The Lancet Psychiatry

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

Corresponding Author:	Joanna Maselko University of North Carolina at Chapel Hill Chapel Hill, North Carolina UNITED STATES
First Author:	Joanna Maselko
Order of Authors:	Joanna Maselko
	Elizabeth L Turner
	Lisa M Bates
	Ikhlaq Ahmad
	Najia Atif
	Victoria Baranov
	Bhalotra Sonia
	Biroli Pietro
	Esther Chung
	John A Gallis
	Ashley Hagaman
	Katherine LeMasters
	Karen O'Donnell
	Elissa Scherer
	Siham Sikander
	Atif Rahman
	Amina Bibi
	Tayyiba Bibi
	Samina Bilal
	Anam Jamil
	Maria Sharif
	Ahmed Waqas
	Ahmed Zaidi
	Shaffaq Zulfiqar
Manuscript Region of Origin:	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 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.

Dear Joan,

We've made the changes below.
Thanks again for everything,
Joanna

We have only one corresponding author: please delete one.
-done

Please include the data sharing statement in the manuscript, at the end of the main text.
-done

References 5 and 8 are the same: please correct and renumber throughout.
-fixed (apologies, we had removed it and it kept coming back!)

Figure 2. Please do not reload the Powerpoint slide; it doesn't help.
-removed

You have then submitted an .emf file. This is not mentioned in our artwork guidelines. I seem to be able to add lines to it but not edit the words. It looks like a vector file in that it doesn't pixellate when enlarged. Our illustrators are very busy with COVID papers, so I don't want to distract them by asking about this. It would be better if you could provide the original .eps, .ai, or .pdf file version. If that is not possible, then include the emf file as Figure 2 and we'll see.

I have uploaded an .eps file for figure 2. What I'm not sure is how to add the title with the proper font, so I added another word document that has the title and footnote for figure 2. I hope that is ok. I also can generate a pdf version now.

I also do have .eps files for the two figures in the appendix, but I can't get them embedded in the word file. So, currently, I have a pdf version of those images that is embedded, per the instructions for appendix figures.

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

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

9

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

27 **Corresponding Authors**
28 Joanna Maselko, ScD
29 jmaselko@unc.edu
30 Dept. of Epidemiology
31 Carolina Population Center Fellow
32 2105e McGavran-Greenberg Hall
33 Campus Box 7435
34 Gillings School of Global Public Health, UNC
35 Chapel Hill, NC 27599-7435 USA
36
37

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.

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]

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

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

171

Methods

Study design and participants

174 We conducted a stratified cluster-randomized controlled trial in Kallar Syedan, a rural subdistrict of
175 Rawalpindi, Pakistan. The sub-district is a socioeconomically deprived area with poverty rates around
176 50%, female literacy of 40-45%, and a high fertility rate (3·8 births per woman).[21] It is primarily agrarian
177 with close knit communities co-residing in large households (average 6·2 people per household). The
178 sub-district has 11 Union Councils (UC), the smallest administrative unit, each with a population of about
179 22,000-25,000. Each UC is serviced by a Primary Health Care Facility which houses a physician, midwife,
180 vaccinator, dispenser, and village-based Lady Health Workers (LHWs).
181
182 This trial maintained the original cluster criterion, randomization sequence and procedures under the
183 previous trial.[16,22] Pregnant women in the 3rd trimester, aged 18+ years and registered with their LHWs,
184 were eligible. Approximately 95% of the women in the study area were registered with LHWs. All
185 pregnant women were approached by trained research staff either within the pregnant woman's
186 residence or that of their LHW and, if they consented, were assessed. Women who needed immediate
187 medical or psychiatric inpatient care were excluded from the study. All eligible women were invited to
188 be screened for depression using the Urdu version of the Patient Health Questionnaire 9 (PHQ-9), which
189 has been used extensively as a screening tool in the study setting and has an acceptable criterion validity
190 and reliability for this population.[23]   Women screening positive (PHQ-9 score $\geq$ 10) were eligible for
191 enrolment into the trial and follow-up as part of the Bachpan Cohort.[16] One out of every three women
192 who screened with a < 10 score on PHQ were enrolled to participate in the Bachpan Cohort only,
193 resulting in a roughly equal size of prenatally non-depressed and depressed women at the beginning the
194 cohort.[24]
195

Randomization and masking

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

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

206

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

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

213

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

225

226

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

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

244 depression,[27] and providing depressed participants with a leaflet on how and where to seek appropriate
245 health care during pregnancy and beyond.
246

Procedures

248 Assessments were conducted 6 times over the course of the study (in pregnancy, and 3, 6, 12, 24, and
249 36 months postnatal). As originally specified, the current analysis utilizes mother and child outcome
250 data at 36 months.[22] All measures were extensively piloted. Assessments were done at the community
251 level within households of women by trained interviewers blind to the allocation status, all questions
252 were interviewer administered. Assessors inter-rater reliability was ensured through classroom-based
253 training which included role plays, followed by field practice sessions. During these sessions each pair of
254 assessors assessed up to 10 participants jointly and discussed their coding on each item to establish
255 inter-rater reliability prior to start of actual data collection.

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

260

Outcomes

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

270

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

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

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

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

292   **Statistical analyses**

293   For mother outcomes, anticipating a sample of 480 prenatally depressed women, we were powered at
294   90% to detect a remission rate of 65% in the prenatally depressed-intervention versus 45% in prenatally
295   depressed-control at the two-tailed 5% significance level and assuming a conservative ICC of 0.07 in
296   intervention arm and 0.05 in control.[22] For child outcomes, with this sample size, we were powered at
297   90% to detect a difference between treatment arms of 3 points on the SDQ-TD at the two-tailed 5%
298   significance level assuming a standard deviation of 5.2 points and ICC of 0.04-0.08.[22]

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

305   The primary analyses were designed as intention-to-treat. Data from prenatally depressed and non-
306   depressed participants were analyzed jointly using linear mixed effects models so that all comparisons
307   of interest could be estimated from the same model. The identity link was used for continuous
308   outcomes to estimate differences in mean outcomes. Standardized mean differences (SMDs) and their
309   95% confidence intervals (CIs) were obtained using the method of Hedges.[37] In the primary model, we
310   included a random intercept for cluster and fixed effects for treatment arm (depressed intervention,
311   depressed control, non-depressed), union council (11 levels; the stratification variable), and variables
312   found to be imbalanced by loss to follow-up or at baseline (determined using p<0.10; see Appendix p.9-
313   14). Mixed models assume missing at random conditional on the covariates included in the statistical
314   model. Therefore, we include variables predictive of loss to follow-up to account for missing data.
315   We used restricted maximum likelihood (REML). The between-within method was used to apply small-
316   sample bias corrections to the intervention effect standard errors in the mixed effects framework.[38]
317   These models also generated the intra-cluster correlation (ICC) values.

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

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

329

330   **Role of Funding Source**

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

334

335   **Results**

336   From Oct 15, 2014 to Feb 25, 2016, we identified and randomly selected 40 village clusters out of 46 and
337   randomly assigned 20 village clusters each to intervention (THPP+) and control (EUC) arms. In all we
338   approached 1910 pregnant women; 287 prenatally depressed women in the control arm and 283 in the
339   intervention arm completed the baseline questionnaire. Of the prenatally non-depressed women
340   approached, 584 were enrolled, yielding comparable numbers of prenatally depressed and prenatally
341   non-depressed women in each of the 40 village clusters.
342   Of the 1,154 participants enrolled at baseline, 889 (77.0%) were successfully interviewed at 36 months:
343   206 (72.5%) intervention, 216 (75.3%) EUC, and 467 (80.0%) prenatally non-depressed (Figure 1). There
344   was no differential loss to follow up by treatment arm and no differences in adverse events
345   (Appendix p.9-14).

346

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

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

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

365

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

9

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

375

376 The prevalence of major depressive episodes (SCID-based) among the three groups of women (the
377 prenatally depressed intervention arm, prenatally depressed control arm, and prenatally non-depressed
378 women) became increasingly similar in the proportion depressed by 36 months postnatal (Figure 2).
379 The THPP+ intervention arm showed higher convergence with the prenatally non-depressed women at
380 36 months compared to the control arm, so much so that there was not a statistically significant
381 difference in the probability of being depressed (using SCID) between the intervention arm women (at
382 16.5%) and the prenatally non-depressed (9.0% of whom were depressed at 36 months, RR 0·74 (0·46 to
383 1·18).

384

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

388

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

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

395
396

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

408

409 **Discussion**

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

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

418

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

442
443

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

456

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

462 about their illness status improves outcomes,[49] which raises important methodological (study design)
463 issues for trials that are embedded within community settings. Perhaps future trials, using similar active
464 control arms ought to consider equivalence or non-inferiority trial designs to avoid a nonsignificant
465 superiority trial being wrongly interpreted, as proof of no difference between the two active
466 comparisons.[50]
467

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

472

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

480

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

489

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

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

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

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

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

547

548 **Data Sharing Statement**

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

551

552 **Authors' contributions**

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

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

561

562 **Conflict of Interest**

563 We declare we have no conflict of interest.

564

14

References

565

566   1.   Gelaye B, Rondon MB, Araya R, Williams MA. Epidemiology of maternal depression, risk factors,
567   and child outcomes in low-income and middle-income countries. *The Lancet Psychiatry* 2016; **3**(10): 973-
568   82.10.1016/S2215-0366(16)30284-X
569   2.   Goodman SH, Rouse MH, Connell AM, Broth MR, Hall CM, Heyward D. Maternal Depression and
570   Child Psychopathology: A Meta-Analytic Review. *Clin Child Fam Psychol Rev* 2011; **14**(1): 1-
571   27.10.1007/s10567-010-0080-1
572   3.   Patel V, Saxena S, Lund C, Thornicroft G, Baingana F, Bolton P, Chisholm D, Collins PY, Cooper JL,
573   Eaton J, Herrman H, Herzallah MM, Huang Y, Jordans MJD, Kleinman A, Medina-Mora ME, Morgan E,
574   Niaz U, Omigbodun O, Prince M, Rahman A, Saraceno B, Sarkar BK, De Silva M, Singh I, Stein DJ, Sunkel
575   C, UnÜtzer J. The Lancet Commission on global mental health and sustainable development. *The Lancet*
576   2018; **392**(10157): 1553-98.10.1016/S0140-6736(18)31612-X
577   4.   Maselko J, Sikander S, Bhalotra S, Bangash O, Ganga N, Mukherjee S, Egger H, Franz L, Bibi A,
578   Liaqat R, Kanwal M, Abbasi T, Noor M, Ameen N, Rahman A. Effect of an early perinatal depression
579   intervention on long-term child development outcomes: follow-up of the Thinking Healthy Programme
580   randomised controlled trial. *Lancet Psychiatry* 2015; **2**(7): 609-17.10.1016/s2215-0366(15)00109-1
581   5.   Rahman A, Fisher J, Bower P, Luchters S, Tran T, Yasamy MT, Saxena S, Waheed W. Interventions
582   for common perinatal mental disorders in women in low- and middle-income countries: a systematic
583   review and meta-analysis. *Bulletin of the World Health Organization* 2013; **91**(8): 593-
584   601I.10.2471/BLT.12.109819
585   6.   Tomlinson M, Rotheram-Borus MJ, Scheffler A, le Roux I. Antenatal depressed mood and child
586   cognitive and physical growth at 18-months in South Africa: a cluster randomised controlled trial of
587   home visiting by community health workers. *Epidemiol Psychiatr Sci* 2018; **27**(6): 601-
588   10.10.1017/s2045796017000257
589   7.   Goodman SH, Garber J. Evidence-Based Interventions for Depressed Mothers and Their Young
590   Children. *Child Dev* 2017; **88**(2): 368-77.10.1111/cdev.12732
591   8.   Baranov V, Bhalotra S, Biroli P, Maselko J. Maternal Depression, Women's Empowerment, and
592   Parental Investment: Evidence from a Randomized Control Trial. *American Economic Review* 2020;
593   **110**(3): 824-59
594   9.   Gunlicks ML, Weissman MM. Change in child psychopathology with improvement in parental
595   depression: A systematic review. *Journal of the American Academy of Child and Adolescent Psychiatry*
596   2008; **47**(4): 379-89.10.1097/CHI.0b013e3181640805
597   10.   Cuijpers P, Weitz E, Karyotaki E, Garber J, Andersson G. The effects of psychological treatment of
598   maternal depression on children and parental functioning: a meta-analysis. *Eur Child Adolesc Psych*
599   2015; **24**(2): 237-45.10.1007/s00787-014-0660-6
600   11.   Goodman SH, Cullum KA, Dimidjian S, River LM, Kim CY. Opening windows of opportunities:
601   Evidence for interventions to prevent or treat depression in pregnant women being associated with
602   changes in offspring's developmental trajectories of psychopathology risk. *Development and*
603   *Psychopathology* 2018; **30**(3): 1179-96.10.1017/S0954579418000536
604   12.   Loechner J, Starman K, Galuschka K, Tamm J, Schulte-Körne G, Rubel J, Platt B. Preventing
605   depression in the offspring of parents with depression: A systematic review and meta-analysis of
606   randomized controlled trials. *Clinical Psychology Review* 2018; **60**: 1-
607   14.https://doi.org/10.1016/j.cpr.2017.11.009
608   13.   Milgrom J, Holt CJ, Bleker LS, Holt C, Ross J, Ericksen J, Glover V, O'Donnell KJ, de Rooij SR,
609   Gemmill AW. Maternal antenatal mood and child development: an exploratory study of treatment
610   effects on child outcomes up to 5 years. *J Dev Orig Health Dis* 2019; **10**(2): 221-
611   31.10.1017/s2040174418000739

612   14.   Kersten-Alvarez LE, Hosman CMH, Riksen-Walraven JM, van Doesum KTM, Hoefnagels C. Long-
613 term effects of a home-visiting intervention for depressed mothers and their infants. *Journal of Child*
614 *Psychology and Psychiatry* 2010; **51**(10): 1160-70.10.1111/j.1469-7610.2010.02268.x
615   15.   Rotheram-Borus MJ, Arfer KB, Christodoulou J, Comulada WS, Stewart J, Tubert JE, Tomlinson
616 M. The association of maternal alcohol use and paraprofessional home visiting with children's health: A
617 randomized controlled trial. *J Consult Clin Psychol* 2019; **87**(6): 551-
618 62.10.1037/ccp0000408PMC6775769
619   16.   Sikander S, Ahmad I, Atif N, Zaidi A, Vanobberghen F, Weiss HA, Nisar A, Tabana H, Ain QU, Bibi
620 A, Bilal S, Bibi T, Liaqat R, Sharif M, Zulfiqar S, Fuhr DC, Price LN, Patel V, Rahman A. Delivering the
621 Thinking Healthy Programme for perinatal depression through volunteer peers: a cluster randomised
622 controlled trial in Pakistan. *Lancet Psychiatry* 2019; **6**(2): 128-39.10.1016/s2215-0366(18)30467-x
623   17.   Vanobberghen F, Weiss HA, Fuhr DC, Sikander S, Afonso E, Ahmad I, Atif N, Bibi A, Bibi T, Bilal S,
624 De Sa A, D'Souza E, Joshi A, Korgaonkar P, Krishna R, Lazarus A, Liaqat R, Sharif M, Weobong B, Zaidi A,
625 Zuliqar S, Patel V, Rahman A. Effectiveness of the Thinking Healthy Programme for perinatal depression
626 delivered through peers: Pooled analysis of two randomized controlled trials in India and Pakistan. *J*
627 *Affect Disord* 2019.https://doi.org/10.1016/j.jad.2019.11.110
628   18.   Fuhr DC, Weobong B, Lazarus A, Vanobberghen F, Weiss HA, Singla DR, Tabana H, Afonso E, De
629 Sa A, D'Souza E, Joshi A, Korgaonkar P, Krishna R, Price LN, Rahman A, Patel V. Delivering the Thinking
630 Healthy Programme for perinatal depression through peers: an individually randomised controlled trial
631 in India. *Lancet Psychiatry* 2019; **6**(2): 115-27.10.1016/s2215-0366(18)30466-8
632   19.   Atif N, Bibi A, Nisar A, Zulfiqar S, Ahmed I, LeMasters K, Hagaman A, Sikander S, Maselko J,
633 Rahman A. Delivering maternal mental health through peer volunteers: a 5-year report. *International*
634 *Journal of Mental Health Systems* 2019; **13**(1): 62.10.1186/s13033-019-0318-3
635   20.   National Institute of Population Studies - NIPS/Pakistan & ICF Intternational. Pakistan
636 Demographic and Health Survey 2012-2013. Islamabad: NIPS/Pakistan and ICF International, 2012-3.
637   21.   National Institute of Population Studies Islamabad. Pakistan Demographic and Health Survey.
638 Islamabad, Pakistan: National Institute of Population Studies, 2013.
639   22.   Turner EL, Sikander S, Bangash O, Zaidi A, Bates L, Gallis J, Ganga N, O'Donnell K, Rahman A,
640 Maselko J. The effectiveness of the peer-delivered Thinking Healthy PLUS (THPP+) Program for maternal
641 depression and child socioemotional development in Pakistan: study protocol for a randomized
642 controlled trial. *Trials* 2016; **17**(1): 442.10.1186/s13063-016-1530-yPMC5017048
643   23.   Gallis J, Maselko J, O'Donnell K, Song KE, Saqib K, Turner EL, Sikander S. Criterion-related validity
644 and reliability of the Urdu version of the patient health questionnaire in community-based women in
645 Pakistan. *PeerJ* 2018; **6**(e5185).https://doi.org/10.7717/peerj.5185
646   24.   Sikander S, Ahmad I, Bates LM, Gallis J, Hagaman A, O'Donnell K, Turner EL, Zaidi A, Rahman A,
647 Maselko J. Cohort Profile: Perinatal depression and child socioemotional development ; the Bachpan
648 cohort study from rural Pakistan. *BMJ Open* 2019; **9**(5): e025644.10.1136/bmjopen-2018-025644
649   25.   Sikander S, Lazarus A, Bangash O, Fuhr DC, Weobong B, Krishna RN, Ahmad I, Weiss HA, Price L,
650 Rahman A, Patel V. The effectiveness and cost-effectiveness of the peer-delivered Thinking Healthy
651 Programme for perinatal depression in Pakistan and India: the SHARE study protocol for randomised
652 controlled trials. *Trials* 2015; **16**: 14.10.1186/s13063-015-1063-9
653   26.   Atif N, Nisar A, Bibi A, Khan S, Zulfiqar S, Ahmad I, Sikander S, Rahman A. Scaling-up
654 psychological interventions in resource-poor settings: training and supervising peer volunteers to deliver
655 the 'Thinking Healthy Programme' for perinatal depression in rural Pakistan. *Global mental health*
656 *(Cambridge, England)* 2019; **6**: e4.10.1017/gmh.2019.4PMC6521132
657   27.   Organization. WH. WHO Guidelines Approved by the Guidelines Review Committee. mhGAP
658 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

706   44.   Baqui AH, El-Arifeen S, Darmstadt GL, Ahmed S, Williams EK, Seraji HR, Mannan I, Rahman SM,
707   Shah R, Saha SK, Syed U, Winch PJ, Lefevre A, Santosham M, Black RE. Effect of community-based
708   newborn-care intervention package implemented through two service-delivery strategies in Sylhet
709   district, Bangladesh: a cluster-randomised controlled trial. *The Lancet* 2008; **371**(9628): 1936-
710   44.10.1016/S0140-6736(08)60835-1
711   45.   Tripathy P, Nair N, Barnett S, Mahapatra R, Borghi J, Rath S, Rath S, Gope R, Mahto D, Sinha R,
712   Lakshminarayana R, Patel V, Pagel C, Prost A, Costello A. Effect of a participatory intervention with
713   women's groups on birth outcomes and maternal depression in Jharkhand and Orissa, India: a cluster-
714   randomised controlled trial. *The Lancet* 2010; **375**(9721): 1182-92.10.1016/S0140-6736(09)62042-0
715   46.   Yousafzai AK, Rasheed MA, Rizvi A, Armstrong R, Bhutta ZA. Parenting Skills and Emotional
716   Availability: An RCT. *Pediatrics* 2015; **135**(5): e1247-e57.10.1542/peds.2014-2335
717   47.   Hafeez A, Mohamud B, Shiekh M, Shah I, Jooma R. Lady health workers programme in Pakistan:
718   challenges, achievements and the way forward. *JPMA* 2011; **61**(210):
719   48.   Rahman A, Malik A, Sikander S, Roberts C, Creed F. Cognitive behaviour therapy-based
720   intervention by community health workers for mothers with depression and their infants in rural
721   Pakistan: a cluster-randomised controlled trial. *Lancet* 2008; **372**(9642): 902-9
722   49.   Paterick TE, Patel N, Tajik AJ, Chandrasekaran K. Improving health outcomes through patient
723   education and partnerships with patients. *Proc (Bayl Univ Med Cent)* 2017; **30**(1): 112-
724   3.10.1080/08998280.2017.11929552
725   50.   Lesaffre E. Superiority, equivalence, and non-inferiority trials. *Bulletin of the NYU hospital for
726   joint diseases* 2008; **66**(2): 150-4
727   51.   Netsi E, Pearson RM, Murray L, Cooper P, Craske MG, Stein A. Association of Persistent and
728   Severe Postnatal Depression With Child OutcomesAssociation of Persistent and Severe Postnatal
729   Depression With Child OutcomesAssociation of Persistent and Severe Postnatal Depression With Child
730   Outcomes. *JAMA psychiatry* 2018; **75**(3): 247-53.10.1001/jamapsychiatry.2017.4363
731   52.   Robinson R, Lahti-Pulkkinen M, Heinonen K, Reynolds RM, Räikkönen K. Fetal programming of
732   neuropsychiatric disorders by maternal pregnancy depression: a systematic mini review. *Pediatric
733   Research* 2019; **85**(2): 134-45.10.1038/s41390-018-0173-y
734   53.   O'Donnell KJ, Glover V, Barker ED, O'Connor TG. The persisting effect of maternal mood in
735   pregnancy on childhood psychopathology. *Development and Psychopathology* 2014; **26**(2): 393-
736   403.10.1017/S0954579414000029
737   54.   Stein A, Netsi E, Lawrence PJ, Granger C, Kempton C, Craske MG, Nickless A, Mollison J, Stewart
738   DA, Rapa E, West V, Scerif G, Cooper PJ, Murray L. Mitigating the effect of persistent postnatal
739   depression on child outcomes through an intervention to treat depression and improve parenting: a
740   randomised controlled trial. *Lancet Psychiatry* 2018; **5**(2): 134-44.10.1016/s2215-0366(18)30006-3
741   55.   Frankenhuis WE, Del Giudice M. When Do Adaptive Developmental Mechanisms Yield
742   Maladaptive Outcomes? *Developmental Psychology* 2012; **48**(3): 628-42.10.1037/a0025629

743

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

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

9

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

27 **Corresponding Authors**
28 Joanna Maselko, ScD
29 jmaselko@unc.edu
30 Dept. of Epidemiology
31 Carolina Population Center Fellow
32 2105e McGavran-Greenberg Hall
33 Campus Box 7435
34 Gillings School of Global Public Health, UNC
35 Chapel Hill, NC 27599-7435 USA
36
37 ~~Siham Sikander, PhD~~
38 ~~siham.sikander@hdrfoundation.org, sihamsikander@gmail.com~~
39 ~~Health Services Academy~~
40 ~~H.No 55, Street 55, F-7/4~~
41 ~~Islamabad Pakistan~~
42

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 Meand 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 interventions 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.

76 **Research in context**

77 **Evidence before this study.**

78 Recent systematic reviews of psychotherapy interventions for depression have highlighted the limited
79 evidence on long-term effects of psychotherapy on either maternal mental health or child outcomes.
80 We conducted a search to identify studies designed to evaluate interventions for perinatal depression,
81 whose intervention either lasted beyond 12 months postnatal (e.g. booster sessions), or whose follow-
82 up was more than 12 after the completion of the intervention, in years 2002-April 2020. We limited our
83 search to randomized clinical trials or meta-analyses. We did not place restrictions on language or
84 country. We used Pubmed and Web of Science, with the following search terms: ((maternal depression)
85 or (perinatal depression) or (postpartum depression) or (postpartum depression)) AND ((treatment) or
86 (therapy) or (intervention) or (psychotherapy) or (cognitive behavioral therapy)) AND ((longer-term) or
87 (longer) or (booster)). We identified six RCTs specific to perinatal depression with the longer follow-up
88 period, ranging from 1.5 to 7 years. None utilized an extended duration design (booster sessions) that
89 continued past 12 months postnatal; two studies included a non-depressed comparison group. Most
90 common intervention models were Cognitive Behavioral Therapy (CBT) and Interpersonal Therapy (IPT).
91 Evidence generally showed that interventions improved outcomes which then weakened over time so
92 that, overall, there is limited evidence of intervention effects of perinatal depression interventions that
93 persist beyond the perinatal period. Of the six studies, two reported some lasting impact. One study of
94 884 mother-child dyads assessed maternal and child outcomes 7 years after the end of a CBT
95 intervention and found a 5 point percent lower rates of depression among mothers who received the
96 intervention, but no significant effects on child outcomes. With this one exception, sample sizes were
97 small, with studies having fewer than 60 participants per group available at follow-up.

98

99 **Added value of this study.**

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

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

115

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]

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

165

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

177

178 **Methods**

179 **Study design and participants**

180 We conducted a stratified cluster-randomized controlled trial in Kallar Syedan, a rural subdistrict of
181 Rawalpindi, Pakistan. The sub-district is a socioeconomically deprived area with poverty rates around
182 50%, female literacy of 40-45%, and a high fertility rate (3·8 births per woman).[21] It is primarily agrarian
183 with close knit communities co-residing in large households (average 6·2 people per household). The
184 sub-district has 11 Union Councils (UC), the smallest administrative unit, each with a population of about
185 22,000-25,000. Each UC is serviced by a Primary Health Care Facility which houses a physician, midwife,
186 vaccinator, dispenser, and village-based Lady Health Workers (LHWs).

187

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

201

202 **Randomization and masking**

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

212

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

214 The intervention arm received the longer-duration peer-delivered psychosocial intervention (THPP+). It
215 consisted of 18 group-based "booster" sessions (from 7th to 36th month postnatal). Of these, the first 6
216 sessions were delivered monthly, then bi-monthly until 36 months. These sessions built on the shorter
217 duration intervention and were delivered by the same peers. The peers were lay married women who
218 lived in the same community as that of the depressed women and volunteered their time.
219
220 The key features of this psychosocial intervention, delivered by non-specialists, were peer-support,
221 behavioural activation, and problem-solving in a culturally sensitive, non-medicalized format, and
222 developmental activities for children up to the 36th month (See Appendix p.1-3 for the overall structure
223 of the intervention and peer characteristics.[19] A cascaded model of training and supervision was used
224 which included frequent competence assessments.[26] A competency checklist was used for supervision,
225 including direct observation of the peers, and to provide feedback. Each peer received this assessment
226 six times over the course of the program. During the feedback meetings, the supervisor discussed
227 checklist information along with session logs maintained by the peers. Competence was assessed
228 through observations and the checklist captured whether content was delivered as intended. Refresher
229 trainings were done 6 and 18 months after the initial training. The two-day training included re-
230 orientation to the intervention and its principles, as well as training on materials and use of job aids.
231
232
233 The intervention group sessions provided a safe environment for women to voice their problems, share
234 experiences of childcare, and provide support to women. Peer volunteers were trained to use culturally
235 grounded vignettes that served as tools to deliver health and well-being messages. The sessions aimed
236 for maternal well-being but also child-care and development by encouraging mother-infant interaction
237 and play. The intervention provided examples of age-appropriate activities, derived from the
238 UNICEF/WHO's Care for Development Package and encouraged demonstration of these activities during
239 the sessions. While these 'booster' group sessions did not focus on specific strategies to address
240 depression, the peer could still draw on her prior knowledge and skills of specific psychotherapeutic
241 elements such as behavioural activation when required. Additional details of the intervention are
242 reported elsewhere.[19,26] We defined overall treatment completion as attending 10 (out of 14) sessions
243 from pregnancy through 6 months postnatal (Phase 1: THPP, individual and group sessions) and 12 (out
244 of 18) sessions from 7 to 36 months postnatal (Phase 2: THPP+, booster group sessions).

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

248 women's depression status at enrollment, training all the 11 primary care facility-based physicians in the
249 subdistrict on the mental health Gap Action Programme (mhGAP) treatment guidelines for maternal
250 depression,[27] and providing depressed participants with a leaflet on how and where to seek appropriate
251 health care during pregnancy and beyond.
252

Procedures

254 Assessments were conducted 6 times over the course of the study (in pregnancy, and 3, 6, 12, 24, and
255 36 months postnatal). As originally specified, the current analysis utilizes mother and child outcome
256 data at 36 months.[22]All measures were extensively piloted. Assessments were done at the community
257 level within households of women by trained interviewers blind to the allocation status, all questions
258 were interviewer administered. Assessors inter-rater reliability was ensured through classroom-based
259 training which included role plays, followed by field practice sessions. During these sessions each pair of
260 assessors assessed up to 10 participants jointly and discussed their coding on each item to establish
261 inter-rater reliability prior to start of actual data collection.

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

266

Outcomes

268 The primary maternal outcome was depression symptoms assessed using the Patient Health
269 Questionnaire-9 (PHQ-9) and analyzed as symptom severity (total score) and remission (score < 510).
270 The secondary maternal outcomes were disability, assessed using WHO's Disability Assessment
271 Schedule, WHO-DAS[28], and current major depressive episode based on the *Structured Clinical Interview*
272 *for DSM-IV (SCID) Disor*ders.[29] Since it provides a clinically salient diagnostic outcome, the SCID was also
273 included to increase measurement robustness. Process data regarding THPP+ sessions attended,
274 competence scores of peers, duration of sessions, and peers' supervision attendance are described
275 elsewhere[30] and in the appendix (p.4-8).
276

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

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

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

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

298   **Statistical analyses**

299   For mother outcomes, anticipating a sample of 480 prenatally depressed women, we were powered at
300   90% to detect a remission rate of 65% in the prenatally depressed-intervention versus 45% in prenatally
301   depressed-control at the two-tailed 5% significance level and assuming a conservative ICC of 0.07 in
302   intervention arm and 0.05 in control.[22] For child outcomes, with th~~e~~is sample ~~sample~~ size, we were
303   powered at 90% to detect a difference between treatment arms of 3 points on the SDQ-TD at the two-
304   tailed 5% significance level assuming a standard deviation of 5.2 points and ICC of 0.04-0.08.[22] ~~We were~~
305   ~~also well-powered to test for equivalence of SDQ-TD score between children of prenatally depressed~~
306   ~~mothers in the treatment arm and children of prenatally non-depressed mothers in the EUC arm, with~~
307   ~~equivalence defined as the 95% confidence interval of the mean difference being between -2 and 2~~
308   ~~units.~~

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

315   The primary analyses were designed as intention-to-treat. Data from prenatally depressed and non-
316   depressed participants were analyzed jointly using linear mixed effects models so that all comparisons
317   of interest could be estimated from the same model. The identity link was used for continuous
318   outcomes to estimate differences in mean outcomes. Standardized mean differences (SMDs) and their
319   95% confidence intervals (CIs) were obtained using the method of Hedges.[37] In the primary model, we
320   included a random intercept for cluster and fixed effects for treatment arm (depressed intervention,
321   depressed control, non-depressed), union council (11 levels; the stratification variable), and variables
322   found to be imbalanced by loss to follow-up or at baseline (determined using p<0.10; see Appendix p.9-
323   14). Mixed models assume missing at random conditional on the covariates included in the statistical
324   model. Therefore, ~~W~~we include variables predictive of loss to follow-up to account for missing data.
325   ~~Mixed models assume missing at random, since we include baseline data in the model and adjusting for~~
326   ~~variables lost to follow-up will help account for missing data.~~ We used restricted maximum likelihood
327   (REML). The between-within method was used to apply small-sample bias corrections to the
328   intervention effect standard errors in the mixed effects framework.[38] These models also generated the
329   intra-cluster correlation (ICC) values.

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

8

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

340   ClinicalTrials.gov Identifier: NCT02658994 (registered on 21 January 2016).
341

342   **Role of Funding Source**

346

347   **Results**

348   From Oct 15, 2014 to Feb 25, 2016, we identified and randomly selected 40 village clusters out of 46 and
349   randomly assigned 20 village clusters each to intervention (THPP+) and control (EUC) arms. In all we
350   approached 1910 pregnant women; 287 prenatally depressed women in the control arm and 283 in the
351   intervention arm completed the baseline questionnaire. Of the prenatally non-depressed women
352   approached, 584 were enrolled, yielding comparable numbers of prenatally depressed and prenatally
353   non-depressed women in each of the 40 village clusters.
354   Of the 1,154 participants enrolled at baseline, 889 (77.0%) were successfully interviewed at 36 months:
355   206 (72.5%) intervention, 216 (75.3%) EUC, and 467 (80.0%) prenatally non-depressed (Figure 1). There
356   was no differential loss to follow up by treatment arm and no differences in adverse events
357   (Appendix p.9-14).
358

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

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

370   Just over two-thirds of prenatally depressed women in the intervention arm completed the THPP+
371   intervention (Appendix p.5). Only 63% of women completed treatment during Phase 2 compared to
372   nearly 80% treatment completion in Phase 1. The competence levels of the peers declined over the
373   implementation period, particularly in the period after time point 3 (i.e. at 12 months postnatal) where
374   we introduced new content for the booster sessions and the frequency of peer supervisions dropped
375   from monthly to every two months (Aappendix p.6-8) show high and low scoring peers and their
376   competence levels over time
377

378   For all prenatally depressed women, depression scores dropped meaningfully, regardless of their arm
379   allocation (Figure 2). There were no significant differences in depression outcomes between arms
380   (THPP+ vs EUC) at 36 months postnatal (Table 2). The adjusted standardized mean difference (SMD) in
381   depressive symptom severity (PHQ-9) between arms was -0.13 (95% CI -0.33 to 0.07) and the risk ratio
382   (RR) for depression remission (PHQ-9<10) was 1.00 (95% CI 1.13 to 0.97). Turning to the secondary
383   outcomes, we observed a relatively larger difference between arms on the secondary outcomes of SCID-
384   based major depressive episode at 36 months (22% control vs 16.5% intervention) (RR=0.67, 95% CI 0.43
385   to 1.05). The intervention effects on disability (SMD=-0.12, 95% CI -0.33 to 0.09) were not statistically
386   significant.
387
388   The prevalence of ~~SCID-based~~ major depressive episodes (SCID-based) among the three groups of
389   women (~~:~~ the prenatally depressed intervention arm, prenatally depressed control arm, and prenatally
390   non-depressed women) became increasingly similar in the proportion depressed by 36 months
391   postnatal (F̶figure 2). The THPP+ intervention arm showed higher convergence with the prenatally non-
392   depressed women at 36 months compared to the control arm, so much so that there was not a
393   statisticaly significant difference in the probability of being depressed (using SCID) between the
394   intervention arm women (at 16.5%) and the prenatally non-depressed (9.0% of whom were depressed
395   at 36 months, RR 0·74 (0·46 to 1·18).
396
397   We also examined intervention arm differences in depression severity (PHQ-9 scores) at 36 months
398   postnatal by potential moderators assessed at baseline. There was no strong evidence of meaningful
399   moderation of the intervention effect by any of these factors (Appendix p.18).
400
401   For the child primary outcome of SDQ-TD, the mean adjusted difference between intervention and
402   control arms was -0·10; 95% CI -1·39 to 1·19 (Table 2). There were also no meaningful differences
403   between the two arms in the secondary outcomes of Receptive Language and Fine Motor scores (from
404   BSITD) (Table 2),

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

407
408
409   Children of prenatally non-depressed mothers had somewhat better SDQ-TD scores than the
410   intervention or control arm children (e.g. SDQ-TD=13.7 among the children of the prenatally non-
411   depressed vs. 14.7 in the intervention arm, p-value=0.07, Table 2). In other words, the prenatal
412   depression episode predicted slightly worse SDQ-TD scores at 36 months of age, independent of the
413   intervention. Exploratory analyses of the five sub-scales of the SDQ separately showed that this overall
414   difference was driven by the hyperactivity and conduct problems sub-scales, with negligible differences
415   by prenatal depression status for the peer problems, emotional problems, and the pro-social scales
416   (Appendix p.19).   As an example, the adjusted mean difference in the hyperactivity scores between the
417   non-depressed and the control arm was 0·31 (95% CI -0·62, 0·01,). The receptive language, fine motor
418   and physical growth indicators did not meaningfully differ between the children of prenatally non-
419   depressed and depressed mothers, regardless of intervention arm.

420

421   **Discussion**

10

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

431

432 The overall small effect sizes and lack of statistical significance on maternal outcomes might be
433 attributable to several factors. First, the intervention was a non-specialist, lay peer delivered
434 psychosocial intervention. The lay peers were housewives from rural villages, without prior training or
435 work experience. They were trained and supervised by non-specialists using a cascaded model (with no
436 direct specialist contact).[26] Second, this lay peer delivered approach was used to inform scaling-up of
437 maternal mental health services through existing health systems and community resources. It is possible
438 that this non-specialist, low-intensity design, coupled with longer-duration implementation, led to what
439 Chambers and colleagues refer to as 'voltage drop' (the intervention loses some degree of its potency
440 (or fidelity) when moving from efficacy to effectiveness in the real world) and 'program drift' (the
441 intervention deviates from its manualized or implementation protocols).[41] We saw a substantial drop in
442 women attending the maintenance group sessions delivered every two months beyond the 6th month
443 postnatal period. This implementation challenge of attendance is reported in other community based
444 interventions targeting maternal outcomes.[42] Women reported that the ~~womeny~~ lost interest in
445 attending these group sessions or found it demanding on their time. This challenge of poor attendance
446 and how to best address it came up regularly in peer supervision meetings. In addition to attempts to
447 add more interesting content to the sessions, we made sure that community health workers reminded
448 women about the sessions and followed up with those who missed a session. Addressing sustaining
449 participant interest and limited time is seminal in community-based program success and has been
450 highlighted as an important challenge in other low resource settings.[43] Finally, it is possible that, since
451 the booster group sessions did not focus on specific strategies to address depression as mentioned
452 earlier, the intervention arm was not so different from the EUC arm. There were no detectable
453 treatment effects at the 6 month postnatal wave.[16]The boosters introduced after the 6 month~~s~~ mark did
454 not change this - we continue to see no intervention effect at 36 months.
455
456
457 Maintaining the competence levels and motivation of the peers over the multi-year long
458 implementation period was challenging. Perhaps the cascaded model of supervision via non-specialists
459 led to dropped potency (or fidelity) of sessions. We explored the perceptions of the peers about this
460 intervention in a nested qualitative study at six months postnatal.[26] We did not find any negative
461 perceptions towards the intervention which might have led to change in motivation or competence
462 levels.. The reduced frequency of supervision sessions of the peers from every month to every two
463 months seems to have contributed to the drop in competence levels. This drop in supervisory intensity
464 has been shown to reduce effectiveness of known approaches.[44,45] Another contributing factor to the
465 drop in competence levels was the addition of new content, beyond the 6th month postnatal period:
466 both the high and low scoring peers experienced a drop in competence levels when more content was

467 added. Finding the right balance of content vs capacity of peers, and maintaining fidelity, is an
468 implementation challenge reported in other programs.[46,47]
469
470 Finally, we saw a substantial improvement in the control arm (lowered rates and higher recovery from
471 depression). The control arm recovery rates were higher compared to our previous studies from 2008
472 and 2020.[8,48] This could be attributable to "regression to the mean" or to spillover of the intervention
473 through the LHWs who regularly interact with LHWs responsible for women residing in control arm sites.
474 The active control arm received enhanced usual care and some evidence suggests that informing people
475 about their illness status improves outcomes,[49] which raises important methodological (study design)
476 issues for trials that are embedded within community settings. Perhaps future trials, using similar active
477 control arms ought to consider equivalence or non-inferiority trial designs to avoid a nonsignificant
478 superiority trial being wrongly interpreted, as proof of no difference between the two active
479 comparisons.[50]
480
481 We found no clear indications of sub-group differences. This points to the challenge of intervention
482 targeting especially given that approximately one-fifth of women did not respond to this treatment,
483 indicating a need for different interventions. A collaborative care model where non-responders can be
484 detected earlier and connected to more specialized care may be needed.
485
486 The findings suggesting an association (albeit weak) between prenatal depression and child
487 socioemotional outcomes at age 3, which is not mitigated by the intervention, mirrors results from our
488 previous trial with a different sample, designed to examine a similar intervention's effect on children at
489 age seven.[4] Children's socioemotional development may be less likely to 'catch up' during the resolution
490 phase of depression within the postnatal period. If confirmed in other studies, differences in
491 socioemotional, but not other developmental outcomes, may point to specific mechanisms in maternal-
492 child intergenerational transmission of risk.[11]
493
494 While chronic depression likely has the greatest effect on child development,[51] the majority of children
495 in our study were exposed to varying maternal depression levels, including periods of low or no
496 depression symptoms over the study period. An association between the prenatal episode and child
497 socioemotional outcomes at 36 months would be consistent with 'foetal programming,'[52] and is
498 supported by evidence of linking prenatal depression to a number of child outcomes regardless of
499 postnatal depression.[53] However, inconsistent with foetal programming, another study concluded that a
500 reduction in maternal depression levels postnatally lead to 'near normal levels' of child behavioural
501 problems.[54]
502
503 A possible explanation for the lack of intervention effects on child outcomes is that, unlike the trained
504 specialist delivery in Stein et al's study, this intervention was delivered by lay peers in a low-income
505 country.[54] As mentioned above, we experienced a 'voltage drop' over time in terms of reductions in
506 both session frequency and attendance, posing significant challenges to sustained implementation.[41]
507 Another potential explanation of no group differences in socioemotional outcomes is that the variability
508 in maternal depression symptoms in the first 3 years could itself be a risk factor. Given that depression is

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

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

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

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

13

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

560

561 **Data Sharing Statement**

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

564

565 **Authors' contributions**

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

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

574

575 **Conflict of Interest**

576 We declare we have no conflict of interest.

577

References

578

579   1.   Gelaye B, Rondon MB, Araya R, Williams MA. Epidemiology of maternal depression, risk factors,
580 and child outcomes in low-income and middle-income countries. *The Lancet Psychiatry* 2016; **3**(10): 973-
581 82.10.1016/S2215-0366(16)30284-X
582   2.   Goodman SH, Rouse MH, Connell AM, Broth MR, Hall CM, Heyward D. Maternal Depression and
583 Child Psychopathology: A Meta-Analytic Review. *Clin Child Fam Psychol Rev* 2011; **14**(1): 1-
584 27.10.1007/s10567-010-0080-1
585   3.   Patel V, Saxena S, Lund C, Thornicroft G, Baingana F, Bolton P, Chisholm D, Collins PY, Cooper JL,
586 Eaton J, Herrman H, Herzallah MM, Huang Y, Jordans MJD, Kleinman A, Medina-Mora ME, Morgan E,
587 Niaz U, Omigbodun O, Prince M, Rahman A, Saraceno B, Sarkar BK, De Silva M, Singh I, Stein DJ, Sunkel
588 C, UnÜtzer J. The Lancet Commission on global mental health and sustainable development. *The Lancet*
589 2018; **392**(10157): 1553-98.10.1016/S0140-6736(18)31612-X
590   4.   Maselko J, Sikander S, Bhalotra S, Bangash O, Ganga N, Mukherjee S, Egger H, Franz L, Bibi A,
591 Liaqat R, Kanwal M, Abbasi T, Noor M, Ameen N, Rahman A. Effect of an early perinatal depression
592 intervention on long-term child development outcomes: follow-up of the Thinking Healthy Programme
593 randomised controlled trial. *Lancet Psychiatry* 2015; **2**(7): 609-17.10.1016/s2215-0366(15)00109-1
594   5.   Rahman A, Fisher J, Bower P, Luchters S, Tran T, Yasamy MT, Saxena S, Waheed W. Interventions
595 for common perinatal mental disorders in women in low- and middle-income countries: a systematic
596 review and meta-analysis. *Bulletin of the World Health Organization* 2013; **91**(8): 593-
597 601I.10.2471/BLT.12.109819
598   6.   Tomlinson M, Rotheram-Borus MJ, Scheffler A, le Roux I. Antenatal depressed mood and child
599 cognitive and physical growth at 18-months in South Africa: a cluster randomised controlled trial of
600 home visiting by community health workers. *Epidemiol Psychiatr Sci* 2018; **27**(6): 601-
601 10.10.1017/s2045796017000257
602   7.   Goodman SH, Garber J. Evidence-Based Interventions for Depressed Mothers and Their Young
603 Children. *Child Dev* 2017; **88**(2): 368-77.10.1111/cdev.12732
604   8.   Baranov V, Bhalotra S, Biroli P, Maselko J. Maternal Depression, Women's Empowerment, and
605 Parental Investment: Evidence from a Randomized Control Trial. *American Economic Review* 2020;
606 **110**(3): 824-59
607   9.   Gunlicks ML, Weissman MM. Change in child psychopathology with improvement in parental
608 depression: A systematic review. *Journal of the American Academy of Child and Adolescent Psychiatry*
609 2008; **47**(4): 379-89.10.1097/CHI.0b013e3181640805
610   10.   Cuijpers P, Weitz E, Karyotaki E, Garber J, Andersson G. The effects of psychological treatment of
611 maternal depression on children and parental functioning: a meta-analysis. *Eur Child Adolesc Psych*
612 2015; **24**(2): 237-45.10.1007/s00787-014-0660-6
613   11.   Goodman SH, Cullum KA, Dimidjian S, River LM, Kim CY. Opening windows of opportunities:
614 Evidence for interventions to prevent or treat depression in pregnant women being associated with
615 changes in offspring's developmental trajectories of psychopathology risk. *Development and*
616 *Psychopathology* 2018; **30**(3): 1179-96.10.1017/S0954579418000536
617   12.   Loechner J, Starman K, Galuschka K, Tamm J, Schulte-Körne G, Rubel J, Platt B. Preventing
618 depression in the offspring of parents with depression: A systematic review and meta-analysis of
619 randomized controlled trials. *Clinical Psychology Review* 2018; **60**: 1-
620 14.https://doi.org/10.1016/j.cpr.2017.11.009
621   13.   Milgrom J, Holt CJ, Bleker LS, Holt C, Ross J, Ericksen J, Glover V, O'Donnell KJ, de Rooij SR,
622 Gemmill AW. Maternal antenatal mood and child development: an exploratory study of treatment
623 effects on child outcomes up to 5 years. *J Dev Orig Health Dis* 2019; **10**(2): 221-
624 31.10.1017/s2040174418000739

625   14.   Kersten-Alvarez LE, Hosman CMH, Riksen-Walraven JM, van Doesum KTM, Hoefnagels C. Long-
626   term effects of a home-visiting intervention for depressed mothers and their infants. *Journal of Child*
627   *Psychology and Psychiatry* 2010; **51**(10): 1160-70.10.1111/j.1469-7610.2010.02268.x
628   15.   Rotheram-Borus MJ, Arfer KB, Christodoulou J, Comulada WS, Stewart J, Tubert JE, Tomlinson
629   M. The association of maternal alcohol use and paraprofessional home visiting with children's health: A
630   randomized controlled trial. *J Consult Clin Psychol* 2019; **87**(6): 551-
631   62.10.1037/ccp0000408PMC6775769
632   16.   Sikander S, Ahmad I, Atif N, Zaidi A, Vanobberghen F, Weiss HA, Nisar A, Tabana H, Ain QU, Bibi
633   A, Bilal S, Bibi T, Liaqat R, Sharif M, Zulfiqar S, Fuhr DC, Price LN, Patel V, Rahman A. Delivering the
634   Thinking Healthy Programme for perinatal depression through volunteer peers: a cluster randomised
635   controlled trial in Pakistan. *Lancet Psychiatry* 2019; **6**(2): 128-39.10.1016/s2215-0366(18)30467-x
636   17.   Vanobberghen F, Weiss HA, Fuhr DC, Sikander S, Afonso E, Ahmad I, Atif N, Bibi A, Bibi T, Bilal S,
637   De Sa A, D'Souza E, Joshi A, Korgaonkar P, Krishna R, Lazarus A, Liaqat R, Sharif M, Weobong B, Zaidi A,
638   Zuliqar S, Patel V, Rahman A. Effectiveness of the Thinking Healthy Programme for perinatal depression
639   delivered through peers: Pooled analysis of two randomized controlled trials in India and Pakistan. *J*
640   *Affect Disord* 2019.https://doi.org/10.1016/j.jad.2019.11.110
641   18.   Fuhr DC, Weobong B, Lazarus A, Vanobberghen F, Weiss HA, Singla DR, Tabana H, Afonso E, De
642   Sa A, D'Souza E, Joshi A, Korgaonkar P, Krishna R, Price LN, Rahman A, Patel V. Delivering the Thinking
643   Healthy Programme for perinatal depression through peers: an individually randomised controlled trial
644   in India. *Lancet Psychiatry* 2019; **6**(2): 115-27.10.1016/s2215-0366(18)30466-8
645   19.   Atif N, Bibi A, Nisar A, Zulfiqar S, Ahmed I, LeMasters K, Hagaman A, Sikander S, Maselko J,
646   Rahman A. Delivering maternal mental health through peer volunteers: a 5-year report. *International*
647   *Journal of Mental Health Systems* 2019; **13**(1): 62.10.1186/s13033-019-0318-3
648   20.   National Institute of Population Studies - NIPS/Pakistan & ICF Intternational. Pakistan
649   Demographic and Health Survey 2012-2013. Islamabad: NIPS/Pakistan and ICF International, 2012-3.
650   21.   National Institute of Population Studies Islamabad. Pakistan Demographic and Health Survey.
651   Islamabad, Pakistan: National Institute of Population Studies, 2013.
652   22.   Turner EL, Sikander S, Bangash O, Zaidi A, Bates L, Gallis J, Ganga N, O'Donnell K, Rahman A,
653   Maselko J. The effectiveness of the peer-delivered Thinking Healthy PLUS (THPP+) Program for maternal
654   depression and child socioemotional development in Pakistan: study protocol for a randomized
655   controlled trial. *Trials* 2016; **17**(1): 442.10.1186/s13063-016-1530-yPMC5017048
656   23.   Gallis J, Maselko J, O'Donnell K, Song KE, Saqib K, Turner EL, Sikander S. Criterion-related validity
657   and reliability of the Urdu version of the patient health questionnaire in community-based women in
658   Pakistan. *PeerJ* 2018; **6**(e5185).https://doi.org/10.7717/peerj.5185
659   24.   Sikander S, Ahmad I, Bates LM, Gallis J, Hagaman A, O'Donnell K, Turner EL, Zaidi A, Rahman A,
660   Maselko J. Cohort Profile: Perinatal depression and child socioemotional development ; the Bachpan
661   cohort study from rural Pakistan. *BMJ Open* 2019; **9**(5): e025644.10.1136/bmjopen-2018-025644
662   25.   Sikander S, Lazarus A, Bangash O, Fuhr DC, Weobong B, Krishna RN, Ahmad I, Weiss HA, Price L,
663   Rahman A, Patel V. The effectiveness and cost-effectiveness of the peer-delivered Thinking Healthy
664   Programme for perinatal depression in Pakistan and India: the SHARE study protocol for randomised
665   controlled trials. *Trials* 2015; **16**: 14.10.1186/s13063-015-1063-9
666   26.   Atif N, Nisar A, Bibi A, Khan S, Zulfiqar S, Ahmad I, Sikander S, Rahman A. Scaling-up
667   psychological interventions in resource-poor settings: training and supervising peer volunteers to deliver
668   the 'Thinking Healthy Programme' for perinatal depression in rural Pakistan. *Global mental health*
669   *(Cambridge, England)* 2019; **6**: e4.10.1017/gmh.2019.4PMC6521132
670   27.   Organization. WH. WHO Guidelines Approved by the Guidelines Review Committee. mhGAP
671   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

719   44.   Baqui AH, El-Arifeen S, Darmstadt GL, Ahmed S, Williams EK, Seraji HR, Mannan I, Rahman SM,
720   Shah R, Saha SK, Syed U, Winch PJ, Lefevre A, Santosham M, Black RE. Effect of community-based
721   newborn-care intervention package implemented through two service-delivery strategies in Sylhet
722   district, Bangladesh: a cluster-randomised controlled trial. *The Lancet* 2008; **371**(9628): 1936-
723   44.10.1016/S0140-6736(08)60835-1
724   45.   Tripathy P, Nair N, Barnett S, Mahapatra R, Borghi J, Rath S, Rath S, Gope R, Mahto D, Sinha R,
725   Lakshminarayana R, Patel V, Pagel C, Prost A, Costello A. Effect of a participatory intervention with
726   women's groups on birth outcomes and maternal depression in Jharkhand and Orissa, India: a cluster-
727   randomised controlled trial. *The Lancet* 2010; **375**(9721): 1182-92.10.1016/S0140-6736(09)62042-0
728   46.   Yousafzai AK, Rasheed MA, Rizvi A, Armstrong R, Bhutta ZA. Parenting Skills and Emotional
729   Availability: An RCT. *Pediatrics* 2015; **135**(5): e1247-e57.10.1542/peds.2014-2335
730   47.   Hafeez A, Mohamud B, Shiekh M, Shah I, Jooma R. Lady health workers programme in Pakistan:
731   challenges, achievements and the way forward. *JPMA* 2011; **61**(210):
732   48.   Rahman A, Malik A, Sikander S, Roberts C, Creed F. Cognitive behaviour therapy-based
733   intervention by community health workers for mothers with depression and their infants in rural
734   Pakistan: a cluster-randomised controlled trial. *Lancet* 2008; **372**(9642): 902-9
735   49.   Paterick TE, Patel N, Tajik AJ, Chandrasekaran K. Improving health outcomes through patient
736   education and partnerships with patients. *Proc (Bayl Univ Med Cent)* 2017; **30**(1): 112-
737   3.10.1080/08998280.2017.11929552
738   50.   Lesaffre E. Superiority, equivalence, and non-inferiority trials. *Bulletin of the NYU hospital for
739   joint diseases* 2008; **66**(2): 150-4
740   51.   Netsi E, Pearson RM, Murray L, Cooper P, Craske MG, Stein A. Association of Persistent and
741   Severe Postnatal Depression With Child OutcomesAssociation of Persistent and Severe Postnatal
742   Depression With Child OutcomesAssociation of Persistent and Severe Postnatal Depression With Child
743   Outcomes. *JAMA psychiatry* 2018; **75**(3): 247-53.10.1001/jamapsychiatry.2017.4363
744   52.   Robinson R, Lahti-Pulkkinen M, Heinonen K, Reynolds RM, Räikkönen K. Fetal programming of
745   neuropsychiatric disorders by maternal pregnancy depression: a systematic mini review. *Pediatric
746   Research* 2019; **85**(2): 134-45.10.1038/s41390-018-0173-y
747   53.   O'Donnell KJ, Glover V, Barker ED, O'Connor TG. The persisting effect of maternal mood in
748   pregnancy on childhood psychopathology. *Development and Psychopathology* 2014; **26**(2): 393-
749   403.10.1017/S0954579414000029
750   54.   Stein A, Netsi E, Lawrence PJ, Granger C, Kempton C, Craske MG, Nickless A, Mollison J, Stewart
751   DA, Rapa E, West V, Scerif G, Cooper PJ, Murray L. Mitigating the effect of persistent postnatal
752   depression on child outcomes through an intervention to treat depression and improve parenting: a
753   randomised controlled trial. *Lancet Psychiatry* 2018; **5**(2): 134-44.10.1016/s2215-0366(18)30006-3
754   55.   Frankenhuis WE, Del Giudice M. When Do Adaptive Developmental Mechanisms Yield
755   Maladaptive Outcomes? *Developmental Psychology* 2012; **48**(3): 628-42.10.1037/a0025629

756

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

Figure 1. CONSORT sample flow chart

46 village clusters assessed for eligibility

6 village clusters randomly selected for exclusion as per sample size calculations

40 village clusters included (1910 potential participants)

Excluded due to ineligibility = 154
- In 1st or 2nd trimester of pregnancy = 110
- Age <18 years =8
- Does not intend to remain in area for duration of study=11
- Does not speak Urdu, Punjabi, Potohari or English=22
- Needed imminent care= 3
Refusals = 25

Screening

288 assigned to enhanced usual care (EUC) control group
Mean cluster size 14.4 (SD 0.8)

284 women assigned to THPP intervention group
Mean cluster size 14.2 (SD 0.4)

1,159 prenatally non-depressed women
Mean cluster size 29.0 (SD 7.7)

Refusals (n=1).

Refusals (n=1).

Prenatally non-depressed women not invited to participate in the cohort (n=566).
Refusals (n=9).

Baseline

EUC control group:
287 women with data
Mean cluster size 14.4 (SD 0.7)

THPP intervention group:
283 women with data
Mean cluster size 14.2 (SD 0.4)

Prenatally non-depressed group:
584 women with data
Mean cluster size 14.6 (SD 1.4)

Lost to follow-up: moved (n=9), temporarily absent (n=37), loss of child (n=26), child illness (n=3).
Refusals (n=1).

Lost to follow-up: moved (n=7), temporarily absent (n=21), loss of child (n=25), child illness (n=4).
Refusals (n=3).

Lost to follow-up: moved (n=6), temporarily absent (n=78), loss of child (n=42), child illness (n=4).
Refusals (n=1).

3-month follow-up

EUC control group:
221 dyads with follow-up data
Mean cluster size 10.6 (SD 2.5)

THPP intervention group:
223 dyads with follow-up data
Mean cluster size 11.2 (SD 1.2)

Prenatally non-depressed group:
451 dyads with follow-up data
Mean cluster size 11.3 (SD 2.2)

Lost to follow-up: moved (n=2), temporarily absent (n=16), loss of child (n=1).
Refusals (n=1).

Lost to follow-up: moved (n=2), temporarily absent (n=11).
Refusals (n=1).

Lost to follow-up: moved (n=3), temporarily absent (n=51), loss of child (n=2).
Refusals (n=2).

6-month follow-up

EUC control group:
228 dyads with follow-up data
Mean cluster size 11.4 (SD 2.0)

THPP intervention group:
230 dyads with follow-up data
Mean cluster size 11.5 (SD 1.6)

Prenatally non-depressed group:
471 dyads with follow-up data
Mean cluster size 11.8 (SD 1.9)

Lost to follow-up: temporarily absent (n=10), loss of child (n=3).
Refusals (n=3).

Lost to follow-up: moved (n=3), temporarily absent (n=9), mother was ill (n=1), loss of child (n=2), mother died (n=1). **Refusals** (n=2).

Lost to follow-up: moved (n=2), temporarily absent (n=22), mother was ill (n=1), loss of child (n=2), mother died (n=1), child hernia (n=1), child disabled (n=1)
Refusals (n=3).

12-month follow-up

EUC control group:
228 dyads with follow-up data
Mean cluster size 11.4 (SD 1.7)

THPP+ intervention group:
223 dyads with follow-up data
Mean cluster size 11.2 (SD 1.5)

Prenatally non-depressed group:
489 dyads with follow-up data
Mean cluster size 12.2 (SD 1.8)

Lost to follow-up: moved (n=2), temporarily absent (n=12), loss of child (n=2), child illness (n=1).
Refusals (n=2).

Lost to follow-up: moved (n=4), temporarily absent (n=15), loss of child (n=1).

Lost to follow-up: moved (n=5), temporarily absent (n=23), child illness (n=1).
Refusals (n=4).

24-month follow-up

EUC control group:
220 dyads with follow-up data
Mean cluster size 11.0 (SD 1.5)

THPP+ intervention group:
210 dyads with follow-up data
Mean cluster size 10.5 (SD 1.6)

Prenatally non-depressed group:
473 dyads with follow-up data
Mean cluster size 11.8 (SD 2.0)

Lost to follow-up: moved (n=8), temporarily absent (n=4), child illness (n=1), loss of child (n=2), child developmentally delayed (n=1), mom death (n=1).
Refusals (n=3).

Lost to follow-up: moved (n=9), temporarily absent (n=10), child illness (n=2), loss of child (n=1), child developmentally delayed (n=1).
Refusals (n=1).

Lost to follow-up: moved (n=16), temporarily absent (n=19), loss of child (n=1), child developmentally delayed (n=2).
Refusals (n=6).

36-month follow-up

EUC control group:
216 dyads with follow-up data
Mean cluster size 10.8 (SD 1.8)

THPP+ intervention group:
206 dyads with follow-up data
Mean cluster size 10.3 (SD 1.6)

Prenatally non-depressed group:
467 dyads with follow-up data
Mean cluster size 11.7 (SD 1.8)

Figure 2

Figure 2 Title
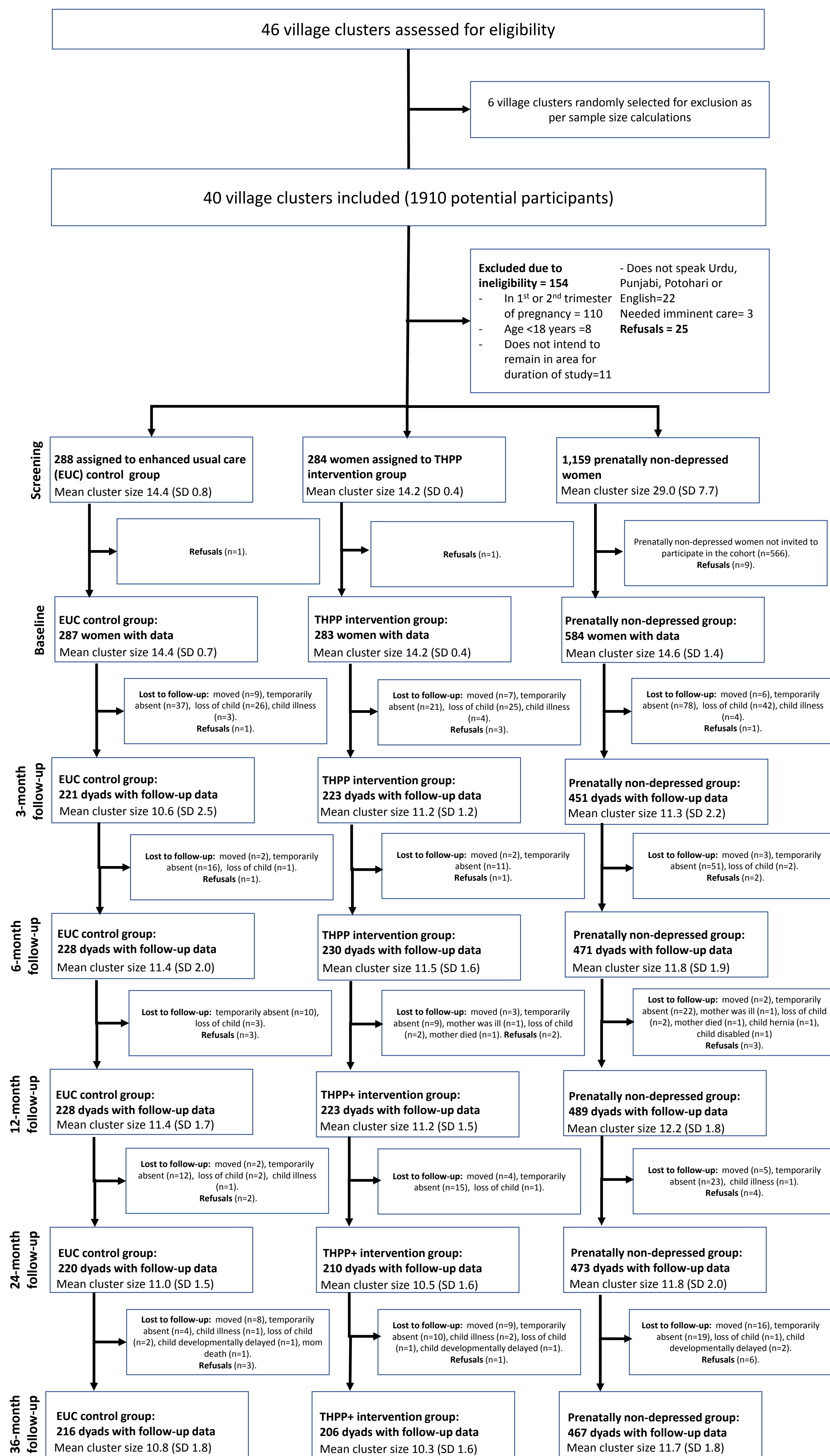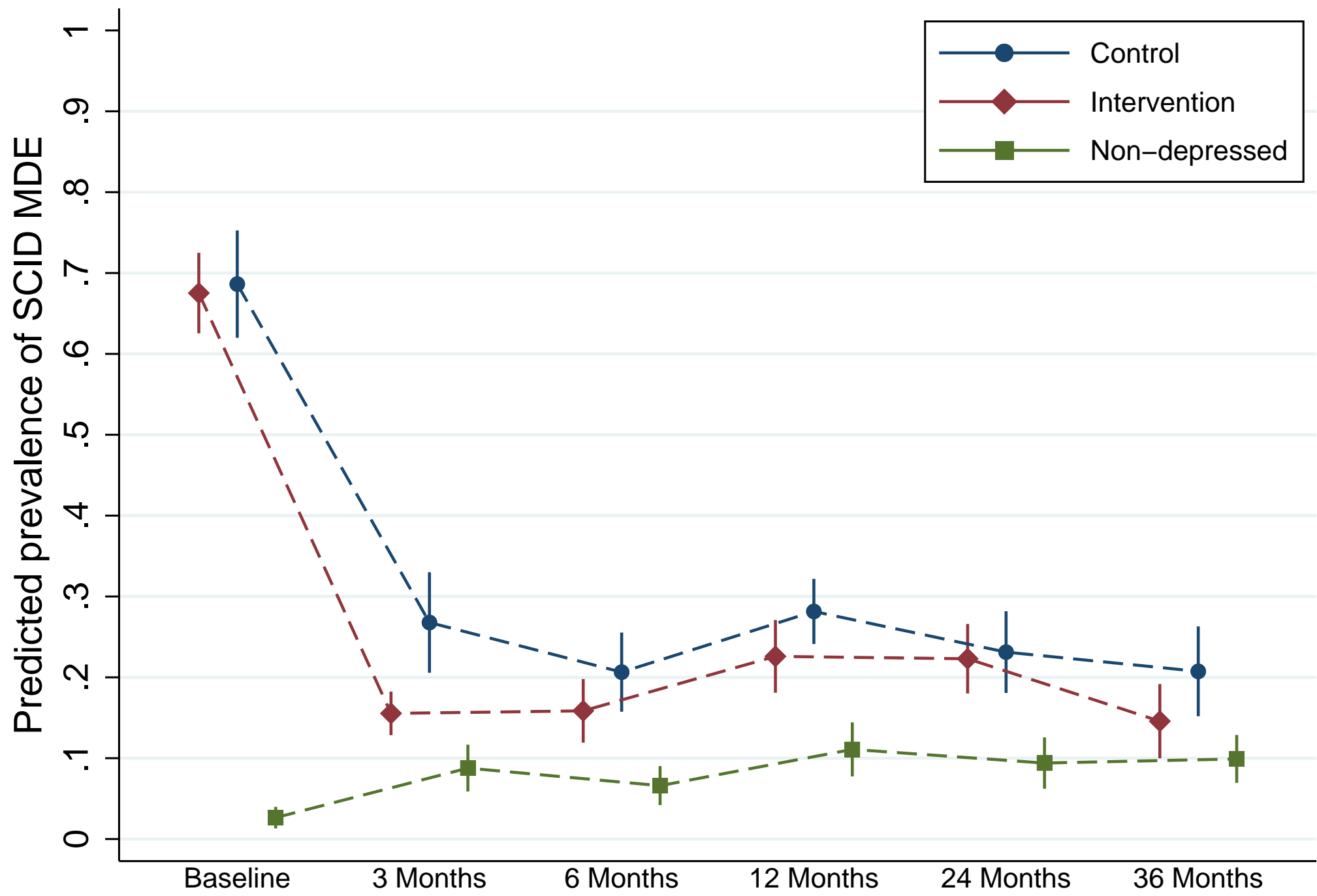
Figure 2 title

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.

Appendix tables and figures
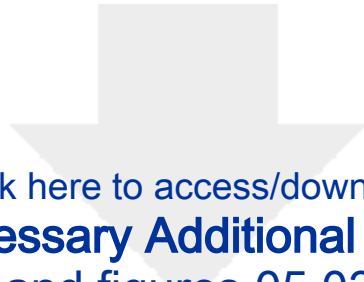
Click here to access/download
Necessary Additional Data
Appendix tables and figures 05.03 CLEAN v2.docx

Trials

CrossMark

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).

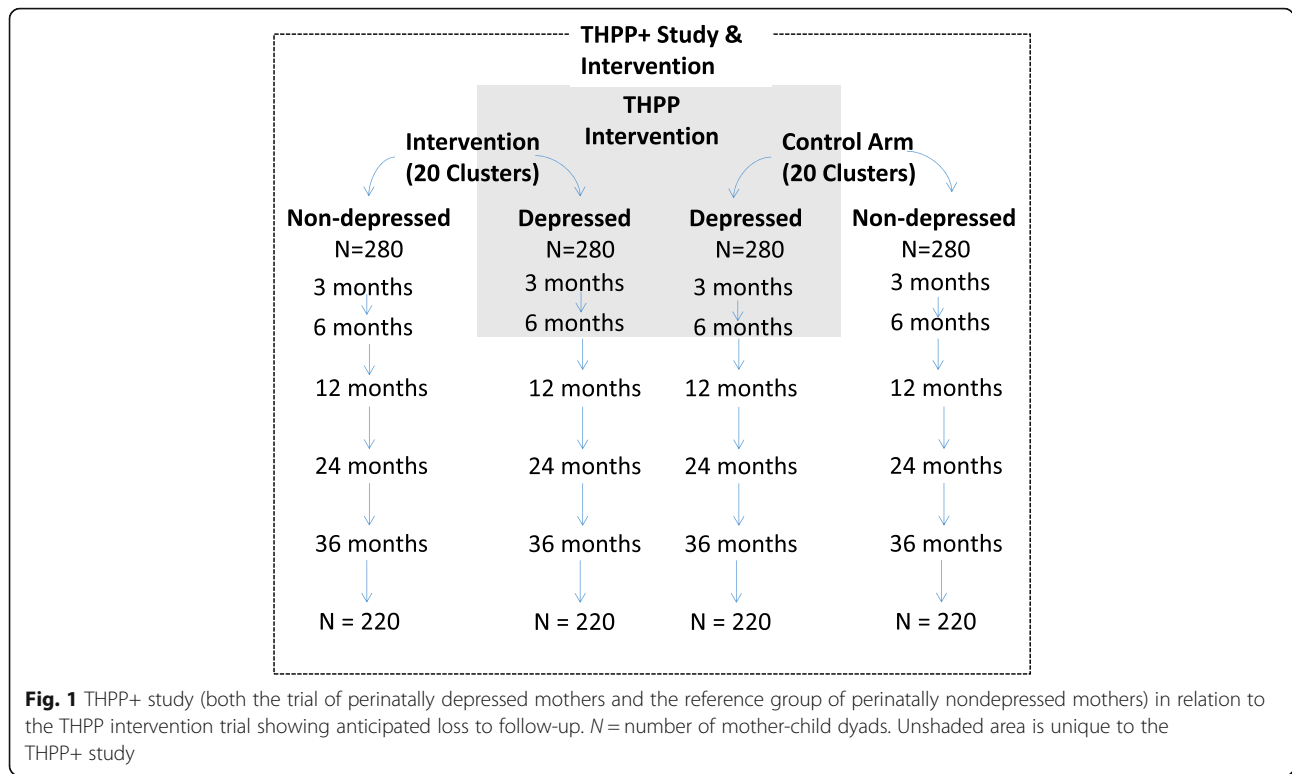#### Recruitment of study participants: inclusion and exclusion criteria

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+)

Outcomes	Source of data	Postnatal months				
	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
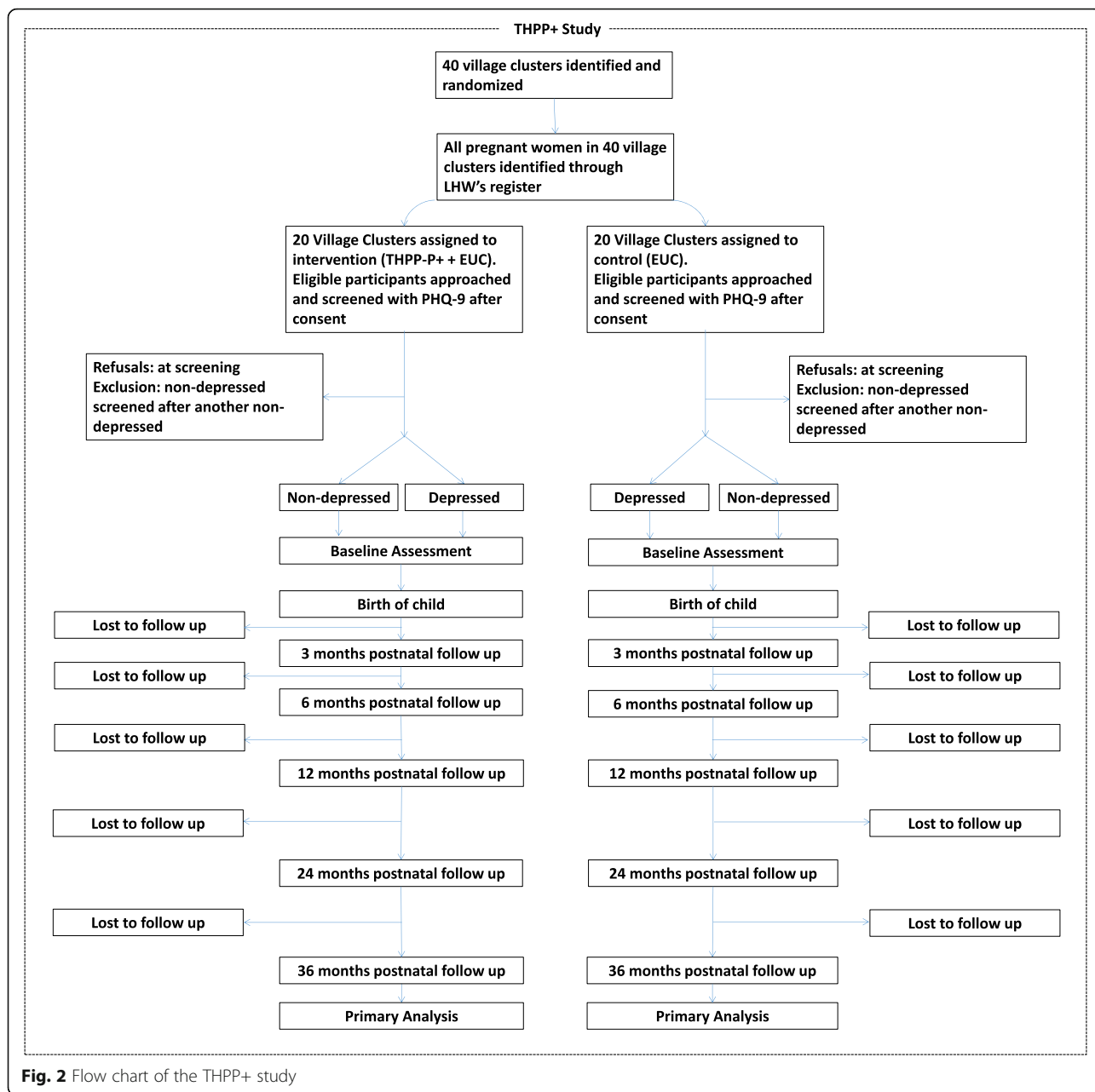
Fig. 2 Flow chart of the THPP+ study

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

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.

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.