BMJ Open

Design and rationale of a matched cohort study to assess the effectiveness of a combined household-level piped water and sanitation intervention in rural Odisha, India

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2016-012719
Article Type:	Protocol
Date Submitted by the Author:	19-May-2016

Complete List of Authors:	Reese, Heather; Emory University School of Public Health, Department of Environmental Health Routray, Parimita; London School of Hygiene and Tropical Medicine Torondel, Belen; London School of Hygiene and Tropical Medicine Sclar, Gloria; Emory University School of Public Health, Department of Environmental Health Delea, Maryann; Emory University School of Public Health; London School of Hygiene and Tropical Medicine Sinharoy, Sheela; Emory University Laney Graduate School, Nutrition and Health Sciences Program Zambrano, Laura; Emory University School of Public Health, Department of Environmental Health Caruso, Bethany; Emory University School of Public Health, Department of Environmental Health Mishra, Samir; KIIT University, School of Biotechnology Chang, Howard; Emory University, Department of Biostatistics and Bioinformatics Clasen, Thomas; London School of Hygiene & Tropical Medicine; Emory University School of Public Health, Department of Environmental Health

Primary Subject Heading:	Public health
Secondary Subject Heading:	Epidemiology
Keywords:	sanitation, piped water supply, diarrheal diseases

SCHOLARONE™ Manuscripts

Design and rationale of a matched cohort study to assess the effectiveness of a combined household-level piped water and sanitation intervention in rural Odisha, India

Heather Reese,[1] Parimita Routray,[2] Belen Torondel,[2] Gloria Sclar,[1] Maryann G. Delea,[1,2] Sheela S. Sinharoy,[3] Laura Zambrano,[1] Bethany Caruso,[1] Samir R. Mishra,[4] Howard H. Chang,[5] Thomas Clasen[1,2]

[1] Department of Environmental Health, Rollins School of Public Health, Emory University, Atlanta, Georgia, United States of America

[2] London School of Hygiene and Tropical Medicine, London, United Kingdom

[3] Nutrition and Health Sciences Program, Laney Graduate School, Emory University, Atlanta, Georgia, United States of America

[4] School of Biotechnology, KIIT University, Bhubaneswar, Odisha, India

[5] Department of Biostatistics and Bioinformatics, Rollins School of Public Health, Emory University, Atlanta, Georgia, United States of America

Contact information for corresponding author, Heather Reese:

Department of Environmental Health, Rollins School of Public Health, Emory University

1518 Clifton Rd, Atlanta, GA 30322

Email: heather.e.reese@emory.edu, Phone: (619) 379-8963

Keywords: rural, piped water, sanitation

Word count: 4422

ABSTRACT

Introduction*:* Government efforts to address massive shortfalls in rural water and sanitation in India have centered on construction of community water sources and toilets for selected households. However, deficiencies with water quality and quantity at the household level, and community coverage and actual use of toilets has led Gram Vikas, a local NGO in Odisha, India, to develop an approach that provides household-level piped water connections contingent on full community-level toilet coverage.

Methods: This matched cohort study was designed to assess the impact of a combined piped water and sanitation intervention. Households with children under five years in 45 randomly selected intervention villages and 45 matched control villages will be followed over 16 months. The primary outcome is prevalence of diarrheal diseases; secondary health outcomes include soil-transmitted helminth infection, nutritional status, seroconversion to enteric pathogens, urogenital infections, and environmental enteric dysfunction. In addition, intervention effects on sanitation and water coverage, access and use, environmental fecal contamination, women's empowerment, as well as collective efficacy, and intervention cost and cost-effectiveness will be assessed.

Ethics and dissemination: The study protocol has been reviewed and approved by the ethics boards of the London School of Hygiene and Tropical Medicine, U.K. and KIIT University, Bhubaneswar, India. Findings will be disseminated via peer-reviewed literature and presentation to stakeholders, government officials, implementers and researchers.

Trial registration identifier: NCT02441699

STRENGTHS AND LIMITATIONS OF THIS STUDY

- The study assesses a combined household-level piped water and sanitation intervention that requires complete community-level compliance.
- The intervention was not randomly allocated; but, controls are selected through a restriction process to limit possible partial exposure to the intervention through spillover, and matched to intervention villages using pre-intervention data.
- The study uses a holistic definition of health to assess intervention impacts on physical, mental and social well-being, including more novel outcomes such as seroconversion, environmental enteric dysfunction, and sanitation insecurity. It also assesses intervention coverage, cost-effectiveness, and collective efficacy.
- The time lapse between intervention completion and the beginning of the evaluation process prevents baseline comparison or assessment of immediate intervention impacts. However, it allows for a biologically plausible length of time for die-off of even the most persistent pathogens in the environment, and provides time for children to have be born into this environment.

INTRODUCTION

Of the one billion people practicing open defecation worldwide, over half live in India[1]. While international and national pressure on improving sanitation conditions in India has led to over 350 thousand people gaining access to improved toilets since 1990, it has barely kept up with population growth[1,2]. Recent studies show that even in areas with access to household-level improved sanitation, use of these toilets is low[3–5]. This may be due in part to a mismatch between the culturally acceptable pour-flush toilets and the level of water access. Coverage of

improved water sources, usually community-level pumps or taps, is relatively high even in rural areas in India, but it may not be sufficient for flushing purposes on top of other daily water needs[1,6].

Although the effectiveness of water, sanitation and hygiene (WASH) interventions vary, meta-analyses have found that individual or combined WASH interventions decrease diarrheal disease prevalence by up to 48%[7–11]. While combined interventions would be expected to have a greater influence on multiple exposure pathways and thus a greater combined impact on health, there is limited evidence of additive benefits[12]. This may be due to poor uptake, inconsistent use, or an incomplete understanding of relevant pathways[8–10]. In India, a combined water and sanitation intervention may be more critical than just interrupting multiple transmission pathways for enteric infection; evidence suggests that household-level water access is integral to the use of improved sanitation in this context[13].

While the intent of improved sanitation facilities is to separate human feces from human contact, most of the focus is on constructing household toilets to increase improved sanitation coverage—the primary metric used in monitoring progress toward international targets. However, studies in India have further shown that toilet construction does not translate into toilet use in this context[5,14–16]. Moreover, with the interdependence between members of households and households within communities, safe water and sanitation is a community-level issue. There is growing emphasis on assessing health risk from poor water and sanitation conditions not simply due to individual or even household-level risk factors, but also from conditions in the community environment[17]. There is evidence that even households without toilets, and households which do not filter drinking water, showed decreased health risk if they live in communities with high levels of coverage and use[18–20].

A main risk of poor WASH conditions is enteric infection, caused by a diverse array of bacteria, viruses, protozoa, and parasites, including soil-transmitted helminths. These infections may cause diarrhea, the second leading cause of mortality for children under five years worldwide and in India, a leading cause of mortality regardless of age[21,22]. There is also growing evidence that asymptomatic enteric infections may pose a similar risk, with repeat enteric infections contributing to chronic malnutrition, environmental enteric dysfunction, poor cognitive outcomes, and poor vaccine uptake[23–28]. Poor WASH conditions are also linked to increased risk of respiratory infection, the leading cause of mortality for children under five years worldwide[21,29,30]. Poor water and sanitation access can also affect the social, physical and mental well-being of women, acting through pathways ranging from unsafe menstrual hygiene management practices and increased risk of violence[31–33].

Description of the intervention

Over the past decades there has been a global commitment to determine water and sanitation interventions with demonstrated effectiveness, not just efficacy[34]. Gram Vikas, a non-governmental organization based in Odisha, India (http://www.gramvikas.org/), has responded by implementing its MANTRA (Movement and Action Network for Transformation of Rural Areas) water and sanitation program in more than 1000 villages since 2002[35]. This approach includes both household-level piped water connections, and community-level mobilization for culturally appropriate household toilets. A previous interrupted time series analysis of the MANTRA intervention reported it to be protective against diarrheal diseases[36]. However, in addition to limitations of design, this study relied on outcome data collected and reported by

Gram Vikas, the intervention implementer, and did not assess intervention coverage or impacts on environmental fecal contamination.

The MANTRA water and sanitation intervention is rolled out in a three-phase process over an average three to five years (Figure 1). During the first, or Motivational, phase, representatives of Gram Vikas visit the identified village several times to assess village interest and progress towards a set of Gram Vikas requirements, including: 1) the commitment of every household to participate, 2) creation of a village corpus fund from contributions from every household, and 3) development of village guidelines for maintenance and use of facilities.

Once this set of requirements is achieved, the village progresses into the second, or Operational, phase of the intervention. Each household constructs a pour-flush toilet with two soak-pits and a separate bathing room. The households hire a local, skilled mason and provide their own unskilled labor and locally available materials to complete the superstructure. Gram Vikas provides external materials such as PVC pipes and porcelain pans. At the same time, a water tank, community meeting space, and piped water distribution system connected to every household, with taps in the toilet and bathing rooms and a separate tap in the kitchen, is constructed through a similar collaborative process.

All households must construct a toilet and bathing room for the village to progress into the final, or Completed, phase of the intervention, in which the water system is turned on. Notably, this three-phase process only allows each household access to piped water once every household in the village has a toilet and bathing room. This model contrasts with most previous water and sanitation interventions, including those implemented under India's Total Sanitation Campaign and other government programs, which do not require community-level sanitation compliance and do not provide a piped water supply at the household level[37].

Study aims

The primary objective of this study is to evaluate the effectiveness of the combined household-level water supply and sanitation intervention, as implemented by Gram Vikas in Odisha, India. Toward that objective, this study aims to:

1) Assess the effectiveness of the intervention in improving water and sanitation infrastructure coverage, access, and use and to assess fecal sludge management practices in intervention communities.
2) Assess the effectiveness of the intervention in reducing environmental fecal contamination.
3) Assess the effectiveness of the intervention in improving health. This includes reported diarrheal disease in children under 5 years (primary outcome), acute respiratory infection, infection with soil-transmitted helminthes, nutritional status, environmental enteric dysfunction, seroconversion for selected enteric pathogens, and urogenital diseases associated with menstrual hygiene management practices. Mental and social well-being will be explored through assessment of sanitation insecurity and women's empowerment.
4) Assess the cost and cost-effectiveness of the intervention.
5) Develop and assess a theoretically-grounded, empirically informed collective efficacy scale; and determine the effect of collective efficacy on intervention effectiveness.

METHODS

Setting

The study is located in Ganjam and Gajapati districts located in eastern Odisha, India (Figure 2). These two contiguous districts were a single district until 1992. Over 44% of the population in these districts is recognized by the Government of India as being below the poverty line (BPL)[38]. As of 2008, a majority of households in both districts had access to an improved, likely community-level, drinking water source, with over 23% of households in Ganjam having access to any sanitation facility, compared to only 8% of households in Gajapati[38]. The area is primarily rural and agrarian, and the climate is characterized by a monsoon season from June to September, with an average rainfall of ~1400 mms/year.

Study design

This study uses a matched cohort design with data collected across four study rounds from June 2015 to September 2016. As described below, control villages were matched to randomly selected intervention villages through a multi-step restriction, genetic matching, and exclusion process using the following eligibility criteria (Figure 3).

Eligibility criteria for villages

1. Restriction. Intervention villages were randomly selected from a list of Gram Vikas villages in Ganjam and Gajapati districts provided by the NGO, after restriction to villages with a Motivation phase start date between 2002-2006 and a Construction phase start date no earlier than 2003. Since the intervention process takes on average three to five years, the criteria for the Motivation start date helped to identify those villages with ongoing interventions at the same time. In addition, this allowed the use of the Government of India Census 2001 and the Below

Poverty Line (BPL) Survey 2002 data to characterize baseline characteristics in both intervention and control villages.

Eligible control villages include all villages without a Gram Vikas intervention within the study districts which: 1) are not within the same Gram Panchayat (a political subdivision with some administrative responsibility for water and sanitation) as a Gram Vikas village, or bordering a Gram Vikas village, and 2) had not received a Motivation visit from the Gram Vikas NGO. These criteria serve to limit the possibility of previous partial exposure to the intervention through spillover from adjacent villages or direct contact with the NGO. These criteria also increase strength of the counterfactual provided by the control villages, i.e. if they had received a motivation visit from Gram Vikas, the control villages would have been equally as likely as the intervention villages to demand the intervention.

In addition, to be eligible for inclusion both intervention and control villages must: 1) appear in the Government of India Census 2001 and the BPL Survey 2002, 2) have a population of at least 20 households, and 3) be within approximately three hours travel from the study office in Brahmapur, Ganjam District. This last criterion is due to logistical constraints.

2. Matching. After restriction, genetic matching was used to match potential control villages to the randomly selected intervention villages without replacement[5,39,40]. Villages were exact matched on district to limit any political or large scale geographic variation between district populations, and were also matched on pre-intervention demographic, socioeconomic, sanitation, and water access characteristics[5]. These village level matching variables were selected due to their theorized association with the primary outcome, diarrheal diseases, as well as data availability.

3. Exclusion. The field team visited matched potential control villages and intervention villages to assess suitability for the study through a rapid assessment interview with village leadership and to ensure accessibility. Villages were excluded if they are not within three hours travel of the field office in Brahmapur, had sustained major infrastructure damage due to a natural disaster, or if there was a current or planned sanitation or water intervention by an organization external to the village in the next 12 months. In addition, villages were excluded if there were fewer than three children under five years old. As villages were removed from the pool of prospective control villages, the matching and exclusion processes were repeated.

After matching and exclusion, covariate balance was assessed for all matching variables through examination of balance measures[41–43]. Matching resulted in an improvement in balance as assessed through comparison of several measures including q-q plots, Kolmogorov-Smirnov bootstrap p-values, and standardized differences. After matching, there were no significant differences between intervention and control groups (Table 1).

Eligibility criteria for households

Households within selected intervention and control villages will be eligible if they have at least one child under 5 years old at time of enrollment, verified with birth or immunization card, and expect to reside in the village for the duration of the study. If there are more than 40 eligible households within a village, 40 will be randomly selected to be enrolled. Informed consent will be obtained from the male and/or female head of the selected households. All children under five years within each enrolled household are eligible and will not age-out over the course of the

Table 1. Pre-intervention characteristics and balance diagnostics before and after matching and exclusion process.

Variable	**Intervention** (n=45)	**Control (all eligible)** (n=1580)	Std Diff (all eligible)	**Control (study)** (n=45)	Std Diff (study)
Number of households	157.9	215.5	0.37	148.1	0.06
Population under 6 years (%)	16.2	16.9	0.19	16.3	0.02
Household income score ($\bar{x}$)	2.9	3.1**	0.26	2.9	0.01
Household goods owned ($\bar{x}$)	1.1	1.2*	0.27	1.1	0.02
Pucca house (%)	59.2	61.6	0.09	60.5	0.05
≥2 meals a day (%)	57.7	63.7	0.19	57.8	0.01
Scheduled caste (%)	11.5	18.7**	0.46	11.8	0.01
Scheduled tribe (%)	33.4	19.1*	0.31	29.8	0.08
Female literacy (%)	30.9	29.8	0.07	30.9	0.00
Open defecation (%)	95.6	95.2*	0.04	95.8	0.01
Improved drinking water source ǂ (%)	38.6	42.5	0.10	37.2	0.02
Water source <500m and 50m elevation (%)	81.5	72.2	0.31	81.7	0.01

All eligible: villages that are eligible for the matching process after restriction
Std Diff (absolute standardized difference): a value greater than 0.1 is considered meaningful imbalance [41]
ǂ Ganjam villages only; no data available for Gajapati villages
Kolmogorov-Smirnov bootstrap p-values: * <0.05 ** <0.01

study. Households with newborn children will be enrolled as they become eligible on an ongoing basis throughout the study in villages with fewer than 40 enrolled households.

Sample Size

Sample size was determined through a simulation estimating the log odds of diarrheal disease (the primary outcome) through a multilevel random effects model and parameterized with data from a previous study in a neighboring district in Odisha[16]. Sample size estimates were also checked with G*Power[44]. The simulation assumes a longitudinal 7-day period prevalence for diarrhea of 8.8% in children under five years, a heterogeneity variance between villages of 0.07, a heterogeneity variance between households of 0.57, and four study rounds. An effect size of 0.20 was selected for public health importance and based on estimates of effect from systematic reviews of water and sanitation studies[45]. Assuming at least 80% power, 0.05 significance level, 10% for loss to follow up, and at least one child per household, we estimate a sample size of 45 villages per study arm and 26 households per village.

Outcome Measurement

Outcomes will be measured through surveys, interviews, or through the collection and analysis of environmental, stool or dried blood spot samples. All survey questions will be translated into the primary local language, Odia, and back-translated to confirm wording. Household surveys will be verbally administered by trained field workers to the mother or primary caregiver of the youngest child under five in each household, unless otherwise specified below. Community surveys will be verbally administered to the *sarpanch* (village head) or other member of village leadership. Survey data will be collected on mobile phones using Open Data Kit[46]. GPS coordinates for households, water sources and other relevant sites will be collected using Garmin eTrex 10 or 20 devices (Garmin Ltd., Olathe, KS, USA).

Diarrheal Diseases

The primary outcome for this study is prevalence of diarrheal diseases, recorded as both daily point prevalence over the previous three days and seven-day period prevalence, for all household members in each of the four rounds. Although self-reported diarrhea is a subjective outcome with a well-established risk of bias, three-day recall reduces recall bias[47,48]. Diarrheal disease will be measured using the World Health Organization (WHO) definition of three or more loose stools in a 24-hour period, with or without the presence of blood. Field workers will use a simple calendar as a visual aid to help respondents with recall. Each household member will be asked to recall his or her own disease status and the mother or primary caregiver will be asked to report disease for children.

Respiratory infection

Prevalence of respiratory infections will be recorded as both daily point prevalence over the previous three days and seven-day period prevalence for all household members in each round. Respiratory infection is defined as the presence of cough and/or shortness of breath/difficulty breathing according to WHO's Integrated Management of Childhood Illness (IMCI)[49]. The full IMCI case definition for acute lower respiratory infection also includes measurement respiratory rate and observation of chest indrawing, stridor and other danger signs; these criteria were excluded from our definition as there is concern about the technical support required to produce consistent and accurate data within this context[49]. Our definition will provide a broad assessment of respiratory illness burden.

Nutritional Status

Children's height and weight will be measured in all rounds using standard anthropometric measurement methods as established by WHO[50,51]. Field workers will be trained and standardized[51]. Weight will be measured for all children under five years of age using Seca 385 digital scales, with 20g increment for weight below 20kg and a 50g increment for weight between 20 and 50kg. Recumbent length will be measured for children less than two years of age using Seca 417 measuring boards with 1mm increment. Standing height will be measured for children two to five years of age using Seca 213 portable stadiometers with 1mm increment. Height and weight will be used to calculate height-for-age z-scores (HAZ) and weight-for-height z-scores (WHZ) based on WHO reference standards. A random subset of 10% of households will receive back check visits each day to repeat height/length measurements to ensure inter-observer reliability.

Soil-transmitted helminth infection

Stool samples will be collected in rounds 2 and 4 from all household members in a randomly selected subset of households, and used to assess the presence and intensity of soil-transmitted helminth (STH) infection. Formalin ether concentration and microscopy will be used to quantify worms and ova for hookworms, *Ascaris lumbricoides*, and *Tricuris trichura*[52]. Quality assurance will include independent duplicate assessment of all positive and 10% of negative samples. After stool collection, each participant will be offered a single dose of Albendazole, a broad-spectrum antihelmenthic drug recommended by the Ministry of Health and Family Welfare, Government of India. Stools collected in round 2 will allow for comparison of STH infection prevalence between intervention and control villages, while the stool samples collected approximately 8 months later in round 4 will provide a measure of re-infection rate.

Environmental enteric dysfunction

Stools from randomly selected children under two years old, collected in rounds 2 and 4, will be used to assess environmental enteric dysfunction (EED) through quantification of biomarkers of intestinal inflammation and permeability. Fecal myeloperoxidase (MPO), alpha-1-antitrypsin (AAT), and neopterin (NEO), markers for neutrophil activity, intestinal permeability and TH1 immune activation, respectively, were selected for this study based on evidence of association with EED, subsequent linear growth deficits, and household environmental fecal contamination[23,24,53].

Seroconversion for enteric pathogens

Serological assays that assess antibody production against various enteric pathogens can provide an objective measure of exposure to enteric infections[54]. Enrolling children aged 6 to 18 months will reduce the potential for interference from maternally acquired antibodies and permit analysis of seroconversion data in a critical window for young children who experience higher diarrheal disease morbidity and mortality before two years of age[55–60]. Children who are 6 to 12 months during round 2 will have capillary blood drawn by fingerstick or heelstick, as appropriate, and will be visited again during round 4 for a second capillary blood sample. All blood samples will be preserved on TropBio (Sydney, Australia) filter discs and stored within 7 days of collection at -20°C. Seroconversion against markers for norovirus, *Giardia intestinalis*, *Cryptosporidium parvum*, *Entamoeba histolytica*, enterotoxigenic *E. coli* heat-labile enterotoxin (ETEC-LT), *Salmonella* spp., *Campylobacter jejuni*, *Vibrio cholera*, and *Toxoplasma* spp. will be compared using multiplex immunoassay technology on the Luminex xMAP platform[61].

Environmental fecal contamination

Field workers will collect samples of household stored drinking water and source water from a random subset of households in each village in all study rounds, and child hand rinses in rounds 2 and 4. All water and hand rinse samples will be stored on ice during transport and analyzed within 6 hours of collection using membrane filtration. Three assays will be used: 1) plating on m-Coli Blue 24 (Millipore, Billerica, MA) for *E.coli* according to EPA Method 10029, 2) alkaline peptone water enrichment prior to plating on thiosulfate citrate bile salts sucrose agar and slide agglutination serotyping for *V. cholerae*, and 3) plating on xylose lysine desoxycholate agar, and slide agglutination serotyping for *Shigella* spp.[62–64]. Source and stored water samples will be assayed for *E. coli*, *Vibrio cholerae* and *Shigella* spp., and hand rinse samples will be assayed for *E. coli* and *Shigella* spp. *E. coli* was selected as a standard non-human specific indicator of fecal contamination, though the limitations of this indicator are well-established[65–67]. In order to better characterize human fecal contamination of the household environment, *Vibrio cholerae* and *Shigella* spp. were selected based on prevalence in southern Asia, evidence of public health importance, and field laboratory limitations[68–70].

Cost and cost-effectiveness

Costs and potential cost savings (i.e., averted costs) associated with the intervention will be assessed through an economic costing approach that recognizes and quantifies costs and benefits from a societal perspective[71]. Data on program and point-of-delivery inputs will be collected at household, community, and implementer levels. Field workers will administer community surveys to a village leader, and household surveys to the household decision-maker for toilet installation, in 20 randomly selected households in half of the intervention and control villages. Surveys will collect data on household- and village-level inputs related to materials and labor required to construct household toilets and wash rooms, the community water tank and distribution system, and household water connections; longer-term water supply and toilet maintenance costs; and financing required for this infrastructure as well as perceived benefits, including averted social opportunity costs. Implementer inputs from Gram Vikas will be collected through an enumeration exercise, interviews, and examination of the implementer's financial records.

Collective efficacy

Collective efficacy (CE) is a latent construct comprised of the structural and cognitive components that facilitate a community's shared belief in its ability to come together and execute actions related to a common goal[72]. A review of the literature and established conceptual frameworks will be performed to define the CE construct. A sequential exploratory mixed qualitative and quantitative design will be used to develop and refine a scale to measure CE and test hypotheses. Field workers will administer the refined, multi-item, Likert-type CE scale to one randomly selected household member aged 18 years or older in each household.

Women's empowerment

Four dimensions of women's empowerment will be measured: group participation, leadership, decision-making and freedom of movement. Group participation and leadership will be measured using modules from the Women's Empowerment in Agriculture Index, which has been tested in South Asia[73]. Decision-making and freedom of movement will be measured using questions from the women's status module of Demographic and Health Surveys. These measures were selected based on the importance of women's empowerment for child nutrition[74,75]. Women's empowerment is conceptualized as both an outcome and a potential mediator along the pathway between the Gram Vikas intervention and child health outcomes.

Menstrual hygiene management

Menstrual hygiene management practices vary worldwide and depend on personal preference, socioeconomic status, local traditions and beliefs, and access to water and sanitation resources[76]. Unhygienic washing practices are common in rural India and among women and girls in lower socioeconomic groups, and may increase risk of urogenital infection[77–79]. However, the link between access to water and sanitation, menstrual hygiene management and urogenital infections has been poorly studied. Household surveys will capture self-reported urogenital infection, defined as at least one of the following symptoms: 1) abnormal vaginal discharge (unusual texture and color/more abundant than normal), 2) burning or itching in the genitalia, 3) burning or itching when urinating, or 4) genital sores[78].

Sanitation Insecurity

This study will assess the associations between sanitation access and sanitation insecurity with mental health among women. In previous research in Odisha, a contextually specific definition and measure for sanitation insecurity was developed, with associations between facets of sanitation insecurity and mental health independent of sanitation facility access[80]. This previously developed measure will be used to determine if levels of sanitation insecurity differ between intervention and control villages and how it may be associated with mental health outcomes, specifically well-being, anxiety, depression, and distress.

Fecal sludge management.

In sanitation systems where sewerage is not feasible, such as the household toilets constructed as part of the MANTRA intervention, safe management of fecal waste is necessary. Although there is growing emphasis on safe fecal sludge management (FSM), research has mainly focused on urban settings[81,82]. Preliminary research in Odisha suggests that fecal sludge management in this rural setting is a substantial challenge, and may impact household use of toilets. Household surveys and spot checks of toilets in intervention villages will be used to assess toilet use and fecal sludge management practices.

STATISTICAL ANALYSES

The effect of the intervention on infrastructure coverage, access, and use (aim 1), and the effect of the intervention on improving health (aim 3), will be analyzed using logistic, linear, or negative binomial multilevel regression depending on the outcome, to compare intervention versus control villages. The hierarchical structure of the data will be accounted for using random effects. Unadjusted models will be presented along with models adjusting for covariates. Potential mediation will be assessed using structured equation modeling or regression, as appropriate.

The impact of intervention on reducing environmental fecal contamination (aim 2), will be assessed through two methods. First, hierarchical logistic and negative binomial multilevel regression will be used to compare intervention versus control villages. Unadjusted models will be presented along with models adjusting for covariates. Second, a stochastic microbial risk framework will be used to assess differential fecal environmental contamination between intervention and control villages.

The cost and cost-effectiveness of the intervention (aim 4) will be assessed in two steps. Incremental intervention benefits will be ascertained by combining health benefit data, from analysis of health outcome data and established averted cost data, with other averted social opportunity costs. An incremental cost-effectiveness ratio, expressed in cost per disease-specific DALY, will be calculated by dividing the incremental intervention costs by the incremental intervention benefits.

The collective efficacy scale will be analyzed using factor and psychometric analyses to identify an appropriate factor solution. Once a CE factor solution is identified, multilevel regression will be used to assess the associations between CE and intervention effectiveness (aim 5).

DISCUSSION

This matched cohort study is one of the first to evaluate the effect of a rural combined household-level piped water and sanitation intervention, implemented at the community level, on a large scale. The matched design provides a rigorous means for estimating causal effects given that randomization was not feasible due to the several year implementation process[5]. By focusing on a completed intervention, it also avoids the risk presented by randomized controlled trials, where the intervention has little uptake, an especially important study challenge given interdependence of exposure and outcomes within communities, and a problem that has characterized previous trials of sanitation interventions in India[15,16].

A strength of this study is assessment of health impacts using the holistic WHO definition of health, including not just disease status, but also mental, social, and physical well-being[83]. Outcomes along the causal chain include standard, but more subjective measures, such as reported diarrheal diseases and respiratory infection, as well as more objective measures such as fecal environmental contamination, soil transmitted helminth infection, and anthropometry. Although there is risk of response bias for reported outcomes, it is unlikely to be differential by intervention status since the study team is not directly linked to Gram Vikas. Even though field workers may be aware of village intervention status, lab staff analyzing water, hand rinse, stool, and blood samples will be blinded. In addition, this study includes the more novel use of seroconversion for enteric pathogens, biomarkers of environmental enteric dysfunction, and measures of collective efficacy in an evaluation assessment. While there are limitations inherent to observational studies, the matched study design and multivariate modeling analysis plan reduces the potential for confounding. However, there is still the potential for residual unmeasured confounding.

Ethics and Dissemination. This study has been reviewed and approved by the Ethics Committee of the London School of Hygiene and Tropical Medicine, U.K (No. 9071) and Institute Ethics Committee of the Kalinga Institute of Medical Sciences of KIIT University, Bhubaneswar, India (KIMS/KIIT/IEC/053/2015). Efforts will be made to communicate the central findings and implications with study communities, the implementing organization and government officials in India. The results of this study will be submitted for publication in peer reviewed journals and presented at conferences. The data collected in the study will be publicly available, with personal identifiable data redacted, following the publication of the primary results within 24 months of the final data collection date.

Funding. This study is supported by a grant from the Bill & Melinda Gates Foundation to the London School of Hygiene & Tropical Medicine (OPP1008048) and to Emory University. (OOP1125067).

Competing Interests: None declared.

Contributions from authors: TC, HR, PR, BT, and HC contributed to study design. HR, LZ and BT developed laboratory protocols. HR, BT, GS, MD, SS, LZ, and BC developed data collection tools. All authors contributed to editing and revising the manuscript.

Figure 1. Description of the three phases of the Gram Vikas MANTRA water and sanitation intervention.

Figure 2. Study sites in Ganjam and Gajapati districts, Odisha, India with intervention villages in black and control villages in white. Inset shows location of districts in India.

Figure 3. Restriction, matching and exclusion process for selection of intervention and control villages (1), and timeline for study data collection (2).

REFERENCES

1 UNICEF, WHO. Progress and Sanitation and Drinking Water: 2015 Update and MDG Assessment. 2015.

2 Water and Sanitation Program. A Decade of the Total Sanitation Campaign: Rapid Assessment of Processes and Outcomes. 2010.

3 Coffey D, Gupta A, Hathi P, *et al.* Culture and the health transition: Understanding sanitation behavior in rural north India. 2015.

4 Clasen T, Pruss-Ustun A, Mathers CD, *et al.* Estimating the impact of unsafe water, sanitation and hygiene on the global burden of disease: evolving and alternative methods. *Trop Med Int Heal* 2014;**19**:884–93. doi:10.1111/tmi.12330

5 Arnold BF, Khush RS, Ramaswamy P, *et al.* Causal inference methods to study nonrandomized, preexisting development interventions. *Proc Natl Acad Sci U S A* 2010;**107**:22605–10. doi:10.1073/pnas.1008944107

6 Ercumen A, Arnold BF, Kumpel E, *et al.* Upgrading a Piped Water Supply from Intermittent to Continuous Delivery and Association with Waterborne Illness: A Matched Cohort Study in Urban India. *PLOS Med* 2015;**12**:e1001892. doi:10.1371/journal.pmed.1001892

7 Fewtrell L, Kaufmann RB, Kay D, *et al.* Water, sanitation, and hygiene interventions to reduce diarrhoea in less developed countries: a systematic review and meta-analysis. *Lancet Infect Dis* 2005;**5**:42–52. doi:10.1016/S1473-3099(04)01253-8

8 Engell RE, Lim SS. Does clean water matter? An updated meta-analysis of water supply and sanitation interventions and diarrhoeal diseases. *Lancet* 2013;**381**:S44. doi:10.1016/S0140-6736(13)61298-2

9 Clasen T, Bostoen K, Schmidt W, *et al.* Interventions to improve disposal of human excreta for preventing diarrhoea. *Cochrane database Syst Rev* Published Online First: 2010.http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD007180.pub2/pdf/standard (accessed 4 Jan2015).

10 Clasen T, Schmidt W-P, Rabie T, *et al.* Interventions to improve water quality for preventing diarrhoea: systematic review and meta-analysis. *BMJ* 2007;**334**:782. doi:10.1136/bmj.39118.489931.BE

11 Cairncross S, Hunt C, Boisson S, *et al.* Water, sanitation and hygiene for the prevention of diarrhoea. *Int J Epidemiol* 2010;**39 Suppl 1**:i193–205. doi:10.1093/ije/dyq035

12 Clasen TF, Alexander KT, Sinclair D, *et al.* Interventions to improve water quality for preventing diarrhoea. *Cochrane database Syst Rev* 2015;**10**:CD004794. doi:10.1002/14651858.CD004794.pub3

13 Routray P, Schmidt W-P, Boisson S, *et al.* Socio-cultural and behavioural factors constraining latrine adoption in rural coastal Odisha: an exploratory qualitative study. *BMC Public Health* 2015;**15**:880. doi:10.1186/s12889-015-2206-3

14 Barnard S, Routray P, Majorin F, *et al.* Impact of Indian Total Sanitation Campaign on Latrine Coverage and Use: A Cross-Sectional Study in Orissa Three Years following Programme Implementation. *PLoS One* 2013;**8**. doi:10.1371/journal.pone.0071438

15 Patil SR, Arnold BF, Salvatore AL, *et al.* The effect of India's total sanitation campaign on defecation behaviors and child health in rural Madhya Pradesh: a cluster randomized controlled trial. *PLoS Med* 2014;**11**:e1001709. doi:10.1371/journal.pmed.1001709

16 Clasen T, Boisson S, Routray P, *et al.* Effectiveness of a rural sanitation programme on diarrhoea, soil-transmitted helminth infection, and child malnutrition in Odisha, India: a

cluster-randomised trial. *Lancet Glob Heal* 2014;**2**:e645–53. doi:10.1016/S2214-109X(14)70307-9

17 Eisenberg JNS, Trostle J, Sorensen RJD, *et al.* Towards a Systems Approach to Enteric Pathogen Transmission: From Individual Independence to Community Interdependence. *Annu Rev Public Health* 2013;**33**:239–57. doi:10.1016/j.micinf.2011.07.011.Innate

18 Root GPM. Sanitation, Community Environments, and Childhoold Diarrhea in Rural Zimbabwe. *J Heal Popul Nutr* 2001;**19**:73–82.

19 Bateman OM, Smith S. A Comparison of the Health Effects of Water Supply and Sanitation in Urban and Rural Guatemala. 1999.

20 Huq A, Yunus M, Sohel SS, *et al.* Simple Sari Cloth Filtration of Water Is Sustainable and Continues to Protect Villagers from Cholera in Matlab, Bangladesh. *MBio* 2010;**1**:1–5. doi:10.1128/mBio.00034-10.Invited

21 UN Inter-agency Group for Child Mortality Estimation. Levels and trends in child mortality: Report 2015. 2015.

22 World Health Organization. Estimated total deaths by cause, sex and WHO Member State, 2008. Geneva, Switzerland: 2011.

23 Lin A, Arnold BF, Afreen S, *et al.* Household environmental conditions are associated with enteropathy and impaired growth in rural Bangladesh. *Am J Trop Med Hyg* 2013;**89**:130–7. doi:10.4269/ajtmh.12-0629

24 Korpe PS, Petri WA. Environmental enteropathy: critical implications of a poorly understood condition. *Trends Mol Med* 2012;**18**:328–36. doi:10.1016/j.molmed.2012.04.007

25 Guerrant RL, Oriá RB, Moore SR, *et al.* Malnutrition as an enteric infectious disease with

long-term effects on child development. 2008;**66**:487–505. doi:10.1111/j.1753-4887.2008.00082.x.Malnutrition

26 Crane RJ, Jones KDJ, Berkley JA. Environmental enteric dysfunction : An overview. 2015;**36**:76–87.

27 Humphrey JH. Child undernutrition, tropical enteropathy, toilets, and handwashing. *Lancet* 2009;**374**:1032–5. doi:10.1016/S0140-6736(09)60950-8

28 Mbuya MNN, Humphrey JH. Preventing environmental enteric dysfunction through improved water, sanitation and hygiene: An opportunity for stunting reduction in developing countries. *Matern Child Nutr* 2015;**12**:1–15. doi:10.1111/mcn.12220

29 Budge PJ, Griffin MR, Edwards KM, *et al.* Impact of home environment interventions on the risk of influenza-associated ARI in Andean children: observations from a prospective household-based cohort study. *PLoS One* 2014;**9**:e91247. doi:10.1371/journal.pone.0091247

30 Aiello AE, Coulborn RM, Perez V, *et al.* Effect of hand hygiene on infectious disease risk in the community setting: a meta-analysis. *Am J Public Health* 2008;**98**:1372–81. doi:10.2105/AJPH.2007.124610

31 Caruso BA, Sevilimedu V, Fung IC-H, *et al.* Gender disparities in water, sanitation, and global health. *Lancet (London, England)* 2015;**386**:650–1. doi:10.1016/S0140-6736(15)61497-0

32 Sahoo KC, Hulland KR, Caruso B a., *et al.* Sanitation-related psychosocial stress: a grounded theory study of women across the life-course in Odisha, India. *Soc Sci Med* 2015;**139**:80–9. doi:10.1016/j.socscimed.2015.06.031

33 Hulland KRS, Chase RP, Caruso BA, *et al.* Sanitation, stress, and life stage: A systematic

data collection study among women in Odisha, India. *PLoS One* 2015;**10**. doi:10.1371/journal.pone.0141883

34 World Health Organization/The United Nations Children's Fund (UNICEF). Ending Preventable Child Deaths from Pneumonia and Diarrhoea by 2025: The integrated Global Action Plan for Pneumonia and Diarrhoea (GAPPD). 2013.

35 Gram Vikas. Annual Report 2013-2014. 2014.

36 Duflo E, Greenstone M, Guiteras R, *et al.* Toilets Can Work : Short and Medium Run Health Impacts of Addressing Complementarities and Externalities in Water and Sanitation.

37 Rosa G, Clasen T. Estimating the scope of household water treatment in low- and medium-income countries. *Am J Trop Med Hyg* 2010;**82**:289–300. doi:10.4269/ajtmh.2010.09-0382

38 International Institute for Population Sciences (IIPS). District Level Household and Facility Survey (DLHS-3), 2007-08: India, Orissa. Mumbai: 2010.

39 Brady H, Caughey D, Dehejia R, *et al.* Genetic Matching for Estimating Causal Effects : 2012.

40 Sekhon JS. "Multivariate and Propensity Score Matching Software. *J Stat Softw* 2011;**42**.http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Multivariate+and+Propensity+Score+Matching#9 (accessed 30 May2014).

41 Austin PC. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Stat Med* 2009;**28**:3083–107. doi:10.1002/sim

42 Rubin DB. Using propensity scores to help design observational studies: Application to

the tobacco litigation. *Health Serv Outcomes Res Methodol* 2002;**2**:169–88.

43 Harder VS, Stuart EA, Anthony JC. Propensity score techniques and the assessment of measured covariate balance to test causal associations in psychological research. *Psychol Methods* 2010;**15**:997–1003. doi:10.1016/j.biotechadv.2011.08.021.Secreted

44 Faul F, Erdfelder E, Buchner A, *et al.* Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. *Behav Res Methods* 2009;**41**:1149–60. doi:10.3758/BRM.41.4.1149

45 Wolf J, Prüss-Ustün A, Cumming O, *et al.* Systematic review: Assessing the impact of drinking water and sanitation on diarrhoeal disease in low- and middle-income settings: Systematic review and meta-regression. *Trop Med Int Heal* 2014;**19**:928–42. doi:10.1111/tmi.12331

46 Carl Hartung YAWBALCTGB. Open Data Kit: Tools to Build Information Services for Developing Regions. http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.176.8017

47 Feikin DR, Audi a., Olack B, *et al.* Evaluation of the optimal recall period for disease symptoms in home-based morbidity surveillance in rural and urban Kenya. *Int J Epidemiol* 2010;**39**:450–8. doi:10.1093/ije/dyp374

48 Arnold BF, Galiani S, Ram PK, *et al.* Optimal Recall Period for Caregiver-reported Illness in Risk Factor and Intervention Studies: A Multicountry Study. *Am J Epidemiol* 2013;**177**:361–70. doi:10.1093/aje/kws281

49 World Health Organization. Integrated Management of Childhood Illness: Chart Booklet.

50 Cogill B. Anthropometric indicators measurement guide. Revised edition. *Washington, DC, Acad Educ Dev [AED], Food Nutr Tech Assist Proj* 2003;:92.http://www.developmentgateway.org/download/202582/anthro_2003.pdf

51 de Onis M, Onyango AW, Van den Broeck J, *et al.* Measurement and standardization protocols for anthropometry used in the construction of a new international growth reference. *Food Nutr Bull* 2004;**25**:S27–36.

52 Truant a. L, Elliott SH, Kelly MT, *et al.* Comparison of formalin-ethyl ether sedimentation, formalin-ethyl acetate sedimentation, and zinc sulfate flotation techniques for detection of intestinal parasites. *J Clin Microbiol* 1981;**13**:882–4.

53 Kosek M, Haque R, Lima A, *et al.* Fecal markers of intestinal inflammation and permeability associated with the subsequent acquisition of linear growth deficits in infants. *Am J Trop Med Hyg* 2013;**88**:390–6. doi:10.4269/ajtmh.2012.12-0549

54 Crump J a, Mendoza CE, Priest JW, *et al.* Comparing serologic response against enteric pathogens with reported diarrhea to assess the impact of improved household drinking water quality. *Am J Trop Med Hyg* 2007;**77**:136–41.

55 Steinberg EB, Mendoza CE, Glass R, *et al.* Prevalence of Infection with Waterbourne Pathogens: a Seroepidemiologic Study in Children 6-36 Months Old in San Juan Sacatepquez, Guatemala. 2004;**70**:83–8.

56 Brussow H, Sidoti J, Link H, *et al.* Age-specific prevalence of antibody to enterotoxigenic Escherichia coli in Ecuadorian and German children. *J Infect Dis* 1990;**162**:974–7.

57 Khanna B, Cutler A, Israel N, *et al.* Use caution with serologic testing for Helicobacter pylori infection in children. *J Infect Dis* 1998;**178**:460–5.

58 Lindkvist P, Asrat D, Nilsson I. Age at acquisition of Helicobacter pylori infection: comparison of a high and a low prevalence country. *Scand J Infect Dis* 1996;**28**:181–4.

59 Ungar B, Gilman R, Lanata C, *et al.* Seroepidemiology of Cryptosporidium infection in two Latin American populations. *J Infect Dis* 1988;**157**:551–6.

60 Vitral C, Yoshida C, Lemos E, *et al.* Age-specific prevalence of antibodies to hepatitis A in children and adolescents from Rio de Janeiro, Brazil, 1978 and 1995: relationship of prevalence to environmental factors. *Mem Inst Oswaldo Cruz* 1998;**93**:1–5.

61 Lammie PJ, Moss DM, Brook Goodhew E, *et al.* Development of a new platform for neglected tropical disease surveillance. *Int J Parasitol* 2012;**42**:797–800. doi:10.1016/j.ijpara.2012.07.002

62 Centers for Disease Control and Prevention. Laboratory Methods for the Diagnosis of Vibio cholerae. Atlanta, Georgia: 1999.

63 Centers for Disease Control and Prevention. Isolation and Identification of Shigella. Atlanta, Georgia:

64 United States Environmental Protection Agency. Coliforms—Total and E. coli, Membrane Filtration Method 10029. 1999.

65 Gronewold AD, Borsuk ME, Wolpert RL, *et al.* An Assessment of Fecal Indicator Bacteria-Based Water Quality Standards. *Environ Sci Technol* 2008;**42**:4676–82. doi:10.1021/es703144k

66 Gruber JS, Ercumen A, Colford JM. Coliform bacteria as indicators of diarrheal risk in household drinking water: systematic review and meta-analysis. *PLoS One* 2014;**9**:e107429. doi:10.1371/journal.pone.0107429

67 Levy K, Nelson KL, Hubbard A, *et al.* Rethinking indicators of microbial drinking water quality for health studies in tropical developing countries: case study in northern coastal Ecuador. *Am J Trop Med Hyg* 2012;**86**:499–507. doi:10.4269/ajtmh.2012.11-0263

68 Livio S, Strockbine N a, Panchalingam S, *et al.* Shigella Isolates From the Global Enteric Multicenter Study Inform Vaccine Development. *Clin Infect Dis* 2014;**59**.

doi:10.1093/cid/ciu468

69 Kotloff KL, Nataro JP, Blackwelder WC, *et al.* Burden and aetiology of diarrhoeal disease in infants and young children in developing countries (the Global Enteric Multicenter Study, GEMS): a prospective, case-control study. *Lancet* 2013;**382**:209–22. doi:10.1016/S0140-6736(13)60844-2

70 Kanungo S, Sah BK, Lopez a. L, *et al.* Cholera in India: An analysis of reports, 1997-2006. *Bull World Health Organ* 2010;**88**:185–91. doi:10.2471/BLT.09.073460

71 Cellini SR, Kee JE. Cost-Effectiveness and Cost-Benefit Analysis. In: Wholey JS, Hatry HP, Newcomer KE, eds. *Handbook of Practical Program Evaluation*. San Francisco: : John Wiley & Sons, Inc. 2010.

72 Bandura A. *Self-Efficacy: The Exercise of Control*. Worth Publishers 1997.

73 Malapit HJ, Sproule K, Kovarik C, *et al.* Measuring progress toward empowerment Women's Empowerment in Agriculture Index: Baseline Report. *IfpriOrg* 2014;:1–60.

74 Black RE, Victora CG, Walker SP, *et al.* Maternal and child undernutrition and overweight in low-income and middle-income countries. *Lancet* 2013;**382**:427–51. doi:10.1016/S0140-6736(13)60937-X

75 Smith LC, Ramakrishnan U, Ndiaye A, *et al.* The Importance of Women's Status for Child Nutrition in Developing Countries. 2003. http://www.ifpri.org/publication/importance-womens-status-child-nutrition-developing-countries

76 Sumpter C, Torondel B. A Systematic Review of the Health and Social Effects of Menstrual Hygiene Management. *PLoS One* 2013;**8**:e62004. doi:10.1371/journal.pone.0062004

77 Dasgupta A, Sarkar M. Menstrual Hygiene: How Hygienic is the Adolescent Girl? *Indian J Community Med* 2008;**33**:77–80. doi:10.4103/0970-0218.40872

78 Das P, Baker KK, Dutta A, *et al.* Menstrual Hygiene Practices, WASH Access and the Risk of Urogenital Infection in Women from Odisha, India. *PLoS One* 2015;**10**:e0130777. doi:10.1371/journal.pone.0130777

79 van Eijk AM, Sivakami M, Thakkar MB, *et al.* Menstrual hygiene management among adolescent girls in India: a systematic review and meta-analysis. *BMJ Open* 2016;**6**:e010290. doi:10.1136/bmjopen-2015-010290

80 Caruso B. *Sanitation Insecurity: Definition, Measurement, and Associations with Women's Mental Health in Rural Orissa, India.* 2015.http://pid.emory.edu/ark:/25593/rfhnt

81 Peal A, Evans B, Blackett I, *et al.* Fecal sludge management (FSM): analytical tools for assessing FSM in cities. *J Water, Sanit Hyg Dev* 2014;**4**:371. doi:10.2166/washdev.2014.139

82 Peal A, Evans B, Blackett I, *et al.* Fecal Sludge Management: a comparative analysis of 12 cities. *J Water, Sanit Hyg Dev* 2014;**4**:563–75.

83 World Health Organization. WHO definition of health. http://www.who.int/about/definition/en/print.html (accessed 26 Apr2016).

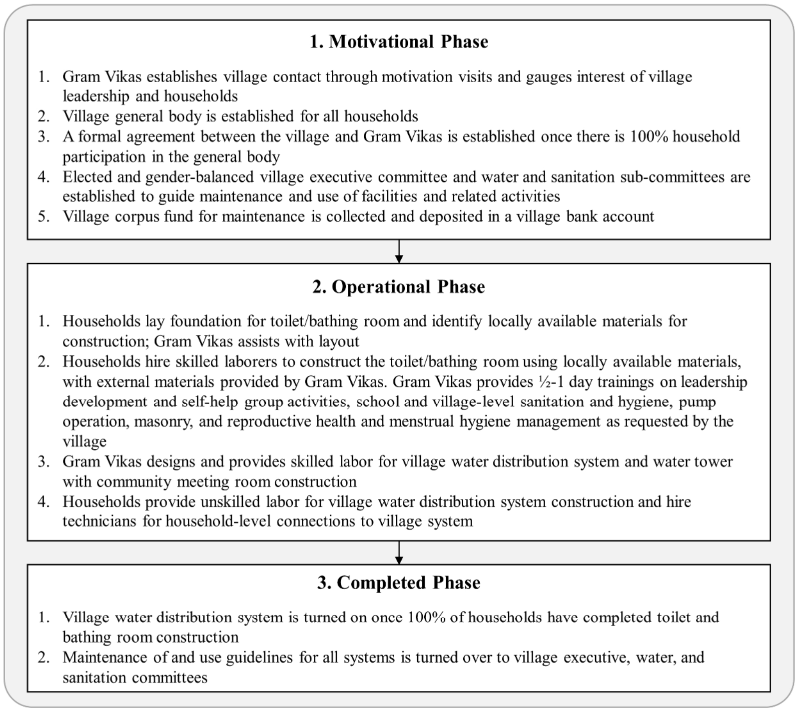

Figure 1. Description of the three phases of the Gram Vikas MANTRA water and sanitation intervention.
Figure 1
254x190mm (300 x 300 DPI)

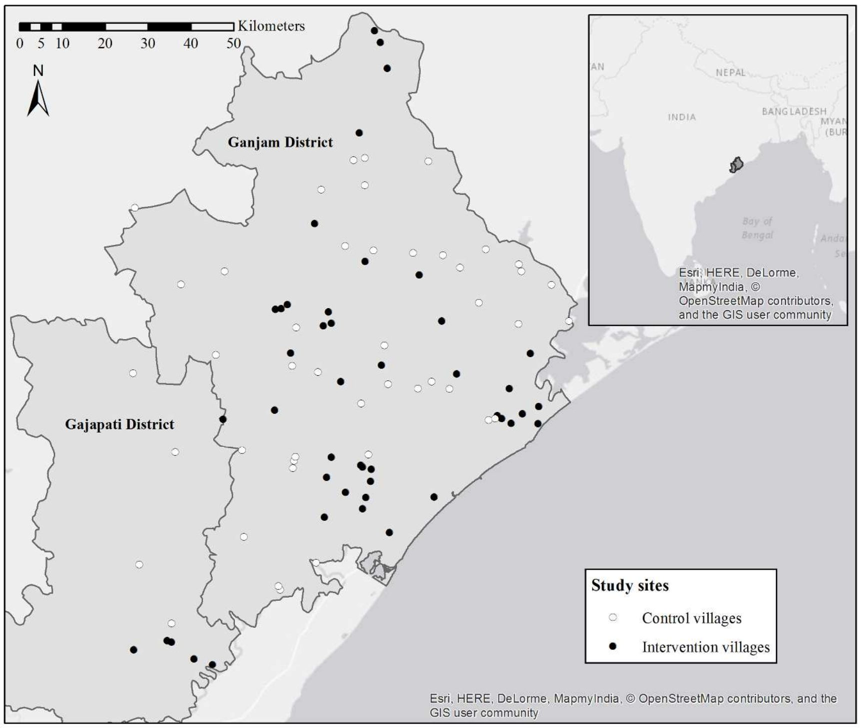

Figure 1. Study sites in Ganjam and Gajapati districts, Odisha, India with intervention villages in black and control villages in white. Inset shows location of districts in India.
Figure 1
215x279mm (300 x 300 DPI)

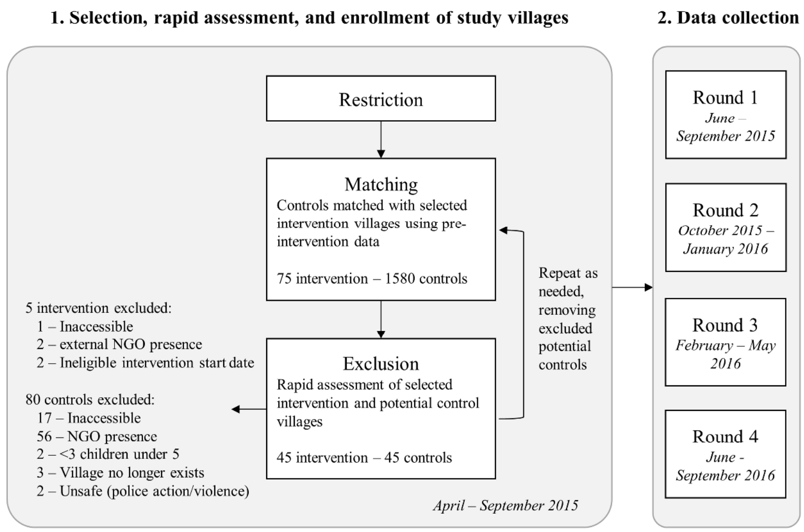

Figure 3. Restriction, matching and exclusion process for selection of intervention and control villages (1), and timeline for study data collection (2).
Figure 3
254x190mm (300 x 300 DPI)

STROBE Statement—checklist of items that should be included in reports of observational studies

	Item No.	Recommendation	Page No.	Relevant text from manuscript
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	1	
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	2	
Introduction				
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	3-5	
Objectives	3	State specific objectives, including any prespecified hypotheses	7	
Methods				
Study design	4	Present key elements of study design early in the paper	8	
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	8	
Participants	6	(*a*) *Cohort study*—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study*—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study*—Give the eligibility criteria, and the sources and methods of selection of participants	8-11	
		(*b*) *Cohort study*—For matched studies, give matching criteria and number of exposed and unexposed *Case-control study*—For matched studies, give matching criteria and the number of controls per case	8-11	
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	12-19	
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	12-19	
Bias	9	Describe any efforts to address potential sources of bias	NA	
Study size	10	Explain how the study size was arrived at	12	

Continued on next page

Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	NA
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	NA (overview 19-20)
		(*b*) Describe any methods used to examine subgroups and interactions	NA
		(*c*) Explain how missing data were addressed	NA
		(*d*) *Cohort study*—If applicable, explain how loss to follow-up was addressed *Case-control study*—If applicable, explain how matching of cases and controls was addressed *Cross-sectional study*—If applicable, describe analytical methods taking account of sampling strategy	NA
		(*e*) Describe any sensitivity analyses	NA
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	NA
		(b) Give reasons for non-participation at each stage	NA
		(c) Consider use of a flow diagram	NA
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	NA
		(b) Indicate number of participants with missing data for each variable of interest	NA
		(c) *Cohort study*—Summarise follow-up time (eg, average and total amount)	NA
Outcome data	15*	*Cohort study*—Report numbers of outcome events or summary measures over time	NA
		Case-control study—Report numbers in each exposure category, or summary measures of exposure	
		Cross-sectional study—Report numbers of outcome events or summary measures	
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	NA
		(*b*) Report category boundaries when continuous variables were categorized	NA
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	NA

Continued on next page

Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	NA
Discussion			
Key results	18	Summarise key results with reference to study objectives	NA
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	20-21
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	NA
Generalisability	21	Discuss the generalisability (external validity) of the study results	NA
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	22

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.

BMJ Open

Design and rationale of a matched cohort study to assess the effectiveness of a combined household-level piped water and sanitation intervention in rural Odisha, India

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2016-012719.R1
Article Type:	Protocol
Date Submitted by the Author:	24-Aug-2016
Complete List of Authors:	Reese, Heather; Emory University School of Public Health, Department of Environmental Health Routray, Parimita; London School of Hygiene and Tropical Medicine Torondel, Belen; London School of Hygiene and Tropical Medicine Sclar, Gloria; Emory University School of Public Health, Department of Environmental Health Delea, Maryann; Emory University School of Public Health; London School of Hygiene and Tropical Medicine Sinharoy, Sheela; Emory University Laney Graduate School, Nutrition and Health Sciences Program Zambrano, Laura; Emory University School of Public Health, Department of Environmental Health Caruso, Bethany; Emory University School of Public Health, Department of Environmental Health Mishra, Samir; KIIT University, School of Biotechnology Chang, Howard; Emory University, Department of Biostatistics and Bioinformatics Clasen, Thomas; London School of Hygiene & Tropical Medicine; Emory University School of Public Health, Department of Environmental Health
Primary Subject Heading:	Public health
Secondary Subject Heading:	Epidemiology
Keywords:	sanitation, piped water supply, diarrheal diseases

Design and rationale of a matched cohort study to assess the effectiveness of a combined household-level piped water and sanitation intervention in rural Odisha, India

Heather Reese,[1] Parimita Routray,[2] Belen Torondel,[2] Gloria Sclar,[1] Maryann G. Delea,[1,2] Sheela S. Sinharoy,[3] Laura Zambrano,[1] Bethany Caruso,[1] Samir R. Mishra,[4] Howard H. Chang,[5] Thomas Clasen[1,2]

[1] Department of Environmental Health, Rollins School of Public Health, Emory University, Atlanta, Georgia, United States of America

[2] London School of Hygiene and Tropical Medicine, London, United Kingdom

[3] Nutrition and Health Sciences Program, Laney Graduate School, Emory University, Atlanta, Georgia, United States of America

[4] School of Biotechnology, KIIT University, Bhubaneswar, Odisha, India

[5] Department of Biostatistics and Bioinformatics, Rollins School of Public Health, Emory University, Atlanta, Georgia, United States of America

Contact information for corresponding author, Heather Reese:

Department of Environmental Health, Rollins School of Public Health, Emory University

1518 Clifton Rd, Atlanta, GA 30322

Email: heather.e.reese@emory.edu, Phone: (619) 379-8963

Keywords: rural, piped water, sanitation

Word count: 5442

ABSTRACT

Introduction*:* Government efforts to address massive shortfalls in rural water and sanitation in India have centered on construction of community water sources and toilets for selected households. However, deficiencies with water quality and quantity at the household level, and community coverage and actual use of toilets has led Gram Vikas, a local NGO in Odisha, India, to develop an approach that provides household-level piped water connections contingent on full community-level toilet coverage.

Methods: This matched cohort study was designed to assess the effectiveness of a combined piped water and sanitation intervention. Households with children under five years in 45 randomly selected intervention villages and 45 matched control villages will be followed over 17 months. The primary outcome is prevalence of diarrheal diseases; secondary health outcomes include soil-transmitted helminth infection, nutritional status, seroconversion to enteric pathogens, urogenital infections, and environmental enteric dysfunction. In addition, intervention effects on sanitation and water coverage, access and use, environmental fecal contamination, women's empowerment, as well as collective efficacy, and intervention cost and cost-effectiveness will be assessed.

Ethics and dissemination: The study protocol has been reviewed and approved by the ethics boards of the London School of Hygiene and Tropical Medicine, U.K. and KIIT University, Bhubaneswar, India. Findings will be disseminated via peer-reviewed literature and presentation to stakeholders, government officials, implementers and researchers.

Trial registration identifier: NCT02441699

STRENGTHS AND LIMITATIONS OF THIS STUDY

- The study assesses a combined household-level piped water and sanitation intervention that requires complete community-level compliance.
- The intervention was not randomly allocated; but, controls are selected through a restriction process to limit possible partial exposure to the intervention through spillover, and matched to intervention villages using pre-intervention data.
- The study uses a holistic definition of health to assess intervention impacts on physical, mental and social well-being, including more novel outcomes such as seroconversion to enteric pathogens, environmental enteric dysfunction, and sanitation insecurity. It also assesses intervention coverage, cost-effectiveness, and collective efficacy.
- The time lapse between intervention completion and the beginning of the evaluation process prevents baseline comparison or assessment of immediate intervention impacts. However, it allows for a biologically plausible length of time for die-off of even the most persistent pathogens in the environment, and provides time for children to have be born into this environment.

INTRODUCTION

Of the one billion people practicing open defecation worldwide, over half live in India[1]. While international and national pressure on improving sanitation conditions in India has led to over 350 thousand people gaining access to improved toilets since 1990, it has barely kept up with population growth[1,2]. Recent studies show that even in areas with access to household-level improved sanitation, use of these toilets is low[3–5]. This may be due in part to a mismatch between the culturally acceptable pour-flush toilets and the level of water access. Coverage of

improved water sources, usually community-level pumps or taps, is relatively high even in rural areas in India, but it may not be sufficient for flushing purposes on top of other daily water needs[1,6].

Although the effectiveness of water, sanitation and hygiene (WASH) interventions vary, meta-analyses have found that individual or combined WASH interventions decrease diarrheal disease prevalence by up to 48%[7–11]. While combined interventions would be expected to have a greater influence on multiple exposure pathways and thus a greater combined impact on health, there is limited evidence of additive benefits[12]. This may be due to poor uptake, inconsistent use, or an incomplete understanding of relevant pathways[8–10]. In India, combining water and sanitation interventions may be more critical than just interrupting multiple transmission pathways for enteric infection; evidence suggests that household-level water access is integral to the use of improved sanitation in this context[13].

While the intent of improved sanitation facilities is to separate human feces from human contact, most of the focus is on constructing household toilets to increase improved sanitation coverage—the primary metric used in monitoring progress toward international targets. However, studies in India have further shown that toilet construction does not translate into toilet use in this context[5,14–16]. Moreover, with the interdependence between members of households and households within communities, safe water and sanitation is a community-level issue. There is growing emphasis on assessing health risk from poor water and sanitation conditions not simply due to individual or even household-level risk factors, but also from conditions in the community environment[17]. There is evidence that even households without toilets, and households which do not filter drinking water, show decreased health risk if they live in communities with high levels of coverage and use[18–20].

Moreover, the effectiveness of community interventions may be higher in communities with positive perceptions of their collective ability to come together to improve their conditions. Collective efficacy, a latent construct comprised of the structural and cognitive components that facilitate a community's shared belief in its ability to come together and execute actions related to a common goal, may explain some variance in intervention effectiveness across communities receiving WASH interventions[21].

A main risk of poor WASH conditions is enteric infection, caused by a diverse array of bacteria, viruses, protozoa, and parasites, including soil-transmitted helminths. These infections may cause diarrhea, the second leading cause of mortality for children under five years worldwide and in India, a leading cause of mortality regardless of age[22,23]. There is also growing evidence that asymptomatic enteric infections may pose a similar risk, with repeat enteric infections contributing to chronic malnutrition, environmental enteric dysfunction, poor cognitive outcomes, and poor vaccine uptake[24–29]. Poor WASH conditions are also linked to increased risk of respiratory infection, the leading cause of mortality for children under five years worldwide[22,30,31]. Poor water and sanitation access can also affect the social, physical and mental well-being of women, acting through pathways ranging from unsafe menstrual hygiene management practices and increased risk of violence[32–34].

Description of the intervention

Over the past decades there has been a global commitment to determine water and sanitation interventions with demonstrated effectiveness, not just efficacy[35]. Gram Vikas, a non-governmental organization (NGO) based in Odisha, India (http://www.gramvikas.org/), has responded by implementing its MANTRA (Movement and Action Network for Transformation of Rural Areas) water and sanitation program in more than 1000 villages since 2002[36]. This approach includes both household-level piped water connections, and community-level mobilization for culturally appropriate household toilets. A previous interrupted time series analysis of the MANTRA intervention reported it to be protective against diarrheal diseases[37]. However, in addition to limitations of design, this study relied on outcome data collected and reported by Gram Vikas, the intervention implementer, and did not assess intervention coverage or impacts on environmental fecal contamination.

The MANTRA water and sanitation intervention is rolled out in a three-phase process over an average three to five years (Figure 1). During the first, or Motivational, phase, representatives of Gram Vikas visit the identified village several times to assess village interest and progress towards a set of Gram Vikas requirements, including: 1) the commitment of every household to participate, 2) creation of a village corpus fund from contributions from every household, and 3) development of village guidelines for maintenance and use of facilities.

Once this set of requirements is achieved, the village progresses into the second, or Operational, phase of the intervention. Each household constructs a pour-flush toilet with two soak-pits and a separate bathing room. The households hire a local, skilled mason and provide their own unskilled labor and locally available materials to complete the superstructure. Gram Vikas provides external materials such as PVC pipes and porcelain pans. At the same time, a water tank, community meeting space, and piped water distribution system connected to every household, with taps in the toilet and bathing rooms and a separate tap in the kitchen, is constructed through a similar collaborative process.

All households must construct a toilet and bathing room for the village to progress into the final, or Completed, phase of the intervention, in which the water system is turned on.

Notably, this three-phase process only allows each household access to piped water once every household in the village has a toilet and bathing room. This model contrasts with most previous water and sanitation interventions, including those implemented under India's Total Sanitation Campaign and other government programs, which do not require community-level sanitation compliance and do not provide a piped water supply at the household level[38].

Study aims

The primary objective of this study is to evaluate the effectiveness of the combined household-level water supply and sanitation intervention, as implemented by Gram Vikas in Odisha, India. Toward that objective, this study aims to:

1) Assess the effectiveness of the intervention in improving water and sanitation infrastructure coverage, access, and use, and to assess fecal sludge management practices in intervention communities.
2) Assess the effectiveness of the intervention in reducing environmental fecal contamination.
3) Assess the effectiveness of the intervention in improving health. This includes reported diarrheal disease in children under 5 years (primary outcome), acute respiratory infection, infection with soil-transmitted helminthes, nutritional status, environmental enteric dysfunction, seroconversion for selected enteric pathogens, and urogenital diseases associated with menstrual hygiene management practices. Mental and social well-being will be explored through assessment of sanitation insecurity and women's empowerment.
4) Assess the cost and cost-effectiveness of the intervention.

5) Develop and assess a theoretically-grounded, empirically informed collective efficacy scale; and determine the effect of collective efficacy on intervention effectiveness.

METHODS

Setting

The study is located in Ganjam and Gajapati districts located in eastern Odisha, India (Figure 2). These two contiguous districts were a single district until 1992. Over 44% of the population in these districts is recognized by the Government of India as being below the poverty line (BPL)[39]. As of 2008, a majority of households in both districts had access to an improved, likely community-level, drinking water source, with over 23% of households in Ganjam having access to any sanitation facility, compared to only 8% of households in Gajapati[39]. The area is primarily rural and agrarian, and the climate is characterized by a monsoon season from June to September, with an average rainfall of ~1400 mms/year.

Study design

This study uses a matched cohort design to assess the effectiveness of a completed intervention with data collected across four study rounds from June 2015 to October 2016 (Figure 3). Data are collected continuously across all study rounds for diarrhea, acute respiratory infection, nutritional status, and stored and source water outcomes to assess seasonality. Data are collected in rounds 2 and 4 for environmental enteric dysfunction, seroconversion, and hand-rinses, and cross-sectionally in one or more rounds for the remaining outcomes. As described below, control villages were matched to randomly selected intervention villages through a multi-step restriction, genetic matching, and exclusion process using the following eligibility criteria.

Eligibility criteria for villages

1. Restriction. Intervention villages were randomly selected from a list of Gram Vikas villages in Ganjam and Gajapati districts provided by the NGO, after restriction to villages with a Motivation phase start date between 2002-2006 and a Construction phase start date no earlier than 2003. Since the intervention process takes on average three to five years, the criteria for the Motivation start date helped to identify those villages with ongoing interventions at the same time. In addition, this allowed the use of the Government of India Census 2001 and the Below Poverty Line (BPL) Survey 2002 data to characterize baseline characteristics used in the matching process in both intervention and control villages.

Eligible control villages include all villages without a Gram Vikas intervention within the study districts which: 1) are not within the same Gram Panchayat (a political subdivision with some administrative responsibility for water and sanitation comprised of several villages) as a Gram Vikas village, or bordering a Gram Vikas village, and 2) had not received a Motivation visit from the Gram Vikas NGO. These criteria serve to limit the possibility of previous partial exposure to the intervention through spillover from adjacent villages or direct contact with the NGO. These criteria also increase the strength of the comparison provided by the control villages, i.e. it increases the likelihood that if they had received a motivation visit from Gram Vikas, the control villages would have been equally as likely as the intervention villages to demand the intervention.

In addition, to be eligible for inclusion both intervention and control villages must: 1) appear in the Government of India Census 2001 and the BPL Survey 2002, 2) have a population of at least 20 households, and 3) be within approximately three hours travel from the study office in Brahmapur, Ganjam District. This last criterion is due to logistical constraints.

2. Matching. After restriction, genetic matching was used to match potential control villages to the randomly selected intervention villages without replacement[5,40,41]. Villages were exact matched on district to limit any political or large scale geographic variation between district populations, and were also matched on pre-intervention demographic, socioeconomic, sanitation, and water access characteristics listed in Table 1[5]. These village-level matching variables were selected due to their theorized association with the primary outcome, diarrheal diseases, as well as data availability.

3. Exclusion. The field team visited matched potential control villages and intervention villages to assess suitability for the study through a rapid assessment interview with village leadership and to ensure accessibility. Villages were excluded if they are not within three hours travel of the field office in Brahmapur, had sustained major infrastructure damage due to a natural disaster, or if there was a current or planned sanitation or water intervention by an organization external to the village in the next 12 months as determined through the rapid assessment interview with village leadership. In addition, villages were excluded if there were fewer than three households with children under five years old. As villages were removed from the pool of prospective control villages, the matching process was repeated for all intervention villages and remaining eligible control villages, and balance measures were assessed. The matching and exclusion processes were repeated as necessary.

After the iterative matching and exclusion process was complete, covariate balance was assessed for all matching variables for the final set of intervention and control villages through examination of balance measures[42–44]. Matching resulted in an improvement in balance as assessed through comparison of several measures including q-q plots, Kolmogorov-Smirnov bootstrap p-values, and standardized differences. After matching, there were no significant differences between intervention and control groups (Table 1).

Eligibility criteria for households

Households within selected intervention and control villages will be eligible if they have at least one child under 5 years old at time of enrollment, verified with birth or immunization card, and expect to reside in the village for the duration of the study. If there are more than 40 eligible households within a village, 40 will be randomly selected to be enrolled. Informed consent will be obtained from the male and/or female head of the selected households. All children under five years within each enrolled household are eligible and will not age-out over the course of the

Table 1. Pre-intervention characteristics used in matching, and balance diagnostics before and after matching and exclusion process.

Variable	**Intervention** (n=45)	**Control (all eligible)** (n=1580)	Std Diff (all eligible)	**Control (study)** (n=45)	Std Diff (study)
Number of households	157.9	215.5	0.37	148.1	0.06
Population under 6 years (%)	16.2	16.9	0.19	16.3	0.02
Household income score ($\bar{x}$)	2.9	3.1**	0.26	2.9	0.01
Household goods owned ($\bar{x}$)	1.1	1.2*	0.27	1.1	0.02
Pucca house (%)	59.2	61.6	0.09	60.5	0.05
≥2 meals a day (%)	57.7	63.7	0.19	57.8	0.01
Scheduled caste (%)	11.5	18.7**	0.46	11.8	0.01

Scheduled tribe (%)	33.4	19.1*	0.31	29.8	0.08
Female literacy (%)	30.9	29.8	0.07	30.9	0.00
Open defecation (%)	95.6	95.2*	0.04	95.8	0.01
Improved drinking water source[ǂ] (%)	38.6	42.5	0.10	37.2	0.02
Water source <500m and 50m elevation (%)	81.5	72.2	0.31	81.7	0.01

All eligible: all villages that are eligible for the matching process after restriction
Std Diff (absolute standardized difference): a value greater than 0.1 is considered meaningful imbalance [42]
ǂ Ganjam villages only; no data available for Gajapati villages
Kolmogorov-Smirnov bootstrap p-values: * <0.05 ** <0.01

study. Households with newborn children will be enrolled as they become eligible on an ongoing basis throughout the study, in villages with fewer than 40 enrolled households.

Sample Size

Sample size was determined through a simulation estimating the log odds of diarrheal disease (the primary outcome) through a multilevel random effects model and parameterized with data from a previous study in a neighboring district in Odisha[16]. Sample size estimates were also checked with G*Power[45]. The simulation assumes a longitudinal 7-day period prevalence for diarrhea of 8.8% in children under five years, a heterogeneity variance between villages of 0.07, a heterogeneity variance between households of 0.57, and four study rounds[16]. An effect size of 0.20 was selected for public health importance and based on estimates of effect from systematic reviews of water and sanitation studies[46]. Assuming at least 80% power, 0.05 significance level, 10% for loss to follow up, and at least one child per household, we estimate a sample size of 45 villages per study arm and 26 households per village. This estimate was the most conservative compared to sample size estimates for secondary outcomes, and was therefore used for the broader study population.

Outcome Measurement

Outcomes, and individual-, household-, and community-level risk factors, will be measured through surveys, interviews, or through the collection and analysis of environmental, stool or dried blood spot samples. All survey questions will be translated into the primary local language, Odia, and back-translated to confirm wording. Household surveys will be verbally administered by trained field workers to the mother or primary caregiver of the youngest child under five in each household, unless otherwise specified below. Community surveys will be verbally administered to the *sarpanch* (village head) or other member of village leadership. Survey data will be collected on mobile phones using Open Data Kit[47]. GPS coordinates for households, water sources and other relevant sites will be collected using Garmin eTrex 10 or 20 devices (Garmin Ltd., Olathe, KS, USA).

Coverage, access and use of sanitation, water and hygiene infrastructure

Coverage, access and use of WASH infrastructure will be assessed in all four rounds. Presence of and access to toilets, water sources and hand-washing stations will be assessed through standard questions from the Demographic and Health Surveys and confirmed through spot observations. Spot observations of household toilets and hand-washing stations will be further used to assess indicators of functionality, maintenance, recent use. Reported water and sanitation practices, including child feces disposal practices, will be captured through household survey questions.

Diarrheal Diseases

The primary outcome for this study is prevalence of diarrheal diseases, recorded as both daily point prevalence over the previous three days and seven-day period prevalence, for all household members in all four rounds. Although self-reported diarrhea is a subjective outcome with a well-established risk of bias, three-day recall reduces recall bias[48,49]. Diarrheal disease will be measured using the World Health Organization (WHO) definition of three or more loose stools in a 24-hour period, with or without the presence of blood. Field workers will use a simple calendar as a visual aid to help respondents with recall. Each household member will be asked to recall his or her own disease status and the mother or primary caregiver will be asked to report disease for children.

Respiratory infection

Prevalence of respiratory infections will be recorded as both daily point prevalence over the previous three days and seven-day period prevalence for all household members in all four rounds. Respiratory infection is defined as the presence of cough and/or shortness of breath/difficulty breathing according to WHO's Integrated Management of Childhood Illness (IMCI)[50]. The full IMCI case definition for acute lower respiratory infection also includes measurement of respiratory rate and observation of chest indrawing, stridor and other danger signs; these criteria were excluded from our definition as there is concern about the technical support required to produce consistent and accurate data within this context[50]. Our definition will provide a broad assessment of respiratory illness burden. Each household member will be asked to recall his or her own disease status and the mother or primary caregiver will be asked to report disease for children.

Nutritional Status

Anthropometric data will be collected for children under age five in all four rounds using standard methods as established by WHO[51,52]. Field workers will be trained and standardized in line with WHO protocols to reduce measurement error [52]. Weight will be measured for all children under five years of age using Seca 385 digital scales, with 20g increment for weight below 20kg and a 50g increment for weight between 20 and 50kg. Recumbent length will be measured for children under two years of age using Seca 417 measuring boards with 1mm increment. Standing height will be measured for children two to five years of age using Seca 213 portable stadiometers with 1mm increment. Height and weight will be used to calculate height-for-age z-scores (HAZ) and weight-for-height z-scores (WHZ) based on WHO reference standards. A random subset of 10% of households will receive back check visits each day to repeat height/length measurements to ensure inter-observer reliability.

Soil-transmitted helminth infection

Stool samples will be collected in rounds 2 and 4 from all household members in a randomly selected subset of 500 households, and used to assess the presence and intensity of soil-transmitted helminth (STH) infection. Formalin ether concentration and microscopy will be used to quantify worms and ova for hookworms, *Ascaris lumbricoides*, and *Tricuris trichura*[53]. Quality assurance will include independent duplicate assessment of all positive and 10% of negative samples. After stool collection, each participant will be offered a single dose of Albendazole, a broad-spectrum antihelmenthic drug recommended by the Ministry of Health and Family Welfare, Government of India. Stools collected in round 2 will allow for comparison of

STH infection prevalence between intervention and control villages, while the stool samples collected approximately 8 months later in round 4 will provide a measure of re-infection rate.

Environmental enteric dysfunction

Stools from a randomly selected subset of 200 children under two years old, collected in rounds 2 and 4, will be used to assess environmental enteric dysfunction (EED) through quantification of biomarkers of intestinal inflammation and permeability. Fecal myeloperoxidase (MPO), alpha-1-antitrypsin (AAT), and neopterin (NEO), markers for neutrophil activity, intestinal permeability and TH1 immune activation, respectively, were selected for this study based on evidence of association with EED, subsequent linear growth deficits, and household environmental fecal contamination[24,25,54].

Seroconversion for enteric pathogens

Serological assays that assess antibody production against various enteric pathogens can provide an objective measure of exposure to enteric infections[55]. Enrolling children aged 6 to 18 months will reduce the potential for interference from maternally acquired antibodies and permit analysis of seroconversion data in a critical window for young children who experience higher diarrheal disease morbidity and mortality before two years of age[56–61]. Children who are 6 to 12 months during round 2 will have capillary blood drawn by fingerstick or heelstick, as appropriate, and will be visited again during round 4 for a second capillary blood sample. All blood samples will be preserved on TropBio (Sydney, Australia) filter discs and stored within 7 days of collection at -20°C. Seroconversion against markers for norovirus, *Giardia intestinalis*, *Cryptosporidium parvum*, *Entamoeba histolytica*, enterotoxigenic *E. coli* heat-labile enterotoxin

(ETEC-LT), *Salmonella* spp., *Campylobacter jejuni*, *Vibrio cholera*, and *Toxoplasma* spp. will be assessed using multiplex immunoassay technology on the Luminex xMAP platform[62].

Environmental fecal contamination

Field workers will collect samples of household stored drinking water and source water from a random subset of 500 households in all four rounds, and child hand rinses in rounds 2 and 4. All water and hand rinse samples will be stored on ice during transport and analyzed within 6 hours of collection using membrane filtration. Three assays will be used: 1) plating on m-Coli Blue 24 (Millipore, Billerica, MA) for *E.coli* according to EPA Method 10029, 2) alkaline peptone water enrichment prior to plating on thiosulfate citrate bile salts sucrose agar and slide agglutination serotyping for *V. cholerae*, and 3) plating on xylose lysine desoxycholate agar, and slide agglutination serotyping for *Shigella* spp.[63–65]. Source and stored water samples will be assayed for *E. coli*, *Vibrio cholerae* and *Shigella* spp., and hand rinse samples will be assayed for *E. coli* and *Shigella* spp. *E. coli* was selected as a standard non-human specific indicator of fecal contamination, though the limitations of this indicator are well-established[66–68]. In order to better characterize human fecal contamination of the household environment, *Vibrio cholerae* and *Shigella* spp. were selected based on prevalence in southern Asia, evidence of public health importance, and field laboratory limitations[69–71].

Cost and cost-effectiveness

Costs and potential cost savings (i.e., averted costs) associated with the intervention will be assessed through an economic costing approach that recognizes and quantifies costs and benefits from a societal perspective[72]. Data on program and point-of-delivery inputs will be collected at household, community, and implementer levels in round 3. Field workers will administer community surveys to a village leader, and household surveys to the household decision-maker for toilet installation, in 20 randomly selected households in half of the intervention and control villages. Given cost-effectiveness analyses require the effect of the intervention to be measured against a counterfactual, and the intervention of interest is a community-based intervention, cost and effectiveness measures will be summarized at the village level [73]. Surveys will collect data on household- and community-level inputs related to materials and labor required to construct household toilets and wash rooms, the community water tank and distribution system, and household water connections; longer-term water supply and toilet maintenance costs; and financing required for this infrastructure as well as perceived benefits, including averted social opportunity costs. Implementer inputs from Gram Vikas will be collected through an enumeration exercise, interviews, and examination of the implementer's financial records.

Collective efficacy

Collective efficacy (CE) is a latent construct comprised of the structural and cognitive components that facilitate a community's shared belief in its ability to come together and execute actions related to a common goal[21]. A review of the literature and established conceptual frameworks will be performed to define the CE construct. A sequential exploratory mixed qualitative and quantitative design will be used to develop and refine a scale to measure CE and test hypotheses. Field workers will administer the refined, multi-item, Likert-type CE scale to one randomly selected household member aged 18 years or older in each household in round 3.

Women's empowerment

Four dimensions of women's empowerment will be measured in rounds 3 and 4: group participation, leadership, decision-making and freedom of movement. Group participation and leadership will be measured using modules from the Women's Empowerment in Agriculture Index (WEAI), which has been tested in South Asia[74]. Decision-making will be measured using questions from the women's status module of the Demographic and Health Surveys. Freedom of movement will be measured using questions from the project-level Women's Empowerment in Agriculture Index (pro-WEAI). These measures will be collected for the primary female caregiver of the youngest child under 5, and were selected based on the importance of women's empowerment for child nutrition[75,76]. Women's empowerment is conceptualized as both an outcome and a potential mediator along the pathway between the Gram Vikas intervention and child health outcomes.

Menstrual hygiene management

Menstrual hygiene management practices vary worldwide and depend on personal preference, socioeconomic status, local traditions and beliefs, and access to water and sanitation resources[77]. Unhygienic washing practices are common in rural India and among women and girls in lower socioeconomic groups, and may increase risk of urogenital infection[78–80]. However, the link between access to water and sanitation, menstrual hygiene management and urogenital infections has been poorly studied. Household surveys will be administered in round 4 to a randomly selected woman aged 18 or older, in a subset of 800 households, and will capture self-reported urogenital infection, defined as at least one of the following symptoms: 1) abnormal vaginal discharge (unusual texture and color/more abundant than normal), 2) burning or itching in the genitalia, 3) burning or itching when urinating, or 4) genital sores[79].

Sanitation Insecurity

This study will assess the associations between sanitation access and sanitation insecurity with mental health among women. In previous research in Odisha, a contextually specific definition and measure for sanitation insecurity was developed, with associations between facets of sanitation insecurity and mental health independent of sanitation facility access[81]. This previously developed measure will be used to determine if levels of sanitation insecurity differ between intervention and control villages and how it may be associated with mental health outcomes, specifically well-being, anxiety, depression, and distress. Household surveys will be administered in round 4 to a randomly selected woman aged 18 or older, in a random subset of 800 households.

Fecal sludge management.

In sanitation systems where sewerage is not feasible, such as the household toilets constructed as part of the MANTRA intervention, safe management of fecal waste is necessary. Although there is growing emphasis on safe fecal sludge management (FSM), research has mainly focused on urban settings[82,83]. Preliminary research in Odisha suggests that fecal sludge management in this rural setting is a substantial challenge, and may impact household use of toilets. In round 3, household surveys and spot checks of toilets in intervention villages will be used to assess toilet use and fecal sludge management practices.

STATISTICAL ANALYSES

The effect of the intervention on infrastructure coverage, access, and use (aim 1), and the effect of the intervention on improving health (aim 3), will be analyzed using logistic, linear, log binomial, or negative binomial multilevel regression depending on the outcome, to compare intervention versus control villages. Prevalence of fecal sludge management practices in intervention communities will be assessed using multilevel regression (aim 1). For all models, the hierarchical structure of the data will be accounted for using random effects. Estimation of relative risks through Poisson regression or binary regression methods for binary outcomes will be considered to ensure robustness of results. Mediation of the potential association between intervention and nutritional status outcomes by women's empowerment will be assessed using multilevel structural equation modeling, and statistical approaches to reduce bias will be explored as needed[84].

The impact of intervention on reducing environmental fecal contamination (aim 2), will be assessed through two methods. First, logistic and negative binomial multilevel regression to estimate intervention effects on the relative scale will be used to compare intervention versus control villages. Estimation of relative risks through Poisson regression or binary regression methods for binary outcomes will be considered to ensure robustness of results. Second, a stochastic microbial risk framework will be used to assess differential fecal environmental contamination between intervention and control villages.

The cost and cost-effectiveness of the intervention (aim 4) will be assessed in two steps. Incremental intervention benefits will be ascertained by combining health benefit data, from analysis of health outcome data and established averted cost data, with other averted social opportunity costs. An incremental cost-effectiveness ratio, expressed in cost per disease-specific

DALY, will be calculated by dividing the incremental intervention costs by the incremental intervention benefits.

The collective efficacy scale will be analyzed using a psychometric approach in which factor analytics are employed to identify an appropriate factor solution and test the reliability and validity of the CE scores. Once a CE factor solution and an empirically derived multilevel data structure have been identified, the association between CE and intervention effectiveness will be analyzed using multilevel generalized linear mixed models to estimate relative risks (aim 5) [85,86].

For all outcomes, variables used in the matching process may be considered as covariates, as needed, in addition to individual-, household-, and community-level risk factors. Covariates that are statistically associated with outcomes of interest in bivariate analyses will be considered for inclusion in final multivariate models, following standard stepwise model-building approaches. Secondary analyses may also evaluate models for effect modification as relevant, including exposure-mediator interaction for mediation models and cross-level interaction, by assessing changes in parameter values based on potential effect modifiers. Potential effect modifiers may include breastfeeding for seroconversion outcomes, and climate factors and population density for environmental fecal contamination and health outcomes. However, this study was not designed to assess effect modification and therefore is not specifically powered for these analyses. For all outcomes, unadjusted models will be presented along with models adjusting for covariates.

DISCUSSION

This matched cohort study is one of the first to evaluate the effect of a rural combined household-level piped water and sanitation intervention, implemented at the community level, on a large scale. The matched design provides a rigorous means for estimating causal effects given that randomization to intervention group was not feasible due to the several year implementation process[5]. By focusing on an intervention where the implementation process is complete, it also limits the risk presented by randomized controlled trials, where the intervention has little uptake, an especially important study challenge given interdependence of exposure and outcomes within communities, and a problem that has characterized previous trials of sanitation interventions in India[15,16].

A strength of this study is the assessment of health impacts using the holistic WHO definition of health, including not just disease status, but also mental, social, and physical well-being[87]. Outcomes along the causal chain include standard, but more subjective measures, such as reported diarrheal diseases and respiratory infection, as well as more objective measures such as fecal environmental contamination, soil transmitted helminth infection, and anthropometry. Although there is risk of response bias for reported outcomes, it is unlikely to be differential by intervention status since the study team is not directly linked to Gram Vikas. Even though field workers may be aware of village intervention status, lab staff analyzing water, hand rinse, stool, and blood samples will be blinded. In addition, this study includes the more novel use of seroconversion for enteric pathogens, biomarkers of environmental enteric dysfunction, and measures of collective efficacy in an evaluation assessment. While there are limitations inherent to observational studies, the matched study design and multivariate modeling analysis plan reduce the potential for confounding. However, there is still the potential for residual unmeasured confounding.

Ethics and Dissemination. This study has been reviewed and approved by the Ethics Committee of the London School of Hygiene and Tropical Medicine, U.K (No. 9071) and Institute Ethics Committee of the Kalinga Institute of Medical Sciences of KIIT University, Bhubaneswar, India (KIMS/KIIT/IEC/053/2015). Efforts will be made to communicate the central findings and implications with study communities, the implementing organization and government officials in India. The results of this study will be submitted for publication in peer reviewed journals and presented at conferences. The data collected in the study will be publicly available, with personal identifiable data redacted, following the publication of the primary results within 24 months of the final data collection date.

Funding. This study is supported by a grant from the Bill & Melinda Gates Foundation to the London School of Hygiene & Tropical Medicine (OPP1008048) and to Emory University. (OOP1125067).

Competing Interests: None declared.

Contributions from authors: TC, HR, PR, BT, and HC contributed to study design. HR, LZ and BT developed laboratory protocols. HR, BT, GS, MD, SS, LZ, and BC developed data collection tools. All authors contributed to editing and revising the manuscript.

535 **Figure 1.** Description of the three phases of the Gram Vikas MANTRA water and sanitation intervention.

538 **Figure 2.** Study sites in Ganjam and Gajapati districts, Odisha, India with intervention villages in black and control villages in white. Inset shows location of districts in India.

541 **Figure 3.** Restriction, matching and exclusion process for selection of intervention and control villages (1), and timeline for study rounds and outcome data collection (2).

543 **REFERENCES**

544 1 UNICEF, WHO. Progress and Sanitation and Drinking Water: 2015 Update and MDG
545 Assessment. 2015.

546 2 Water and Sanitation Program. A Decade of the Total Sanitation Campaign: Rapid
547 Assessment of Processes and Outcomes. 2010.

548 3 Coffey D, Gupta A, Hathi P, *et al.* Culture and the health transition: Understanding
549 sanitation behavior in rural north India. 2015.

550 4 Clasen T, Pruss-Ustun A, Mathers CD, *et al.* Estimating the impact of unsafe water,
551 sanitation and hygiene on the global burden of disease: evolving and alternative methods.
552 *Trop Med Int Heal* 2014;**19**:884–93. doi:10.1111/tmi.12330

553 5 Arnold BF, Khush RS, Ramaswamy P, *et al.* Causal inference methods to study
554 nonrandomized, preexisting development interventions. *Proc Natl Acad Sci U S A*
555 2010;**107**:22605–10. doi:10.1073/pnas.1008944107

556 6 Ercumen A, Arnold BF, Kumpel E, *et al.* Upgrading a Piped Water Supply from
557 Intermittent to Continuous Delivery and Association with Waterborne Illness: A Matched
558 Cohort Study in Urban India. *PLOS Med* 2015;**12**:e1001892.
559 doi:10.1371/journal.pmed.1001892

560 7 Fewtrell L, Kaufmann RB, Kay D, *et al.* Water, sanitation, and hygiene interventions to
561 reduce diarrhoea in less developed countries: a systematic review and meta-analysis.
562 *Lancet Infect Dis* 2005;**5**:42–52. doi:10.1016/S1473-3099(04)01253-8

563 8 Engell RE, Lim SS. Does clean water matter? An updated meta-analysis of water supply
564 and sanitation interventions and diarrhoeal diseases. *Lancet* 2013;**381**:S44.
565 doi:10.1016/S0140-6736(13)61298-2

566 9 Clasen T, Bostoen K, Schmidt W, *et al.* Interventions to improve disposal of human excreta for preventing diarrhoea. *Cochrane database Syst Rev* Published Online First: 2010.http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD007180.pub2/pdf/standard (accessed 4 Jan2015).

570 10 Clasen T, Schmidt W-P, Rabie T, *et al.* Interventions to improve water quality for preventing diarrhoea: systematic review and meta-analysis. *BMJ* 2007;**334**:782. doi:10.1136/bmj.39118.489931.BE

573 11 Cairncross S, Hunt C, Boisson S, *et al.* Water, sanitation and hygiene for the prevention of diarrhoea. *Int J Epidemiol* 2010;**39 Suppl 1**:i193–205. doi:10.1093/ije/dyq035

575 12 Clasen TF, Alexander KT, Sinclair D, *et al.* Interventions to improve water quality for preventing diarrhoea. *Cochrane database Syst Rev* 2015;**10**:CD004794. doi:10.1002/14651858.CD004794.pub3

578 13 Routray P, Schmidt W-P, Boisson S, *et al.* Socio-cultural and behavioural factors constraining latrine adoption in rural coastal Odisha: an exploratory qualitative study. *BMC Public Health* 2015;**15**:880. doi:10.1186/s12889-015-2206-3

581 14 Barnard S, Routray P, Majorin F, *et al.* Impact of Indian Total Sanitation Campaign on Latrine Coverage and Use: A Cross-Sectional Study in Orissa Three Years following Programme Implementation. *PLoS One* 2013;**8**. doi:10.1371/journal.pone.0071438

584 15 Patil SR, Arnold BF, Salvatore AL, *et al.* The effect of India's total sanitation campaign on defecation behaviors and child health in rural Madhya Pradesh: a cluster randomized controlled trial. *PLoS Med* 2014;**11**:e1001709. doi:10.1371/journal.pmed.1001709

587 16 Clasen T, Boisson S, Routray P, *et al.* Effectiveness of a rural sanitation programme on diarrhoea, soil-transmitted helminth infection, and child malnutrition in Odisha, India: a

cluster-randomised trial. *Lancet Glob Heal* 2014;**2**:e645–53. doi:10.1016/S2214-109X(14)70307-9

17 Eisenberg JNS, Trostle J, Sorensen RJD, *et al.* Towards a Systems Approach to Enteric Pathogen Transmission: From Individual Independence to Community Interdependence. *Annu Rev Public Health* 2013;**33**:239–57. doi:10.1016/j.micinf.2011.07.011.Innate

18 Root GPM. Sanitation, Community Environments, and Childhoold Diarrhea in Rural Zimbabwe. *J Heal Popul Nutr* 2001;**19**:73–82.

19 Bateman OM, Smith S. A Comparison of the Health Effects of Water Supply and Sanitation in Urban and Rural Guatemala. 1999.

20 Huq A, Yunus M, Sohel SS, *et al.* Simple Sari Cloth Filtration of Water Is Sustainable and Continues to Protect Villagers from Cholera in Matlab, Bangladesh. *MBio* 2010;**1**:1–5. doi:10.1128/mBio.00034-10.Invited

21 Bandura A. *Self-Efficacy: The Exercise of Control*. Worth Publishers 1997.

22 UN Inter-agency Group for Child Mortality Estimation. Levels and trends in child mortality: Report 2015. 2015.

23 World Health Organization. Estimated total deaths by cause, sex and WHO Member State, 2008. Geneva, Switzerland: 2011.

24 Lin A, Arnold BF, Afreen S, *et al.* Household environmental conditions are associated with enteropathy and impaired growth in rural Bangladesh. *Am J Trop Med Hyg* 2013;**89**:130–7. doi:10.4269/ajtmh.12-0629

25 Korpe PS, Petri WA. Environmental enteropathy: critical implications of a poorly understood condition. *Trends Mol Med* 2012;**18**:328–36. doi:10.1016/j.molmed.2012.04.007

26 Guerrant RL, Oriá RB, Moore SR, *et al.* Malnutrition as an enteric infectious disease with long-term effects on child development. 2008;**66**:487–505. doi:10.1111/j.1753-4887.2008.00082.x.Malnutrition

27 Crane RJ, Jones KDJ, Berkley JA. Environmental enteric dysfunction : An overview. 2015;**36**:76–87.

28 Humphrey JH. Child undernutrition, tropical enteropathy, toilets, and handwashing. *Lancet* 2009;**374**:1032–5. doi:10.1016/S0140-6736(09)60950-8

29 Mbuya MNN, Humphrey JH. Preventing environmental enteric dysfunction through improved water, sanitation and hygiene: An opportunity for stunting reduction in developing countries. *Matern Child Nutr* 2015;**12**:1–15. doi:10.1111/mcn.12220

30 Budge PJ, Griffin MR, Edwards KM, *et al.* Impact of home environment interventions on the risk of influenza-associated ARI in Andean children: observations from a prospective household-based cohort study. *PLoS One* 2014;**9**:e91247. doi:10.1371/journal.pone.0091247

31 Aiello AE, Coulborn RM, Perez V, *et al.* Effect of hand hygiene on infectious disease risk in the community setting: a meta-analysis. *Am J Public Health* 2008;**98**:1372–81. doi:10.2105/AJPH.2007.124610

32 Caruso BA, Sevilimedu V, Fung IC-H, *et al.* Gender disparities in water, sanitation, and global health. *Lancet (London, England)* 2015;**386**:650–1. doi:10.1016/S0140-6736(15)61497-0

33 Sahoo KC, Hulland KR, Caruso B a., *et al.* Sanitation-related psychosocial stress: a grounded theory study of women across the life-course in Odisha, India. *Soc Sci Med* 2015;**139**:80–9. doi:10.1016/j.socscimed.2015.06.031

34 Hulland KRS, Chase RP, Caruso BA, *et al.* Sanitation, stress, and life stage: A systematic data collection study among women in Odisha, India. *PLoS One* 2015;**10**. doi:10.1371/journal.pone.0141883

35 World Health Organization/The United Nations Children's Fund (UNICEF). Ending Preventable Child Deaths from Pneumonia and Diarrhoea by 2025: The integrated Global Action Plan for Pneumonia and Diarrhoea (GAPPD). 2013.

36 Gram Vikas. Annual Report 2013-2014. 2014.

37 Duflo E, Greenstone M, Guiteras R, *et al.* Toilets Can Work : Short and Medium Run Health Impacts of Addressing Complementarities and Externalities in Water and Sanitation.

38 Rosa G, Clasen T. Estimating the scope of household water treatment in low- and medium-income countries. *Am J Trop Med Hyg* 2010;**82**:289–300. doi:10.4269/ajtmh.2010.09-0382

39 International Institute for Population Sciences (IIPS). District Level Household and Facility Survey (DLHS-3), 2007-08: India, Orissa. Mumbai: 2010.

40 Brady H, Caughey D, Dehejia R, *et al.* Genetic Matching for Estimating Causal Effects : 2012.

41 Sekhon JS. "Multivariate and Propensity Score Matching Software. *J Stat Softw* 2011;**42**.http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Multivariate+and+Propensity+Score+Matching#9 (accessed 30 May2014).

42 Austin PC. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Stat Med* 2009;**28**:3083–107. doi:10.1002/sim

43 Rubin DB. Using propensity scores to help design observational studies: Application to the tobacco litigation. *Health Serv Outcomes Res Methodol* 2002;**2**:169–88.

44 Harder VS, Stuart EA, Anthony JC. Propensity score techniques and the assessment of measured covariate balance to test causal associations in psychological research. *Psychol Methods* 2010;**15**:997–1003. doi:10.1016/j.biotechadv.2011.08.021.Secreted

45 Faul F, Erdfelder E, Buchner A, *et al.* Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. *Behav Res Methods* 2009;**41**:1149–60. doi:10.3758/BRM.41.4.1149

46 Wolf J, Prüss-Ustün A, Cumming O, *et al.* Systematic review: Assessing the impact of drinking water and sanitation on diarrhoeal disease in low- and middle-income settings: Systematic review and meta-regression. *Trop Med Int Heal* 2014;**19**:928–42. doi:10.1111/tmi.12331

47 Carl Hartung YAWBALCTGB. Open Data Kit: Tools to Build Information Services for Developing Regions. http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.176.8017

48 Feikin DR, Audi a., Olack B, *et al.* Evaluation of the optimal recall period for disease symptoms in home-based morbidity surveillance in rural and urban Kenya. *Int J Epidemiol* 2010;**39**:450–8. doi:10.1093/ije/dyp374

49 Arnold BF, Galiani S, Ram PK, *et al.* Optimal Recall Period for Caregiver-reported Illness in Risk Factor and Intervention Studies: A Multicountry Study. *Am J Epidemiol* 2013;**177**:361–70. doi:10.1093/aje/kws281

50 World Health Organization. Integrated Management of Childhood Illness: Chart Booklet.

51 Cogill B. Anthropometric indicators measurement guide. Revised edition. *Washington, DC, Acad Educ Dev [AED], Food Nutr Tech Assist Proj*

2003;:92.http://www.developmentgateway.org/download/202582/anthro_2003.pdf

52 de Onis M, Onyango AW, Van den Broeck J, *et al.* Measurement and standardization protocols for anthropometry used in the construction of a new international growth reference. *Food Nutr Bull* 2004;**25**:S27–36.

53 Truant a. L, Elliott SH, Kelly MT, *et al.* Comparison of formalin-ethyl ether sedimentation, formalin-ethyl acetate sedimentation, and zinc sulfate flotation techniques for detection of intestinal parasites. *J Clin Microbiol* 1981;**13**:882–4.

54 Kosek M, Haque R, Lima A, *et al.* Fecal markers of intestinal inflammation and permeability associated with the subsequent acquisition of linear growth deficits in infants. *Am J Trop Med Hyg* 2013;**88**:390–6. doi:10.4269/ajtmh.2012.12-0549

55 Crump J a, Mendoza CE, Priest JW, *et al.* Comparing serologic response against enteric pathogens with reported diarrhea to assess the impact of improved household drinking water quality. *Am J Trop Med Hyg* 2007;**77**:136–41.

56 Steinberg EB, Mendoza CE, Glass R, *et al.* Prevalence of Infection with Waterbourne Pathogens: a Seroepidemiologic Study in Children 6-36 Months Old in San Juan Sacatepquez, Guatemala. 2004;**70**:83–8.

57 Brussow H, Sidoti J, Link H, *et al.* Age-specific prevalence of antibody to enterotoxigenic Escherichia coli in Ecuadorian and German children. *J Infect Dis* 1990;**162**:974–7.

58 Khanna B, Cutler A, Israel N, *et al.* Use caution with serologic testing for Helicobacter pylori infection in children. *J Infect Dis* 1998;**178**:460–5.

59 Lindkvist P, Asrat D, Nilsson I. Age at acquisition of Helicobacter pylori infection: comparison of a high and a low prevalence country. *Scand J Infect Dis* 1996;**28**:181–4.

60 Ungar B, Gilman R, Lanata C, *et al.* Seroepidemiology of Cryptosporidium infection in

two Latin American populations. *J Infect Dis* 1988;**157**:551–6.

61 Vitral C, Yoshida C, Lemos E, *et al.* Age-specific prevalence of antibodies to hepatitis A in children and adolescents from Rio de Janeiro, Brazil, 1978 and 1995: relationship of prevalence to environmental factors. *Mem Inst Oswaldo Cruz* 1998;**93**:1–5.

62 Lammie PJ, Moss DM, Brook Goodhew E, *et al.* Development of a new platform for neglected tropical disease surveillance. *Int J Parasitol* 2012;**42**:797–800. doi:10.1016/j.ijpara.2012.07.002

63 Centers for Disease Control and Prevention. Laboratory Methods for the Diagnosis of Vibio cholerae. Atlanta, Georgia: 1999.

64 Centers for Disease Control and Prevention. Isolation and Identification of Shigella. Atlanta, Georgia:

65 United States Environmental Protection Agency. Coliforms—Total and E. coli, Membrane Filtration Method 10029. 1999.

66 Gronewold AD, Borsuk ME, Wolpert RL, *et al.* An Assessment of Fecal Indicator Bacteria-Based Water Quality Standards. *Environ Sci Technol* 2008;**42**:4676–82. doi:10.1021/es703144k

67 Gruber JS, Ercumen A, Colford JM. Coliform bacteria as indicators of diarrheal risk in household drinking water: systematic review and meta-analysis. *PLoS One* 2014;**9**:e107429. doi:10.1371/journal.pone.0107429

68 Levy K, Nelson KL, Hubbard A, *et al.* Rethinking indicators of microbial drinking water quality for health studies in tropical developing countries: case study in northern coastal Ecuador. *Am J Trop Med Hyg* 2012;**86**:499–507. doi:10.4269/ajtmh.2012.11-0263

69 Livio S, Strockbine N a, Panchalingam S, *et al.* Shigella Isolates From the Global Enteric

Multicenter Study Inform Vaccine Development. *Clin Infect Dis* 2014;**59**. doi:10.1093/cid/ciu468

70 Kotloff KL, Nataro JP, Blackwelder WC, *et al.* Burden and aetiology of diarrhoeal disease in infants and young children in developing countries (the Global Enteric Multicenter Study, GEMS): a prospective, case-control study. *Lancet* 2013;**382**:209–22. doi:10.1016/S0140-6736(13)60844-2

71 Kanungo S, Sah BK, Lopez a. L, *et al.* Cholera in India: An analysis of reports, 1997-2006. *Bull World Health Organ* 2010;**88**:185–91. doi:10.2471/BLT.09.073460

72 Cellini SR, Kee JE. Cost-Effectiveness and Cost-Benefit Analysis. In: Wholey JS, Hatry HP, Newcomer KE, eds. *Handbook of Practical Program Evaluation*. San Francisco: : John Wiley & Sons, Inc. 2010.

73 Edejer TT, Baltussen R, Adam T, *et al.* Making Choices in Health: WHO Guide to Cost-Effectiveness Analysis. 2003.

74 Malapit HJ, Sproule K, Kovarik C, *et al.* Measuring progress toward empowerment Women's Empowerment in Agriculture Index: Baseline Report. *IfpriOrg* 2014;:1–60.

75 Black RE, Victora CG, Walker SP, *et al.* Maternal and child undernutrition and overweight in low-income and middle-income countries. *Lancet* 2013;**382**:427–51. doi:10.1016/S0140-6736(13)60937-X

76 Smith LC, Ramakrishnan U, Ndiaye A, *et al.* The Importance of Women's Status for Child Nutrition in Developing Countries. 2003. http://www.ifpri.org/publication/importance-womens-status-child-nutrition-developing-countries

77 Sumpter C, Torondel B. A Systematic Review of the Health and Social Effects of

Menstrual Hygiene Management. *PLoS One* 2013;**8**:e62004. doi:10.1371/journal.pone.0062004

78 Dasgupta A, Sarkar M. Menstrual Hygiene: How Hygienic is the Adolescent Girl? *Indian J Community Med* 2008;**33**:77–80. doi:10.4103/0970-0218.40872

79 Das P, Baker KK, Dutta A, *et al.* Menstrual Hygiene Practices, WASH Access and the Risk of Urogenital Infection in Women from Odisha, India. *PLoS One* 2015;**10**:e0130777. doi:10.1371/journal.pone.0130777

80 van Eijk AM, Sivakami M, Thakkar MB, *et al.* Menstrual hygiene management among adolescent girls in India: a systematic review and meta-analysis. *BMJ Open* 2016;**6**:e010290. doi:10.1136/bmjopen-2015-010290

81 Caruso B. *Sanitation Insecurity: Definition, Measurement, and Associations with Women's Mental Health in Rural Orissa, India.* 2015.http://pid.emory.edu/ark:/25593/rfhnt

82 Peal A, Evans B, Blackett I, *et al.* Fecal sludge management (FSM): analytical tools for assessing FSM in cities. *J Water, Sanit Hyg Dev* 2014;**4**:371. doi:10.2166/washdev.2014.139

83 Peal A, Evans B, Blackett I, *et al.* Fecal Sludge Management: a comparative analysis of 12 cities. *J Water, Sanit Hyg Dev* 2014;**4**:563–75.

84 Richiardi L, Bellocco R, Zugna D. Mediation analysis in epidemiology: Methods, interpretation and bias. *Int J Epidemiol* 2013;**42**:1511–9. doi:10.1093/ije/dyt127

85 McNutt LA, Wu C, Xue X, *et al.* Estimating the relative risk in cohort studies and clinical trials of common outcomes. *Am J Epidemiol* 2003;**157**:940–3. doi:10.1093/aje/kwg074

86 Thompson ML, Myers JE, Kriebel D. Prevalence odds ratio or prevalence ratio in the

analysis of cross sectional data: what is to be done? *Occup Environ Med* 1998;**55**:272–7. doi:10.1136/oem.55.4.272

87 World Health Organization. WHO definition of health. http://www.who.int/about/definition/en/print.html (accessed 26 Apr2016).

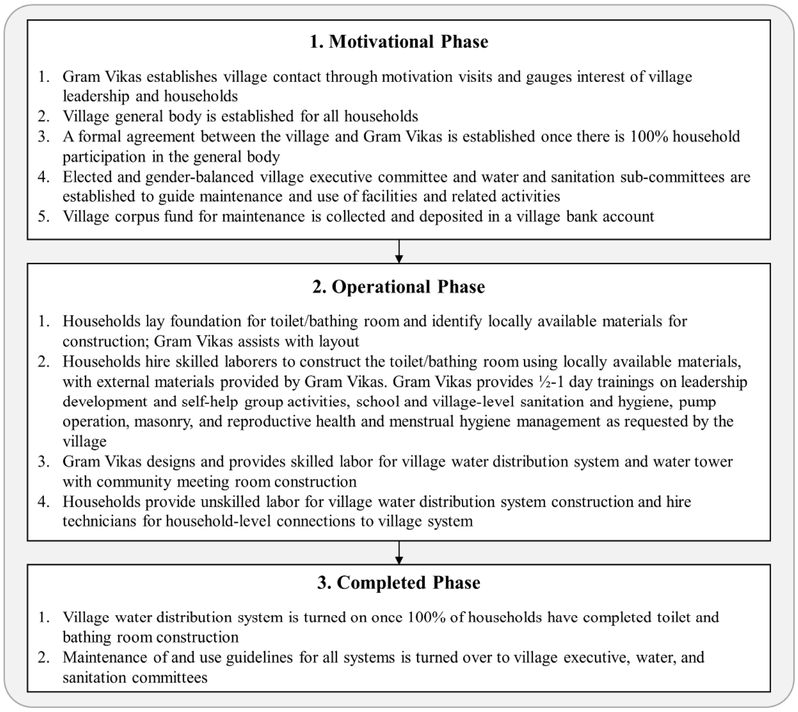

Figure 1. Description of the three phases of the Gram Vikas MANTRA water and sanitation intervention.
Figure 1
254x190mm (300 x 300 DPI)

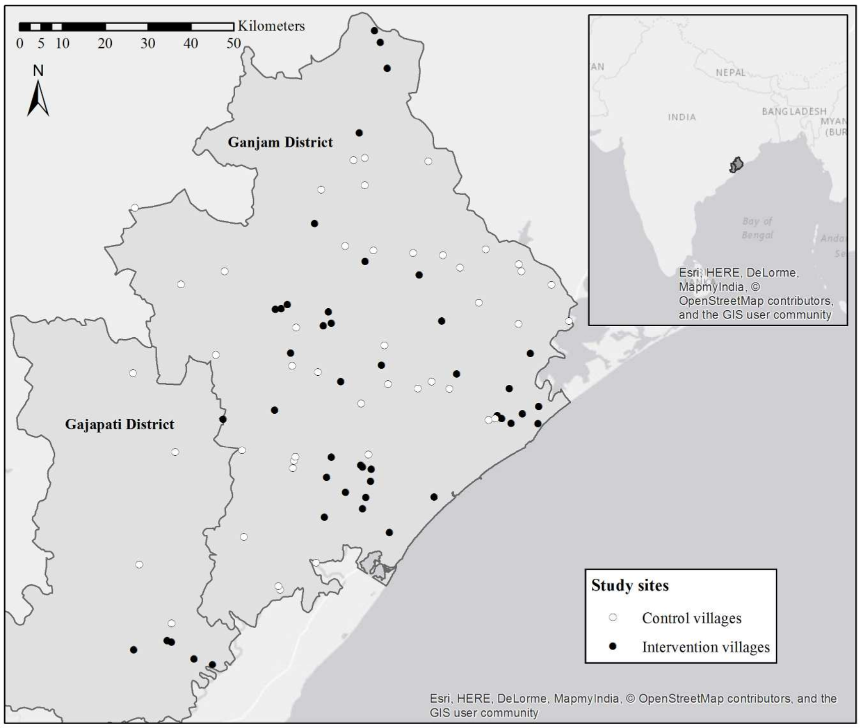

Figure 2. Study sites in Ganjam and Gajapati districts, Odisha, India with intervention villages in black and control villages in white. Inset shows location of districts in India.
Figure 2
215x279mm (300 x 300 DPI)

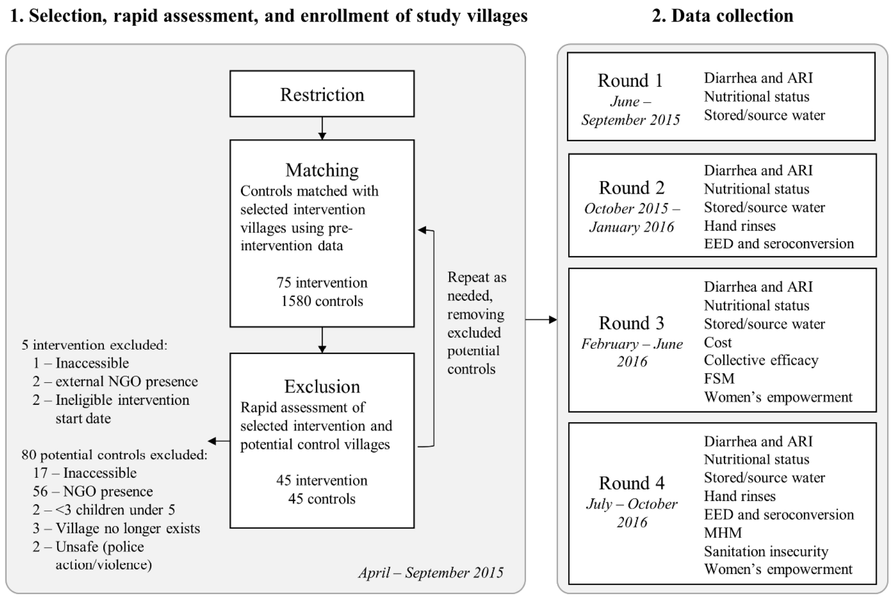

Figure 3. Restriction, matching and exclusion process for selection of intervention and control villages (1), and timeline for study rounds and outcome data collection (2).
Figure 3
254x190mm (300 x 300 DPI)

STROBE Statement—checklist of items that should be included in reports of observational studies

	Item No.	Recommendation	Page No.	Relevant text from manuscript
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	1	
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	2	
Introduction				
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	3-5	
Objectives	3	State specific objectives, including any prespecified hypotheses	7	
Methods				
Study design	4	Present key elements of study design early in the paper	8	
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	8	
Participants	6	(*a*) *Cohort study*—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study*—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study*—Give the eligibility criteria, and the sources and methods of selection of participants	8-11	
		(*b*) *Cohort study*—For matched studies, give matching criteria and number of exposed and unexposed *Case-control study*—For matched studies, give matching criteria and the number of controls per case	8-11	
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	12-19	
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	12-19	
Bias	9	Describe any efforts to address potential sources of bias	NA	
Study size	10	Explain how the study size was arrived at	12	

Continued on next page

Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	NA
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	NA (overview 19-20)
		(*b*) Describe any methods used to examine subgroups and interactions	NA
		(*c*) Explain how missing data were addressed	NA
		(*d*) *Cohort study*—If applicable, explain how loss to follow-up was addressed *Case-control study*—If applicable, explain how matching of cases and controls was addressed *Cross-sectional study*—If applicable, describe analytical methods taking account of sampling strategy	NA
		(*e*) Describe any sensitivity analyses	NA
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	NA
		(b) Give reasons for non-participation at each stage	NA
		(c) Consider use of a flow diagram	NA
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	NA
		(b) Indicate number of participants with missing data for each variable of interest	NA
		(c) *Cohort study*—Summarise follow-up time (eg, average and total amount)	NA
Outcome data	15*	*Cohort study*—Report numbers of outcome events or summary measures over time	NA
		Case-control study—Report numbers in each exposure category, or summary measures of exposure	
		Cross-sectional study—Report numbers of outcome events or summary measures	
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	NA
		(*b*) Report category boundaries when continuous variables were categorized	NA
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	NA

Continued on next page

Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	NA
Discussion			
Key results	18	Summarise key results with reference to study objectives	NA
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	20-21
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	NA
Generalisability	21	Discuss the generalisability (external validity) of the study results	NA
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	22

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.

BMJ Open

Design and rationale of a matched cohort study to assess the effectiveness of a combined household-level piped water and sanitation intervention in rural Odisha, India

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2016-012719.R2
Article Type:	Protocol
Date Submitted by the Author:	25-Oct-2016
Complete List of Authors:	Reese, Heather; Emory University School of Public Health, Department of Environmental Health Routray, Parimita; London School of Hygiene and Tropical Medicine Torondel, Belen; London School of Hygiene and Tropical Medicine Sclar, Gloria; Emory University School of Public Health, Department of Environmental Health Delea, Maryann; Emory University School of Public Health; London School of Hygiene and Tropical Medicine Sinharoy, Sheela; Emory University Laney Graduate School, Nutrition and Health Sciences Program Zambrano, Laura; Emory University School of Public Health, Department of Environmental Health Caruso, Bethany; Emory University School of Public Health, Department of Environmental Health Mishra, Samir; KIIT University, School of Biotechnology Chang, Howard; Emory University, Department of Biostatistics and Bioinformatics Clasen, Thomas; London School of Hygiene & Tropical Medicine; Emory University School of Public Health, Department of Environmental Health
Primary Subject Heading:	Public health
Secondary Subject Heading:	Epidemiology
Keywords:	sanitation, piped water supply, diarrheal diseases

SCHOLARONE™ Manuscripts

Design and rationale of a matched cohort study to assess the effectiveness of a combined household-level piped water and sanitation intervention in rural Odisha, India

Heather Reese,[1] Parimita Routray,[2] Belen Torondel,[2] Gloria Sclar,[1] Maryann G. Delea,[1,2] Sheela S. Sinharoy,[3] Laura Zambrano,[1] Bethany Caruso,[1] Samir R. Mishra,[4] Howard H. Chang,[5] Thomas Clasen[1,2]

[1] Department of Environmental Health, Rollins School of Public Health, Emory University, Atlanta, Georgia, United States of America

[2] London School of Hygiene and Tropical Medicine, London, United Kingdom

[3] Nutrition and Health Sciences Program, Laney Graduate School, Emory University, Atlanta, Georgia, United States of America

[4] School of Biotechnology, KIIT University, Bhubaneswar, Odisha, India

[5] Department of Biostatistics and Bioinformatics, Rollins School of Public Health, Emory University, Atlanta, Georgia, United States of America

Contact information for corresponding author, Heather Reese:

Department of Environmental Health, Rollins School of Public Health, Emory University

1518 Clifton Rd, Atlanta, GA 30322

Email: heather.e.reese@emory.edu, Phone: (619) 379-8963

Keywords: rural, piped water, sanitation

Word count: 5442

ABSTRACT

Introduction: Government efforts to address massive shortfalls in rural water and sanitation in India have centered on construction of community water sources and toilets for selected households. However, deficiencies with water quality and quantity at the household level, and community coverage and actual use of toilets has led Gram Vikas, a local NGO in Odisha, India, to develop an approach that provides household-level piped water connections contingent on full community-level toilet coverage.

Methods: This matched cohort study was designed to assess the effectiveness of a combined piped water and sanitation intervention. Households with children under five years in 45 randomly selected intervention villages and 45 matched control villages will be followed over 17 months. The primary outcome is prevalence of diarrheal diseases; secondary health outcomes include soil-transmitted helminth infection, nutritional status, seroconversion to enteric pathogens, urogenital infections, and environmental enteric dysfunction. In addition, intervention effects on sanitation and water coverage, access and use, environmental fecal contamination, women's empowerment, as well as collective efficacy, and intervention cost and cost-effectiveness will be assessed.

Ethics and dissemination: The study protocol has been reviewed and approved by the ethics boards of the London School of Hygiene and Tropical Medicine, U.K. and KIIT University, Bhubaneswar, India. Findings will be disseminated via peer-reviewed literature and presentation to stakeholders, government officials, implementers and researchers.

Trial registration identifier: NCT02441699

STRENGTHS AND LIMITATIONS OF THIS STUDY

- The study assesses a combined household-level piped water and sanitation intervention that requires complete community-level compliance.
- The intervention was not randomly allocated; but, controls are selected through a restriction process to limit possible partial exposure to the intervention through spillover, and matched to intervention villages using pre-intervention data.
- The study uses a holistic definition of health to assess intervention impacts on physical, mental and social well-being, including more novel outcomes such as seroconversion to enteric pathogens, environmental enteric dysfunction, and sanitation insecurity. It also assesses intervention coverage, cost-effectiveness, and collective efficacy.
- The time lapse between intervention completion and the beginning of the evaluation process prevents baseline comparison or assessment of immediate intervention impacts. However, it allows for a biologically plausible length of time for die-off of even the most persistent pathogens in the environment, and provides time for children to have be born into this environment.

INTRODUCTION

Of the one billion people practicing open defecation worldwide, over half live in India[1]. While international and national pressure on improving sanitation conditions in India has led to over 350 thousand people gaining access to improved toilets since 1990, it has barely kept up with population growth[1,2]. Recent studies show that even in areas with access to household-level improved sanitation, use of these toilets is low[3–5]. This may be due in part to a mismatch between the culturally acceptable pour-flush toilets and the level of water access. Coverage of

improved water sources, usually community-level pumps or taps, is relatively high even in rural areas in India, but it may not be sufficient for flushing purposes on top of other daily water needs[1,6].

Although the effectiveness of water, sanitation and hygiene (WASH) interventions vary, meta-analyses have found that individual or combined WASH interventions decrease diarrheal disease prevalence by up to 48%[7–11]. While combined interventions would be expected to have a greater influence on multiple exposure pathways and thus a greater combined impact on health, there is limited evidence of additive benefits[12]. This may be due to poor uptake, inconsistent use, or an incomplete understanding of relevant pathways[8–10]. In India, combining water and sanitation interventions may be more critical than just interrupting multiple transmission pathways for enteric infection; evidence suggests that household-level water access is integral to the use of improved sanitation in this context[13].

While the intent of improved sanitation facilities is to separate human feces from human contact, most of the focus is on constructing household toilets to increase improved sanitation coverage—the primary metric used in monitoring progress toward international targets. However, studies in India have further shown that toilet construction does not translate into toilet use in this context[5,14–16]. Moreover, with the interdependence between members of households and households within communities, safe water and sanitation is a community-level issue. There is growing emphasis on assessing health risk from poor water and sanitation conditions not simply due to individual or even household-level risk factors, but also from conditions in the community environment[17]. There is evidence that even households without toilets, and households which do not filter drinking water, showed decreased health risk if they live in communities with high levels of coverage and use[18–20].

Moreover, the effectiveness of community interventions may be higher in communities with positive perceptions of their collective ability to come together to improve their conditions. Collective efficacy, a latent construct comprised of the structural and cognitive components that facilitate a community's shared belief in its ability to come together and execute actions related to a common goal, may explain some variance in intervention effectiveness across communities receiving WASH interventions[21].

A main risk of poor WASH conditions is enteric infection, caused by a diverse array of bacteria, viruses, protozoa, and parasites, including soil-transmitted helminths. These infections may cause diarrhea, the second leading cause of mortality for children under five years worldwide and in India, a leading cause of mortality regardless of age[22,23]. There is also growing evidence that asymptomatic enteric infections may pose a similar risk, with repeat enteric infections contributing to chronic malnutrition, environmental enteric dysfunction, poor cognitive outcomes, and poor vaccine uptake[24–29]. Poor WASH conditions are also linked to increased risk of respiratory infection, the leading cause of mortality for children under five years worldwide[22,30,31]. Poor water and sanitation access can also affect the social, physical and mental well-being of women, acting through pathways ranging from unsafe menstrual hygiene management practices and increased risk of violence[32–34].

Description of the intervention

Over the past decades there has been a global commitment to determine water and sanitation interventions with demonstrated effectiveness, not just efficacy[35]. Gram Vikas, a non-governmental organization based in Odisha, India (http://www.gramvikas.org/), has responded by implementing its MANTRA (Movement and Action Network for Transformation of Rural

Areas) water and sanitation program in more than 1000 villages since 2002[36]. This approach includes both household-level piped water connections, and community-level mobilization for culturally appropriate household toilets. A previous interrupted time series analysis of the MANTRA intervention reported it to be protective against diarrheal diseases[37]. However, in addition to limitations of design, this study relied on outcome data collected and reported by Gram Vikas, the intervention implementer, and did not assess intervention coverage or impacts on environmental fecal contamination.

The MANTRA water and sanitation intervention is rolled out in a three-phase process over an average of three years. During the first, or Motivational, phase (approximately 8-12 mo), representatives of Gram Vikas visit the identified village several times to assess village interest and progress towards a set of Gram Vikas requirements, including: 1) the commitment of every household to participate, 2) creation of a village corpus fund from contributions from every household, and 3) development of village guidelines for maintenance and use of facilities.

Once this set of requirements is achieved, the village progresses into the second, or Operational, phase of the intervention (approximately 17-35 mo). Each household constructs a pour-flush toilet with two soak-pits and a separate bathing room. The households hire a local, skilled mason and provide their own unskilled labor and locally available materials to complete the superstructure. Gram Vikas provides external materials such as PVC pipes and porcelain pans. At the same time, a water tank, community meeting space, and piped water distribution system connected to every household, with taps in the toilet and bathing rooms and a separate tap in the kitchen, is constructed through a similar collaborative process.

All households must construct a toilet and bathing room for the village to progress into the final, or Completed, phase of the intervention, in which the water system is turned on.

Notably, this three-phase process only allows each household access to piped water once every household in the village has a toilet and bathing room. This model contrasts with most previous water and sanitation interventions, including those implemented under India's Total Sanitation Campaign and other government programs, which do not require community-level sanitation compliance and do not provide a piped water supply at the household level[38].

Study aims

The primary objective of this study is to evaluate the effectiveness of the combined household-level water supply and sanitation intervention, as implemented by Gram Vikas in Odisha, India. Toward that objective, this study aims to:

1) Assess the effectiveness of the intervention in improving water and sanitation infrastructure coverage, access, and use, and to assess fecal sludge management practices in intervention communities.
2) Assess the effectiveness of the intervention in reducing environmental fecal contamination.
3) Assess the effectiveness of the intervention in improving health. This includes reported diarrheal disease in children under 5 years (primary outcome), acute respiratory infection, infection with soil-transmitted helminthes, nutritional status, environmental enteric dysfunction, seroconversion for selected enteric pathogens, and urogenital diseases associated with menstrual hygiene management practices. Mental and social well-being will be explored through assessment of sanitation insecurity and women's empowerment.
4) Assess the cost and cost-effectiveness of the intervention.

5) Develop and assess a theoretically-grounded, empirically informed collective efficacy scale; and determine the effect of collective efficacy on intervention effectiveness.

METHODS

Setting

The study is located in Ganjam and Gajapati districts in eastern Odisha, India (Figure 1). These two contiguous districts were a single district until 1992. Over 44% of the population in these districts is recognized by the Government of India as being below the poverty line (BPL)[39]. As of 2008, a majority of households in both districts had access to an improved, likely community-level, drinking water source, with over 23% of households in Ganjam having access to any sanitation facility, compared to only 8% of households in Gajapati[39]. The area is primarily rural and agrarian, and the climate is characterized by a monsoon season from June to September, with an average rainfall of ~1400 mms/year.

Study design

This study uses a matched cohort design to assess the effectiveness of a completed intervention with data collected across four study rounds from June 2015 to October 2016 (Figure 2). Data was collected in all study rounds for diarrhea, acute respiratory infection, nutritional status, and stored and source water outcomes to assess seasonality. Data was collected in rounds 2 and 4 for environmental enteric dysfunction, seroconversion, and hand-rinses, and cross-sectionally in one or more rounds for the remaining outcomes. As described below, control villages were matched to randomly selected intervention villages through a multi-step restriction, genetic matching, and exclusion process using the following eligibility criteria.

Eligibility criteria for villages

1. Restriction. Intervention villages were randomly selected from a list of Gram Vikas villages in Ganjam and Gajapati districts provided by the NGO, after restriction to villages with a Motivation phase start date between 2002-2006 and a Construction phase start date no earlier than 2003. Since the intervention process takes on average three years, the criteria for the Motivation start date helped to identify those villages with ongoing interventions at the same time. In addition, this allowed the use of the Government of India Census 2001 and the Below Poverty Line (BPL) Survey 2002 data to characterize baseline characteristics used in the matching process in both intervention and control villages.

Eligible control villages include all villages without a Gram Vikas intervention within the study districts which: 1) are not within the same Gram Panchayat (a political subdivision with some administrative responsibility for water and sanitation comprised of several villages) as a Gram Vikas village, or bordering a Gram Vikas village, and 2) had not received a Motivation visit from the Gram Vikas NGO. These criteria serve to limit the possibility of previous partial exposure to the intervention through spillover from adjacent villages or direct contact with the NGO. These criteria also increase the strength of the comparison provided by the control villages, i.e. it increases the likelihood that if they had received a motivation visit from Gram Vikas, the control villages would have been equally as likely as the intervention villages to demand the intervention.

In addition, to be eligible for inclusion both intervention and control villages must: 1) appear in the Government of India Census 2001 and the BPL Survey 2002, 2) have a population of at least 20 households, and 3) be within approximately three hours travel from the study office in Brahmapur, Ganjam District. This last criterion is due to logistical constraints.

2. Matching. After restriction, genetic matching was used to match potential control villages to the randomly selected intervention villages without replacement[5,40,41]. Villages were exact matched on district to limit any political or large scale geographic variation between district populations, and were also matched on pre-intervention demographic, socioeconomic, sanitation, and water access characteristics listed in Table 1[5]. These village-level matching variables were selected due to their theorized association with the primary outcome, diarrheal diseases, as well as data availability.

3. Exclusion. The field team visited matched potential control villages and intervention villages to assess suitability for the study through a rapid assessment interview with village leadership and to ensure accessibility. Villages were excluded if they are not within three hours travel of the field office in Brahmapur, had sustained major infrastructure damage due to a natural disaster, or if there was a current or planned sanitation or water intervention by an organization external to the village in the next 12 months as determined through the rapid assessment interview with village leadership. In addition, villages were excluded if there were fewer than three households with children under five years old. As villages were removed from the pool of prospective control villages, the matching process was repeated for all intervention villages and remaining eligible control villages, and balance measures were assessed. The matching and exclusion processes were repeated as necessary.

After the iterative matching and exclusion process was complete, covariate balance was assessed for all matching variables for the final set of intervention and control villages through examination of balance measures[42–44]. Matching resulted in an improvement in balance as assessed through comparison of several measures including q-q plots, Kolmogorov-Smirnov bootstrap p-values, and standardized differences. After matching, there were no significant differences between intervention and control groups (Table 1).

Eligibility criteria for households

Households within selected intervention and control villages were eligible if they had at least one child under 5 years old at time of enrollment, verified with birth or immunization card, and expected to reside in the village for the duration of the study. If there were more than 40 eligible households within a village, 40 were randomly selected to be enrolled. Informed consent was be obtained from the male and/or female head of the selected households. All children under five years within each enrolled household were eligible and do not age-out over the course of the

Table 1. Pre-intervention characteristics used in matching, and balance diagnostics before and after matching and exclusion process.

Variable	**Intervention** (n=45)	**Control (all eligible)** (n=1580)	Std Diff (all eligible)	**Control (study)** (n=45)	Std Diff (study)
Number of households	157.9	215.5	0.37	148.1	0.06
Population under 6 years (%)	16.2	16.9	0.19	16.3	0.02
Household income score ($\bar{x}$)	2.9	3.1**	0.26	2.9	0.01
Household goods owned ($\bar{x}$)	1.1	1.2*	0.27	1.1	0.02
Pucca house (%)	59.2	61.6	0.09	60.5	0.05
≥2 meals a day (%)	57.7	63.7	0.19	57.8	0.01
Scheduled caste (%)	11.5	18.7**	0.46	11.8	0.01

Scheduled tribe (%)	33.4	19.1*	0.31	29.8	0.08
Female literacy (%)	30.9	29.8	0.07	30.9	0.00
Open defecation (%)	95.6	95.2*	0.04	95.8	0.01
Improved drinking water sourceǂ (%)	38.6	42.5	0.10	37.2	0.02
Water source <500m and 50m elevation (%)	81.5	72.2	0.31	81.7	0.01

All eligible: all villages that are eligible for the matching process after restriction
Std Diff (absolute standardized difference): a value greater than 0.1 is considered meaningful imbalance [42]
ǂ Ganjam villages only; no data available for Gajapati villages
Kolmogorov-Smirnov bootstrap p-values: * <0.05 ** <0.01

study. Households with newborn children were enrolled as they became eligible on an ongoing basis throughout the study, in villages with fewer than 40 enrolled households.

Sample Size

Sample size was determined through a simulation estimating the log odds of diarrheal disease (the primary outcome) through a multilevel random effects model and parameterized with data from a previous study in a neighboring district in Odisha[16]. Sample size estimates were also checked with G*Power[45]. The simulation assumes a longitudinal 7-day period prevalence for diarrhea of 8.8% in children under five years, a heterogeneity variance between villages of 0.07, a heterogeneity variance between households of 0.57, and four study rounds[16]. An effect size of 0.20 was selected for public health importance and based on estimates of effect from systematic reviews of water and sanitation studies[46]. Assuming at least 80% power, 0.05 significance level, 10% for loss to follow up, and at least one child per household, we estimate a sample size of 45 villages per study arm and 26 households per village. This estimate was the most conservative compared to sample size estimates for secondary outcomes, and was therefore used for the broader study population.

Outcome Measurement

Outcomes, and individual, household, and community-level risk factors, will be measured through surveys, interviews, or through the collection and analysis of environmental, stool or dried blood spot samples. All survey questions will be translated into the primary local language, Odia, and back-translated to confirm wording. Household surveys include household and individual factors and will be verbally administered by trained field workers to the mother or primary caregiver of the youngest child under five in each household, unless otherwise specified below. Community surveys will be verbally administered to the *sarpanch* (village head) or other member of village leadership. Survey data will be collected on mobile phones using Open Data Kit[47]. GPS coordinates for households, water sources and other relevant sites will be collected using Garmin eTrex 10 or 20 devices (Garmin Ltd., Olathe, KS, USA).

Coverage, access and use of sanitation, water and hygiene infrastructure

Coverage, access and use of WASH infrastructure will be assessed in all four rounds. Presence of and access to toilets, water sources and hand-washing stations will be assessed through standard questions from the Demographic and Health Surveys (DHS) and confirmed through spot observations. Spot observations of household toilets and hand-washing stations will be further used to assess indicators of functionality, maintenance, recent use. Reported water and sanitation practices, including child feces disposal practices, will be captured through household survey questions.

Diarrheal Diseases

The primary outcome for this study is prevalence of diarrheal diseases, recorded as both daily point prevalence over the previous three days and seven-day period prevalence, for all household members in all four rounds. Although self-reported diarrhea is a subjective outcome with a well-established risk of bias, three-day recall reduces recall bias[48,49]. Diarrheal disease will be measured using the World Health Organization (WHO) definition of three or more loose stools in a 24-hour period, with or without the presence of blood. Field workers will use a simple calendar as a visual aid to help respondents with recall. Each household member will be asked to recall his or her own disease status and the mother or primary caregiver will be asked to report disease for children.

Respiratory infection

Prevalence of respiratory infections will be recorded as both daily point prevalence over the previous three days and seven-day period prevalence for all household members in all four rounds. Respiratory infection is defined as the presence of cough and/or shortness of breath/difficulty breathing according to WHO's Integrated Management of Childhood Illness (IMCI)[50]. The full IMCI case definition for acute lower respiratory infection also includes measurement of respiratory rate and observation of chest indrawing, stridor and other danger signs; these criteria were excluded from our definition as there was concern about the technical support required to produce consistent and accurate data within this context[50]. Our definition provides a broad assessment of respiratory illness burden. Each household member will be asked to recall his or her own disease status and the mother or primary caregiver will be asked to report disease for children.

Nutritional Status

Anthropometric data will be collected for children under age five in all four rounds using standard methods as established by WHO[51,52]. Field workers will be trained and standardized in line with WHO protocols to reduce measurement error [52]. Weight will be measured for all children under five years of age using Seca 385 digital scales, with 20g increment for weight below 20kg and a 50g increment for weight between 20 and 50kg. Recumbent length will be measured for children under two years of age using Seca 417 measuring boards with 1mm increment. Standing height will be measured for children two to five years of age using Seca 213 portable stadiometers with 1mm increment. Height and weight will be used to calculate height-for-age z-scores (HAZ) and weight-for-height z-scores (WHZ) based on WHO reference standards. A random subset of 10% of households will receive back check visits each day to repeat height/length measurements to ensure inter-observer reliability.

Soil-transmitted helminth infection

Stool samples will be collected in rounds 2 and 4 from all household members in a randomly selected subset of 500 households, and used to assess the presence and intensity of soil-transmitted helminth (STH) infection. Formalin ether concentration and microscopy will be used to quantify worms and ova for hookworms, *Ascaris lumbricoides*, and *Tricuris trichura*[53]. Quality assurance includes independent duplicate assessment of all positive and 10% of negative samples. After stool collection, each participant will be offered a single dose of Albendazole, a broad-spectrum antihelmenthic drug recommended by the Ministry of Health and Family Welfare, Government of India. Stools collected in round 2 will allow for comparison of STH infection prevalence between intervention and control villages, while the stool samples collected approximately 8 months later in round 4 will provide a measure of re-infection rate.

Environmental enteric dysfunction

Stools from a randomly selected subset of 200 children under two years old, collected in rounds 2 and 4, will be used to assess environmental enteric dysfunction (EED) through quantification of biomarkers of intestinal inflammation and permeability. Fecal myeloperoxidase (MPO), alpha-1-antitrypsin (AAT), and neopterin (NEO), markers for neutrophil activity, intestinal permeability and TH1 immune activation, respectively, were selected for this study based on evidence of association with EED, subsequent linear growth deficits, and household environmental fecal contamination[24,25,54].

Seroconversion for enteric pathogens

Serological assays that assess antibody production against various enteric pathogens can provide an objective measure of exposure to enteric infections[55]. Enrolling children aged 6 to 18 months will reduce the potential for interference from maternally acquired antibodies and permit analysis of seroconversion data in a critical window for young children who experience higher diarrheal disease morbidity and mortality before two years of age[56–61]. Children who are 6 to 12 months during round 2 will have capillary blood drawn by fingerstick or heelstick, as appropriate, and will be visited again during round 4 for a second capillary blood sample. All blood samples will be preserved on TropBio (Sydney, Australia) filter discs and stored within 7 days of collection at -20°C. Seroconversion against markers for norovirus, *Giardia intestinalis*, *Cryptosporidium parvum*, *Entamoeba histolytica*, enterotoxigenic *E. coli* heat-labile enterotoxin

(ETEC-LT), *Salmonella* spp., *Campylobacter jejuni*, *Vibrio cholera*, and *Toxoplasma* spp. willbe assessed using multiplex immunoassay technology on the Luminex xMAP platform[62].

Environmental fecal contamination

Field workers will collect samples of household stored drinking water and source water from a random subset of 500 households in all four rounds, and child hand rinses in rounds 2 and 4. All water and hand rinse samples will be stored on ice during transport and analyzed within 6 hours of collection using membrane filtration. Three assays will be used: 1) plating on m-Coli Blue 24 (Millipore, Billerica, MA) for *E.coli* according to EPA Method 10029, 2) alkaline peptone water enrichment prior to plating on thiosulfate citrate bile salts sucrose agar and slide agglutination serotyping for *V. cholerae*, and 3) plating on xylose lysine desoxycholate agar, and slide agglutination serotyping for *Shigella* spp.[63–65]. Source and stored water samples will be assayed for *E. coli*, *Vibrio cholerae* and *Shigella* spp., and hand rinse samples will be assayed for *E. coli* and *Shigella* spp. *E. coli* was selected as a standard non-human specific indicator of fecal contamination, though the limitations of this indicator are well-established[66–68]. In order to better characterize human fecal contamination of the household environment, *Vibrio cholerae* and *Shigella* spp. were selected based on prevalence in southern Asia, evidence of public health importance, and field laboratory limitations[69–71].

Cost and cost-effectiveness

Costs and potential cost savings (i.e., averted costs) associated with the intervention will be assessed through an economic costing approach that recognizes and quantifies costs and benefits from a societal perspective[72]. Data on program and point-of-delivery inputs will be collected at household, community, and implementer levels in round 3. Field workers will administer community surveys to a village leader, and household surveys to the household decision-maker for toilet installation, in 20 randomly selected households in twenty matched intervention and control villages. Given cost-effectiveness analyses require the effect of the intervention to be measured against a counterfactual, and the intervention of interest is a community-based intervention, cost and effectiveness measures will be summarized at the village level [73]. Surveys will collect data on household- and community-level inputs related to materials and labor required to construct household toilets and wash rooms, the community water tank and distribution system, and household water connections; longer-term water supply and toilet maintenance costs; and financing required for this infrastructure as well as perceived benefits, including averted social opportunity costs. Implementer inputs from Gram Vikas will be collected through an enumeration exercise, interviews, and examination of the implementer's financial records.

Collective efficacy

Collective efficacy (CE) is a latent construct comprised of the structural and cognitive components that facilitate a community's shared belief in its ability to come together and execute actions related to a common goal[21]. A review of the literature and established conceptual frameworks will be performed to define the CE construct. A sequential exploratory mixed qualitative and quantitative design will be used to develop and refine a scale to measure CE and test hypotheses. Field workers will administer the refined, multi-item, Likert-type CE scale to one randomly selected household member aged 18 years or older in each household in round 3.

Women's empowerment

Four dimensions of women's empowerment will be measured in rounds 3 and 4: group participation, leadership, decision-making and freedom of movement. Group participation and leadership will be measured using modules from the Women's Empowerment in Agriculture Index (WEAI), which has been tested in South Asia[74]. Decision-making will be measured using questions from the women's status module of Demographic and Health Surveys. Freedom of movement will be measured using questions from the project-level Women's Empowerment in Agriculture Index (pro-WEAI). These measures will be collected for the primary female caregiver of the youngest child under 5, and were selected based on the importance of women's empowerment for child nutrition[75,76]. Women's empowerment is conceptualized as both an outcome and a potential mediator along the pathway between the Gram Vikas intervention and child health outcomes.

Menstrual hygiene management

Menstrual hygiene management practices vary worldwide and depend on personal preference, socioeconomic status, local traditions and beliefs, and access to water and sanitation resources[77]. Unhygienic washing practices are common in rural India and among women and girls in lower socioeconomic groups, and may increase risk of urogenital infection[78–80]. However, the link between access to water and sanitation, menstrual hygiene management and urogenital infections has been poorly studied. Household surveys will be administered in round 4 to a randomly selected woman aged 18 or older, in a subset of 800 households, and will capture self-reported urogenital infection, defined as at least one of the following symptoms: 1) abnormal vaginal discharge (unusual texture and color/more abundant than normal), 2) burning or itching in the genitalia, 3) burning or itching when urinating, or 4) genital sores[79].

Sanitation Insecurity

This study will assess the associations between sanitation access and sanitation insecurity with mental health among women. In previous research in Odisha, a contextually specific definition and measure for sanitation insecurity was developed, with associations between facets of sanitation insecurity and mental health independent of sanitation facility access[81]. This previously developed measure will be used to determine if levels of sanitation insecurity differ between intervention and control villages and how it may be associated with mental health outcomes, specifically well-being, anxiety, depression, and distress. Household surveys will be administered in round 4 to a randomly selected woman aged 18 or older, in a random subset of 800 households.

Fecal sludge management.

In sanitation systems where sewerage is not feasible, such as the household toilets constructed as part of the MANTRA intervention, safe management of fecal waste is necessary. Although there is growing emphasis on safe fecal sludge management (FSM), research has mainly focused on urban settings[82,83]. Preliminary research in Odisha suggests that fecal sludge management in this rural setting is a substantial challenge, and may impact household use of toilets. In round 3, household surveys and spot checks of toilets in intervention villages will be used to assess toilet use and fecal sludge management practices.

STATISTICAL ANALYSES

The effect of the intervention on infrastructure coverage, access, and use (aim 1), and the effect of the intervention on improving health (aim 3), will be analyzed using logistic, linear, log binomial, or negative binomial multilevel regression depending on the outcome, to compare intervention versus control villages. Prevalence of fecal sludge management practices in intervention communities will be assessed using multilevel regression (aim 1). For all models, the hierarchical structure of the data will be accounted for using random effects. Estimation of relative risks through Poisson regression or binary regression methods for binary outcomes will be considered to ensure robustness of results. Mediation of the potential association between intervention and nutritional status outcomes by women's empowerment will be assessed using multilevel structural equation modeling, and statistical approaches to reduce bias will be explored as needed[84].

The impact of intervention on reducing environmental fecal contamination (aim 2), will be assessed through two methods. First, hierarchical logistic and negative binomial multilevel regression to estimate intervention effects on the relative scale will be used to compare intervention versus control villages. Estimation of relative risks through Poisson regression or binary regression methods for binary outcomes will be considered to ensure robustness of results. Second, a stochastic microbial risk framework will be used to assess differential fecal environmental contamination between intervention and control villages.

The cost and cost-effectiveness of the intervention (aim 4) will be assessed in two steps. Incremental intervention benefits will be ascertained by combining health benefit data, from analysis of health outcome data and established averted cost data, with other averted social opportunity costs. An incremental cost-effectiveness ratio, expressed in cost per disease-specific

DALY, will be calculated by dividing the incremental intervention costs by the incremental intervention benefits.

The collective efficacy scale will be analyzed using a psychometric approach in which factor analytics are employed to identify an appropriate factor solution and test the reliability and validity of the CE scores. Once a CE factor solution and an empirically derived multilevel data structure have been identified, the association between CE and intervention effectiveness will be analyzed using multilevel generalized linear mixed models to estimate relative risks[85,86]. (aim 5).For all outcomes, variables used in the matching process may be considered as covariates, as needed, in addition to individual, household, and community-level risk factors. Covariates that are statistically associated with outcomes of interest in bivariate analyses will be considered for inclusion in final multivariable models, following standard stepwise model-building approaches. Secondary analyses may also evaluate models for effect modification as relevant, including exposure-mediator interaction for mediation models and cross-level interaction, by assessing changes in parameter values based on potential effect modifiers. Potential effect modifiers may include breastfeeding for seroconversion outcomes, and climate factors and population density for environmental fecal contamination and health outcomes. However, this study was not designed to assess effect modification and therefore is not specifically powered for these analyses. For all outcomes, unadjusted models will be presented along with models adjusting for covariates.

DISCUSSION

This matched cohort study is one of the first to evaluate the effect of a rural combined household-level piped water and sanitation intervention, implemented at the community level, on a large scale. The matched design provides a rigorous means for estimating causal effects given that randomization to intervention group was not feasible due to the several year implementation process[5]. By focusing on an intervention where the implementation process is complete, it also limits the risk presented by randomized controlled trials, where the intervention has little uptake, an especially important study challenge given interdependence of exposure and outcomes within communities, and a problem that has characterized previous trials of sanitation interventions in India[15,16].

A strength of this study is the assessment of health impacts using the holistic WHO definition of health, including not just disease status, but also mental, social, and physical well-being[87]. Outcomes along the causal chain include standard, but more subjective measures, such as reported diarrheal diseases and respiratory infection, as well as more objective measures such as fecal environmental contamination, soil transmitted helminth infection, and anthropometry. Although there is risk of response bias for reported outcomes, it is unlikely to be differential by intervention status since the study team is not directly linked to Gram Vikas. Even though field workers may be aware of village intervention status, lab staff analyzing water, hand rinse, stool, and blood samples will be blinded. In addition, this study includes the more novel use of seroconversion for enteric pathogens, biomarkers of environmental enteric dysfunction, and measures of collective efficacy in an evaluation assessment. While there are limitations inherent to observational studies, the matched study design and multivariable modeling analysis plan reduce the potential for confounding. However, there is still the potential for residual unmeasured confounding.

Ethics and Dissemination. This study has been reviewed and approved by the Ethics Committee of the London School of Hygiene and Tropical Medicine, U.K (No. 9071) and Institute Ethics Committee of the Kalinga Institute of Medical Sciences of KIIT University, Bhubaneswar, India (KIMS/KIIT/IEC/053/2015). Efforts will be made to communicate the central findings and implications with study communities, the implementing organization and government officials in India. The results of this study will be submitted for publication in peer reviewed journals and presented at conferences. The data collected in the study will be publicly available, with personal identifiable data redacted, following the publication of the primary results within 24 months of the final data collection date.

Funding. This study is supported by a grant from the Bill & Melinda Gates Foundation to the London School of Hygiene & Tropical Medicine (OPP1008048) and to Emory University. (OOP1125067).

Competing Interests: None declared.

Contributions from authors: TC, HR, PR, BT, and HC contributed to study design. HR, LZ and BT developed laboratory protocols. HR, BT, GS, MD, SS, LZ, and BC developed data collection tools. All authors contributed to editing and revising the manuscript.

Figure 1. Study sites in Ganjam and Gajapati districts, Odisha, India with intervention villages in black and control villages in white. Inset shows location of districts in India.

Figure 2. Restriction, matching and exclusion process for selection of intervention and control villages (1), and timeline for study rounds and outcome data collection (2).

REFERENCES

1 UNICEF, WHO. Progress and Sanitation and Drinking Water: 2015 Update and MDG Assessment. 2015.

2 Water and Sanitation Program. A Decade of the Total Sanitation Campaign: Rapid Assessment of Processes and Outcomes. 2010.

3 Coffey D, Gupta A, Hathi P, *et al.* Culture and the health transition: Understanding sanitation behavior in rural north India. 2015.

4 Clasen T, Pruss-Ustun A, Mathers CD, *et al.* Estimating the impact of unsafe water, sanitation and hygiene on the global burden of disease: evolving and alternative methods. *Trop Med Int Heal* 2014;**19**:884–93. doi:10.1111/tmi.12330

5 Arnold BF, Khush RS, Ramaswamy P, *et al.* Causal inference methods to study nonrandomized, preexisting development interventions. *Proc Natl Acad Sci U S A* 2010;**107**:22605–10. doi:10.1073/pnas.1008944107

6 Ercumen A, Arnold BF, Kumpel E, *et al.* Upgrading a Piped Water Supply from Intermittent to Continuous Delivery and Association with Waterborne Illness: A Matched Cohort Study in Urban India. *PLOS Med* 2015;**12**:e1001892. doi:10.1371/journal.pmed.1001892

7 Fewtrell L, Kaufmann RB, Kay D, *et al.* Water, sanitation, and hygiene interventions to reduce diarrhoea in less developed countries: a systematic review and meta-analysis. *Lancet Infect Dis* 2005;**5**:42–52. doi:10.1016/S1473-3099(04)01253-8

8 Engell RE, Lim SS. Does clean water matter? An updated meta-analysis of water supply and sanitation interventions and diarrhoeal diseases. *Lancet* 2013;**381**:S44. doi:10.1016/S0140-6736(13)61298-2

9 Clasen T, Bostoen K, Schmidt W, *et al.* Interventions to improve disposal of human excreta for preventing diarrhoea. *Cochrane database Syst Rev* Published Online First: 2010.http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD007180.pub2/pdf/standard (accessed 4 Jan2015).

10 Clasen T, Schmidt W-P, Rabie T, *et al.* Interventions to improve water quality for preventing diarrhoea: systematic review and meta-analysis. *BMJ* 2007;**334**:782. doi:10.1136/bmj.39118.489931.BE

11 Cairncross S, Hunt C, Boisson S, *et al.* Water, sanitation and hygiene for the prevention of diarrhoea. *Int J Epidemiol* 2010;**39 Suppl 1**:i193–205. doi:10.1093/ije/dyq035

12 Clasen TF, Alexander KT, Sinclair D, *et al.* Interventions to improve water quality for preventing diarrhoea. *Cochrane database Syst Rev* 2015;**10**:CD004794. doi:10.1002/14651858.CD004794.pub3

13 Routray P, Schmidt W-P, Boisson S, *et al.* Socio-cultural and behavioural factors constraining latrine adoption in rural coastal Odisha: an exploratory qualitative study. *BMC Public Health* 2015;**15**:880. doi:10.1186/s12889-015-2206-3

14 Barnard S, Routray P, Majorin F, *et al.* Impact of Indian Total Sanitation Campaign on Latrine Coverage and Use: A Cross-Sectional Study in Orissa Three Years following Programme Implementation. *PLoS One* 2013;**8**. doi:10.1371/journal.pone.0071438

15 Patil SR, Arnold BF, Salvatore AL, *et al.* The effect of India's total sanitation campaign on defecation behaviors and child health in rural Madhya Pradesh: a cluster randomized controlled trial. *PLoS Med* 2014;**11**:e1001709. doi:10.1371/journal.pmed.1001709

16 Clasen T, Boisson S, Routray P, *et al.* Effectiveness of a rural sanitation programme on diarrhoea, soil-transmitted helminth infection, and child malnutrition in Odisha, India: a

cluster-randomised trial. *Lancet Glob Heal* 2014;**2**:e645–53. doi:10.1016/S2214-109X(14)70307-9

17 Eisenberg JNS, Trostle J, Sorensen RJD, *et al.* Towards a Systems Approach to Enteric Pathogen Transmission: From Individual Independence to Community Interdependence. *Annu Rev Public Health* 2013;**33**:239–57. doi:10.1016/j.micinf.2011.07.011.Innate

18 Root GPM. Sanitation, Community Environments, and Childhoold Diarrhea in Rural Zimbabwe. *J Heal Popul Nutr* 2001;**19**:73–82.

19 Bateman OM, Smith S. A Comparison of the Health Effects of Water Supply and Sanitation in Urban and Rural Guatemala. 1999.

20 Huq A, Yunus M, Sohel SS, *et al.* Simple Sari Cloth Filtration of Water Is Sustainable and Continues to Protect Villagers from Cholera in Matlab, Bangladesh. *MBio* 2010;**1**:1–5. doi:10.1128/mBio.00034-10.Invited

21 Bandura A. *Self-Efficacy: The Exercise of Control*. Worth Publishers 1997.

22 UN Inter-agency Group for Child Mortality Estimation. Levels and trends in child mortality: Report 2015. 2015.

23 World Health Organization. Estimated total deaths by cause, sex and WHO Member State, 2008. Geneva, Switzerland: 2011.

24 Lin A, Arnold BF, Afreen S, *et al.* Household environmental conditions are associated with enteropathy and impaired growth in rural Bangladesh. *Am J Trop Med Hyg* 2013;**89**:130–7. doi:10.4269/ajtmh.12-0629

25 Korpe PS, Petri WA. Environmental enteropathy: critical implications of a poorly understood condition. *Trends Mol Med* 2012;**18**:328–36. doi:10.1016/j.molmed.2012.04.007

26 Guerrant RL, Oriá RB, Moore SR, *et al.* Malnutrition as an enteric infectious disease with long-term effects on child development. 2008;**66**:487–505. doi:10.1111/j.1753-4887.2008.00082.x.Malnutrition

27 Crane RJ, Jones KDJ, Berkley JA. Environmental enteric dysfunction : An overview. 2015;**36**:76–87.

28 Humphrey JH. Child undernutrition, tropical enteropathy, toilets, and handwashing. *Lancet* 2009;**374**:1032–5. doi:10.1016/S0140-6736(09)60950-8

29 Mbuya MNN, Humphrey JH. Preventing environmental enteric dysfunction through improved water, sanitation and hygiene: An opportunity for stunting reduction in developing countries. *Matern Child Nutr* 2015;**12**:1–15. doi:10.1111/mcn.12220

30 Budge PJ, Griffin MR, Edwards KM, *et al.* Impact of home environment interventions on the risk of influenza-associated ARI in Andean children: observations from a prospective household-based cohort study. *PLoS One* 2014;**9**:e91247. doi:10.1371/journal.pone.0091247

31 Aiello AE, Coulborn RM, Perez V, *et al.* Effect of hand hygiene on infectious disease risk in the community setting: a meta-analysis. *Am J Public Health* 2008;**98**:1372–81. doi:10.2105/AJPH.2007.124610

32 Caruso BA, Sevilimedu V, Fung IC-H, *et al.* Gender disparities in water, sanitation, and global health. *Lancet (London, England)* 2015;**386**:650–1. doi:10.1016/S0140-6736(15)61497-0

33 Sahoo KC, Hulland KR, Caruso B a., *et al.* Sanitation-related psychosocial stress: a grounded theory study of women across the life-course in Odisha, India. *Soc Sci Med* 2015;**139**:80–9. doi:10.1016/j.socscimed.2015.06.031

34 Hulland KRS, Chase RP, Caruso BA, *et al.* Sanitation, stress, and life stage: A systematic data collection study among women in Odisha, India. *PLoS One* 2015;**10**. doi:10.1371/journal.pone.0141883

35 World Health Organization/The United Nations Children's Fund (UNICEF). Ending Preventable Child Deaths from Pneumonia and Diarrhoea by 2025: The integrated Global Action Plan for Pneumonia and Diarrhoea (GAPPD). 2013.

36 Gram Vikas. Annual Report 2013-2014. 2014.

37 Duflo E, Greenstone M, Guiteras R, *et al.* Toilets Can Work : Short and Medium Run Health Impacts of Addressing Complementarities and Externalities in Water and Sanitation.

38 Rosa G, Clasen T. Estimating the scope of household water treatment in low- and medium-income countries. *Am J Trop Med Hyg* 2010;**82**:289–300. doi:10.4269/ajtmh.2010.09-0382

39 International Institute for Population Sciences (IIPS). District Level Household and Facility Survey (DLHS-3), 2007-08: India, Orissa. Mumbai: 2010.

40 Brady H, Caughey D, Dehejia R, *et al.* Genetic Matching for Estimating Causal Effects : 2012.

41 Sekhon JS. "Multivariate and Propensity Score Matching Software. *J Stat Softw* 2011;**42**.http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Multivariate+and+Propensity+Score+Matching#9 (accessed 30 May2014).

42 Austin PC. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Stat Med* 2009;**28**:3083–107. doi:10.1002/sim

43 Rubin DB. Using propensity scores to help design observational studies: Application to the tobacco litigation. *Health Serv Outcomes Res Methodol* 2002;**2**:169–88.

44 Harder VS, Stuart EA, Anthony JC. Propensity score techniques and the assessment of measured covariate balance to test causal associations in psychological research. *Psychol Methods* 2010;**15**:997–1003. doi:10.1016/j.biotechadv.2011.08.021.Secreted

45 Faul F, Erdfelder E, Buchner A, *et al.* Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. *Behav Res Methods* 2009;**41**:1149–60. doi:10.3758/BRM.41.4.1149

46 Wolf J, Prüss-Ustün A, Cumming O, *et al.* Systematic review: Assessing the impact of drinking water and sanitation on diarrhoeal disease in low- and middle-income settings: Systematic review and meta-regression. *Trop Med Int Heal* 2014;**19**:928–42. doi:10.1111/tmi.12331

47 Carl Hartung YAWBALCTGB. Open Data Kit: Tools to Build Information Services for Developing Regions. http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.176.8017

48 Feikin DR, Audi a., Olack B, *et al.* Evaluation of the optimal recall period for disease symptoms in home-based morbidity surveillance in rural and urban Kenya. *Int J Epidemiol* 2010;**39**:450–8. doi:10.1093/ije/dyp374

49 Arnold BF, Galiani S, Ram PK, *et al.* Optimal Recall Period for Caregiver-reported Illness in Risk Factor and Intervention Studies: A Multicountry Study. *Am J Epidemiol* 2013;**177**:361–70. doi:10.1093/aje/kws281

50 World Health Organization. Integrated Management of Childhood Illness: Chart Booklet.

51 Cogill B. Anthropometric indicators measurement guide. Revised edition. *Washington, DC, Acad Educ Dev [AED], Food Nutr Tech Assist Proj*

2003;:92.http://www.developmentgateway.org/download/202582/anthro_2003.pdf

52 de Onis M, Onyango AW, Van den Broeck J, *et al.* Measurement and standardization protocols for anthropometry used in the construction of a new international growth reference. *Food Nutr Bull* 2004;**25**:S27–36.

53 Truant a. L, Elliott SH, Kelly MT, *et al.* Comparison of formalin-ethyl ether sedimentation, formalin-ethyl acetate sedimentation, and zinc sulfate flotation techniques for detection of intestinal parasites. *J Clin Microbiol* 1981;**13**:882–4.

54 Kosek M, Haque R, Lima A, *et al.* Fecal markers of intestinal inflammation and permeability associated with the subsequent acquisition of linear growth deficits in infants. *Am J Trop Med Hyg* 2013;**88**:390–6. doi:10.4269/ajtmh.2012.12-0549

55 Crump J a, Mendoza CE, Priest JW, *et al.* Comparing serologic response against enteric pathogens with reported diarrhea to assess the impact of improved household drinking water quality. *Am J Trop Med Hyg* 2007;**77**:136–41.

56 Steinberg EB, Mendoza CE, Glass R, *et al.* Prevalence of Infection with Waterbourne Pathogens: a Seroepidemiologic Study in Children 6-36 Months Old in San Juan Sacatepquez, Guatemala. 2004;**70**:83–8.

57 Brussow H, Sidoti J, Link H, *et al.* Age-specific prevalence of antibody to enterotoxigenic Escherichia coli in Ecuadorian and German children. *J Infect Dis* 1990;**162**:974–7.

58 Khanna B, Cutler A, Israel N, *et al.* Use caution with serologic testing for Helicobacter pylori infection in children. *J Infect Dis* 1998;**178**:460–5.

59 Lindkvist P, Asrat D, Nilsson I. Age at acquisition of Helicobacter pylori infection: comparison of a high and a low prevalence country. *Scand J Infect Dis* 1996;**28**:181–4.

60 Ungar B, Gilman R, Lanata C, *et al.* Seroepidemiology of Cryptosporidium infection in

two Latin American populations. *J Infect Dis* 1988;**157**:551–6.

61 Vitral C, Yoshida C, Lemos E, *et al.* Age-specific prevalence of antibodies to hepatitis A in children and adolescents from Rio de Janeiro, Brazil, 1978 and 1995: relationship of prevalence to environmental factors. *Mem Inst Oswaldo Cruz* 1998;**93**:1–5.

62 Lammie PJ, Moss DM, Brook Goodhew E, *et al.* Development of a new platform for neglected tropical disease surveillance. *Int J Parasitol* 2012;**42**:797–800. doi:10.1016/j.ijpara.2012.07.002

63 Centers for Disease Control and Prevention. Laboratory Methods for the Diagnosis of Vibio cholerae. Atlanta, Georgia: 1999.

64 Centers for Disease Control and Prevention. Isolation and Identification of Shigella. Atlanta, Georgia:

65 United States Environmental Protection Agency. Coliforms—Total and E. coli, Membrane Filtration Method 10029. 1999.

66 Gronewold AD, Borsuk ME, Wolpert RL, *et al.* An Assessment of Fecal Indicator Bacteria-Based Water Quality Standards. *Environ Sci Technol* 2008;**42**:4676–82. doi:10.1021/es703144k

67 Gruber JS, Ercumen A, Colford JM. Coliform bacteria as indicators of diarrheal risk in household drinking water: systematic review and meta-analysis. *PLoS One* 2014;**9**:e107429. doi:10.1371/journal.pone.0107429

68 Levy K, Nelson KL, Hubbard A, *et al.* Rethinking indicators of microbial drinking water quality for health studies in tropical developing countries: case study in northern coastal Ecuador. *Am J Trop Med Hyg* 2012;**86**:499–507. doi:10.4269/ajtmh.2012.11-0263

69 Livio S, Strockbine N a, Panchalingam S, *et al.* Shigella Isolates From the Global Enteric

Multicenter Study Inform Vaccine Development. *Clin Infect Dis* 2014;**59**. doi:10.1093/cid/ciu468

70 Kotloff KL, Nataro JP, Blackwelder WC, *et al.* Burden and aetiology of diarrhoeal disease in infants and young children in developing countries (the Global Enteric Multicenter Study, GEMS): a prospective, case-control study. *Lancet* 2013;**382**:209–22. doi:10.1016/S0140-6736(13)60844-2

71 Kanungo S, Sah BK, Lopez a. L, *et al.* Cholera in India: An analysis of reports, 1997-2006. *Bull World Health Organ* 2010;**88**:185–91. doi:10.2471/BLT.09.073460

72 Cellini SR, Kee JE. Cost-Effectiveness and Cost-Benefit Analysis. In: Wholey JS, Hatry HP, Newcomer KE, eds. *Handbook of Practical Program Evaluation*. San Francisco: : John Wiley & Sons, Inc. 2010.

73 Edejer TT, Baltussen R, Adam T, *et al.* Making Choices in Health: WHO Guide to Cost-Effectiveness Analysis. 2003.

74 Malapit HJ, Sproule K, Kovarik C, *et al.* Measuring progress toward empowerment Women's Empowerment in Agriculture Index: Baseline Report. *IfpriOrg* 2014;:1–60.

75 Black RE, Victora CG, Walker SP, *et al.* Maternal and child undernutrition and overweight in low-income and middle-income countries. *Lancet* 2013;**382**:427–51. doi:10.1016/S0140-6736(13)60937-X

76 Smith LC, Ramakrishnan U, Ndiaye A, *et al.* The Importance of Women's Status for Child Nutrition in Developing Countries. 2003. http://www.ifpri.org/publication/importance-womens-status-child-nutrition-developing-countries

77 Sumpter C, Torondel B. A Systematic Review of the Health and Social Effects of

Menstrual Hygiene Management. *PLoS One* 2013;**8**:e62004. doi:10.1371/journal.pone.0062004

78 Dasgupta A, Sarkar M. Menstrual Hygiene: How Hygienic is the Adolescent Girl? *Indian J Community Med* 2008;**33**:77–80. doi:10.4103/0970-0218.40872

79 Das P, Baker KK, Dutta A, *et al.* Menstrual Hygiene Practices, WASH Access and the Risk of Urogenital Infection in Women from Odisha, India. *PLoS One* 2015;**10**:e0130777. doi:10.1371/journal.pone.0130777

80 van Eijk AM, Sivakami M, Thakkar MB, *et al.* Menstrual hygiene management among adolescent girls in India: a systematic review and meta-analysis. *BMJ Open* 2016;**6**:e010290. doi:10.1136/bmjopen-2015-010290

81 Caruso B. *Sanitation Insecurity: Definition, Measurement, and Associations with Women's Mental Health in Rural Orissa, India.* 2015.http://pid.emory.edu/ark:/25593/rfhnt

82 Peal A, Evans B, Blackett I, *et al.* Fecal sludge management (FSM): analytical tools for assessing FSM in cities. *J Water, Sanit Hyg Dev* 2014;**4**:371. doi:10.2166/washdev.2014.139

83 Peal A, Evans B, Blackett I, *et al.* Fecal Sludge Management: a comparative analysis of 12 cities. *J Water, Sanit Hyg Dev* 2014;**4**:563–75.

84 Richiardi L, Bellocco R, Zugna D. Mediation analysis in epidemiology: Methods, interpretation and bias. *Int J Epidemiol* 2013;**42**:1511–9. doi:10.1093/ije/dyt127

85 McNutt LA, Wu C, Xue X, *et al.* Estimating the relative risk in cohort studies and clinical trials of common outcomes. *Am J Epidemiol* 2003;**157**:940–3. doi:10.1093/aje/kwg074

86 Thompson ML, Myers JE, Kriebel D. Prevalence odds ratio or prevalence ratio in the

analysis of cross sectional data: what is to be done? *Occup Environ Med* 1998;**55**:272–7. doi:10.1136/oem.55.4.272

87 World Health Organization. WHO definition of health. http://www.who.int/about/definition/en/print.html (accessed 26 Apr2016).

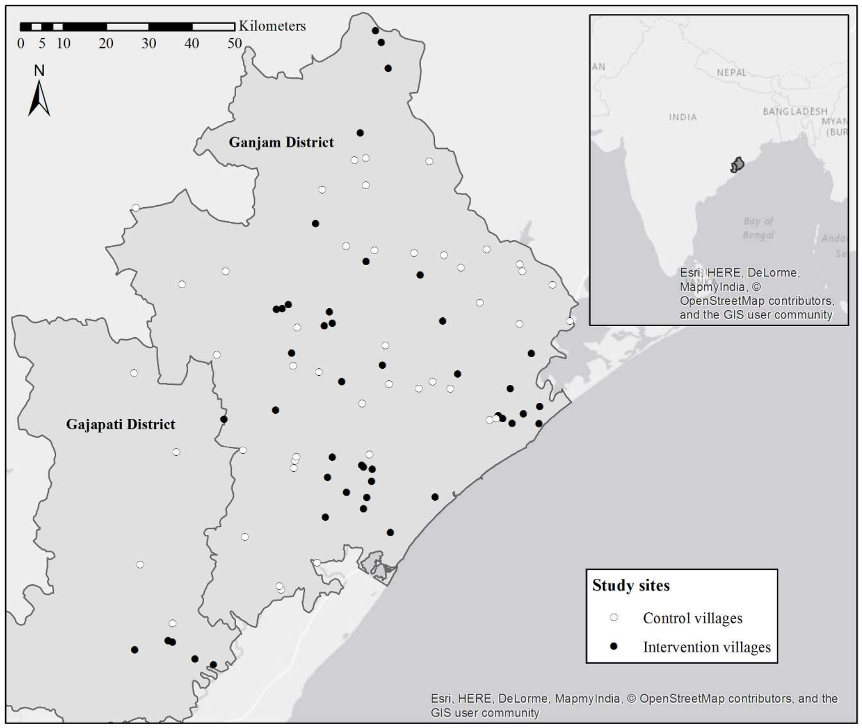

Figure 1. Study sites in Ganjam and Gajapati districts, Odisha, India with intervention villages in black and control villages in white. Inset shows location of districts in India.
Figure 1
215x279mm (300 x 300 DPI)

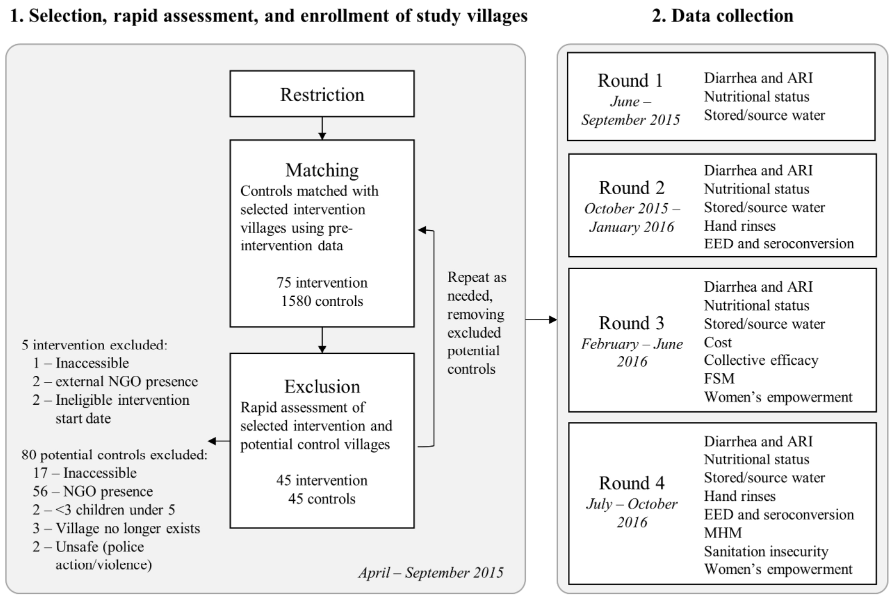

Figure 2. Restriction, matching and exclusion process for selection of intervention and control villages (1), and timeline for study rounds and outcome data collection (2).
Figure 2
254x190mm (300 x 300 DPI)

STROBE Statement—checklist of items that should be included in reports of observational studies

	Item No.	Recommendation	Page No.	Relevant text from manuscript
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	1	
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	2	
Introduction				
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	3-5	
Objectives	3	State specific objectives, including any prespecified hypotheses	7	
Methods				
Study design	4	Present key elements of study design early in the paper	8	
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	8	
Participants	6	(*a*) *Cohort study*—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study*—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study*—Give the eligibility criteria, and the sources and methods of selection of participants	8-11	
		(*b*) *Cohort study*—For matched studies, give matching criteria and number of exposed and unexposed *Case-control study*—For matched studies, give matching criteria and the number of controls per case	8-11	
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	12-19	
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	12-19	
Bias	9	Describe any efforts to address potential sources of bias	NA	
Study size	10	Explain how the study size was arrived at	12	

Continued on next page

Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	NA
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	NA (overview 19-20)
		(*b*) Describe any methods used to examine subgroups and interactions	NA
		(*c*) Explain how missing data were addressed	NA
		(*d*) *Cohort study*—If applicable, explain how loss to follow-up was addressed *Case-control study*—If applicable, explain how matching of cases and controls was addressed *Cross-sectional study*—If applicable, describe analytical methods taking account of sampling strategy	NA
		(*e*) Describe any sensitivity analyses	NA
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	NA
		(b) Give reasons for non-participation at each stage	NA
		(c) Consider use of a flow diagram	NA
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	NA
		(b) Indicate number of participants with missing data for each variable of interest	NA
		(c) *Cohort study*—Summarise follow-up time (eg, average and total amount)	NA
Outcome data	15*	*Cohort study*—Report numbers of outcome events or summary measures over time	NA
		Case-control study—Report numbers in each exposure category, or summary measures of exposure	
		Cross-sectional study—Report numbers of outcome events or summary measures	
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	NA
		(*b*) Report category boundaries when continuous variables were categorized	NA
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	NA

Continued on next page

Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	NA
Discussion			
Key results	18	Summarise key results with reference to study objectives	NA
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	20-21
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	NA
Generalisability	21	Discuss the generalisability (external validity) of the study results	NA
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	22

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.