# Pre-analysis plan: Effect of a behaviour change and hardware intervention on safe child feces management practices in rural Odisha, India: a cluster-randomized controlled trial

\*N.B. The template for this pre-analysis plan (PAP) was adapted from a version developed by 3ie for their TW14 grantees. We chose to use that PAP as a template since this impact evaluation also assesses a sanitation behaviour as the primary outcome. In addition, much of the information provided in sections 1 to 4 is adapted or comes directly from our protocol paper (Sclar et al., under review) for consistency. Lastly, this pre-analysis plan was developed after data collection had finished but before starting analysis. However, we did examine attrition rate to determine how many participants surveyed at endline had also been surveyed at baseline (this is discussed further in section 5.2.3).

## Intervention

## Theoretical framework

**Please describe the underlying behavioural theory, which will be used to guide your strategy for eliciting behaviour change through your intervention.**

The aim of this intervention is to improve safe child feces management (CFM) practices. However, the intervention specifically focuses on increasing two practices: **caregivers safely dispose** of their child’s feces into a latrine and **children use the latrine** as soon as they are able to.

We applied the **Risks, Attitudes, Norms, Ability, and Self-regulation (RANAS)** behavioural theory in order to understand these two behaviors (Mosler, 2012). The RANAS theory is also part of a larger approach for designing interventions for systematic behaviour change, which includes four phases. We followed these four phases in order to design the CFM intervention. We also applied **user-centered design (UCD)** to co-develop novel hardware with caregivers that would help aid them in adopting the behaviors. The resultant hardware distributed in the intervention includes a wash basin and bucket with lid for safely washing and storing soiled cloths/nappies and a latrine training mat with removable tray for aiding safe disposal and latrine training.

## Intervention summary

**Please summarise your intervention.**

\*To be consistent, here we present the intervention summary from our protocol paper (Sclar et al., under review).

The resultant CFM intervention includes six BC strategies (i.e. program activities), which consist of small group meetings and household visits that employ specific BCTs. The target participant for the strategies are primary caregivers, typically mothers, of children <5 years. However, in some of the strategies other household members, such as fathers and grandmothers, are also engaged to foster a supportive household environment towards the new CFM behaviors. A team of mobilizers with the NGO Gram Vikas will implement the activities across the 37 trial intervention villages. Due to the nature of the intervention, participant blinding will not be possible. There is no plan to implement the intervention in control villages at the end of the trial. However, Gram Vikas may choose to do so, or implement a modified version of the intervention based on study findings.

The intervention starts with an opening meeting that uses risk, attitude, and norm BCTs and also includes distribution of CFM hardware as an ability BCT. The intervention then shifts to alternating household visits and caregiver group meetings, which primarily use ability and self-regulation BCTs. The final strategy is a closing celebratory meeting that uses norm BCTs. Throughout the intervention, caregivers receive behavioral messaging on how to safely manage their child’s feces based on their child’s current developmental stage. Caregivers will also receive messaging on what to do as their child grows, such as transitioning from safe disposal to latrine training. Each BC strategy is further described below:

* **Hardware and Action Knowledge Opening Meeting:** The meeting starts with a discussion on typical CFM practices and why they are unsafe, followed by a video that tells the story of two mothers; one mother safely manages her child’s feces and another does not, illustrating messages related to health risks, costs and benefits, and the needs at different child development stages. The Gram Vikas mobilizer then uses a banner with illustrations to explain how to use the CFM hardware to safely dispose of children’s feces or teach them how to use the latrine. Volunteers are called upon to demonstrate the new information and then certain hardware is distributed to each caregiver depending on her child’s age (wash basin and bucket with lid for 0 to <7 months old; latrine training mat with tray for 7 to <48 months old). The meeting closes with a group commitment to use the new hardware and practice safe disposal and/or child latrine training.
* **Building Self-Efficacy and Goal Setting Household Visits:** The Gram Vikas mobilizer then visits each caregiver at her home and consults with them on their new practice, tailored to safe disposal and/or latrine training. During the visit the caregiver demonstrates her current practice, discusses any challenges she is facing and creates a barrier plan or is given tips, and creates a ‘goal tracker’ to monitor her progress in reaching the behavior change goal. The visit ends with the Gram Vikas mobilizer inviting other household members to express their approval of safe disposal/latrine training and to explain how they will support the caregiver. The second household visit is similar to the first but involves checking on the ‘goal tracker’ and having the caregiver positively self-reflect on her change.
* **Caregiver Support Group Meeting:** Facilitated groupmeeting is held in-between the household visits to allow caregivers to reflect on their progress, re-commit to their goal of practicing safe disposal/latrine training, and provide strategies to fellow caregivers on how to address common challenges and offer words of encouragement to each other.
* **Celebrating ‘Safe CFM Families’ Closing Meeting:** The final activity is a celebratory meeting that invites caregivers, their family members, and important village stakeholders (i.e. Anganwadi worker, VWSC members) to come together and share their experiences with adopting the safe CFM practices and its importance. The village stakeholders then give each caregiver a certificate to acknowledge her and her household’s achievement.

## Primary Outcome, Evaluation Questions, and Hypotheses

* 1. **What is the primary outcome?**

The primary outcome is a **binary measure of safe disposal of child feces***,* as defined by the WHO/UNICEF Joint Monitoring Programme for Water Supply and Sanitation (JMP). The updated JMP definition describes safe disposal as encompassing two distinct behaviors based on the last time the child defecated: (1) caregiver put/rinsed the child’s feces into the toilet/latrine or (2) child used the toilet/latrine (JMP, 2018).

In the baseline and endline trial surveys, the primary caregiver of the child <5 years old is asked “The last time [NAME OF CHILD] defecated, where did [NAME OF CHILD] defecate?” If the caregiver reports anything except “in latrine” then the following question is asked “Where was [NAME OF CHILD]’s feces disposed?” If the child defecated in the latrine or the caregiver disposed of the child’s feces into the latrine then the household is assigned a “1” for safe disposal, otherwise the household is assigned a “0” for unsafe disposal.

With regard to the child age of <5 years old, for baseline this means the child is <5 years old at the time of the survey while at endline this means the child was <5 years old at the start of the intervention (which is determined based on intervention implementation records and the child’s date of birth as provided by the caregiver).

* + 1. **What are the secondary outcomes?**

In secondary analyses, we will separately examine the two behaviours that comprise “safe disposal.” We will examine these behaviors —child latrine use and caregiver disposal in latrine —using the following metrics:

**Child latrine use** is a binary measure where a “1” is assigned if the child defecated in the latrine the last time they defecated and a “0” is assigned if the child defecated anywhere else.

**Caregiver disposal in latrine** is a binary measure where a “1” is assigned if the caregiver safely disposed of the child’s feces into the latrine the last time the child defecated and a “0” is assigned if the caregiver disposed of the child’s feces anywhere else. (\*This outcome is examined only among those caregivers whose child did **not** use the latrine the last time they defecated)

Other secondary analyses will also examine:

* **behavioral factors related to child latrine training and caregiver disposal in latrine**
* **different CFM practices along the “child feces exposure pathway”**
* **latrine use among household members >5 years old**
* **visibility of feces in the household compound**

These are described in more detail in section 5.1.1.

* 1. **What are the main evaluation question(s) the study seeks to answer?**

Main research question (RQ1): What is the effect of the intervention on **safe disposal of child feces** when comparing children among the intervention versus control arms?

Secondary research question (RQ2):

* RQ1a: What is the effect of the intervention on **caregiver disposal in latrine** when comparing children among the intervention versus control arms? (sub-analysis: stratified examinations by child sex and child age)
* RQ1b: What is the effect of the intervention on **child latrine use** when comparing children among the intervention versus control arms? (sub-analysis: stratified examinations by child sex and child age)

Secondary research question (RQ3):

* RQ3a. Are there behavioral factors that were changed by the intervention and, if so, which of these factors in turn influenced **caregiver disposal in latrine**? In other words, which behavioral factors mediate the effect of the intervention on **caregiver disposal in latrine** (taking into account factors and behavior at baseline)?
* RQ3b: Are there behavioral factors that were changed by the intervention and, if so, which of these factors in turn influenced **child latrine use**? In other words, which behavioral factors mediate the effect of the intervention on **child latrine use** (taking into account factors and behavior at baseline)?

Secondary research question (RQ4):

* RQ4a: Does the prevalence of **other CFM practices** along the “child feces exposure pathway” (such as where wash water is disposed, if and how hardware is used, whether caregiver washes her hands after, etc.) differ between households that received the intervention compared to households that did not?
* RQ4b: Does the prevalence of latrine use among **household members ≥ 6 years old differ** between households that received the intervention compared to households that did not?
* RQ4c: Does the prevalence of **visible feces** in the household compound differ between households that received the intervention compared to households that did not?

Sensitivity analysis (RQ5):

What is the effect of the intervention on **safe disposal of child feces** when comparing intervention versus control among caregivers who received ***both*** baseline and endline surveys (i.e. 2 surveys)? What is the effect of the intervention among caregivers who received ***only*** the endline survey (i.e. 1 survey)? How do these effects compare?

* RQ5a: What is the effect of the intervention on **caregiver disposal in latrine** when comparing intervention versus control among caregivers who received ***both*** baseline and endline surveys (i.e. 2 surveys)? What is the effect of the intervention among caregivers who received ***only*** the endline survey (i.e. 1 survey)? How do these effects compare?
* RQ5b: What is the effect of the intervention on **child latrine use** when comparing intervention versus control among caregivers who received ***both*** baseline and endline surveys (i.e. 2 surveys)? What is the effect of the intervention among caregivers who received ***only*** the endline survey (i.e. 1 survey)? How do these effects compare?
  1. **What are the hypotheses to be tested throughout the causal chain?**

Main research question (RQ1):

H1: The prevalence of safe disposal among children that reside in intervention villages will be significantly higher compared to children that reside in control villages.

Secondary research question (RQ2):

H2a: The prevalence of caregiver disposal in latrine among children that reside in intervention villages will be significantly higher compared to children that reside in control villages.

H2b: The prevalence of child latrine use among children that reside in intervention villages will be significantly higher compared to children that reside in control villages.

Secondary research question (RQ3):

H3a: The intervention will influence attitudes, norms, ability and self-regulation factors related to caregiver disposal in latrine (/child latrine training) and these factors will in turn be significantly associated with an increase in caregiver disposal in latrine (/child latrine use).

H3a: The intervention will influence attitudes, norms, ability and self-regulation factors related to caregiver disposal in latrine (/child latrine training) and these factors will in turn be significantly associated with an increase in caregiver disposal in latrine (/child latrine use).

Secondary research question (RQ4):

H4a: Intervention households will have a higher prevalence of safer CFM practices along the “child feces exposure pathway” compared to control households.

H4b: Intervention households will report higher prevalence of latrine use among household members ≥ 6 years old compared to control households.

H4c: Intervention households will have a lower prevalence of visible feces observed in the household compound compared to control households.

Sensitivity analysis (RQ5):

H5: The effect of the intervention on safe disposal among caregivers who received both baseline and endline (i.e. 2 surveys) will not be significantly different from the effect of the intervention among caregivers who received only endline (i.e. 1 survey).

H5a: The effect of the intervention on caregiver disposal in latrine among caregivers who received both baseline and endline (i.e. 2 surveys) will not be significantly different from the effect of the intervention among caregivers who received only endline (i.e. 1 survey).

H5b: The effect of the intervention on child latrine use among caregivers who received both baseline and endline (i.e. 2 surveys) will not be significantly different from the effect of the intervention among caregivers who received only endline (i.e. 1 survey).

## Sampling

* 1. **What is the eligible population for the study? Please list inclusion and exclusion criteria for the eligible population.**

**Village eligibility:**

Villages were eligible for the trial if they were located in Ganjam and Gajapati districts and had previously completed a community-based water and sanitation program called “MANTRA,” led by the implementing NGO Gram Vikas. Additional eligibility inclusion criteria for study villages included: village size was between 35 to 250 households, 75% of households had access to a latrine, the community water tank was functional, village had its own *Anganwadi* center (government-run daycare and preschool center), and there was no programming planned in the village during the study period by the implementing NGO Gram Vikas. The rationale behind these criteria are explained in the protocol paper (Sclar et al., under review).

**Household eligibility:**

All households with a latrine and at least one child <5 years old residing in the home were eligible for inclusion. For the baseline trial survey, the child had to be <5 years old at the time of the survey. For the endline trial survey, the child had to be <5 years old at the time of intervention delivery.

**Target respondent:**

The target respondent from the household was the primary caregiver of the child <5 years old. If there were more than one primary caregiver of children <5 years old in the household, then the primary caregiver of the *youngest* child was surveyed. If the primary caregiver was not available at the time of the survey, then a secondary caregiver of surveyed. All respondents had to verbally consent to participate and be at least 18 years old.

* 1. **What is the expected sample size?**

We planned to engage 74 trial villages based on the sample size calculation, with 37 assigned to intervention and 37 assigned to control. We estimated an average of 9 eligible households per village (i.e. households with a latrine and at least one child <5 years) based on data from a previous study conducted in the same area (Reese et al., 2019). As such, the expected sample size was around 650 households.

* 1. **Is there any reason to believe that the sample differs from the population? If so, how does it differ?**

It is reasonable to believe that the study sample does not reflect the larger population in Ganjam and Gajapati districts due to the village-level eligibility criteria. All study villages had to have high levels of sanitation and water access and to have previously participated in the MANTRA water and sanitation program. As such, the majority of households in these villages had access to quality water and sanitation infrastructure, which typically included a twin-pit pour flush latrine with attaching bathing room and piped water supply to the home (often to the kitchen, latrine, and bathing room). This ensured eligible households already had the enabling environment they needed to perform safe CFM practices.

Accordingly, the results from this study will only be generalizable to villages in rural Odisha with similar strong WASH infrastructure.

* 1. **Please describe the anticipated subgroups, if relevant, which will be studied.**

*Note: Since behaviour change interventions require village-level clustering to prevent spillovers, studies will likely not be adequately powered to conduct sub-group analysis, and subgroup analysis is not expected. Proposals to do subgroup analysis should be accompanied by an explanation of how studies will be able to detect differences between subgroups.*

Based on our research questions outlined in section 2.2, we will examine a few different subgroups:

1. RQ2a = Caregiver disposal in latrine is an outcome that will be examined only *among* caregivers whose child did not use the latrine the last they defecated.
2. RQ3 = Change in behavioral factors (whether for child latrine training or caregiver disposal in latrine) will be examined only among caregivers who were surveyed at baseline *and* endline (see section 5.2.3 on attrition).
3. RQ5 = The sensitivity analysis will examine the target behaviors (safe disposal, caregiver disposal in latrine, and child latrine use) among two different sub-groups: caregivers that were only surveyed at endline (i.e. surveyed once) and caregivers that were surveyed at both baseline and endline (i.e. surveyed twice). The purpose of this sensitivity analysis is to determine if multiple surveying led to response bias, where there would be a higher prevalence of the desired behaviour reported by those surveyed twice compared to those only surveyed once.
   1. **Statistical power**
      1. **What is the effect size that you will be able to detect?**

We are powered to detect an absolute increase of 16% in safe disposal prevalence among the intervention group.

For the sample size calculation, we used a version of formula five from Rutterford et al. (2015) for a standard parallel-group, two-arm study design with a binary outcome and cluster-level analysis. Formula parameters came from Gram Vikas provided data and previous sanitation studies conducted in Odisha (Bauza et al., 2019; Clasen et al., 2012).

The parameters used in our sample size calculation were as follows:

1. Average number of eligible households in trial villages = 9
2. Baseline prevalence of safe disposal = 0.407
3. Village­level ICC = 0.103
4. Alpha = 0.05
5. Beta = 0.80
6. Effect size = Absolute increase of 15% in safe disposal prevalence among intervention group
7. Attrition = 10% loss to follow-up between baseline and follow-up.

However, we realized during baseline data collection that the cluster sizes greatly varied. We accounted for this variation in cluster size by repeatedly resampling 37 villages per arm from the selected 74 villages. The median effect size that can be detected with 80% power is **16%** across simulations.

* + - 1. **What are your assumptions for the sample size calculation?**

We assumed the following:

* We would be able to survey on average at least 9 eligible households in each study village
* Baseline prevalence of safe disposal is 40.7%
* Village-level intraclass correlation is 0.103
  + - 1. **How many clusters will you have?**

We have 74 clusters (i.e. villages); 37 received the intervention, 37 acted as controls.

* + - 1. **How many people will you have in each cluster?**

We anticipated an average of 9 eligible households in each cluster, however this cluster size varies by village. Some of the villages have a small population and thus only a few eligible households while other villages have a much larger population and thus many eligible households.

* + 1. **If you plan to include covariates in your analysis, what share of variance do you expect to predict with your co-variates?**

*Note: It is not required that you include covariates*

We plan to include two covariates in the adjusted models for the main research question (RQ1): baseline prevalence of safe disposal and the stratification variable used in randomization. We also plan to include any variables found to be severely imbalanced at baseline. However, we did not calculate our sample size based on assumptions about shared variance from covariates and as such, we are using a more conservative approach.

* 1. **Assignment to treatment**
     1. **How will individuals be assigned to treatment and control conditions?**

We randomly selected 74 villages from a sample frame of 501 MANTRA villages in Ganjam and Gajapati districts. Following baseline data collection, we used stratified randomization to allocate villages into intervention or control arm on a 1:1 basis. A computer-based random number generator was used to complete the allocation. Villages were categorized into one of five distinct geographic-demographic (‘geo-demo’) groups, which served as the stratification variable to ensure balance on a wide range of factors. Villages in each ‘geo-demo’ group have a distinct set of characteristics with regard to their geography (coastal, hilly, tableland, etc.), predominant religion and caste, village size, market access and other characteristics. The ‘geo-demo’ groups mostly align with the ‘block’ level administrative unit (a district subdivision), with several blocks falling under each group. As such, three of the ‘geo-demo’ groups are located entirely in Ganjam district while the remaining two are in Gajapati district. We used replacement re-randomization to ensure balance on two *a priori* specified variables: number of eligible households per village and village-level baseline reported safe disposal of child feces. We used the Mann-Whitney U Test to test if the distribution of villages on these two variables was not significantly different between the study arms. We also used an independent samples t-test to test if the mean baseline reported safe disposal of child feces was not significantly different between the study arms. The *a priori* balance criteria were met with the first randomization attempt. Participants and study investigators were not blinded to village treatment assignment.

* + 1. **How will you check that individuals in the treatment condition received treatment as anticipated?**

In the endline trial survey, we included questions for intervention households that asked about their exposure to the intervention. The intervention included provision of hardware and five activities. Accordingly, the survey includes a question about what hardware the household received, if any, and a question for each activity that asks whether or not the respondent or someone in their household attended that activity.

We also conducted a mixed methods process evaluation to examine the fidelity of intervention delivery, as well as participant attendance. However, this data will be used for process evaluation analyses rather than the impact analysis.

## Data Collection

* 1. **Primary data collection instruments**
     1. **What data collection instruments will you employ for analysis?**
     2. **What (groups of) indicators will each instrument cover?**
     3. **How will each instrument be developed?**

The data collection instrument we used was a **structured survey**. This survey was administered to eligible households before (“baseline”) and after (“endline”) delivery of the intervention. The baseline survey took place between December 2019 – February 2020, the intervention was delivered between December 2020 – April 2021, and the endline survey was conducted between July – September 2021.

The baseline and endline surveys were very similar. Both surveys had almost all the same sub-questionnaires. The endline survey included an additional sub-questionnaire on intervention awareness and participation (one version of this sub-questionnaire was administered to control households to assess potential “spillover” of the intervention and another version was administered to intervention households). In addition, some of the questions throughout the endline survey were revised based on findings and lessons learned from the baseline survey.

The relevant sub-questionnaires from the baseline and endline trial surveys are presented below in Table 1. The table includes descriptions of the different indicators/questions and information on how each sub-questionnaire was developed.

**Table 1.** Overview of baseline/endline trial survey

|  |  |  |  |
| --- | --- | --- | --- |
| **Sub-questionnaire** | **Description of indicators/questions** | **Analysis purpose** | **How was sub-questionnaire developed?** |
| **Eligibility & Consent** | To verify household’s eligibility and obtain respondent’s consent to the survey | -Record # consents  -Record # HHs that dropped out due to eligibility issue | NA |
| **Household and Respondent Demographics** | To capture household demographics, such as SES, caste, and religion, and demographics about the respondent, such as education level and age. | -Relevant covariates to check balance at baseline (if imbalance then potential covariates used to adjust models) | Most questions based on versions used in past sanitation trials conducted in Odisha, India. Questions are well-tested. |
| **CFM Behaviours** | To record current CFM practices of the child and caregiver from defecation site to anal cleansing practices. | -Measure target CFM behaviors  (i.e. Safe disposal, caregiver disposal in latrine, child latrine use)  -Measure all CFM behaviors along “exposure pathway” which were also targeted in program | Questions used to measure target CFM behaviors (primary and secondary outcomes) taken from JMP CORE questionnaire (JMP, 2018). |
| **RANAS Behavioural Factors** | To capture caregiver’s perspective on different behavioral factors likely influencing the adoption of safe CFM practices | -Analyze change in RANAS factors of interest (targeted by program as the intermediate outcomes) | Questions informed by formative qualitative research and the RANAS Practical Guide (Mosler & Contzen, 2016). Revised based on cognitive interviews. |
| **Social Support** | To capture caregiver’s perceived social support around childcare tasks, esp. CFM, and the size/diversity of current social support network | -Analyze change in perceived social support and size/diversity of social support network | Questions adapted from several validated metrics (Barrera et al., 1981; Schwarzer & Schulz, 2013; Vaux et al., 1987; Zimet et al., 1988) and also informed by qualitative formative research. Revised based on cognitive interviews. |
| **Intervention: CFM Program activities & hardware** | To understand if there was spillover of the program activities in control villages; to capture level of exposure to the program among intervention households and condition and use of hardware | -Assess spillover (control HHs)  -Measure level of intervention exposure (intervention HHs)  -Analyze durability and use of hardware provided in program | These were original questions drafted by our research team since it is specific to the CFM intervention. |
| **Household WASH Facilities**  **(self-reported)** | To record availability and condition of the household’s WASH facilities (water, latrine, bathroom) in order to understand the *behavioral environment* in which safe CFM practices would be performed. | -Relevant covariates to check balance at baseline (if imbalance then potential covariates used to adjust models)  -Identification of households that have problematic *behavioral environment* for CFM | Most questions were based on versions used in past sanitation trials conducted by our research team in Odisha, India |
| **WASH Facilities Spot Check**  (observed by surveyor) |

* + 1. **Please comment on the validity and reliability of each instrument, including any anticipated validation checks.**

We ensured validity (accuracy) and reliability (consistency) of the baseline/endline trial surveys through a number of steps.

Both surveys were extensively reviewed by the survey team to ensure questions and response options were understood in a consistent manner by all. The survey team also provided feedback on the Odia translation as needed. The surveys were then pilot-tested in several non-trial villages prior to data collection to check how well participants understood the questions and again check on accuracy of Odia translation. Revisions were made as needed based on the survey team and pilot-testing feedback.

The RANAS Behavioural Factors and Social Support sub-questionnaires were developed based on previously used and validated metrics. Validity was further checked through cognitive interviews with participants. Since these sub-questionnaires tap into more abstract psychosocial concepts, it was important to conduct cognitive interviews to more thoroughly check if questions were measuring the intended concept. Again, revisions were made based on the cognitive interview feedback.

* 1. **What is the data collection procedure for each instrument?**
     1. **Who will be the target respondent for each data collection instrument?**

Baseline and endline data collection was carried out by a team of enumerators and supervised by a local research manager. All team members were fluent speakers of the local language Odia. For endline data collection, the team was not informed of village intervention status, but it may have been possible for them to infer due to the visible components of the intervention (i.e. hardware) or survey questions related to the intervention.

During both baseline and endline data collection, the enumerator teams attempted to survey all eligible households within the study village. The target respondent was the primary caregiver of the child <5 years old. If the primary caregiver was not available then a secondary caregiver was asked to participate. After confirming eligibility and obtaining verbal consent from the participant, the enumerator conducted the survey. Following the survey, the enumerator asked to observe the household’s latrine and bathing room and completed a structured spot-check.

* 1. **Secondary data sources**

**Please describe the anticipated secondary sources of data, if any, which will be used for this study.**

Not applicable. We will not use any secondary data sources.

## Analysis

* 1. **Outcome Variables**
     1. **Please describe the primary and any secondary outcome variables of interest using the following table:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Outcome** | **Description** | **Research Question** | **Level** |
| *Safe disposal*  ***(\*primary outcome)*** | This is a binary outcome based on 2 questions:  Q1: “The last time [NAME OF CHILD] defecated, where did [NAME OF CHILD] defecate?”  Q2: If the caregiver reports anything except “in latrine” or “in latrine using latrine training mat” then the following question is asked “Where was [NAME OF CHILD]’s feces disposed?”  1 = Q1 “in latrine” or “in latrine using latrine training mat” or Q2 “into latrine”  0 = Q2 response is anything except “into latrine” | RQ1  RQ2  RQ3  RQ5 | Individual  (child-level) |
| *Caregiver disposal in latrine* | This is a binary outcome based on 1 question:  Q2: If the caregiver reports anything except “in latrine” or “in latrine using latrine training mat” then the following question is asked “Where was [NAME OF CHILD]’s feces disposed?”  1 = Q2 “into latrine” or “in latrine using latrine training mat”  0 = Q2 response is anything except “into latrine” | RQ1a  RQ2a  RQ3a  RQ5a | Individual  (child-level) |
| *Child latrine use* | This is a binary outcome based on 1 question:  Q1: “The last time [NAME OF CHILD] defecated, where did [NAME OF CHILD] defecate?”  1 = Q1 “in latrine” or “in latrine using latrine training mat”  0 = Q1 response is anything except “in latrine” or “in latrine using latrine training mat” | RQ1b  RQ2  RQ3b  RQ5b | Individual  (child-level) |
| *RANAS behavioral factors for caregiver disposal in latrine* | The RANAS behavioral factors will act as intermediate outcomes for the mediation analysis (intervention status -> change in behavioral factors -> change in behavior).  Each RANAS sub-questionnaire is composed of about 25 questions, with each question examining a specific behavioral factor. Most questions use a 5pt Likert response while a few are dichotomous.  We will aim to combine some questions that examine the same factor into a single metric (if Cronbach’s alpha value is > 0.60).  Each behavioral factor outcome will be calculated by taking the difference between the endline survey response and baseline survey response for the question(s) measuring that factor. | RQ3a | Individual  (caregiver-level) |
| *RANAS behavioral factors for child latrine training* | RQ3b | Individual  (caregiver-level) |
| *CFM practices* | We will compare a variety of different CFM practices along the “child feces exposure pathway” between the intervention and control arm. For example - where child defecated, on what defecated, what materials used to handle feces, anal cleansing practices, hand washing practices, etc.  We will also examine caregivers’ perceived workload with these practices and how many other household members help with CFM practices (including the proportion of caregivers who report male household members providing support).  All of these outcomes are self-reported by the caregiver. | RQ4a | Individual  (child-level) |
| *Latrine use among household members ≥ 6 years old*  *Household total latrine use among members ≥ 6 years old* | The survey includes a household census where the caregiver reports on all members residing in the household, including their age, sex, and whether they defecated in the open or used the latrine the last time they defecated.  The last time this household member defecated, did they defecate in the open or use the latrine?  1 = Open  2 = Latrine  3 = Somewhere else (not open field or water body or roadway or latrine)  **Latrine use by age groups**  We will create a binary variable that records for each household member reported on a “1” for latrine use (answer “2 = latrine” for census question) or “0” for not latrine use (answer “1” or “3” for census question). We will then calculate the proportion of latrine use stratified by specific age groups (e.g. Household members 6 to 10 years old, 11 to 18 years old, 20 to 29 years old, etc.)  **Household total latrine use**  For each surveyed household, we will create a binary variable that records a “1” for total latrine use (i.e. all household members ≥ 6 years oldused the latrine) and a “0” for not total latrine use (i.e. at least one household member ≥ 6 years oldhad a “1” or “3” reported for the census question). | RQ4b | Household |
| *Visible feces in household compound* | This is a binary outcome. During the spot check at the end of the survey, the enumerator is prompted to look for any feces in the household compound:  Observe, presence of feces in household compound?   1. Yes, oxen/cattle 2. Yes, goats/sheep 3. Yes, chickens/ducks/pigeons 4. Yes, human / pig / dog / monkey 5. No feces observed   **Human feces in household compounds**  We will create a binary variable to measure the presence of human feces in compounds:  1 = Response “4. Yes, human/pig/dog/monkey” selected  0 = Response is anything except “4. Yes, human/pig/dog/monkey”  **Any feces in household compounds**  We will also create a binary variable to measure the presence of any (human or animal) feces in compounds:  1 = Response is anything except “5. No feces observed”  0 = Response “5. No feces observed” selected | RQ4c | Household |

* + 1. **If you plan on including covariates in your analysis, please provide a list of covariates that may be included.**

The following covariates will be included in some of the analyses:

* Stratification variable “geo-demo group” used in the randomization procedure (categorical with 5 response options that will be turned into dummy variables)
* Baseline prevalence of safe disposal at the village-level
* Imbalanced baseline covariates (see more details in section 5.2.1)

For RQ1, we will run both unadjusted and adjusted models and compare models as a sensitivity analysis. Specifically, we will run an unadjusted model, an adjusted model with stratification variable and imbalanced baseline covariates (if any), and then a “fully” adjusted model that also includes baseline prevalence of the primary outcome safe disposal.

* + 1. **If you plan to aggregate multiple variables into an index, which variables will you aggregate and how?**

**Household wealth:** Household wealth will be measured by constructing an asset index (scooter/motorcycle, television, telephone [landline or mobile], refrigerator, mattress, cot, chair, table, sewing machine, pressure cooker, watch/clock, electric fan, water pump, animal drawn cart, thresher, tractor, electricity, livestock) and using polychloric principle components analysis (PCA) to categorize households into five wealth quintiles (quintile 1 – least wealthy; quintile 5 – most wealthy) (reference).

\*This variable will be used to check for balance (see section 5.2.1 below)

* 1. **Quantitative Analysis**
     1. **Balance Checks**
        1. **How will you check balance between treatment and control groups?**

We will check balance between the intervention and control arms on a number of individual, household and village-level sociodemographic variables (see Table 2 below). The variables we selected were either determined by our study team to have a potential influence on safe disposal or have been shown in other studies to be associated with safe disposal. We will use the baseline survey data to calculate descriptive statistics and then compute the standardized difference between the arms to compare the two groups. The standardized difference is a metric that expresses the difference between groups in standard deviation units. For continuous covariates, we will calculate the standardized difference (*d*) using sample means and variance while for dichotomous covariates we will use sample prevalence.

In accordance with CONSORT guidelines, we will *not* perform significance testing of between-group differences on these sociodemographic variables when assessing balance. Since the study villages were *randomly* allocated to either intervention or control arm, then any imbalance we might identify can be attributed to chance. Accordingly, the null hypothesis for any significance test is already known to be true.

**Table 2.** Variables that will be examined for balance between intervention and control arm

|  |  |  |
| --- | --- | --- |
| **Variable** | **Type** | **Level** |
| 1. Village size | Continuous | Village |
| 1. Household religion | Categorical | Household |
| 1. Household caste | Categorical  (each category as dichotomous) | Household |
| 1. Household wealth quintile | Continuous | Household |
| 1. Household has only one child <5 years old | Dichotomous | Household |
| 1. Household size | Continuous | Household |
| 1. Caregiver age | Continuous | Caregiver |
| 1. Caregiver education level | Continuous | Caregiver |
| 1. Caregiver latrine use | Dichotomous | Caregiver |
| 1. Caregiver employment | Dichotomous | Caregiver |
| 1. Child age | Continuous | Child |
| 1. Child sex | Dichotomous | Child |
| 1. Household has functional piped water on plot/in household | Dichotomous | Household |
| 1. Latrine near household | Dichotomous (observed) | Household |
| 1. Latrine has two pits | Dichotomous (observed) | Household |
| 1. Latrine structure fully intact | Dichotomous (observed) | Household |
| 1. Latrine has water | Dichotomous (observed) | Household |

* + - 1. **What is the specification that you will run and what variables will you include?**

See section 5.2.1.1 above (we will compute standardized differences and variables are detailed in Table 2).

* + - 1. **If there is an imbalance (between treatment and control groups) in one or more baseline covariates, how do you plan to address this?**

As described in section 5.1.2, we will run an adjusted model for RQ1 that includes covariates that were found to be severely imbalanced between the intervention and control arms at baseline.

* + 1. **Contamination: How will you detect and manage any potential differential contamination between treatment and control groups?**

We aimed to have intervention and control villages geographically separated by at least a few kilometers to help mitigate potential spillover (i.e. contamination). The CFM intervention activities also involved only group meetings and individual household visits, so we did not anticipate spillover (as opposed to community-wide and mass communication strategies). However, to confirm this, we included a series of questions in the endline trial survey that asked control households about their awareness of the intervention. We asked the following:

* F1. Since December, have there been any activities in your village to encourage the safe disposal of child feces into latrine and/or teaching young children how to use the latrine for defecation?
* F2. What activities took place in your village?
* F5. [After showing picture and reading explanation] Have you ever heard of or seen this item called a ‘latrine training mat’?

We will report the results from these three questions alongside the evaluation results so it is apparent the level of spillover, if any.

* + 1. **Attrition / Matching**
       1. **What was your estimated attrition rate and what was your actual attrition rate? If the actual attrition rate is much higher than estimated, please offer some potential reasons for this higher attrition rate.**

We estimated an attrition rate of 10%, which was based on our previous experience conducting sanitation trials in the region. In actuality, about 35% more surveys were conducted at baseline compared to endline which could be viewed as ***no*** attrition. However, when we matched the endline and baseline survey participants we found that only about 45% of the endline participants were also surveyed at baseline (vice versa, this equates to about 60% of the baseline sample was surveyed at endline). If the main research question (RQ1) relied on matching the baseline and endline samples, then in that sense we would have a large attrition rate, but the analysis only uses the endline sample.

The more participants at endline may be due to the following factors:

* + **Timing** = During the baseline survey the data collection team kept a 9-5 schedule. However, during endline data collection the team was more strategic about when they went to villages to ensure caregivers would be available. For example, in some regions of the study area caregivers were more involved in agricultural works and thus were not available until much later in the day.
  + **COVID-19** = Our study area includes a large migrant population. Many households during baseline were locked/absent and this was in part because the family had migrated to a city (typically Surat, Hyderabad, Mumbai) for work. Then families returned after the first COVID wave and stayed. This might be the main reason why endline saw so many more eligible HHs.
  + **Language** = The endline survey team included a team member who was fluent in Sora language which is spoken in some regions of the study area. This may have allowed some caregivers to participate in the endline survey who were unable to during baseline due to language barrier.
    - 1. **How does the actual attrition rate impact your analyses for the research questions outlined in section 2.2?**

There is no impact on the main research question (RQ1) or some of the secondary research questions (RQ2 and RQ4), since these analyses rely only on the endline sample and actually many more participants were surveyed at endline. The mediation analyses for RQ3 do rely on matching the baseline sample to the endline sample but the attrition rate does not result in too small of a sample size for these analyses.

However, since only about 45% of the endline participants were also surveyed at baseline, we decided to conduct some additional analyses to determine whether or not the two samples — those who received endline and baseline surveys vs. those who received only endline — were meaningfully different. This led to the following analytical decisions:

* Decided to conduct the RQ5 sensitivity analysis
* Decided to also compare sociodemographics of the two groups (those who received both endline and baseline vs. those who only received endline)
* No matter the outcome of the sensitivity analyses, we also decided to produce 3 different models for RQ1: unadjusted, adjusted for stratification variable and any imbalanced covariates, adjusted for those variables as well as baseline prevalence of safe disposal. This decision was made since it is possible the baseline sample does not accurately reflect the “true” baseline prevalence of safe disposal for the endline sample. As such, we decided to run different models with and without the baseline prevalence of safe disposal.
  + - 1. **How does expected attrition change your power calculations?**

Our power calculation included a 10% attrition rate. In actuality, the endline sample had about 35% more participants surveyed compared to endline and so this, in essence, *lack of attrition* means our power has increased.

* + - 1. **How will you check balance between attritors and non-attritors? What is the specification that you will run and what variables will you include in these balancing checks?**

As described in 5.2.3.2, we will check balance between those who received only the endline survey and those who received both surveys (baseline and endline)

by comparing their sociodemographic characteristics. We will again use standardized differences and examine the same variables outlined in Table 2.

* + 1. **Missing Data: How will you deal with incomplete or missing data?**

We do not anticipate much missing data since the primary analysis does not rely on matching endline participants to their baseline survey (i.e. no missing data as a result of loss to follow-up). In one of the models for RQ1, we will adjust for baseline prevalence of safe disposal *at the village-level* and as such do not need to match participants. In addition, all questions were programmed to be required so there should be minimal missing data from enumerator error. As such, the main source of missing data will come from participants answering “don’t know” or “refused” to questions or ending the survey early. However, these situations were relatively rate in the baseline survey and we anticipate the same in the endline survey dataset.

However, since GEE is robust to missing data under the MCAR (missing completely at random) mechanism, we will use Little’s MCAR test to determine whether or not the MCAR assumption is met. If it is not met, then we will assume a MAR (missing at random) mechanism and do sensitivity analyses by conducting weighted (inverse-probability weighting) GEE and comparing model results.

* + 1. **Treatment Effects: Intention to Treat**
       1. **How will you estimate the (causal) effect of the offer of the treatment?**

Our RQ1 is the intention-to-treat question examining the causal effect of the CFM intervention on safe disposal (primary outcome). The effect of the intervention on JMP defined safe disposal will be assessed using a log-binomial model, which will yield the prevalence ratio of post-intervention safe disposal among intervention households relative to control households. We will use generalized estimating equations (GEE) with robust standard errors to account for village-level clustering. We will conduct an unadjusted GEE model, a model adjusted for the stratification variable used in randomization (i.e. geo-demo group) and any covariates found to be severely imbalanced between the two arms, and a model adjusted for those variables as well as village-level baseline prevalence of safe disposal.

* + - 1. **What is the specification that you will run and what controls will you include in your specification?**

We will use GEE with robust standard errors to estimate a population average model with the general form

g(.) = link function

= outcome of interest for the jth observation (i.e. child) in the ith cluster (i.e. village)

= a vector of covariates

= a vector of regression coefficients

We will specify an exchangeable correlation matrix as the most plausible and parsimonious choice of working correlation structure. However, we note that GEE with robust estimation yields valid estimates of model coefficients and standard errors when the correlation structure is misspecified (as opposed to multilevel mixed models).

We will use a log-binomial link function, which will yield the prevalence ratio of post-intervention safe disposal among intervention households relative to control households. As noted above, we will conduct three models that differ in their adjusted covariates.

We will specifically estimate and report the following three models:

Model 1: unadjusted

Model 2: adjusted for geo-demo group (village-level stratification variable used in randomization) and imbalanced covariates (if any)

Model 3: adjusted for geo-demo group, village-level baseline prevalence of safe disposal, imbalanced covariates (if any)

* + 1. **Treatment on the Treated**
       1. **How will you estimate the (causal) effect of the receipt of the treatment?**

We will only do a treatment on the treated, or per-protocol analysis, in the following situations:

* The ITT models (described in the section above) show no significant effect
* If a substantial proportion of the endline sample reports they did not participate in at least 3 (of the 5) intervention activities.

If these situations arise then we will define “treatment” as attending at least 3 out of the 5 intervention activities and receiving hardware (if eligible for hardware). In the endline survey caregivers were asked for each intervention activity if they or someone in their household attended the activity. This data will be used to determine the attendance requirement. In addition, caregivers are asked if they received any hardware. Caregivers with a child <7 months were eligible to receive a wash basin and bucket with lid while caregivers with a child between 7 months to <4 years were eligible to receive a latrine training mat with tray. However, in some cases caregivers asked for the other hardware depending on their child’s needs. As such, for the hardware component of the “treatment” definition we will consider it to be per-protocol if a caregiver whose youngest child was <4 years old at the time of the intervention reports receiving any of the intervention hardware.

* + - 1. **What is the specification that you will run and what controls will you include in your specification?**

We will run the exact same models outlined in section 5.2.5.2 but with a sub-sample of the intervention households that meet the “treatment” definition.

* + 1. **Other Treatment Effects: What analyses will you conduct to answer the secondary research questions outlined in section 2.2?**

Secondary research question (RQ2):

* RQ1a: What is the effect of the intervention on **caregiver disposal in latrine** when comparing children among the intervention versus control arms? (sub-analysis: stratified examinations by child sex and child age)
* RQ1b: What is the effect of the intervention on **child latrine use** when comparing children among the intervention versus control arms? (sub-analysis: stratified examinations by child sex and child age)

**Analysis:** In order to answer RQ2,we will conduct the same intention-to-treat GEE model outlined under 5.2.5.2. However, instead of safe disposal being the outcome, we will use caregiver disposal in latrine as the outcome for one set of models and child latrine use as the outcome in another set of models. For both of these outcomes, we will conduct 3 models (unadjusted, adjusted for geo-demo group and any imbalanced covariates, adjusted also for baseline prevalence).

Secondary research question (RQ3):

* RQ3a. Are there behavioral factors that were changed by the intervention and, if so, which of these factors in turn influenced **caregiver disposal in latrine**? In other words, which behavioral factors mediate the effect of the intervention on **caregiver disposal in latrine** (taking into account factors and behavior at baseline)?
* RQ3b: Are there behavioral factors that were changed by the intervention and, if so, which of these factors in turn influenced **child latrine use**? In other words, which behavioral factors mediate the effect of the intervention on **child latrine use** (taking into account factors and behavior at baseline)?

**Analysis:** In order to answer RQ3 and understand *how* the intervention changed behavior**,** we will conduct simple mediation analyses following the approach outlined by Preacher and Hayes (2004). Since we want to examine the effect of the intervention on *change* in behavioral factors, we will conduct the analysis among participants who received both the baseline and endline trial surveys and could be matched. Change in each behavioral factor will be calculated by taking the difference between the endline survey response and baseline survey response for the question(s) measuring that factor. In one mediation analysis, the study arm (intervention or control) will be the independent variable, change in behavioral factors the mediation variables, and caregiver disposal in latrine the dependent variable. In another mediation analysis, child latrine use will be the dependent variable. We will examine both direct and indirect effects.

Secondary research question (RQ4):

* RQ4a: Does the prevalence of **other CFM practices** along the “child feces exposure pathway” (such as where wash water is disposed, if and how hardware is used, whether caregiver washes her hands after, etc.) differ between households that received the intervention compared to households that did not?
* RQ4b: Does the prevalence of latrine use among **household members ≥ 6 years old differ** between households that received the intervention compared to households that did not?
* RQ4c: Does the prevalence of **visible feces** in the household compound differ between households that received the intervention compared to households that did not?

**Analysis:** In order to answer RQ4 and understand what other impacts the intervention may have had related to child feces management (CFM), we will calculate the prevalence/mean among the intervention arm and control arm for a variety of secondary outcomes (other CFM practices, latrine use among household members ≥ 6 years old, and presence of visible feces in the household compound). The calculations will be done among participants who received the endline trial survey (i.e. not restricted to participants who received *both* baseline and endline trial survey). We will then calculate the difference between intervention and control arm for each prevalence/mean and present 95% confidence intervals for the calculated difference.

Sensitivity analysis (RQ5): What is the effect of the intervention on **safe disposal of child feces** when comparing intervention versus control among caregivers who received ***both*** baseline and endline surveys (i.e. 2 surveys)? What is the effect of the intervention among caregivers who received ***only*** the endline survey (i.e. 1 survey)? How do these effects compare?

* RQ5a: What is the effect of the intervention on **caregiver disposal in latrine** when comparing intervention versus control among caregivers who received ***both*** baseline and endline surveys (i.e. 2 surveys)? What is the effect of the intervention among caregivers who received ***only*** the endline survey (i.e. 1 survey)? How do these effects compare?
* RQ5b: What is the effect of the intervention on **child latrine use** when comparing intervention versus control among caregivers who received ***both*** baseline and endline surveys (i.e. 2 surveys)? What is the effect of the intervention among caregivers who received ***only*** the endline survey (i.e. 1 survey)? How do these effects compare?

**Analysis:** In order to answer RQ5,we will conduct the same intention-to-treat GEE model outlined under 5.2.5.2. However, we will conduct the analysis with different study samples. In one analysis, the sample will be participants who received only the endline survey while in another analysis the sample will be participants who received both the baseline and endline surveys. For each of these samples, we will conduct three models: a model for prevalence of safe disposal of child feces, a model for prevalence of caregiver disposal in latrine, and a model for prevalence of child latrine use. As such, we will conduct 6 total models to answer RQ5 and then will compare the model results to determine if there is a significant difference in change in behavior between the two samples.

* 1. **Heterogeneous Effects**
     1. **Which groups do you anticipate will display heterogeneous effects?**

We anticipate caregivers who received both baseline and endline trial survey *may* display heterogenous effects compared to caregivers who received only the endline trial survey. As such, we are conducting a sensitivity analysis (see RQ5).

* + 1. **What is the broad theory of action that leads you to anticipate these effects?**

Studies show that repeated surveying on health and behavioural outcomes can lead to response bias. As such, we want to conduct a sensitivity analysis to see if caregivers who were surveyed twice (received both baseline and endline surveys) report higher prevalence of safe disposal compared to caregivers who were only surveyed once (endline survey). If we detect a difference, then this suggests the potential for response bias. However, there are other confounding factors that could also explain the difference.

* 1. **Standard Error Adjustments**
     1. **How will you address clustering in your data?**

We will employ generalized estimating equations (GEE) with robust standard errors to account for village-level clustering in the outcome.

* + 1. **How will you address false positives from multiple hypothesis testing?**

For the main research question (RQ1), we are fitting only three pre-specified models (unadjusted, adjusted, fully adjusted). If we consider the secondary analyses from RQ2, RQ3 and RQ5 as well, then this does introduce a number of additional hypothesis tests. However, RQ4 will not include any hypothesis testing as we will simply present differences between intervention and control on those outcomes with 95% confidence intervals. Overall, we do not think the number of planned analyses, which we describe here *a priori,* require any form of correction due to issues of multiple hypothesis testing

* + - 1. **If you plan to adjust your standard errors, what adjustment procedure will you use? (e.g., Family Wise Error Rate, False Discovery Rates, etc.)**

Not applicable. We plan to use robust standard errors.

* + - 1. **How will you deal with outcomes with limited variation? For instance, one option could be to decide in advance that outcomes that vary below a certain threshold will be omitted from the analysis.**

We do not anticipate limited variation in the behavioural outcomes (safe disposal, caregiver disposal in latrine, and child latrine use) since past sanitation studies conducted in the area show these are not relatively common practices (i.e. we do not anticipate the vast majority of participants practicing or not practicing them). We also do not anticipate limited variation among the secondary outcomes outlined in RQ4, and even if there is, we will still report on these outcomes since we are simply calculating differences in prevalence between intervention and control.

**References**

Barrera, M. J., Sandler, I. N., & Ramsay, T. B. (1981). Preliminary Development of a Scale of Social Support: Studies on College Students. *American Journal of Community Psychology, 9*(4).

Bauza, V., Reese, H., Routray, P., & Clasen, T. (2019, Apr). Child Defecation and Feces Disposal Practices and Determinants among Households after a Combined Household-Level Piped Water and Sanitation Intervention in Rural Odisha, India. *Am J Trop Med Hyg, 100*(4), 1013-1021. <https://doi.org/10.4269/ajtmh.18-0840>

Clasen, T., Boisson, S., Routray, P., Cumming, O., Jenkins, M., Ensink, J. H., Bell, M., Freeman, M. C., Peppin, S., & Schmidt, W.-P. (2012). The effect of improved rural sanitation on diarrhoea and helminth infection: design of a cluster-randomized trial in Orissa, India. *Emerging themes in epidemiology, 9*(1), 7.

JMP. (2018). *Core questions on drinking water, sanitation and hygiene for household surveys: 2018 update*. <https://www.who.int/water_sanitation_health/monitoring/oms_brochure_core_questionsfinal24608.pdf>

Mosler, H.-J. (2012). A systematic approach to behavior change interventions for the water and sanitation sector in developing countries: a conceptual model, a review, and a guideline. *Int J Environ Health Res, 22*(5), 431-449. <https://doi.org/10.1080/09603123.2011.650156>

Mosler, H.-J., & Contzen, N. (2016). *Systematic Behavior Change in Water, Sanitation, and Hygiene: A practical guide using the RANAS approach. Version 1.1.*

Preacher, K. J., & Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior Research Methods, Instruments, & Computers, 36*(4), 717-731.

Reese, H., Routray, P., Torondel, B., Sinharoy, S. S., Mishra, S., Freeman, M. C., Chang, H. H., & Clasen, T. (2019, Dec 1). Assessing longer-term effectiveness of a combined household-level piped water and sanitation intervention on child diarrhoea, acute respiratory infection, soil-transmitted helminth infection and nutritional status: a matched cohort study in rural Odisha, India. *Int J Epidemiol, 48*(6), 1757-1767. <https://doi.org/10.1093/ije/dyz157>

Rutterford, C., Copas, A., & Eldridge, S. (2015, Jun). Methods for sample size determination in cluster randomized trials. *Int J Epidemiol, 44*(3), 1051-1067. <https://doi.org/10.1093/ije/dyv113>

Schwarzer, R., & Schulz, U. (2013). *Berlin Social Support Scales (BSSS)*. Measurement Instrument Database for the Social Science. [www.midss.ie](file:///Users/gloria/Desktop/OneDrive%20-%20Emory%20University/Box/CFM%20Project%20w_%20Gram%20Vikas/(8)%20Manuscripts/(7)%20Trial%20Impact%20paper/www.midss.ie)

Sclar, G. D., Bauza, V., Mosler, H.-J., Bisoyi, A., & Clasen, T. (under review). Study design and rationale for a cluster randomized trial of a safe child feces management intervention in rural Odisha, India. *BMC public health*.

Vaux, A., Riedel, S., & Stewart, D. (1987). Modes of social support: The Social Support Behaviors (SS-B) Scale. *American Journal of Community Psychology, 15*(2), 209-237.

Zimet, G. D., Dahlem, N. W., Zimet, S. G., & Farley, G. K. (1988). The Multidimensional Scale of Perceived Social Support. *Journal of Personality Assessment, 52*(1), 30-41. <https://doi.org/10.1207/s15327752jpa5201_2>