Bachpan Trial Analysis Plan

Version 5.0

October 14, 2019

Contents

1 Purpose and scope

This document provides the plan of analysis (statistical analysis plan) for the South Asian Hub for Advocacy, Resarch, and Education for mental health (SHARE) Child trial in Pakistan, conducted between October 2014 and August 2019. In Summer of 2018, the trial was renamed Bachpan. This plan covers the primary analysis comparing intervention arms and depressed and non-depressed perinatal women on their depressive symptoms at 36 months of age, in addition to comparing their children on various developmental outcomes at 36 months of age.

2 Description of the Trial

The Thinking Healthy Program (THP) is a community health worker (CHW)-delivered intervention which was developed in Pakistan and shown to have beneficial effects on both maternal depression and short-term child outcomes.

The current trial is described in detail in Turner et al.[1] In brief, the peer-delivered version of the THP was developed because of the lack of resources in many low and middle income (LMIC) countries to have professionals deliver the intervention. The present trial (THPP+) began in October 2014, and is an extension of the Thinking Healthy Program Peer-Delivered (THPP) trial, which is a six-month intervention delivered to women with perinatal depression, with the goal of reducing the burden of perinatal depression.[2] THPP+ is an extension of THPP in two ways: we follow the same sample of perinatally depressed women an additional 30 months; we recruited a sample of perinatally non-depressed women equal in sample size with the perinatally depressed group, and follow them for a total of 36 months, as well.

2.1 Principal research objectives and hypotheses to be addressed

The primary objective of the study is to evaluate the impact of a 36-month perinatal peer-delivered community-based perinatal depression intervention on (a) maternal depression and (b) child development. Our primary hypothesis for the perinatally depressed mothers is that the intervention will result in lower prevalence of depression at three 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 three years of age. Additional child hypotheses address proposed mediators and moderators of the effects of the perinatal depression intervention on child outcomes; however, the mediators will be analyzed as part of a separate paper and so are not included in this document.

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 non-depressed mothers and children as well as, secondarily, to determine whether there are any carryover effects of the intervention to benefit perinatally non-depressed mothers and children.

2.1.1 Primary and secondary outcomes

The primary and secondary outcomes are shown in Table 1. The primary mother outcome is Patient Health Questionnaire (PHQ-9). The primary child outcome is the Total Difficulties (TD) score from the Strengths and Difficulties Questionnaire (SDQ).

2.2 Trial Design

THPP+ is a stratified cluster randomized controlled trial, with 40 village clusters allocated in a 1:1 ratio to the intervention and control arms within 11 strata defined by Union Councils. The unit of randomization is a village cluster (a population of 2400–3600 with 2-3 adjacent catchment areas of LHWs). Participants are recruited from the community. The reason for choosing a village cluster as the unit of randomization is to minimize contamination between trial participants, since the delivery of the intervention is community based.

Approximately 14 women who screened depressed on the PHQ-9 (≥10) in each of the 40 village clusters were invited to participate in the trial. Bachpan also contains an embedded cohort of non-depressed women. In each of the 40 village clusters, an equal number of women who screened non-depressed on the PHQ-9 (i.e., approximately 14 per cluster) were invited to participate.

2.3 Randomization

Within each of the 11 union councils of the sub-district, village clusters were randomized to Intervention or Usual Care control using a 1:1 allocation ratio. Randomization was done before the participants were recruited. Research teams responsible for identifying, obtaining consent and recruiting trial participants were blind to the allocation status (to minimize the post-randomization recruitment bias). Randomization was stratified by union councils, to help minimize the imbalance between arms by factors likely to be associated with the primary outcome and reduce the between-cluster variability, hence increasing the power of the study. Allocation of clusters was carried out by an independent statistician based at the London School of Hygiene and Tropical Medicine (LSHTM) using a computerized randomization sequence.

The trial PI, site PIs/coordinators, trial statisticians and members of the Trial Steering Committee were blinded to the allocation status until the analysis of the six-month data for the parent trial, SHARE.

2.4 Sample size calculation

Much of the following is copied or adapted from the THPP+ protocol paper.[1]

The primary power calculations for the THPP+ study are for the c-RCT comparisons of perinatally depressed women and their children in the control vs. intervention arms at 36 months postnatally at the 5% two-tailed significance level. As with the THPP trial,[2] we assume 40 village clusters randomized in a 1:1 allocation ratio within 11 union councils 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 non-depressed women per village cluster for a total of 560 perinatally non-depressed 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 non-depressed women at 36 months will be 20%. Therefore, the total sample size available at 36 months is anticipated to be 480 perinatally depressed and 480 perinatally non-depressed women and their children. Using standard formulae[3,4] for a cluster randomized design and assuming intra-cluster 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 depression remission of 65% in perinatally depressed-Intervention vs. 45% in 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 total difficulties (TD) score (range 0-40) of 3 points for children of perinatally depressed mothers using plausible estimates for intra-cluster correlations of 0.04-0.08[5] and 5.2 for standard deviation for the TD score among 3 year olds.[6]

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 non-depressed mothers in the control arm, we will conclude equivalence if the 95% 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 [6,7]. 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 ICC of 0.04, and no difference between the groups, we will have 83% power to conclude equivalence.[3,8] For the secondary research question of the community benefit (i.e. carryover) of the intervention for perinatally non-depressed 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 non-depressed-Intervention vs. perinatally non-depressed-Control, Figure 1) for the same assumptions of the primary comparison above.

2.5 Intervention

The intervention for moderate-severe perinatal depression being assessed in these trials was the Thinking Healthy Program Peer-Delivered (THPP). THPP is the adapted version (peer-delivered) of the Thinking Healthy Program (THP) originally developed and evaluated (based on delivery by government-

employed LHWs) in Pakistan.[9] THP has recently been adopted by the World Health Organization (WHO) for global implementation (http://www.who.int/mental_health/maternal-child/thinking healthy/en/).

THPP comprises up to 14 sessions. The core strategies used by the peers are: active listening, collaboration with the family, guided discovery, homework, and behavioral (patient) activation (identifying and replacing unhealthy behaviors with healthy ones and practicing them).
The intervention begins in the third trimester of pregnancy, and includes ten home-based individual face-to-face sessions and four group-based sessions (given in the LHW women's group meetings) until the sixth month postnatally. Part of the task sharing by peers is to assist LHWs in organizing and conducting these mandatory routine women's groups and to encourage trial participants to attend these sessions. The following description comes from Atif et al. (2019).[10]

> The extended intervention began after the child was 6 months old and consisted of 18 group 'booster' sessions which were delivered monthly for 6 months, and subsequently every 2 months, until the child was 3 years old. These group sessions provided a safe environment to women to voice their problems, share experiences of childcare, and provide mutual support. The peer volunteers were trained to use culturally-grounded vignettes that served as tools to deliver health and well-being messages. Derived from our qualitative studies, these vignettes depicted a variety of real-life challenges and situations faced by rural women with young children. They aimed to help participants reflect and gain better insight about their own problems and allowed them to share personal experiences of overcoming similar problems. In addition to the narratives, card games containing pictorial illustrations of personal well-being messages were developed. These aimed to challenge unhelpful thinking patterns and behaviour and to replace them with alternative thoughts and behaviour in a fun and interesting way. The sessions aimed to enhance not just maternal well-being but also child-care and development by encouraging mother-infant interaction and play. The intervention provided examples of age-appropriate activities, derived from the WHO's Care for Development Package and encouraged demonstration of these activities during the sessions. While the 'booster' group sessions did not focus on specific strategies to address depression, the peer could still draw on her knowledge and skills of specific psychotherapeutic elements such as behavioural activation when required. A reference manual of the extended THP-P is available at http://hdrfoundation.org/publications (under "Training Materials").

2.6 Enhanced Usual Care (Control condition)

In Pakistan, there is non-existent usual care for perinatal depression. The majority of cases of maternal depression either remain unidentified or are classified as 'pregnancy blues'. In part, this is because gynecologists receive no mental health training and are overburdened, resulting in low detection and treatment rates. Therefore, usual care was enhanced in the following ways, for participants in both the intervention and control arms:

- Screening and informing participants about their screening results.

- Informing depressed participants about ways to seek appropriate health care
- Providing the primary health-care centers and gynecologists with the adapted WHO mhGAP treatment guidelines for maternal depression.
- Providing an 'information about health and stress during pregnancy' sheet to all pregnant women about how and where to seek health care from, including local Lady Health Workers (LHWs), primary health facilities and tertiary care facilities, both during pregnancy and beyond.

2.7 Time of outcome assessment and visit windows

For all analyses, time will be measured from the date of delivery. Outcome data (depending on the outcome) are collected at 3, 6, 12, 24, and 36 months postnatal. The protocol defined the visit window to be from four weeks (28 days) before to four weeks (28 days) after the scheduled visit date for all mother outcomes. No in-window vs. out-of-window adjustments will be taken into account.

2.8 Data management

Baseline and follow-up outcome data is captured electronically using tablet computers. In addition, process data from peers is also collected. Baseline and outcome data is remotely uploaded as a CSV (comma-separated values) file to the main data server, which run on the software program ODK. Both files are compliant with good clinical practices (including a date and time stamp of original data entry and with an audit trail to document any subsequent changes). Process data is to be collected in paper form, and this is to be manually entered and stored as CSV files using the same data collection platform. THPP quality-related data is also be collected in paper form.

All data, range and consistency checks are performed at weekly intervals separately for each data source. Any queries identified are resolved promptly by the data management team, and the database is updated, maintaining the audit trail. Consistency checks are performed using PostgreSQL, pgAdminIII (ie SQL based queries run on the server). This includes Trial I.D duplication/mismatch, merging two tables in case any changes are made in the questionnaire.

Additional data quality checks are performed by a biostatistician at Duke University.

3 Data Analysis Plan

The main trial paper for both moms and children will focus on the 36-month outcomes solely. Statistical analysis will be conducted according to the CONSORT guidelines.[11,12] A flow chart will show participation of both perinatally depressed (by intervention arm) and perinatally non-depressed

mothers and their children from recruitment in the third trimester through to 36 months postnatally (Figure 2 is an example, but in practice it will be more detailed). Withdrawals and loss to follow-up will be shown at 36-month post-intervention. All analyses will focus on comparison of three groups: perintally depressed intervention, perinatally depressed control, and perinatally non-depressed, although we will test for a differential effect by non-depressed in intervention vs. control as a supplemental analysis.

Baseline characteristics of recruited mothers will be reported by study arm for perinatally depressed women, and overall for perinatally non-depressed women. 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. Distributions of outcomes will also be explored using graphical techniques such as box plots and histograms.

3.1 Analysis of primary and secondary outcomes

The primary analyses are designed as intention-to-treat and will be conducted using Stata version 16. Maternal and child outcomes are listed in Table 1. ***For all outcomes, the 36-month time point is considered primary.*** The primary child outcome is the Total Difficulties score from the Strengths and Difficulties Questionnaire (SDQ-TD). The primary maternal outcome is the Patient Health Questionnaire (PHQ-9). The secondary outcomes are detailed in Table 1. Of note, the Bayley scores (secondary outcome) will be scaled using the provided U.S. normed scales.

Separate outcome analysis will be conducted for mothers and for children. In both cases, data from perinatally depressed (by intervention arm) and perinatally non-depressed participants will be analysed jointly using 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.

Since our main predictor of interest is the treatment arm indicator (with non-depressed as a separate "arm"), which is related to the sampling variable, the primary analyses will be unweighted, and we will use restricted maximum likelihood (REML). If the model fails to converge after 50 iterations, then we use maximum likelihood. In addition, the between-within method will be used to apply small-sample corrections in the mixed effects framework.[13] The reason for not using the sampling weights is that the sampling scheme is related to the main predictor of interest, so weighting should lead to minimal, if any, changes in the main results of interest. In addition, not using the sampling weights simplifies the modeling process and allows us to use estimation procedures (REML) shown to reduce bias, in addition to allowing us to use small-sample corrections on the models.

GEE analyses with a **log-link** and **Poisson distribution** will be used for **binary outcomes** in order to estimate prevalence ratios. The GEE analyses will apply the Kauermann-Carroll small-sample corrections with a t-test for inference.[14,15] The GEE analyses will take into account clustering using an exchangeable working correlation matrix, and robust standard errors to account for potential model misspecification. All mixed models will include village cluster as a random intercept to account for the clustered study design. (Note that a separate random intercept for each group [intervention, control, non-depressed] was considered and formulated in section 3.1.1.1, but model convergence issues led us to revert to a single random intercept for cluster.) Additionally, clustering by peer group within the intervention arm will be potentially taken into account by adding an additional random intercept for each peer group.

Primary analyses of outcomes measured at a single follow-up time point are designed to include the following fixed factor variables: group (non-depressed, depressed intervention, depressed control), strata (11 Union councils), and any variables found to be imbalanced by loss to follow-up or at baseline (determined using $p<0.10$). 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).[16] 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.[17]

Model assumptions will be assessed using examination of the residuals from the model and tests for linearity. In the presence of a ceiling or floor effect, we will also model the outcome using mixed effects tobit regression, which may allow us to more appropriately model the mean in such a situation.

As complementary information, we will also model the child outcomes using quantile regression, adjusted for clustering using the `qreg2` package in Stata, which implements the method of Parente and Santos Silva (2016)[18]. The same types of fixed effects as were included in the linear mixed effects model will be included in the quantile regression model. *A priori*, we will analyze the difference between the three groups at the 25th, 50th, and 75th percentiles.

All outcomes which are psychosocial scales will also be reported as standardized mean differences (SMDs) using the method of Hedges [19] (formula just below formula 18 in the paper). However, we will report none of the child outcomes as SMDs.

3.1.1 Analysis models and code

For a continuous variable measured and/or analyzed at a single time point, we have the following model formulation (assuming no clustering at the peer group level and assuming one random intercept for cluster), which can easily be extended to the tobit framework:

$$Y_{jk} = \alpha_1 + \beta_1 X_{1jk} + \beta_2 X_{2jk} + \beta_3 U_k + \beta_5 \boldsymbol{C_{jk}} + b_{0k} + \epsilon_{jk},$$
$$b_{0k} \sim N(0, \sigma^2_{b0k}),\ \epsilon_{jk} \sim N\left(0, \sigma^2_{\epsilon_{jk}}\right)$$

for the j^{th} mother-child dyad in the k^{th} village cluster.

The fixed effects terms are as follows:

- Y_{jk} is the continuous outcome of interest
- X_{1jk} is a dummy variable for perinatally depressed intervention vs. control (1 if intervention; 0 if control) for mother j in cluster k
- X_{2jk} is a dummy variable for perintally non-depressed vs. perinatally depressed control (1 if non-depressed; 0 if control) for mother j in cluster k
- U_k is the union council within which cluster k is located
- $\boldsymbol{C_{jk}}$ are other adjustors

The two random effects terms are as follows:

- b_{0k} is a random intercept for cluster k
- ϵ_{jk} is the residual measurement error for mother/child j in cluster k

This model can be fit with the following Stata code (with no additional adjustors):

```
mixed sdqTotal i.arm3 i.ucCode || clusterCode:, reml dfmethod(repeated)
```

Estimates of the pre-specified comparisons of interest will be derived from the fitted model.

3.2 Sensitivity analyses

Since we oversampled depressed women for the study, as a sensitivity/supplemental analysis, the linear mixed effects model analyses will be weighted using inverse probability sample weights, in which case maximum likelihood (ML) will be used, since REML is invalid in the presence of sampling weights.

Additionally, we will adjust the models for assessor (to examine the sensitivity of the model results to assessor).

3.3 Moderator analyses

Secondary analyses focus on potential moderators (effect modifiers) of any main associations with the primary outcomes. *A priori*, we hypothesize the following baseline variables may impact the degree to which the intervention affects depression symptoms, with their specific categorizations given in Table 2: socioeconomic status (education, asset index), maternal education (none vs. any) household composition (nuclear family status), child gender, the presence of intimate partner violence, treatment expectations (very/moderately useful vs. somewhat/not useful), maternal age (18-24 vs. 25+), depression severity (PHQ-9 = 10-14 vs. 15+), and depression chronicity (<12 weeks vs. ≥12 weeks). For the maternal outcomes, we will also look at parity (4+ kids vs. <4 kids) as a moderator. We will test for moderation of the effect by including these potential moderators in the model as an interaction between the potential moderator and the intervention indicator. Specifically, for our primary outcomes, we now have:

$$Y_{jk} = \alpha_1 + \beta_1 X_{jk} + \beta_2 M_{jk} + \beta_3 X_{jk} M_{jk} + \beta_4 U_k + \beta_5 \boldsymbol{C_{jk}} + b_{0k} + \epsilon_{jk},$$
$$b_{0k} \sim N\left(0, \sigma^2_{b0k}\right),\ \epsilon_{jk} \sim N\left(0, \sigma^2_{\epsilon_{jk}}\right)$$

where X_{jk} is the indicator of perinatally depressed intervention vs. control (1 if intervention; 0 if control) for mother i in cluster j and M_{jk} is the moderator variable for mother-child dyad j in cluster k. Other parameters are defined as in Section 3.1.1. β_3 tests the strength of the effect modification, and intervention effects by levels of the moderators can be estimated from this model.

4 References

1. Turner EL, Sikander S, Bangash O, et al. 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. *Trials.* 2016;17(1):442.
2. Sikander S, Lazarus A, Bangash O, 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.
3. Julious SA. Tutorial in biostatistics - Sample sizes for clinical trials with normal data. *Stat Med.* 2004;23(12):1921-1986.
4. Hayes R, Moulton L. *Cluster Randomized Trials.* CRC Press; 2009.
5. 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):902-909.
6. SDQ W. SDQ Normative Data http://www.sdqinfo.org/g0.html. Youth in Mind. http://www.sdqinfo.org/g0.html. Published 2012. Accessed Nov 28, 2012, 2012.

7. Mieloo C, Raat H, van Oort F, 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).
8. Hayes R, Moulton L. Sample Size. In: *Cluster Randomised Trials.* CRC Press; 2009.
9. 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. *The Lancet.* 2008;372(9642):902-909.
10. Atif N, Bibi A, Nisar A, et al. Delivering maternal mental health through peer volunteers: a 5-year report. *International Journal of Mental Health Systems.* 2019;13(1):62.
11. Campbell MK, Piaggio G, Elbourne DR, Altman DG. Consort 2010 statement: extension to cluster randomised trials. *BMJ.* 2012;345:e5661.
12. Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMC Med.* 2010;8(1):18.
13. Li P, Redden DT. Comparing denominator degrees of freedom approximations for the generalized linear mixed model in analyzing binary outcome in small sample cluster-randomized trials. *BMC Med Res Methodol.* 2015;15(1):38.
14. Kauermann G, Carroll RJ. A note on the efficiency of sandwich covariance matrix estimation. *J Am Stat Assoc.* 2001;96(456):1387-1396.
15. Li P, Redden DT. Small sample performance of bias-corrected sandwich estimators for cluster-randomized trials with binary outcomes. *Stat Med.* 2015;34(2):281-296.
16. Molenberghs G, Thijs H, Jansen I, et al. Analyzing incomplete longitudinal clinical trial data. *Biostatistics.* 2004;5(3):445-464.
17. Little RJA. Pattern-mixture models for multivariate incomplete data. *J Am Stat Assoc.* 1993;88(421):125.
18. Parente PM, Santos Silva J. Quantile regression with clustered data. *Journal of Econometric Methods.* 2016;5(1):1-15.
19. Hedges LV. Effect sizes in cluster-randomized designs. 2007.
20. 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):606-613.
21. Woerner W, Fleitlich-Bilyk B, Martinussen R, et al. The Strengths and Difficulties Questionnaire overseas: evaluations and applications of the SDQ beyond Europe. *Eur Child Adolesc Psychiatry.* 2004;13(2):ii47-ii54.
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):158-166.
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):421-427.
24. Üstün TB, Chatterji S, Kostanjsek N, et al. Developing the World Health Organization disability assessment schedule 2.0. *Bull World Health Organ.* 2010;88(11):815-823.
25. Bayley N, Reuner G. *Bayley scales of infant and toddler development: Bayley-III.* Vol 7: Harcourt Assessment, Psych. Corporation; 2006.
26. Spitzer RL, Williams JB, Gibbon M, First MB. The structured clinical interview for DSM-III-R (SCID): I: history, rationale, and description. *Arch Gen Psychiatry.* 1992;49(8):624-629.

5 Tables and figures

Table 1. Outcome Assessments

Instrument	Description	Outcome	Contextual validity
		Primary Outcomes	
PHQ-9	Nine-item questionnaire of depressive symptoms assessed on a scale of 0 to 3.	Prevalence of moderate–severe depression; mean total score	Validated in primary care [20]
PHQ-9 dichotomized at 10	Also called "remission".		
SDQ-TD	The SDQ is a parent report of 25 child attributes divided into five subscales: emotional symptoms, conduct problems, hyperactivity, peer problems, and pro-social behavior.	Total difficulties score: Calculated based on four subscales (except pro-social 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]
		Secondary Outcomes	
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 [24]
BSITD-III	An individually administered assessment of the child's achievement of developmental milestones across two areas: fine motor and receptive language. [25]	The total score from each domain, rescaled using the provided age-scaled score.	The standard scores are derived from the U.S. 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.
SCID	Structured clinical interview for depression. [26]	Prevalence of depression	
		Other Outcomes	
Recurrence	Only measured in prentally depressed. This measures in those who recovered by 6 months (based on SCID), how many had a recurrence of depression (based on SCID) by 36 months?		
SCID across time	Prevalence of depression from baseline through 36 months, all fit in one longitudinal model.		
Child Anthropometrics	Weight-for-age and height-for-age z-scores, based on WHO criteria		

Table 2. Moderators

Measure	Description
Socio-economic status (SES)	Asset-based SES index (dichotomized at the bottom 1/3 of the distribution) Maternal education (dichotomized at no education)
Household composition	Nuclear family status (nuclear vs. non-nuclear)
Intimate partner violence (IPV)	Any type of IPV in the past 12 months
Child Gender	
Treatment Expectations	Very/moderately useful vs. somewhat/not useful
Maternal age	18-24 vs. 25+
Depression severity	PHQ-9 = 10-14 vs. 15+
Depression chronicity	<12 weeks vs. ≥12 weeks
Parity*	<4 kids vs. 4+ kids

* Only considered a moderator for maternal outcomes

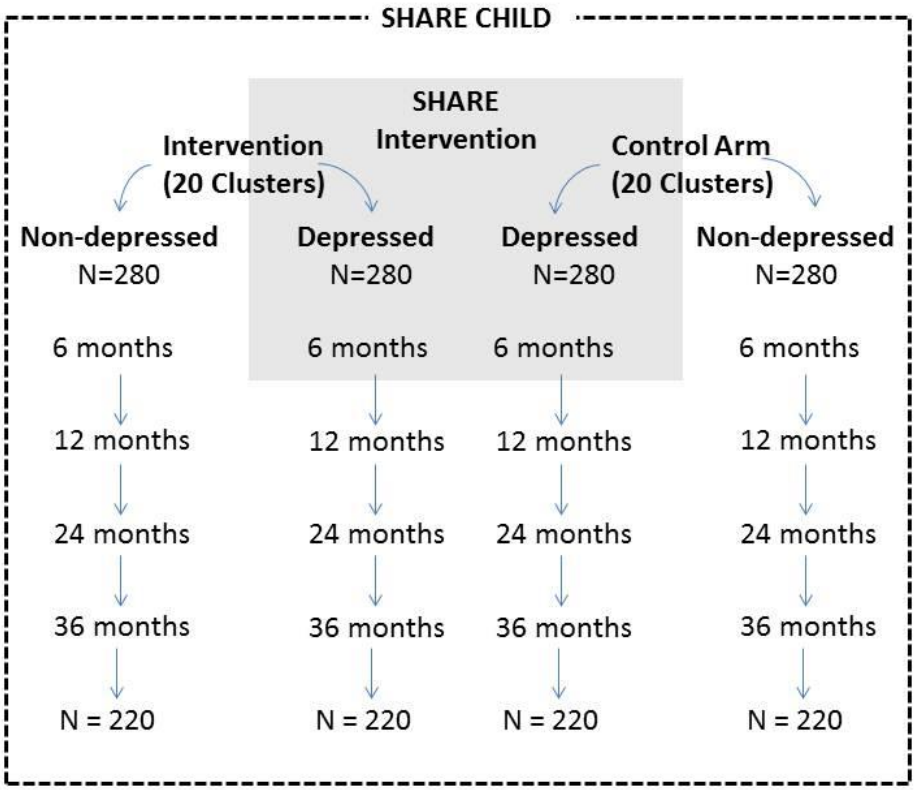
SHARE CHILD
SHARE Intervention
Intervention (20 Clusters)
Control Arm (20 Clusters)
Non-depressed N=280
Depressed N=280
Depressed N=280
Non-depressed N=280
6 months
12 months
24 months
36 months
N = 220

Figure 1.

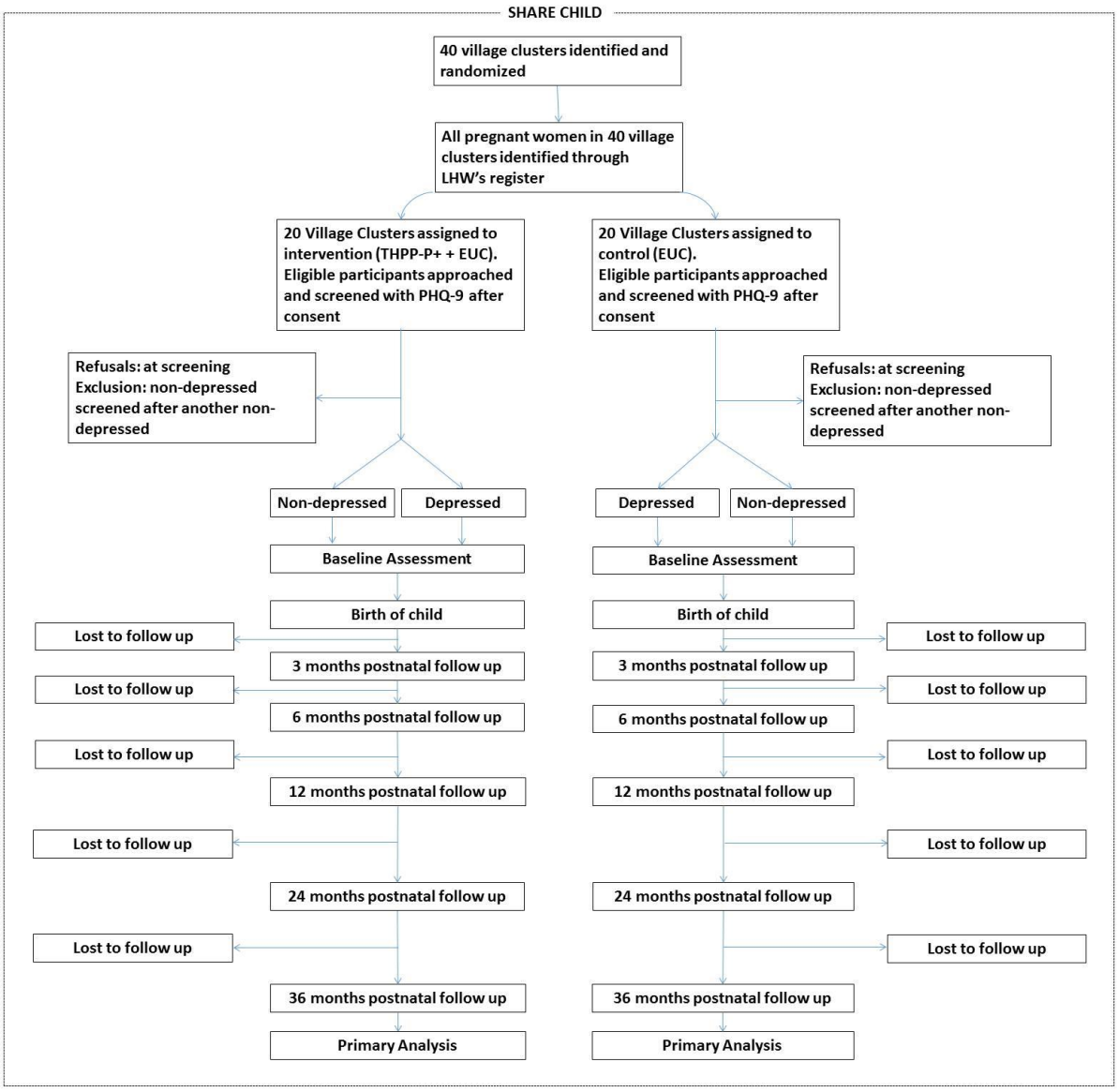
SHARE CHILD
40 village clusters identified and randomized
All pregnant women in 40 village clusters identified through LHW's register
20 Village Clusters assigned to intervention (THPP-P+ + EUC). Eligible participants approached and screened with PHQ-9 after consent
20 Village Clusters assigned to control (EUC). Eligible participants approached and screened with PHQ-9 after consent
Refusals: at screening
Exclusion: non-depressed screened after another non-depressed
Refusals: at screening
Exclusion: non-depressed screened after another non-depressed
Non-depressed
Depressed
Depressed
Non-depressed
Baseline Assessment
Baseline Assessment
Birth of child
Birth of child
Lost to follow up
3 months postnatal follow up
Lost to follow up
3 months postnatal follow up
Lost to follow up
Lost to follow up
6 months postnatal follow up
Lost to follow up
6 months postnatal follow up
Lost to follow up
Lost to follow up
12 months postnatal follow up
Lost to follow up
12 months postnatal follow up
Lost to follow up
Lost to follow up
24 months postnatal follow up
Lost to follow up
24 months postnatal follow up
Lost to follow up
36 months postnatal follow up
36 months postnatal follow up
Primary Analysis
Primary Analysis

Figure 2.