

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

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}

To cite: Reese H, Routray P, Torondel B, *et al.* 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. *BMJ Open* 2017;7:e012719. doi:10.1136/bmjopen-2016-012719

► Prepublication history for this paper is available online. To view these files please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2016-012719>).

Received 19 May 2016
Revised 25 October 2016
Accepted 9 January 2017



CrossMark

For numbered affiliations see end of article.

Correspondence to

Heather Reese;
heather.e.reese@emory.edu

ABSTRACT

Introduction: Government efforts to address massive shortfalls in rural water and sanitation in India have centred 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 have led Gram Vikas, a local non-governmental organization 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 <5 years in 45 randomly selected intervention villages and 45 matched control villages will be followed over 17 months. The primary outcome is prevalence of diarrhoeal 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, UK and KIIT University, Bhubaneswar, India. Findings will be disseminated via peer-reviewed literature and presentation to stakeholders, government officials, implementers and researchers.

Trial registration number: NCT02441699.

INTRODUCTION

Of the one billion people practicing open defecation worldwide, over half of them live

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 spill over, and matched to intervention villages using preintervention 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 been born into this environment.

in India.¹ While international and national pressure on improving sanitation conditions in India has led to over 350 000 people gaining access to improved toilets since 1990, it has barely kept up with population growth.^{1–5} 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 diarrhoeal 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 towards 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 (CE), a latent construct comprised 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 diarrhoea, the second leading cause of mortality for children <5 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 (EED), 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 <5 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 few 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 organisation 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 programme in more than 1000 villages since 2002.³⁶ This approach includes household-level piped water connections and community-level mobilisation for culturally appropriate household toilets. A previous interrupted time series analysis of the MANTRA intervention reported it to be protective against diarrhoeal 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 3 years. During the first, or Motivational, phase (~8–12 months), 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 (~17–35 months). 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 labour 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 programmes, 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. Towards 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 diarrhoeal disease

in children <5 years (primary outcome), acute respiratory infection, infection with soil-transmitted helminthes, nutritional status, EED, 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 CE scale and determine the effect of CE 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 recognised by the Government of India as being below the poverty line

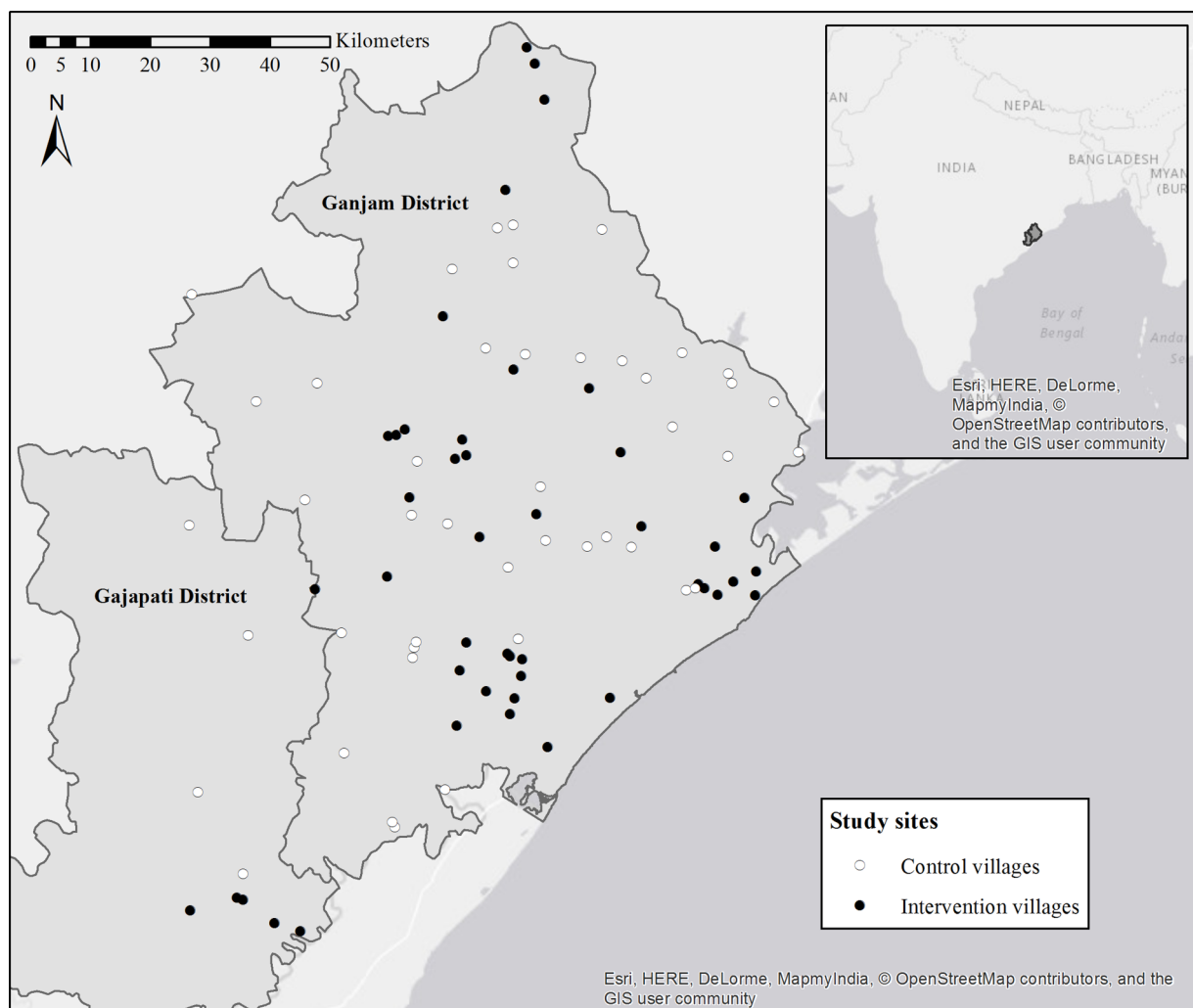


Figure 1 Study sites in Ganjam and Gajapati districts, Odisha, India, with intervention villages in black and control villages in white. Inset shows location of districts in India.

(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 characterised by a monsoon season from June to September, with an average rainfall of ~1400 mm/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 were collected in all study rounds for diarrhoea, acute respiratory infection, nutritional status and stored and source water outcomes to assess seasonality. Data were collected in rounds 2 and 4 for EED, 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 multistep restriction, genetic matching and exclusion process using the following eligibility criteria.

Eligibility criteria for villages

Restriction

Intervention villages were randomly selected from a list of Gram Vikas villages in Ganjam and Gajapati districts provided by the non-governmental organization (NGO), after restriction to villages with a Motivation phase start date between 2002 and 2006 and a Construction phase start date no earlier than 2003. Since the intervention process takes on average 3 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 BPL Survey 2002 data to characterise baseline characteristics used in the matching process in 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 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 spill over from adjacent villages or direct contact with the NGO. These criteria also increase the strength of the comparison provided by the control villages, that is, 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 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 ~3 hours travel from the study office in Brahmapur, Ganjam District. This last criterion is due to logistical constraints.

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

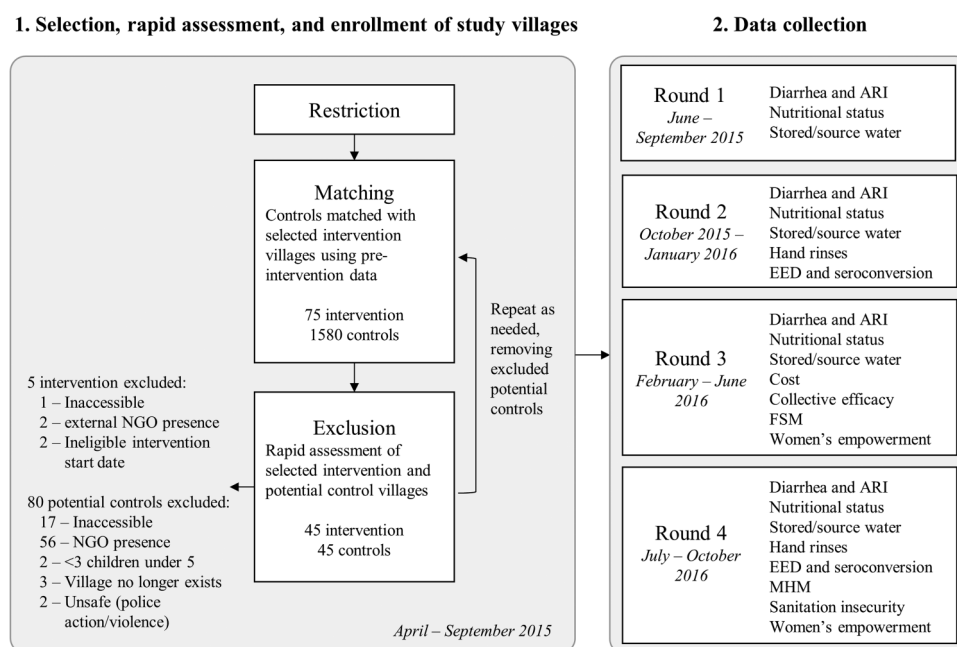


Figure 2 Restriction, matching and exclusion process for selection of intervention and control villages (1), and timeline for study rounds and outcome data collection (2).

large-scale geographic variation between district populations and were also matched on preintervention demographic, socioeconomic, sanitation and water access characteristics listed in table 1.⁵ These village-level matching variables were selected due to their theorised association with the primary outcome, diarrhoeal diseases, as well as data availability.

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 3 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 organisation 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 <5 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 standardised differences. After matching, there were

no significant differences between the 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 <5 years at the time of enrolment, verified with birth or immunisation 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, only 40 were randomly selected to be enrolled. Informed consent was obtained from the male and/or female head of the selected households. All children <5 years within each enrolled household were eligible and do not age-out over the course of the study. Households with newborn children were enrolled as they became eligible on an ongoing basis throughout the study, in villages with <40 enrolled households.

Sample size

Sample size was determined through a simulation estimating the log odds of diarrhoeal disease (the primary outcome) through a multilevel random effects model and parameterised with data from a previous study in a neighbouring district in Odisha.¹⁶ Sample size estimates were also checked with G*Power.⁴⁵ The simulation assumes a longitudinal 7-day period prevalence for diarrhoea of 8.8% in children <5 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

Table 1 Preintervention 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 <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† (%)	38.6	42.5	0.10	37.2	0.02
Water source <500 m and 50 m 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 standardised difference): a value >0.1 is considered meaningful imbalance.⁴²

Kolmogorov-Smirnov bootstrap p values: *<0.05, **<0.01.

†Ganjam villages only; no data available for Gajapati villages.

estimate a sample size of 45 villages per study arm and 26 households per village. This estimate was the most conservative compared with 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 <5 in each household, unless otherwise specified below. Community surveys will be verbally administered to the *sarpanch* (village head) or any 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, Olathe, Kansas, 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.

Diarrhoeal diseases

The primary outcome for this study is prevalence of diarrhoeal diseases, recorded as both daily point prevalence over the previous 3 days and 7-day period prevalence, for all household members in all four rounds. Although self-reported diarrhoea is a subjective outcome with a well-established risk of bias, three-day recall reduces recall bias.^{48 49} Diarrhoeal disease will be measured using the 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 3 days and

7-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 the age of five in all four rounds using standard methods as established by WHO.^{51 52} Field workers will be trained and standardised in line with WHO protocols to reduce measurement error.⁵² Weight will be measured for all children <5 years of age using Seca 385 digital scales, with 20 g increment for weight below 20 kg and a 50 g increment for weight between 20 and 50 kg. Recumbent length will be measured for children <2 years of age using Seca 417 measuring boards with 1 mm increment. Standing height will be measured for children 2–5 years of age using Seca 213 portable stadiometers with 1 mm 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 interobserver 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 antihelminthic 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 ~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 <2 years old, collected in rounds 2 and 4, will be used to

assess EED through quantification of biomarkers of intestinal inflammation and permeability. Fecal myeloperoxidase (MPO), α -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 6–18 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 diarrhoeal disease morbidity and mortality before 2 years of age.^{56–61} Children who are 6–12 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 *Escherichia coli*, heat-labile enterotoxin (ETEC-LT), *Salmonella* spp., *Campylobacter jejuni*, *Vibrio cholerae* 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 analysed within 6 hours of collection using membrane filtration. Three assays will be used: (1) plating on m-Coli Blue 24 (Millipore, Billerica, Massachusetts, USA) for *E. coli* according to EPA Method 10 029, (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*, *V. 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 characterise human fecal contamination of the household environment, *V. 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 (ie, averted costs) associated with the intervention will be assessed through an economic costing approach that recognises and quantifies costs and benefits from a societal perspective.⁷² Data on programme 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 20 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 summarised at the village level.⁷³ Surveys will collect data on household-level and community-level inputs related to materials and labour 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

CE is a latent construct comprised 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 <5 years of age and were selected based on the importance of women's empowerment for child nutrition.^{75 76}

Women's empowerment is conceptualised as 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 colour/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 a 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 FSM 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 FSM 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 analysed using logistic, linear, log binomial or negative binomial multilevel regression depending on the outcome, to compare intervention versus control villages. Prevalence of FSM 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 modelling, 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 CE scale will be analysed 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 analysed using multilevel generalised 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 randomisation to the 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 randomised controlled trials, where the intervention has little uptake, an especially important study challenge given the interdependence of exposure and outcomes within communities, and a problem that has characterised 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, and also mental, social and physical well-being.⁸⁷ Outcomes along the causal chain include standard, but more subjective measures, such as reported diarrhoeal 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, laboratory staff analysing 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 EED and measures of CE in an evaluation assessment. While there are limitations inherent to observational studies, the matched study design and multivariable modelling analysis plan reduce the potential for confounding. However, there is still the potential for residual unmeasured confounding.

Dissemination

Efforts will be made to communicate the central findings and implications with study communities, the implementing organisation 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.

Author affiliations

¹Department of Environmental Health, Rollins School of Public Health, Emory University, Atlanta, Georgia, USA

²London School of Hygiene and Tropical Medicine, London, UK

³Nutrition and Health Sciences Program, Laney Graduate School, Emory University, Atlanta, Georgia, USA

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

⁵Department of Biostatistics and Bioinformatics, Rollins School of Public Health, Emory University, Atlanta, Georgia, USA

Contributors TC, HR, PR, BT and HHC contributed to study design. HR, LZ and BT developed laboratory protocols. HR, BT, GS, MGD, SSS, LZ and BC developed data collection tools. All authors contributed to editing and revising the manuscript.

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.

Ethics approval Ethics Committee of the London School of Hygiene and Tropical Medicine, UK and Institute Ethics Committee of the Kalinga Institute of Medical Sciences of KIIT University, Bhubaneswar, India.

Provenance and peer review Not commissioned; externally peer reviewed.

Open Access This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: <http://creativecommons.org/licenses/by/4.0/>

REFERENCES

1. UNICEF, WHO. *Progress and sanitation and drinking water: 2015 update and MDG assessment*. Geneva: World Health Organization, 2015.
2. Water and Sanitation Program. *A decade of the total sanitation campaign: rapid assessment of processes and outcomes*. World Bank, 2010.
3. Coffey D, Gupta A, Hathi P, *et al*. Culture and the health transition: understanding sanitation behavior in rural north India. International Growth Centre Working Paper, 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.
5. Arnold BF, Khush RS, Ramaswamy P, *et al*. Causal inference methods to study nonrandomized, preexisting development interventions. *Proc Natl Acad Sci USA* 2010;107:22605–10.
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.
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.
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.
9. Clasen T, Bostoen K, Schmidt W, *et al*. Interventions to improve disposal of human excreta for preventing diarrhoea. *Cochrane Database Syst Rev* 2010;(6):CD007180.
10. Clasen T, Schmidt WP, Rabie T, *et al*. Interventions to improve water quality for preventing diarrhoea: systematic review and meta-analysis. *BMJ* 2007;334:782.
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.
12. Clasen TF, Alexander KT, Sinclair D, *et al*. Interventions to improve water quality for preventing diarrhoea. *Cochrane Database Syst Rev* 2015;10:CD004794.
13. Routray P, Schmidt WP, 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.

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:e71438.
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.
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 Health* 2014;2:e645–53.
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* 2012;33:239–57.
18. Root GPM. Sanitation, community environments, and childhood 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*. U.S. Agency for International Development (USAID), 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. pii: e00034-10.
21. Bandura A. *Self-efficacy: the exercise of control*. Worth Publishers, 1997.
22. UN Inter-agency Group for Child Mortality Estimation. (UN IGME). Levels and trends in child mortality: Report 2015. UNICEF, 2015.
23. World Health Organization. *Estimated total deaths by cause, sex and WHO Member State, 2008*. Geneva, Switzerland: WHO, 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.
25. Korpe PS, Petri WA. Environmental enteropathy: critical implications of a poorly understood condition. *Trends Mol Med* 2012;18:328–36.
26. Guerrant RL, Oriá RB, Moore SR, *et al.* Malnutrition as an enteric infectious disease with long-term effects on child development. *Nutr Rev* 2008;66:487–505.
27. Crane RJ, Jones KDJ, Berkley JA. Environmental enteric dysfunction: an overview. *Food Nutr Bull* 2015;36(Suppl 1): S76–87.
28. Humphrey JH. Child undernutrition, tropical enteropathy, toilets, and handwashing. *Lancet* 2009;374:1032–5.
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.
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.
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.
32. Caruso BA, Sevilimedu V, Fung ICH, *et al.* Gender disparities in water, sanitation, and global health. *Lancet* 2015;386:650–1.
33. Sahoo KC, Hulland KR, Caruso BA, *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.
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:e0141883.
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. Gram Vikas. 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. NBER Working Paper, 2015:21521.
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.
39. International Institute for Population Sciences (IIPS). *District level household and facility survey (DLHS-3), 2007–08: India, Orissa*. Mumbai: International Institute for Population Studies (IIPS), 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 May 2014).
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.
43. Rubin DB. Using propensity scores to help design observational studies: application to the tobacco litigation. *Health Serv Outcomes Res Methodol* 2001;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;15:234–49.
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.
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.
47. Hartung C, Anokwa Y, Brunette W, *et al.* 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.
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.
50. World Health Organization. *Integrated management of childhood illness: chart booklet*. Geneva: World Health Organization, 2014.
51. Cogill B. *Anthropometric indicators measurement guide*. Revised edn. Food and Nutrition Technical Assistance Project. Washington (DC): Academy for Educational Development (AED), 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 AL, 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.
55. Crump JA, 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 waterborne pathogens: a seroepidemiologic study in children 6–36 months old in San Juan Sacatepequez, Guatemala. *Am J Trop Med Hyg* 2004;70:83–8.
57. Brüssow 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 NR, *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 BL, Gilman RH, Lanata CF, *et al.* Seroepidemiology of Cryptosporidium infection in two Latin American populations. *J Infect Dis* 1988;157:551–6.
61. Vitral CL, Yoshida CF, Lemos ER, *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.
63. Centers for Disease Control and Prevention. *Laboratory methods for the diagnosis of Vibio cholerae*. Atlanta (GA): Centers for Disease Control and Prevention (CDC), 1999.
64. APHA, AWWA, and WEF. Method 9260 E: Shigella. Standard Methods for the Examination of Water and Wastewater. 21st Edition, 2005. <http://www.standardmethods.org/>.

65. APHA, AWWA, and WEF. Method 10029: Total Coliforms and E. coli Membrane Filtration Method with m-ColiBlue24 Broth. Standard Methods for the Examination of Water and Wastewater. 21st Edition, 2005. <http://www.standardmethods.org/>
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.
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.
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.
69. Livio S, Strockbine NA, Panchalingam S, *et al.* Shigella isolates from the global enteric multicenter study inform vaccine development. *Clin Infect Dis* 2014;59:933–41.
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.
71. Kanungo S, Sah BK, Lopez AL, *et al.* Cholera in India: an analysis of reports, 1997–2006. *Bull World Health Organ* 2010;88:185–91.
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, 2010:493–529.
73. Edejer TT, Baltussen R, Adam T, *et al.* *Making choices in health: WHO guide to cost-effectiveness analysis*. Geneva: World Health Organization, 2003.
74. Malapit HJ, Sproule K, Kovarik C, *et al.* *Measuring progress toward empowerment: women's empowerment in agriculture index: baseline report*. IFPRI, 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.
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.
78. Dasgupta A, Sarkar M. Menstrual hygiene: how hygienic is the adolescent girl? *Indian J Community Med* 2008;33:77–80.
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.
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.
81. Caruso B. *Sanitation insecurity: definition, measurement, and associations with women's mental health in rural Orissa, India* [dissertation]. 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.
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;42:1511–19.
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.
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.
87. World Health Organization. WHO definition of health. <http://www.who.int/about/definition/en/print.html> (accessed 26 Apr 2016).