Last updated: October 28, 2025

Pre-Analysis Plan for
Community-based Mental Health Interventions for Elderly Women in Tamil Nadu, India

AEA Trial ID: AEARCTR-0015018
ClinicalTrials.gov ID: NCT05856552

Authors: Abhijit Banerjee (MIT), Esther Duflo (MIT), Camille Falézan (MIT), Madeline McKelway (Dartmouth), Miriam Sequiera (Sangath), Frank Schilbach (MIT), Girija Vaidyanathan (Tamil Nadu Indian Administrative Service – Retired), Jenny Wang (MIT)

Purpose: This document summarizes the main tests that we intend to conduct in the project "Community-based Mental Health Interventions for Elderly Women in Tamil Nadu, India". We have collected the wave 1 immediate endline and are collecting the wave 2 immediate endline and the wave 1 6-month follow-up. We may conduct additional analyses, but we will indicate in the paper which estimations were or were not specified in this pre-analysis plan. For the experimental design, please see the AEA registry and ClinicalTrials.gov registration.

A. Data sources

The data used in this analysis come from:

1. **Baseline survey (prior to intervention)**: Collected prior to randomization and intervention rollout. The survey captures demographics, socio-economic background, mental and physical health status, social participation, and prior exposure to community programs. For wave 2, the baseline survey also includes exploratory questions about the elder's caregiver and their relationship.
2. **Caregiver baseline survey (prior to intervention):** Collected prior to randomization and intervention rollout for wave 2 only with the elder's caregiver. The survey captures caregiver demographics and exploratory questions about caregiver well-being and relationship with the elder.
3. **Intervention fidelity and compliance to assignment (during intervention):** Individual elder attendance to group activities and CBT sessions is tracked during the intervention by the implementing government volunteers. Feedback for each GA session is documented through session logs maintained by the implementing government volunteers. The quality of CBT sessions is assessed by rating audio recording of sessions by the research team using a therapy quality scale. Additionally, the date and time of each GA and CBT session are systematically recorded.
4. **Immediate endline survey (immediately after intervention)**: Collected immediately after the 12-week intervention period. Measures all primary and secondary outcomes, including mental health, loneliness, functional ability, cognition, and physical mobility. The endline survey also includes exploratory questions about life satisfaction (wave 2 only) and the elder's caregiver and their relationship.
5. **Attendance at health camp (immediately after intervention):** Attendance at the health camp will be measured by surveyors during a one-day health camp after the end of the intervention, where subjects can access health measurements such as blood pressure and blood oxygen levels.
6. **Actigraph device data (immediately post-intervention, subset of ~1,000 participants)**: Used to measure objective indicators of sleep quality (total sleep time, sleep efficiency) and mobility (step counts) around the time of the immediate endline survey (collected from 1 month before to 2 months after the intervention).

7. **Caregiver immediate endline survey (immediately after intervention):** Collected immediately after the 12-week intervention period from the elder's caregiver. We interview the baseline caregiver, or both the baseline and endline caregivers if the caregiver changed. The survey captures caregiver demographics and exploratory questions about caregiver well-being and relationship with the elder.
8. **6-month follow-up survey (6 months after intervention)**: Collected 6 months after the end of the intervention period from the elder. Measures all primary and secondary outcomes, including mental health, loneliness, functional ability, cognition, and physical mobility. Additionally, the 6-month follow-up survey includes exploratory measures such as anxiety, life satisfaction (for wave 2 only), and the elder's caregiver and their relationship.
9. **6-month caregiver follow-up survey (6 months after intervention):** Collected 6 months after the end of the intervention period from the elder's caregiver. We interview the endline caregiver, or both the endline and 6-month follow-up caregivers if the caregiver changed. The survey captures caregiver demographics and exploratory questions about caregiver well-being and relationship with the elder.

B. Outcomes

We pre-specify two primary outcomes, six secondary outcomes, and eight exploratory outcomes.

Outcomes are measured at the immediate endline and 6-month follow-up unless indicated otherwise below.

All indices follow the variance-weighted summary index method proposed by Anderson (2008). All indices are constructed so that higher values indicate better outcomes.

B.1 Primary Outcomes

1. Depression Score :
    a. Measured using the Geriatric Depression Scale (GDS-15). Scores are constructed by summing affirmative responses to 15 yes/no questions (range 0–15).

       ⇒ Higher scores indicate more depressive symptoms.

2. Functional Impairment Score:
    a. Measured using the 12-item WHO Disability Assessment Schedule (WHO-DAS 2.0). Each item is rated on a 5-point scale (0=None to 4=Extreme difficulty) and summed (range 0–48).

       ⇒Higher scores indicate more functional impairment.

B.2 Secondary Outcomes

1. Loneliness / Social Connectedness Index:
    a. UCLA Loneliness Scale (ULS-3): sum of 3 items about feeling of lacking companionship, being left out, and isolation (1=Hardly ever, 2=Some of the time, 3=Often).
    b. Direct report of loneliness: whether felt lonely last week (1=Yes, 0=No).
    c. Perceived Social Support Scale: sum of 8 items from the Family and Friends subscales (1=Very strongly disagree to 7=Very strongly agree) about emotional support, decision-making help, and sharing joys/sorrows.

   d. Brief Sense of Community Scale: sum of 2 items from the Emotional Connection subscale assessing connection to the neighborhood and bond with neighbors (1=Strongly disagree to 7=Strongly agree).
   e. Number of close friendships.

   ⇒ Measures are reverse-coded when needed such that higher values indicate greater social connectedness.

2. Agency Index:
   a. Locus of Control from World Values Survey: the subject is asked the extent to which they feel that they have free choice and control over their lives on a scale (1=No choice at all to 10=A great deal of choice).
   b. Generalized Self-Efficacy Scale: sum of 2 items from the GSE that asks about whether the subject feels they can solve most problems with effort and whether they can find several solutions on a scale (0=No completely to 3=Yes completely).

   ⇒ Higher values indicate greater agency.

3. Objective Sleep Index:
   a. Actigraph-measured total sleep time
   b. Actigraph-measured sleep efficiency

   ⇒ These measurements and the index are available only at the immediate endline and for a subset of participants (~N=1,000). Higher values indicate better sleep.

4. Perceived Health Status Index:
   a. Self-rated health: self reported overall health on a 5-point scale (1=Very poor to 5=Very good).
   b. Self-assessed pain: self-assessed physical pain in the last week on a 11-point scale (0=No pain to 10 =Worst pain possible)

   ⇒ Measures are reverse-coded when needed such that higher values of the index indicate better perceived health and lower pain.

5. Physical Mobility Index:
   a. Indicator for leaving home the previous day (1=Yes, 0= No)
   b. Frequency of physical activities: sum of reported frequency of engaging in two activities: walking at a moderate pace and floor/stretching/gentle yoga exercises (1=Hardly ever/never to 5=Every day)
   c. Actigraph-measured step count: a per-day average step count based on Actigraph device data for a subset (~N=1,000). This measurement is only available for the immediate endline for a subset (~N=1,000).

   ⇒ When Actigraph data is available, all three measures are combined into an index where higher values indicate greater mobility. At baseline and at the 6-month follow-up, when Actigraph data is unavailable, the index is constructed using the first two components.

6. Cognition Score:

a. Adapted version of the Hindi Mental State Exam (HMSE) to Tamil (itself an adaptation from the Mini Mental State Exam (MMSE)): sum of 21 questions assessing orientation, memory, attention, language, and visuospatial skills (1= if correctly completed each question or sub-question, 0= otherwise, range 0–30).

⇒ Higher values indicate better cognition.

B.3 Exploratory Outcomes

B.3.i Elder's outcomes

1. Self-reported Sleep Index:
    a. Self-reported sleep hours based on sleep/wake-up times.
    b. Insomnia Severity Index: sum of 7 items assessing in the past two weeks difficulty falling/staying asleep, waking too early (1=None to 5=Very severe), sleep satisfaction (1=Very satisfied to 5=Very dissatisfied), interference with daily function (1=Not at all to 5=Very much), impact on quality of life (1=Not noticeable to 5=Very noticeable), and worry about sleep (1=Not at all to 5=Very much)
    c. Pittsburgh Sleep Quality Index: 1 item on sleep quality in past month (1=Very good to 4=Very bad)

    ⇒ Measures are reverse-coded when needed such that higher values indicate better sleep.

2. Demand for Intervention Activities**:**
    a. Preference for CBT or Group Activities over a cash transfer: an indicator equal to 1 if the subject prefers either intervention (CBT or Group Activities) over a cash transfer in a hypothetical choice, and 0 if they prefer the cash transfer. This comes from a question asking subjects to rank their preference between a Group Activities program, CBT program, or a one-time 1,000 Rs cash transfer.

3. Health Management Behavior Index:
    a. Self-reported use of preventive check-up services for diabetes and blood pressure in the past 12 months (1=Yes, 0=No)
    b. Self-reported interest in attending a one-day health camp hosted by the research team after the end of the intervention, where subjects can access health measurements such as blood pressure and blood oxygen levels. This outcome is only available at the immediate endline.
    c. Actual attendance at the health camp. This outcome is only available at the immediate endline.

    ⇒ At the immediate endline, all three measures are combined into an index. Higher values indicate greater health management behavior. At the 6-month follow-up, only the first measure is available.

4. Anxiety Score
    a. Measured using the Generalized Anxiety Disorder 7-item (GAD-7). Scores are constructed by summing answers to 7 questions using a 4-point scale (0=Not at all to 3=Nearly every day, range 0–21).

    ⇒ Higher scores indicate more symptoms of anxiety. This measurement is only available at

the 6-month follow-up.

5. Life Satisfaction Score
    a. Measured using the OECD's 11-point life satisfaction scale (0=Not at all satisfied to 10=Completely satisfied).

    ⇒ Higher scores indicate greater life satisfaction. The outcome is only available for wave 2.

6. Elder's Caregiver-Elder Relationship Quality Index
    a. Subjective closeness: 1 item from the Subjective Closeness Index (SCI) asking "How close do you feel to [caregiver's name]?" (1=Not at all close to 7=Very close).
    b. Relationship strength: 4 items from the Relationship Closeness Inventory (RCI) strength subscale (1=Strongly disagree to 7=Strongly agree) capturing perceived importance, influence, and mutual understanding in the relationship.
    c. Communication frequency: 1 item on how often the caregiver and elder talk about personal concerns (1=Multiple times a day to 7=Not in the last month).
    d. **Perceived burden:** 3 items on whether the elder worries about being a burden, hides pain or sickness to avoid worrying the caregiver, or avoids treatment to not burden them (1=Not at all to 5 = All the time).

    ⇒ Measures are reverse-coded when needed such that higher values indicate a closer and more comfortable relationship between the caregiver and the elder.

B.3.ii Caregiver's outcomes

7. Caregiver's Caregiver-Elder Relationship Quality Index
    a. Subjective closeness: 1 item from the Subjective Closeness Index (SCI) asking "How close do you feel to [elder's name]?" (1=Not at all close to 7=Very close).
    b. Relationship strength: 4 items from the Relationship Closeness Inventory (RCI) strength subscale (1=Strongly disagree to 7=Strongly agree) capturing perceived importance, influence, and mutual understanding in the relationship.
    c. Communication frequency: 1 item on how often the caregiver and elder talk about personal concerns (1=Multiple times a day to 7=Not in the last month).

    ⇒ Measures are reverse-coded when needed such that higher values indicate a closer and more comfortable relationship between the caregiver and the elder.

8. Caregiver's Well-being Index
    a. Depressive symptoms: total score on the 9-item Patient Health Questionnaire (PHQ-9), summing responses (0="not at all" to 3="nearly every day", range 0–27).
    b. Life satisfaction: self-reported satisfaction (0=Not at all satisfied to 10=Completely satisfied).
    c. Caregiving burden: sum of 6 Zarit Burden Scale items on strain and role interference (1=Never to 6=Always).
    d. Role/work interference: sum of roles for which caregiving interferes with responsibilities and responses to counterfactual work question ("If you did not have to care for [elder], would you work a lot less, a little less, the same, a little more, or a lot more than you do now?"). The counterfactual is coded 0= "the same," 1= "a little less" or "a little more," 2= "a lot less" or "a lot more," with higher values reflecting greater work interference.

⇒ Measures are reverse-coded when needed such that higher values indicate better well-being.

C. Empirical Strategy

1. **Pooled regression: GA effects.** To estimate the effects of GA, we run:

$$Y_{iv}^{(t)} = \beta_0 + \beta_1 CBT_i + \beta_2 GA_v + \beta_3 Y_{iv}^{(0)} + \delta_s + \theta_{e^{(t)}} + \epsilon_{iv}^{(t)} \quad (1)$$

where $Y_{iv}^{(t)}$ denotes the outcome for individual i in village v at endline wave t ($t \in \{1, 2\}$), and $Y_{iv}^{(0)}$ is the corresponding baseline value of the outcome. CBT_i is an indicator equal to one if individual i was assigned to receive cognitive behavioral therapy, and GA_v is an indicator equal to one if village v was assigned to the group activities arm. $\theta_{e^{(t)}}$ represents enumerator fixed effects for survey wave t. δ_s are strata (parts) fixed effects and we cluster standard errors at the village level to account for village-level randomization of the group activities.

2. **Pooled regression: CBT effects.** To estimate the effects of CBT, we run:

$$Y_{iv}^{(t)} = \gamma_0 + \gamma_1 CBT_i + \gamma_2 Y_{iv}^{(0)} + \delta_s + \theta_{e^{(t)}} + \epsilon_{iv}^{(t)} \quad (2)$$

where δ_s are strata (village) fixed effects and standard errors are not clustered.

3. **Separate regressions.** To estimate the effects of CBT, GA, and CBT + GA, we run:

$$Y_{iv}^{(t)} = \alpha_0 + \alpha_1 CBTonly_i + \alpha_2 GAonly_{iv} + \alpha_3(\text{CBT}_i \times \text{GA}_v) + \alpha_4 Y_{iv}^{(0)} + \delta_s + \theta_{e^{(t)}} + \nu_{iv}^{(t)} \quad (3)$$

where $CBTonly_i$ is an indicator equal to one if individual i was assigned to receive cognitive behavioral therapy and no group activities, and $GAonly_{iv}$ is an indicator equal to one if individual i was assigned ot receive group activities but no CBT. The interaction term $\text{CBT}_i \times \text{GA}_v$ is equal to one for individual who were assigned to both interventions. δ_s are strata (parts) fixed effects. Standard errors are clustered at the village level.

As group activities (GA) were changed from a 12-week program to a continuous intervention in Wave 2, we will also present the 6-month follow-up results separately for Wave 1 and Wave 2.

Additionally, to address multiple hypothesis testing, we report effects on three broad outcome domains that group our main outcomes into conceptually related families. The eight primary and secondary outcomes are grouped into three broader domains—psychological well-being, physical health, and cognition—as described below.

1. Psychological Well-being Index, combining:
    a. Depression score (Geriatric Depression Scale, GDS-15)
    b. Functional Impairment Score (WHO-DAS 2.0)
    c. Loneliness / Social Connectedness Index
    d. Agency Index
    ⇒ The first two components will be reverse-coded so that higher values of the index indicate better well-being.

2. Physical Health Index, combining:

   a. Objective Sleep Index (if available)
   b. Perceived Health Status Index
   c. Physical Mobility Index

3. Cognition Score [not an index]

References

Anderson, Michael L, "Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects," *Journal of the American Statistical Association, 2008, 103* (484), 1481–1495.