Interventions

Thinking Healthy Program Peer-delivered (THPP)

The Thinking Healthy Program Peer-delivered (THPP) is an adaptation of the Lady Health Worker-delivered THP, that was adopted by the WHO mhGAP Series [13]. The protocol for the THPP trial has been published [14]. Similar in content to the THP, the peer-delivered version is simplified with additional strategies added for ease of implementation by peers. The intervention focuses on identifying and altering unhealthy behaviors with a focus on behavioral activation to facilitate change. It consists of both individual sessions with the peers as well as

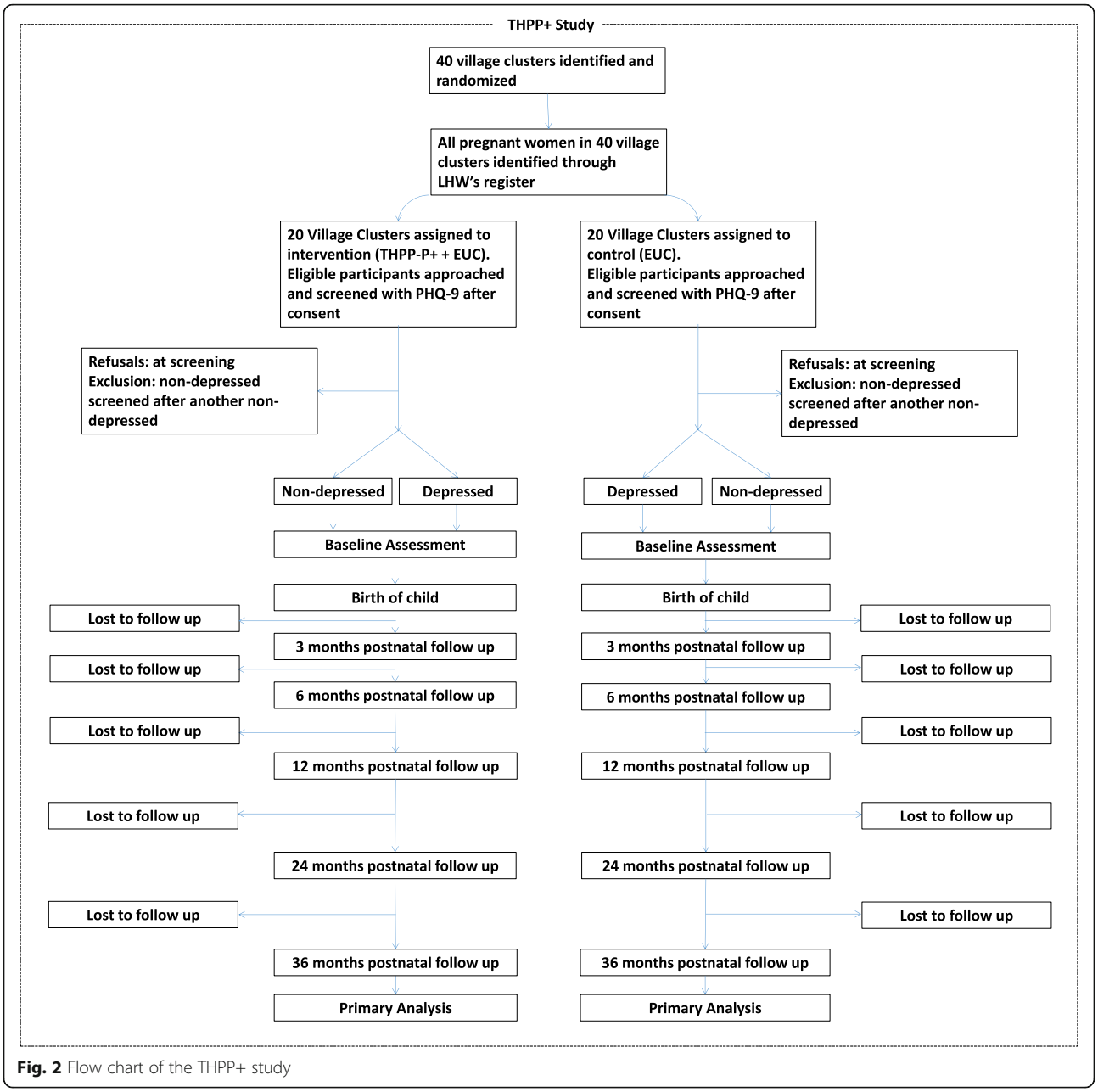

Fig. 2 Flow chart of the THPP+ study

group sessions held at the "Health House," a room in the home of the CHW dedicated for women's group meetings. The THPP begins in the third trimester of pregnancy and finishes at the end of the sixth postnatal month.

Thinking Healthy Program Peer-delivered PLUS (THPP+)

As part of THPP+, the intervention continues from the beginning of the seventh postnatal month through the end of the 36th month, and consists of an additional 30 months of lower-intensity services unique to the THPP+ model. We use the term THPP+ to refer to the combined 6-month THPP intervention and the 30-month THPP+ intervention delivered consecutively through 36 months after the index child's birth. The THPP+ includes group sessions to be held roughly every other month for a total of 18 sessions over the intervention duration. The content is a continuation of the previous THPP sessions with emphasis on self-care and on the baby's health and development. In order to ensure continued participation, peers contact each woman a week prior to the group session and the groups are held in the community where the participants live and are easily accessible. Peers keep session logs which are overseen by the peer supervisors and can be used for calculation of "dose" during analyses.

In case a woman misses attending a session the peer follows up at the household to work out/negotiate with the family to ensure attendance at the next session (so that the "dose" is not missed).

Although the perinatally nondepressed women in the intervention arm do not receive the THPP+ intervention, by recruiting mothers from both intervention and control clusters, we are able to evaluate any effects of the THPP+ on the group that is not directly targeted.

Enhanced Usual Care

Women in the control clusters who were depressed prenatally have been receiving Enhanced Usual Care (EUC). At the time of the screening (and with consent), women, their Lady Health Workers and personnel in their local primary health care facility were informed of the diagnosis; and women were given an information sheet about depression and how to access care. There are no new EUC protocols put in place postnatally as part of the THPP+.

Additional training and supervision of peers

For the THPP, peers were trained in a 5-day classroom-based workshop, followed by a 2-month internship during which they practiced the content of the THPP on nontrial participants [14]. For THPP+, peers will receive an additional 2 days of classroom training after their last session (during the fifth postnatal month) of the THPP to cover the additional content. Competency is assessed by role plays. Peer counselors continue to receive monthly group supervision to maintain high motivation and to address any challenges in the field.

Minimization of contamination

Risk of new contamination between the treatment and control arms is expected to be very low given the low intensity of the intervention and its placement after the end of the more intensive intervention that began prenatally and lasted through to 6 postnatal months. The cluster design makes it less likely that women will exchange information related to the intervention.

Masking of treatment allocation

Although it is not possible to blind study participants from their treatment arm allocation, all project staff, including interviewers, are blind both to a woman's original depression status and to the treatment arm of the village cluster in which she resides. Study participants are instructed to not discuss their depression status or intervention (or lack thereof) with the assessors. The data linking each village cluster with treatment allocation status is kept separate from the remaining outcome dataset until the time of the final analysis.

Fidelity of the intervention

Fidelity of the intervention is assessed through documenting the number of women who attend the meetings in combination with documenting the content covered during the meetings and the duration of each component covered in the session.

Data management

All data capture is performed electronically on tablets and uploaded daily to the main server. Quality checks for consistency, accuracy, missing data and other irregularities are conducted weekly. Any issues are shared with the research team and discussed during a weekly staff meeting to address source of any problems in the field. Data are backed up daily. Data are deidentified/anonymized before being shared with coinvestigators outside of the Human Development Research Foundation (HDRF). At all stages, data are password-protected with multiple layers of authorization.

Outcome evaluation

The primary endpoint is designed to be at 3 years postnatally. The primary comparison tested is between perinatally depressed-intervention versus perinatally depressed-control women in order to evaluate the effectiveness of the THPP+ intervention on long-term outcomes in perinatally depressed mothers and their children born during the intervention period (i.e., "index child"). Secondary comparisons for mothers and their index children are (1) intervention perinatally depressed mothers versus control perinatally nondepressed mothers to assess convergence of outcomes in both mothers and children and (2) intervention perinatally nondepressed versus control perinatally nondepressed mothers to assess whether there are any carryover effects of the intervention that benefit perinatally nondepressed mothers and their index children. For the former, the statistical goal is to demonstrate equivalence of outcomes of control perinatally nondepressed and intervention perinatally depressed mothers and their children. For the latter, the goal is to test the null hypothesis of no difference between groups in the outcomes of interest. The mother and child outcome measures are detailed in Tables 1 and 2.

Mother outcome measure

Patient Health Questionnaire (PHQ-9)

The PHQ-9 is the main indicator of depression symptoms among the women in the study. The PHQ-9 inquires about frequency of depressive symptoms in the last 2 weeks. It has been validated and used extensively in the region [16, 17].

Table 2 Outcome assessments

Instrument	Description	Outcome	Contextual validity
PHQ-9	Nine-item questionnaire assessment of depressive symptoms assessed on a scale of 0 to 3	Prevalence of moderate–severe depression; mean total score	Validated in primary care [37]
WHO-DAS	12-item questionnaire for measuring functional impairment over the last 30 days. In addition, two items assess the number of days the person was unable to work in these 30 days	Total disability score; quality-adjusted life years; number of days out of work	Validated for international use [18]
SDQ-TD	The SDQ is a parent report of 25 child attributes divided into five subscales: emotional symptoms, conduct problems, hyperactivity, peer problems and prosocial behavior	Total Difficulties score: calculated based on four subscales (except prosocial behavior)	The SDQ has previously been translated into Urdu as well as at least 50 other languages and used in low- and middle-income countries [21–23]
ASQ	The ASQ is a widely used, simple set of 30 questions appropriate for 4–60 month-olds that assesses five domains of development	The total score from the five domains, plus the score from an additional domain on the child's socioemotional development	The parent-report-based ASQ assessments have been shown to have good concurrent validity with professionally administered BSITD [24, 38], including internationally [39, 40]
BSITD-III	An individually administered assessment of the child's achievement of developmental milestones across five areas: cognitive, language, motor, social-emotional and adaptive skills [27]	The total score from each domain	The standard scores are derived from the US norms; and, because there are no available Pakistani norms, the scores provide a metric with which to compare groups of children in this Pakistan setting relative to the study hypotheses

ASQ Ages and Stages Questionnaire Socio-Emotional scale, *BSITD-III* Bayley Scales of Infant and Toddler Development, Third Edition, *PHQ* Patient Health Questionnaire, *SDQ-TD* Strengths and Difficulties Questionnaire, *WHO-DAS* WHO Disability Assessment Schedule

WHO Disability Assessment Schedule (WHO-DAS)

The WHO-DAS is a 12-item questionnaire assessing levels of function over the last 30 days. Combined with two items about one's ability to work in the last 30 days, the WHO-DAS generates a total disability score, quality-adjusted life years and number of days the respondent is not able to work [18].

Child outcome measures

Socioemotional development

Our main outcome measure is the Total Difficulties (TD) score derived from the Strengths and Difficulties Questionnaire (SDQ). The SDQ is a parent report of 25 child attributes divided into five subscales: emotional symptoms, conduct problems, hyperactivity, peer problems and prosocial behavior [19]. The TD score is calculated based on four subscales (except prosocial behavior) with a score range of 0–40 points [20]. The SDQ has previously been translated into Urdu as well as at least 50 other languages and used in low- and middle-income countries [21–23].

ASQ

Socioemotional developmental milestones, prior to and including 36 months, are assessed with the Ages and Stages Questionnaire Socio-Emotional scale (ASQ-SE) [24, 25]. The ASQ is a widely used, simple set of 25 questions where parents are asked to report age-appropriate milestones with the help of simple examiner-administered examples, such as whether, at 8 months, the child plays with a toy by banging it up or down on the floor or table [26].

Infant developmental milestone achievement

Bayley Scales of Infant Development The Bayley Scales of Infant and Toddler Development, Third Edition (BSID-III) is an individually administered assessment of the child's achievement of developmental milestones across five areas: cognitive, language, motor, social-emotional and adaptive skills [27]. The evaluations are conducted in the family's home at infant ages 12, 24 and 36 months. Raw scores in each domain are summarized by chronological age-related scaled scores and composite scores for each domain. The standard scores are derived from the US norms; and, because there are no available Pakistani norms, the scores provide a metric with which to compare groups of children in this Pakistan setting relative to the study hypotheses. The evaluators were trained in administration of the BSID-III by the team clinical psychologist (O'Donnell, US-based) and by the local team, which includes a psychiatrist and a physician. Periodic quality assurance is assessed at least quarterly by dyadic testing (evaluator plus team psychologist) and by double scoring by the US-based psychologist.

Physical development

Physical development is assessed using weight-for-age and height-for-age. Weight-for-age is sensitive to weight change over a short time period but fails to distinguish

tall, thin children from those who are short with adequate weight. Height-for-age is useful for identifying children with short stature, a group often vulnerable to longer-term adverse conditions. Based on WHO norms, a measure of 2 standard deviations (SD) below the mean of either weight or height is chosen to indicate poor growth. Head circumference is measured through 24 months. Physical health indicators are recent diarrheal episodes and acute respiratory infections.

Power calculations

The primary power calculations for the THPP+ study are for the c-RCT comparisons of perinatally depressed women and their children in the control versus intervention arms at 36 postnatal months at the 5% two-tailed significance level. As for the THPP trial [14] we assume 40 village clusters randomized in a 1:1 allocation ratio within 11 UCs, with 14 perinatally depressed women per village cluster, to yield a total sample size of 560 perinatally depressed women at baseline. In addition, for THPP+ we recruit 14 perinatally nondepressed women per village cluster for a total of 560 perinatally nondepressed women at baseline. We conservatively estimate that loss to follow-up (including infant mortality and maternal illness and death) of both perinatally depressed and perinatally nondepressed women at 36 months will be 20% (anticipated loss to follow-up in the THP trial was 10% at 6 months and most loss to follow-up is expected in the first 6 months of the study) [12]. Therefore, the total sample size available at 36 months is anticipated to be 480 perinatally depressed and 480 perinatally nondepressed women and their children. Using a standard formula [28, 29] for a cluster randomized design and assuming an intracluster correlation of 0.07 in the intervention arm and 0.05 in the control arm, the trial will have 90% power at 36 months to detect a difference in perinatally depressed remission of 65% in the perinatally depressed-intervention versus 45% in the perinatally depressed-control for the anticipated total sample size of 480 perinatally depressed women at 36 months. For child outcomes, this sample size will yield power of more than 90% to detect a difference between arms in mean TD score (range 0–40) of 3 points for children of perinatally depressed mothers using plausible estimates for intracluster correlations of 0.04–0.08 [12], and 5.2 for SD for the TD score among 3 year-olds [30].

Secondary comparisons mainly focus on child outcomes and are well-powered. For the secondary hypothesis of equivalence between children of perinatally depressed mothers in the intervention arm and perinatally nondepressed mothers in the control arm, we will conclude equivalence if the 95% confidence interval (CI) for the difference between the mean score in the two groups lies between −2 and 2 units. We note that differences of 1.0–2.0 points are often observed between boys and girls [30, 31]. With 220 children in each group and conservatively assuming an overall significance level of 2.5% (corresponding to the 95% CI), an SD of 5.2 and an ICC of 0.04, and no difference between the groups, we will have 83% power to conclude equivalence [28, 32]. For the secondary research question of the community benefit (i.e., carryover) of the intervention for perinatally nondepressed women and their children, we will have 80% power to detect a 1.7 or greater impact of the intervention on mean TD score (groups: perinatally nondepressed-intervention versus perinatally nondepressed-control, Fig. 1) for the same assumptions of the primary comparison above.

Analysis

Statistical analysis will be conducted according to the Consolidated Standards of Reporting Trials (CONSORT) guidelines. A flow chart will show participation of both perinatally depressed and perinatally nondepressed mothers and their children from recruitment in the third trimester through to 36 postnatal months (Fig. 2). Withdrawals and loss to follow-up will be shown at each follow-up (3, 6, 12, 24 and 36 postnatal months). Baseline characteristics of recruited mothers will be reported by study arm, and separately for perinatally depressed and perinatally nondepressed mothers. Continuous variables will be summarized by means and standard deviation (SD), or medians and the 25th and 75th percentile, if needed. Categorical variables will be summarized by counts and percentages.

The primary analyses are designed as intention-to-treat and will be conducted using the latest release of Stata software. Separate outcome analysis will be conducted for mothers and for children. In both cases, data from perinatally depressed and perinatally nondepressed participants will be analyzed jointly using generalized linear mixed-effects models so that all comparisons of interest can be estimated from the same model. The identity link will be used for continuous outcomes in order to estimate differences in mean outcomes. The log-link will be used for binary outcomes in order to estimate prevalence ratios, but if convergence is not achieved we will use the logit link from which prevalence ratios will be estimated. Random intercepts for cluster will be included to account for the clustered study design. For outcomes measured at multiple follow-up time points (e.g., for depression status in both perinatally depressed and perinatally nondepressed mothers, which will be evaluated at all five follow-up time points), random intercepts for person will be added to account for correlation of repeated measures on person. Similarly, in this case, random slopes for both cluster and participant will be considered to allow for heterogeneity

by cluster and participant over time. All random error terms will be assumed independent and zero-mean normally distributed.

Primary analyses of outcomes measured at a single follow-up time point are designed to include the following fixed-factor variables: arm (intervention versus control), strata (11 Union Councils), baseline depression status (perinatally depressed versus perinatally nondepressed) and its interaction with arm. For outcomes measured at multiple follow-up time points, the interactions between study arm, follow-up time point and baseline depression status will be included to allow for different intervention effects at each follow-up time point. Estimates of the prespecified comparisons of interest will be derived from the fitted model. Conclusions about the equivalence of perinatally depressed-intervention and perinatally nondepressed-control will be based on whether the corresponding 95% CI is contained within the equivalence margins (i.e., −2 to 2 for the primary child outcome of the TD score). Model assumptions will be assessed; in the case of non-normally distributed residuals, we will consider bootstrapping or transformations to obtain valid CIs.

Secondary analyses will include any baseline covariates for which there was chance baseline imbalance and for any additional baseline covariates that predict missing outcome data. Under the assumption that those covariates explain the missing data mechanism, we will obtain valid estimates of the intervention effects using the complete case data (i.e., without the need for imputation or an alternative method) [33]. If there are concerns or evidence that covariates cannot explain the nature of the missingness (i.e., if the data are missing not at random), we will perform a series of sensitivity analyses based on the pattern mixture approach [34].

Moderator and mediator analyses

In addition to our main outcomes, auxiliary analyses focus on potential moderators and mediators of any main associations. A-priori variables that might impact the degree to which the intervention affects depression symptoms include socioeconomic status, household composition, and the presence of interpersonal violence. These associations will be examined by including an interaction between the variable of interest and the intervention indicator in the primary outcome model. Potential mediators of interest include maternal responsiveness, the mother-child relationship and social support.

Compliance analysis

We plan to gather information on compliance with the intervention and evaluate whether there is any evidence of contamination between treatment arms.

Trial management

Trial monitoring procedures are a continuation of procedures and infrastructure in place for the THPP. This includes oversight by two committees: the Trial Management Committee (TMC), which is charged with close monitoring of all aspects of the trial and its progress and the Trial Steering Committee (TSC), which will provide additional guidance on the overall trial protocols as well as oversee trial safety issues. The TMC is composed of the principal investigators and the site team (project director, data manager/trial manager, local outcome assessment trainer); it meets weekly. The TSC is composed of the principal investigators, study coinvestigators, the trial manager and the study statistician; the TSC will meet every 6 months.

Ethical considerations

We protect the confidentiality of personal data principally through procedures to separate study data and participant identifiable data. Quantitative data gathered with the tablets for each participant at baseline retain personal identification items to minimize errors in transcribing identities, but these will be removed before transferring the data to Stata for analysis. We monitor the occurrence of a number of specific serious adverse events (SAEs) beyond the THPP trial (among the depressed cases); these include death of the participant or her child due to any cause, suicide attempt, hospital admission due to a psychiatric problem, and hospital admission of participant or infant due to a serious medical emergency. Their detection and appropriate response (involving an independent psychiatrist responding) will be reported to the local Ethics Committee. These SAEs are compiled by the data manager and a blinded summary report is shared with the principal investigators and the TSC.

Discussion

This trial and the parallel reference group of perinatally nondepressed women have the potential to further our understanding of the early developmental lifecourse of children of both women who were, and were not, perinatally depressed and to evaluate whether intervening on mothers' perinatal depression can mitigate the negative effects of maternal depression on child development at 36 months. By beginning our study in the third trimester of pregnancy and following the mother-child dyads with multiple assessments through 36 postnatal months we will be able to analyze the relationship between changes in maternal depressive symptoms and child outcomes. For example, we will be able to analyze the impact of early versus late remission; remission of symptoms followed by recurrence; and new onset of symptoms on child outcomes. With the 3 years of follow-up with multiple assessments,

we will be able to undertake an analysis of potential time-varying mechanisms.

By also enrolling a group of women who were perinatally nondepressed we are additionally be able to address two substantive questions. The first is: How much of the risk due to maternal depression exposure can the intervention mitigate? We ultimately want to know whether the intervention can prevent the intergenerational transmission of negative mental health outcomes. The children of prenatally depressed mothers in both intervention and control arms of the THPP+ intervention study are at high risk for multiple adverse outcomes. We expect that, at the end of the study, the children in the intervention arm will be at lower risk. However, the full impact of the intervention can only be discerned if we know the level of risk remaining – that is, the difference between the reduced level of risk among children (of prenatally depressed mothers) in the intervention arm and the risk among children whose mothers were not depressed to begin with. If outcomes of these two groups are comparable, we can infer that the intervention may prevent the intergenerational transmission of risk. Unlike in high-income country settings, normative data for such a comparison does not exist in many low-resource areas, including Pakistan; hence, the enrollment of nondepressed women [35, 36]. The second substantive question is: Does the intervention have an impact on mothers and children living in the intervention clusters, even if the mother was not depressed prenatally? The community intervention was originally designed to improve outcomes among depressed women. However, we suspect that its design may lead to broader, population-wide effects.

By working in a rural setting in Pakistan and by combining the cohorts of perinatally depressed mothers in the c-RCT and nonperinatally depressed mothers, the THPP+ study offers a unique opportunity to understand, and to potentially help to mitigate, the effects of perinatal depression on both the mother and the child.

Trial status

The THPP trial (and hence the THPP+ trial) began recruitment of participants in October 2014. Based on our previous work and pilot results with an approximately 25% rate of perinatal depression, we expect to recruit the sample by the end of February 2016. The endpoint assessments of all the participants at 36 postnatal months will be completed by end 2018.

Additional file

Additional file 1: SPIRIT study timeline. (DOC 46 kb)

Abbreviations

ASQ: Ages and Stages Questionnaire; BSID-III: Bayley Scales of Infant and Toddler Development (III); CHW: Community health worker; CI: Confidence interval; c-RCT: Cluster randomized controlled trial; EUC: Enhanced Usual Care; HDRF: Human Development Research Foundation; ICC: Intracluster correlation; PHQ: Patient Health Questionnaire; SAE: Serious adverse events; SD: Standard deviation; SDQ: Strengths and Difficulties Questionnaire; TD: Total Difficulties (score of SDQ); THP: Thinking Healthy Program; THPP: Thinking Healthy Program Peer-delivered; THPP+: Thinking Healthy Program Peer-delivered PLUS; TMC: *Trial Management Committee; TSC:* Trial Steering Committee; UC: Union Council; WHO-DAS: WHO Disability Assessment Schedule

Acknowledgements

We thank all of the study participants for their valuable time and the field staff, including Abdul Latif Khan, Abdul Hanif Khan, Umer Javeed and Malik Habib Ahmed.

Funding

The trial has been funded by two agencies of National Institutes of Health, USA, as follows. The National Institute of Mental Health funded the original 6-month THPP trial under award number U19MH95687. The National Institute of Child Health funded the extended follow-up of mother-child dyads for 36 months postpartum under award R01 HD075875. The development and implementation of the THPP+ component of the intervention is funded through HDRF.

Availability of supporting data

Not applicable.

Authors' contributions

ELT contributed to the study conception and study design, performed all power calculations, developed the statistical analysis plan and led the first draft of the manuscript. SS developed the original THPP trial on which the THPP+ trial builds, developed the THPP+ intervention and contributed to the study design and leads all field activities. OB contributed to the study design and coordinates and manages all field activities. AZ developed the electronic data management systems and was in charge of data quality monitoring. JG contributed to the statistical analysis plan and to drafting the tables and figures. NG contributed to the drafting of tables and figures. KOD contributed to the study design and leads the child development measure adaptation and training component. AR contributed to the study design and led the original THP and THPP studies on which this study builds. JM led the conception of the THPP+ study, contributed to the study design, contributed to the writing and editing of the manuscript and led grant funding applications. All authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Consent for publication

Not applicable.

Ethics approval and consent to participate

The study has been granted ethical approvals from the Human Development Research Foundation and Duke University (USA) Institutional Review Boards. Written (or witnessed, if the participant is illiterate) informed consent is mandatory for enrollment. All participants are able to access EUC, representing a higher quality of care than what is available in Pakistan's current primary care set up.

Author details

[1]Duke Global Health Institute, Duke University, Durham, USA. [2]Department of Biostatistics and Bioinformatics, Duke University, Durham, USA. [3]Human Development Research Foundation, Islamabad, Pakistan. [4]Department of Epidemiology, Columbia University School of Public Health, New York, USA. [5]Institute of Psychology, Health and Society, University of Liverpool, Liverpool, UK. [6]Department of Epidemiology, Gillings School of Public Health, University of North Carolina Chapel Hill, Chapel Hill, USA.

Received: 12 February 2016 Accepted: 21 July 2016
Published online: 08 September 2016

References

1. Walker SP, Wachs TD, Gardner JM, Lozoff B, Wasserman GA, Pollitt E, et al. Child development in developing countries 2—Child development: risk factors for adverse outcomes in developing countries. Lancet. 2007;369(9556):145–57.
2. Patel V, Rodrigues M, DeSouza N. Gender, poverty, and postnatal depression: a study of mothers in Goa, India. Am J Psychiatry. 2002;159(1):43–7.
3. Murray L, Cooper PJ. Effects of postnatal depression on infant development. Arch Dis Child. 1997;77(2):99–101.
4. Hussain MFA, Nauman F. Maternal mental distress: a risk factor for infant under nutrition in developing countries. J Pak Med Assoc. 2010;60(4):329.
5. Affonso DD, De AK, Horowitz JA, Mayberry LJ. An international study exploring levels of postpartum depressive symptomatology. J Psychosom Res. 2000;49(3):207–16.
6. Goodman SH, Rouse MH, Connell AM, Broth MR, Hall CM, Heyward D. Maternal depression and child psychopathology: a meta-analytic review. Clin Child Fam Psychol Rev. 2011;14(1):1–27. doi:10.1007/s10567-010-0080-1.
7. Ohara MW, Swain AM. Rates and risk of postpartum depression—A meta-analysis. Int Rev Psychiatry. 1996;8(1):37–54.
8. Almond P. Postnatal depression: a global public health perspective. Perspect Public Health. 2009;129(5):221–7.
9. Rahman A, Patel V, Maselko J, Kirkwood B. The neglected "m" in MCH programmes— why mental health of mothers is important for child nutrition. Trop Med Int Health. 2008;13(4):579–83. doi:10.1111/j.1365-3156.2008.02036.x.
10. Parsons CE, Young KS, Rochat TJ, Kringelbach ML, Stein A. Postnatal depression and its effects on child development: a review of evidence from low- and middle-income countries. Br Med Bull. 2012;101(1):57–79. doi:10.1093/bmb/ldr047.
11. Maselko J, Sikander S, Bhalotra S, Bangash O, Ganga N, Mukherjee S, et al. Effect of an early perinatal depression intervention on long-term child development outcomes: follow-up of the Thinking Healthy Programme randomised controlled trial. Lancet Psychiatry. 2015;2(7):609–17.
12. Rahman A, Malik A, Sikander S, Roberts C, Creed F. Cognitive behaviour therapy-based intervention by community health workers for mothers with depression and their infants in rural Pakistan: a cluster-randomised controlled trial. Lancet. 2008;372(9642):902–9.
13. World Health Organization. Thinking Healthy: a manual for psychosocial management of perinatal depression (WHO generic field-trial version 1.0). Geneva: WHO; 2015.
14. Sikander S, Lazarus A, Bangash O, Fuhr DC, Weobong B, Krishna RN, et al. The effectiveness and cost-effectiveness of the peer-delivered Thinking Healthy Programme for perinatal depression in Pakistan and India: the SHARE study protocol for randomised controlled trials. Trials. 2015;16(1):1–14.
15. Organization PC. District Census Report of Rawalpindi 1008. Islamabad, Pakistan: Population Census Organization; 1998.
16. Fraz K, Khan S, Sikander S. Screening for depression in coronary artery disease patients using PHQ-9. The Health. 2013;4(1):3–6.
17. Patel V, Araya R, Chowdhary N, King M, Kirkwood B, Nayak S, et al. Detecting common mental disorders in primary care in India: a comparison of five screening questionnaires. Psychol Med. 2008;38(02):221–8.
18. Üstün TB, Chatterji S, Kostanjsek N, Rehm J, Kennedy C, Epping-Jordan J, et al. Developing the World Health Organization disability assessment schedule 2.0. Bull World Health Organ. 2010;88(11):815–23.
19. Goodman R. The Strengths and Difficulties Questionnaire: A Research Note. J Child Psychol Psychiatry. 1997;38(5):581–6. doi:10.1111/j.1469-7610.1997.tb01545.x.
20. Goodman A, Lamping DL, Ploubidis GB. When to use broader internalising and externalising subscales instead of the hypothesised five subscales on the Strengths and Difficulties Questionnaire (SDQ): data from British parents, teachers and children. J Abnorm Child Psychol. 2010;38(8):1179–91.
21. Woerner W, Fleitlich-Bilyk B, Martinussen R, Fletcher J, Cucchiaro G, Dalgalarrondo P, et al. The Strengths and Difficulties Questionnaire overseas: evaluations and applications of the SDQ beyond Europe. Eur Child Adolesc Psychiatry. 2004;13(2):ii47–54.
22. Samad L, Hollis C, Prince M, Goodman R. Child and adolescent psychopathology in a developing country: testing the validity of the strengths and difficulties questionnaire (Urdu version). Int J Methods Psychiatr Res. 2005;14(3):158–66.
23. Syed EU, Hussein SA, Mahmud S. Screening for emotional and behavioural problems amongst 5–11-year-old school children in Karachi, Pakistan. Soc Psychiatry Psychiatr Epidemiol. 2007;42(5):421–7.
24. Squires JK, Potter L, Bricker DD, Lamorey S. Parent-completed developmental questionnaires: effectiveness with low and middle income parents. Early Childhood Res Q. 1998;13(2):345–54.
25. Squires J, Bricker D, Potter L. Revision of a parent-completed developmental screening tool: ages and stages questionnaires. J Pediatr Psychol. 1997;22(3):313–28. doi:10.1093/jpepsy/22.3.313.
26. Squires JK, Bricker DD, Twombly E. Ages and Stages Questionnaire: Social-Emotional (ASQ:SE): a parent-completed, child-monitoring system for social-emotional behaviors. Baltimore, MD: Paul H Brookes Publishing; 2002.
27. Bayley N, Reuner G. Bayley scales of infant and toddler development: Bayley-III. 2006. Harcourt Assessment: Psych. Corporation.
28. Julious SA. Tutorial in biostatistics— Sample sizes for clinical trials with normal data. Stat Med. 2004;23(12):1921–86. doi:10.1002/sim.1783.
29. Hayes R, Moulton L. Cluster randomized trials. Boca Raton: CRC Press; 2009.
30. SDQ Website. SDQ Normative Data. Youth in Mind. 2012. http://www.sdqinfo.org/g0.html. Accessed 28 Nov 2012.
31. Mieloo C, Raat H, van Oort F, Bevaart F, Vogel I, Donker M et al. Validity and reliability of the Strengths and Difficulties Questionnaire in 5–6 year olds: differences by gender or by parental education? PLoS One. 2012;7(5). doi:10.1371/journal.pone.0036805.
32. Hayes R, Moulton L. Sample size. Cluster randomised trials: CRC Press; 2009.
33. Molenberghs G, Thijs H, Jansen I, Beunckens C, Kenward MG, Mallinckrodt C, et al. Analyzing incomplete longitudinal clinical trial data. Biostatistics. 2004;5(3):445–64. doi:10.1093/biostatistics/kxh001.
34. Little RJA. Pattern-mixture models for multivariate incomplete data. J Am Stat Assoc. 1993;88(421):125.
35. Nelson CA. A neurobiological perspective on early human deprivation. Child Develop Perspect. 2007;1(1):13–8. doi:10.1111/j.1750-8606.2007.00004.x.
36. Nelson C, Zeanah CH, Fox N, et al. Cognitive recovery in socially deprived young children: The Bucharest Early Intervention Project. Science. 2007;318:1937–40.
37. Kroencke K, Spitzer R, Williams J. The PHQ-9: validity of a brief depression severity measure [Electronic version]. J Gen Intern Med. 2001;16(9):606–13.
38. Gollenberg AL, Lynch CD, Jackson LW, McGuinness BM, Msall ME. Concurrent validity of the parent-completed Ages and Stages Questionnaires, 2nd Ed. with the Bayley Scales of Infant Development II in a low-risk sample. Child Care Health Dev. 2010;36(4):485–90. doi:10.1111/j.1365-2214.2009.01041.x.
39. Kerstjens JM, Bos AF, ten Vergert EMJ, de Meer G, Butcher PR, Reijneveld SA. Support for the global feasibility of the Ages and Stages Questionnaire as developmental screener. Early Hum Dev. 2009;85(7):443–7. doi:10.1016/j.earlhumdev.2009.03.001.
40. Yu LM, Hey E, Doyle LW, Farrell B, Spark P, Altman DG, et al. Evaluation of the Ages and Stages Questionnaires in identifying children with neurosensory disability in the Magpie Trial follow-up study. Acta Paediatr. 2007;96(12):1803–8. doi:10.1111/j.1651-2227.2007.00517.x.