

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:	<i>BMJ Open</i>
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

1
2
3 1 **Design and rationale of a matched cohort study to assess the effectiveness of a combined**
4
5 2 **household-level piped water and sanitation intervention in rural Odisha, India**
6
7
8
9

10 4 Heather Reese,¹ Parimita Routray,² Belen Torondel,² Gloria Sclar,¹ Maryann G. Delea,^{1,2} Sheela
11
12 5 S. Sinharoy,³ Laura Zambrano,¹ Bethany Caruso,¹ Samir R. Mishra,⁴ Howard H. Chang,⁵
13
14 6 Thomas Clasen^{1,2}
15
16
17
18
19

20 8 ¹ Department of Environmental Health, Rollins School of Public Health, Emory University,
21
22 9 Atlanta, Georgia, United States of America
23

24 10 ² London School of Hygiene and Tropical Medicine, London, United Kingdom
25
26

27 11 ³ Nutrition and Health Sciences Program, Laney Graduate School, Emory University, Atlanta,
28
29 12 Georgia, United States of America
30

31 13 ⁴ School of Biotechnology, KIIT University, Bhubaneswar, Odisha, India
32
33

34 14 ⁵ Department of Biostatistics and Bioinformatics, Rollins School of Public Health, Emory
35
36 15 University, Atlanta, Georgia, United States of America
37
38

39 16
40
41 17 Contact information for corresponding author, Heather Reese:

42
43 18 Department of Environmental Health, Rollins School of Public Health, Emory University
44
45 19 1518 Clifton Rd, Atlanta, GA 30322
46
47

48 20 Email: heather.e.reese@emory.edu, Phone: (619) 379-8963
49

50
51 21 Keywords: rural, piped water, sanitation

52
53 22 Word count: 4422
54
55
56
57
58
59
60

1
2
3 23
4
5
6 24**ABSTRACT**

7
8 25 **Introduction:** Government efforts to address massive shortfalls in rural water and sanitation in
9
10 26 India have centered on construction of community water sources and toilets for selected
11
12 27 households. However, deficiencies with water quality and quantity at the household level, and
13
14 28 community coverage and actual use of toilets has led Gram Vikas, a local NGO in Odisha, India,
15
16 29 to develop an approach that provides household-level piped water connections contingent on full
17
18 30 community-level toilet coverage.
19

20
21 31 **Methods:** This matched cohort study was designed to assess the impact of a combined piped
22
23 32 water and sanitation intervention. Households with children under five years in 45 randomly
24
25 33 selected intervention villages and 45 matched control villages will be followed over 16 months.
26
27 34 The primary outcome is prevalence of diarrheal diseases; secondary health outcomes include
28
29 35 soil-transmitted helminth infection, nutritional status, seroconversion to enteric pathogens,
30
31 36 urogenital infections, and environmental enteric dysfunction. In addition, intervention effects on
32
33 37 sanitation and water coverage, access and use, environmental fecal contamination, women's
34
35 38 empowerment, as well as collective efficacy, and intervention cost and cost-effectiveness will be
36
37 39 assessed.
38
39

40 40 **Ethics and dissemination:** The study protocol has been reviewed and approved by the ethics
41
42 41 boards of the London School of Hygiene and Tropical Medicine, U.K. and KIIT University,
43
44 42 Bhubaneswar, India. Findings will be disseminated via peer-reviewed literature and presentation
45
46 43 to stakeholders, government officials, implementers and researchers.
47
48
49
50

51 44 **Trial registration identifier:** NCT02441699
52
53
54
55
56 45
57
58
59
60

46 STRENGTHS AND LIMITATIONS OF THIS STUDY

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

62 INTRODUCTION

63 Of the one billion people practicing open defecation worldwide, over half live in India[1]. While
64 international and national pressure on improving sanitation conditions in India has led to over
65 350 thousand people gaining access to improved toilets since 1990, it has barely kept up with
66 population growth[1,2]. Recent studies show that even in areas with access to household-level
67 improved sanitation, use of these toilets is low[3–5]. This may be due in part to a mismatch
68 between the culturally acceptable pour-flush toilets and the level of water access. Coverage of

1
2
3 69 improved water sources, usually community-level pumps or taps, is relatively high even in rural
4
5 70 areas in India, but it may not be sufficient for flushing purposes on top of other daily water
6
7
8 71 needs[1,6].
9

10 Although the effectiveness of water, sanitation and hygiene (WASH) interventions vary,
11
12 72 meta-analyses have found that individual or combined WASH interventions decrease diarrheal
13
14 73 disease prevalence by up to 48%[7–11]. While combined interventions would be expected to
15
16 74 have a greater influence on multiple exposure pathways and thus a greater combined impact on
17
18 75 health, there is limited evidence of additive benefits[12]. This may be due to poor uptake,
19
20 76 inconsistent use, or an incomplete understanding of relevant pathways[8–10]. In India, a
21
22 77 combined water and sanitation intervention may be more critical than just interrupting multiple
23
24 78 transmission pathways for enteric infection; evidence suggests that household-level water access
25
26 79 is integral to the use of improved sanitation in this context[13].
27
28
29
30
31

32 While the intent of improved sanitation facilities is to separate human feces from human
33
34 82 contact, most of the focus is on constructing household toilets to increase improved sanitation
35
36 83 coverage—the primary metric used in monitoring progress toward international targets.
37
38
39 84 However, studies in India have further shown that toilet construction does not translate into toilet
40
41 85 use in this context[5,14–16]. Moreover, with the interdependence between members of
42
43 86 households and households within communities, safe water and sanitation is a community-level
44
45
46 87 issue. There is growing emphasis on assessing health risk from poor water and sanitation
47
48 88 conditions not simply due to individual or even household-level risk factors, but also from
49
50 89 conditions in the community environment[17]. There is evidence that even households without
51
52 90 toilets, and households which do not filter drinking water, showed decreased health risk if they
53
54
55 91 live in communities with high levels of coverage and use[18–20].
56
57
58
59
60

1
2
3 92 A main risk of poor WASH conditions is enteric infection, caused by a diverse array of
4
5 93 bacteria, viruses, protozoa, and parasites, including soil-transmitted helminths. These infections
6
7
8 94 may cause diarrhea, the second leading cause of mortality for children under five years
9
10 95 worldwide and in India, a leading cause of mortality regardless of age[21,22]. There is also
11
12 96 growing evidence that asymptomatic enteric infections may pose a similar risk, with repeat
13
14 97 enteric infections contributing to chronic malnutrition, environmental enteric dysfunction, poor
15
16 98 cognitive outcomes, and poor vaccine uptake[23–28]. Poor WASH conditions are also linked to
17
18 99 increased risk of respiratory infection, the leading cause of mortality for children under five
19
20 100 years worldwide[21,29,30]. Poor water and sanitation access can also affect the social, physical
21
22 101 and mental well-being of women, acting through pathways ranging from unsafe menstrual
23
24 102 hygiene management practices and increased risk of violence[31–33].
25
26
27
28
29
30
31

32 104 **Description of the intervention**

33
34 105 Over the past decades there has been a global commitment to determine water and sanitation
35
36 106 interventions with demonstrated effectiveness, not just efficacy[34]. Gram Vikas, a non-
37
38 107 governmental organization based in Odisha, India (<http://www.gramvikas.org/>), has responded
39
40 108 by implementing its MANTRA (Movement and Action Network for Transformation of Rural
41
42 109 Areas) water and sanitation program in more than 1000 villages since 2002[35]. This approach
43
44 110 includes both household-level piped water connections, and community-level mobilization for
45
46 111 culturally appropriate household toilets. A previous interrupted time series analysis of the
47
48 112 MANTRA intervention reported it to be protective against diarrheal diseases[36]. However, in
49
50 113 addition to limitations of design, this study relied on outcome data collected and reported by
51
52
53
54
55
56
57
58
59
60

1
2
3 114 Gram Vikas, the intervention implementer, and did not assess intervention coverage or impacts
4
5
6 115 on environmental fecal contamination.
7

8 116 The MANTRA water and sanitation intervention is rolled out in a three-phase process over
9
10 117 an average three to five years (Figure 1). During the first, or Motivational, phase, representatives
11
12 118 of Gram Vikas visit the identified village several times to assess village interest and progress
13
14 119 towards a set of Gram Vikas requirements, including: 1) the commitment of every household to
15
16 120 participate, 2) creation of a village corpus fund from contributions from every household, and 3)
17
18 121 development of village guidelines for maintenance and use of facilities.
19

20 122 Once this set of requirements is achieved, the village progresses into the second, or
21
22 123 Operational, phase of the intervention. Each household constructs a pour-flush toilet with two
23
24 124 soak-pits and a separate bathing room. The households hire a local, skilled mason and provide
25
26 125 their own unskilled labor and locally available materials to complete the superstructure. Gram
27
28 126 Vikas provides external materials such as PVC pipes and porcelain pans. At the same time, a
29
30 127 water tank, community meeting space, and piped water distribution system connected to every
31
32 128 household, with taps in the toilet and bathing rooms and a separate tap in the kitchen, is
33
34 129 constructed through a similar collaborative process.
35
36
37
38
39

40
41 130 All households must construct a toilet and bathing room for the village to progress into
42
43 131 the final, or Completed, phase of the intervention, in which the water system is turned on.
44
45 132 Notably, this three-phase process only allows each household access to piped water once every
46
47 133 household in the village has a toilet and bathing room. This model contrasts with most previous
48
49 134 water and sanitation interventions, including those implemented under India's Total Sanitation
50
51 135 Campaign and other government programs, which do not require community-level sanitation
52
53 136 compliance and do not provide a piped water supply at the household level[37].
54
55
56
57
58
59
60

137

138 Study aims

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

141 Toward that objective, this study aims to:

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

157 METHODS**158 Setting**

1
2
3 159 The study is located in Ganjam and Gajapati districts located in eastern Odisha, India (Figure 2).
4
5
6 160 These two contiguous districts were a single district until 1992. Over 44% of the population in
7
8 161 these districts is recognized by the Government of India as being below the poverty line
9
10 162 (BPL)[38]. As of 2008, a majority of households in both districts had access to an improved,
11
12
13 163 likely community-level, drinking water source, with over 23% of households in Ganjam having
14
15 164 access to any sanitation facility, compared to only 8% of households in Gajapati[38]. The area is
16
17 165 primarily rural and agrarian, and the climate is characterized by a monsoon season from June to
18
19 166 September, with an average rainfall of ~1400 mms/year.
20
21
22 167
23
24
25 168

168 **Study design**

26
27 169 This study uses a matched cohort design with data collected across four study rounds from June
28
29 170 2015 to September 2016. As described below, control villages were matched to randomly
30
31 171 selected intervention villages through a multi-step restriction, genetic matching, and exclusion
32
33 172 process using the following eligibility criteria (Figure 3).
34
35
36 173
37
38

174 **Eligibility criteria for villages**

39 174
40
41 175 *1. Restriction.* Intervention villages were randomly selected from a list of Gram Vikas villages
42
43 176 in Ganjam and Gajapati districts provided by the NGO, after restriction to villages with a
44
45 177 Motivation phase start date between 2002-2006 and a Construction phase start date no earlier
46
47 178 than 2003. Since the intervention process takes on average three to five years, the criteria for the
48
49 179 Motivation start date helped to identify those villages with ongoing interventions at the same
50
51 180 time. In addition, this allowed the use of the Government of India Census 2001 and the Below
52
53
54
55
56
57
58
59
60

1
2
3 181 Poverty Line (BPL) Survey 2002 data to characterize baseline characteristics in both intervention
4
5
6 182 and control villages.

7
8 183 Eligible control villages include all villages without a Gram Vikas intervention within the
9
10 184 study districts which: 1) are not within the same Gram Panchayat (a political subdivision with
11
12 185 some administrative responsibility for water and sanitation) as a Gram Vikas village, or
13
14 186 bordering a Gram Vikas village, and 2) had not received a Motivation visit from the Gram Vikas
15
16 187 NGO. These criteria serve to limit the possibility of previous partial exposure to the intervention
17
18 188 through spillover from adjacent villages or direct contact with the NGO. These criteria also
19
20 189 increase strength of the counterfactual provided by the control villages, i.e. if they had received a
21
22 190 motivation visit from Gram Vikas, the control villages would have been equally as likely as the
23
24 191 intervention villages to demand the intervention.

25
26
27 192 In addition, to be eligible for inclusion both intervention and control villages must: 1)
28
29 193 appear in the Government of India Census 2001 and the BPL Survey 2002, 2) have a population
30
31 194 of at least 20 households, and 3) be within approximately three hours travel from the study office
32
33 195 in Brahmapur, Ganjam District. This last criterion is due to logistical constraints.

34
35
36 196 *2. Matching.* After restriction, genetic matching was used to match potential control villages to
37
38 197 the randomly selected intervention villages without replacement[5,39,40]. Villages were exact
39
40 198 matched on district to limit any political or large scale geographic variation between district
41
42 199 populations, and were also matched on pre-intervention demographic, socioeconomic, sanitation,
43
44 200 and water access characteristics[5]. These village level matching variables were selected due to
45
46 201 their theorized association with the primary outcome, diarrheal diseases, as well as data
47
48 202 availability.

1
2
3 203 3. *Exclusion.* The field team visited matched potential control villages and intervention villages
4
5 204 to assess suitability for the study through a rapid assessment interview with village leadership
6
7
8 205 and to ensure accessibility. Villages were excluded if they are not within three hours travel of the
9
10 206 field office in Brahmapur, had sustained major infrastructure damage due to a natural disaster, or
11
12 207 if there was a current or planned sanitation or water intervention by an organization external to
13
14 208 the village in the next 12 months. In addition, villages were excluded if there were fewer than
15
16 209 three children under five years old. As villages were removed from the pool of prospective
17
18 210 control villages, the matching and exclusion processes were repeated.

21
22 211 After matching and exclusion, covariate balance was assessed for all matching variables
23
24 212 through examination of balance measures[41–43]. Matching resulted in an improvement in
25
26 213 balance as assessed through comparison of several measures including q-q plots, Kolmogorov-
27
28 214 Smirnov bootstrap p-values, and standardized differences. After matching, there were no
29
30 215 significant differences between intervention and control groups (Table 1).
31
32
33
34 216

36 217 **Eligibility criteria for households**

38 218 Households within selected intervention and control villages will be eligible if they have at least
39
40 219 one child under 5 years old at time of enrollment, verified with birth or immunization card, and
41
42 220 expect to reside in the village for the duration of the study. If there are more than 40 eligible
43
44 221 households within a village, 40 will be randomly selected to be enrolled. Informed consent will
45
46 222 be obtained from the male and/or female head of the selected households. All children under five
47
48 223 years within each enrolled household are eligible and will not age-out over the course of the
49
50
51
52
53 224

225 **Table 1.** Pre-intervention characteristics and balance diagnostics before and after matching and
 226 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 [†] (%)	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

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

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

234

235

236

237 **Sample Size**

238 Sample size was determined through a simulation estimating the log odds of diarrheal disease
239 (the primary outcome) through a multilevel random effects model and parameterized with data
240 from a previous study in a neighboring district in Odisha[16]. Sample size estimates were also
241 checked with G*Power[44]. The simulation assumes a longitudinal 7-day period prevalence for
242 diarrhea of 8.8% in children under five years, a heterogeneity variance between villages of 0.07,
243 a heterogeneity variance between households of 0.57, and four study rounds. An effect size of
244 0.20 was selected for public health importance and based on estimates of effect from systematic
245 reviews of water and sanitation studies[45]. Assuming at least 80% power, 0.05 significance
246 level, 10% for loss to follow up, and at least one child per household, we estimate a sample size
247 of 45 villages per study arm and 26 households per village.

249 **Outcome Measurement**

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

259

1
2
3 260 Diarrheal Diseases
4

5 261 The primary outcome for this study is prevalence of diarrheal diseases, recorded as both daily
6
7 262 point prevalence over the previous three days and seven-day period prevalence, for all household
8
9
10 263 members in each of the four rounds. Although self-reported diarrhea is a subjective outcome
11
12 264 with a well-established risk of bias, three-day recall reduces recall bias[47,48]. Diarrheal disease
13
14 265 will be measured using the World Health Organization (WHO) definition of three or more loose
15
16 266 stools in a 24-hour period, with or without the presence of blood. Field workers will use a simple
17
18 267 calendar as a visual aid to help respondents with recall. Each household member will be asked to
19
20 268 recall his or her own disease status and the mother or primary caregiver will be asked to report
21
22 269 disease for children.
23
24
25
26

27 270

28
29 271 Respiratory infection
30

31 272 Prevalence of respiratory infections will be recorded as both daily point prevalence over the
32
33 273 previous three days and seven-day period prevalence for all household members in each round.
34
35 274 Respiratory infection is defined as the presence of cough and/or shortness of breath/difficulty
36
37 275 breathing according to WHO's Integrated Management of Childhood Illness (IMCI)[49]. The
38
39 276 full IMCI case definition for acute lower respiratory infection also includes measurement
40
41 277 respiratory rate and observation of chest indrawing, stridor and other danger signs; these criteria
42
43 278 were excluded from our definition as there is concern about the technical support required to
44
45 279 produce consistent and accurate data within this context[49]. Our definition will provide a broad
46
47 280 assessment of respiratory illness burden.
48
49
50
51
52

53 281

54
55 282 Nutritional Status
56
57
58
59
60

1
2
3 283 Children's height and weight will be measured in all rounds using standard anthropometric
4
5 284 measurement methods as established by WHO[50,51]. Field workers will be trained and
6
7
8 285 standardized[51]. Weight will be measured for all children under five years of age using Seca
9
10 286 385 digital scales, with 20g increment for weight below 20kg and a 50g increment for weight
11
12 287 between 20 and 50kg. Recumbent length will be measured for children less than two years of age
13
14 288 using Seca 417 measuring boards with 1mm increment. Standing height will be measured for
15
16 289 children two to five years of age using Seca 213 portable stadiometers with 1mm increment.
17
18
19 290 Height and weight will be used to calculate height-for-age z-scores (HAZ) and weight-for-height
20
21 291 z-scores (WHZ) based on WHO reference standards. A random subset of 10% of households will
22
23 292 receive back check visits each day to repeat height/length measurements to ensure inter-observer
24
25 293 reliability.
26
27
28
29
30
31

32 295 Soil-transmitted helminth infection

33
34 296 Stool samples will be collected in rounds 2 and 4 from all household members in a randomly
35
36 297 selected subset of households, and used to assess the presence and intensity of soil-transmitted
37
38 298 helminth (STH) infection. Formalin ether concentration and microscopy will be used to quantify
39
40 299 worms and ova for hookworms, *Ascaris lumbricoides*, and *Tricuris trichura*[52]. Quality
41
42 300 assurance will include independent duplicate assessment of all positive and 10% of negative
43
44 301 samples. After stool collection, each participant will be offered a single dose of Albendazole, a
45
46 302 broad-spectrum antihelminthic drug recommended by the Ministry of Health and Family
47
48 303 Welfare, Government of India. Stools collected in round 2 will allow for comparison of STH
49
50 304 infection prevalence between intervention and control villages, while the stool samples collected
51
52 305 approximately 8 months later in round 4 will provide a measure of re-infection rate.
53
54
55
56
57
58
59
60

1
2
3 306
4
5
6 307 Environmental enteric dysfunction
7
8 308 Stools from randomly selected children under two years old, collected in rounds 2 and 4, will be
9
10 309 used to assess environmental enteric dysfunction (EED) through quantification of biomarkers of
11
12 310 intestinal inflammation and permeability. Fecal myeloperoxidase (MPO), alpha-1-antitrypsin
13
14 311 (AAT), and neopterin (NEO), markers for neutrophil activity, intestinal permeability and TH1
15
16 312 immune activation, respectively, were selected for this study based on evidence of association
17
18 313 with EED, subsequent linear growth deficits, and household environmental fecal
19
20 314 contamination[23,24,53].
21
22
23
24
25
26

27 316 Seroconversion for enteric pathogens

28
29 317 Serological assays that assess antibody production against various enteric pathogens can provide
30
31 318 an objective measure of exposure to enteric infections[54]. Enrolling children aged 6 to 18
32
33 319 months will reduce the potential for interference from maternally acquired antibodies and permit
34
35 320 analysis of seroconversion data in a critical window for young children who experience higher
36
37 321 diarrheal disease morbidity and mortality before two years of age[55–60]. Children who are 6 to
38
39 322 12 months during round 2 will have capillary blood drawn by fingerstick or heelstick, as
40
41 323 appropriate, and will be visited again during round 4 for a second capillary blood sample. All
42
43 324 blood samples will be preserved on TropBio (Sydney, Australia) filter discs and stored within 7
44
45 325 days of collection at -20°C. Seroconversion against markers for norovirus, *Giardia intestinalis*,
46
47 326 *Cryptosporidium parvum*, *Entamoeba histolytica*, enterotoxigenic *E. coli* heat-labile enterotoxin
48
49 327 (ETEC-LT), *Salmonella* spp., *Campylobacter jejuni*, *Vibrio cholera*, and *Toxoplasma* spp. will
50
51 328 be compared using multiplex immunoassay technology on the Luminex xMAP platform[61].
52
53
54
55
56
57
58
59
60

329

330 Environmental fecal contamination

331 Field workers will collect samples of household stored drinking water and source water from a

332 random subset of households in each village in all study rounds, and child hand rinses in rounds

333 2 and 4. All water and hand rinse samples will be stored on ice during transport and analyzed

334 within 6 hours of collection using membrane filtration. Three assays will be used: 1) plating on

335 m-Coli Blue 24 (Millipore, Billerica, MA) for *E.coli* according to EPA Method 10029, 2)

336 alkaline peptone water enrichment prior to plating on thiosulfate citrate bile salts sucrose agar

337 and slide agglutination serotyping for *V. cholerae*, and 3) plating on xylose lysine desoxycholate

338 agar, and slide agglutination serotyping for *Shigella* spp.[62–64]. Source and stored water

339 samples will be assayed for *E. coli*, *Vibrio cholerae* and *Shigella* spp., and hand rinse samples

340 will be assayed for *E. coli* and *Shigella* spp. *E. coli* was selected as a standard non-human

341 specific indicator of fecal contamination, though the limitations of this indicator are well-

342 established[65–67]. In order to better characterize human fecal contamination of the household

343 environment, *Vibrio cholerae* and *Shigella* spp. were selected based on prevalence in southern

344 Asia, evidence of public health importance, and field laboratory limitations[68–70].

345

346 Cost and cost-effectiveness

347 Costs and potential cost savings (i.e., averted costs) associated with the intervention will be

348 assessed through an economic costing approach that recognizes and quantifies costs and benefits

349 from a societal perspective[71]. Data on program and point-of-delivery inputs will be collected at

350 household, community, and implementer levels. Field workers will administer community

351 surveys to a village leader, and household surveys to the household decision-maker for toilet

1
2
3 352 installation, in 20 randomly selected households in half of the intervention and control villages.

4
5 353 Surveys will collect data on household- and village-level inputs related to materials and labor

6
7
8 354 required to construct household toilets and wash rooms, the community water tank and

9
10 355 distribution system, and household water connections; longer-term water supply and toilet

11
12 356 maintenance costs; and financing required for this infrastructure as well as perceived benefits,

13
14 357 including averted social opportunity costs. Implementer inputs from Gram Vikas will be

15
16 358 collected through an enumeration exercise, interviews, and examination of the implementer's

17
18
19 359 financial records.

20
21
22 360

23
24 361 Collective efficacy

25
26 362 Collective efficacy (CE) is a latent construct comprised of the structural and cognitive

27
28 363 components that facilitate a community's shared belief in its ability to come together and execute

29
30 364 actions related to a common goal[72]. A review of the literature and established conceptual

31
32 365 frameworks will be performed to define the CE construct. A sequential exploratory mixed

33
34 366 qualitative and quantitative design will be used to develop and refine a scale to measure CE and

35
36 367 test hypotheses. Field workers will administer the refined, multi-item, Likert-type CE scale to

37
38 368 one randomly selected household member aged 18 years or older in each household.

39
40
41 369

42
43 370 Women's empowerment

44
45 371 Four dimensions of women's empowerment will be measured: group participation, leadership,

46
47 372 decision-making and freedom of movement. Group participation and leadership will be measured

48
49 373 using modules from the Women's Empowerment in Agriculture Index, which has been tested in

50
51 374 South Asia[73]. Decision-making and freedom of movement will be measured using questions

1
2
3 375 from the women's status module of Demographic and Health Surveys. These measures were
4
5 376 selected based on the importance of women's empowerment for child nutrition[74,75]. Women's
6
7
8 377 empowerment is conceptualized as both an outcome and a potential mediator along the pathway
9
10 378 between the Gram Vikas intervention and child health outcomes.
11

12
13 379

14 380 Menstrual hygiene management

15 381 Menstrual hygiene management practices vary worldwide and depend on personal preference,
16
17 382 socioeconomic status, local traditions and beliefs, and access to water and sanitation
18
19 383 resources[76]. Unhygienic washing practices are common in rural India and among women and
20
21 384 girls in lower socioeconomic groups, and may increase risk of urogenital infection[77–79].
22
23 385 However, the link between access to water and sanitation, menstrual hygiene management and
24
25 386 urogenital infections has been poorly studied. Household surveys will capture self-reported
26
27 387 urogenital infection, defined as at least one of the following symptoms: 1) abnormal vaginal
28
29 388 discharge (unusual texture and color/more abundant than normal), 2) burning or itching in the
30
31 389 genitalia, 3) burning or itching when urinating, or 4) genital sores[78].
32
33
34
35
36
37
38

39 390

40 391 Sanitation Insecurity

41 392 This study will assess the associations between sanitation access and sanitation insecurity with
42
43 393 mental health among women. In previous research in Odisha, a contextually specific definition
44
45 394 and measure for sanitation insecurity was developed, with associations between facets of
46
47 395 sanitation insecurity and mental health independent of sanitation facility access[80]. This
48
49 396 previously developed measure will be used to determine if levels of sanitation insecurity differ
50
51
52
53
54
55
56
57
58
59
60

397 between intervention and control villages and how it may be associated with mental health
398 outcomes, specifically well-being, anxiety, depression, and distress.

399

400 Fecal sludge management.

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

408

409 **STATISTICAL ANALYSES**

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

417 The impact of intervention on reducing environmental fecal contamination (aim 2), will
418 be assessed through two methods. First, hierarchical logistic and negative binomial multilevel
419 regression will be used to compare intervention versus control villages. Unadjusted models will

1
2
3 420 be presented along with models adjusting for covariates. Second, a stochastic microbial risk
4
5 421 framework will be used to assess differential fecal environmental contamination between
6
7
8 422 intervention and control villages.
9

10 423 The cost and cost-effectiveness of the intervention (aim 4) will be assessed in two steps.
11
12 424 Incremental intervention benefits will be ascertained by combining health benefit data, from
13
14 425 analysis of health outcome data and established averted cost data, with other averted social
15
16 426 opportunity costs. An incremental cost-effectiveness ratio, expressed in cost per disease-specific
17
18 427 DALY, will be calculated by dividing the incremental intervention costs by the incremental
19
20 428 intervention benefits.
21
22
23

24 429 The collective efficacy scale will be analyzed using factor and psychometric analyses to
25
26 430 identify an appropriate factor solution. Once a CE factor solution is identified, multilevel
27
28 431 regression will be used to assess the associations between CE and intervention effectiveness (aim
29
30 432 5).
31
32
33

34 433

35 434 **DISCUSSION**

36
37
38 435 This matched cohort study is one of the first to evaluate the effect of a rural combined
39
40 436 household-level piped water and sanitation intervention, implemented at the community level, on
41
42 437 a large scale. The matched design provides a rigorous means for estimating causal effects given
43
44 438 that randomization was not feasible due to the several year implementation process[5]. By
45
46 439 focusing on a completed intervention, it also avoids the risk presented by randomized controlled
47
48 440 trials, where the intervention has little uptake, an especially important study challenge given
49
50 441 interdependence of exposure and outcomes within communities, and a problem that has
51
52 442 characterized previous trials of sanitation interventions in India[15,16].
53
54
55
56
57
58
59
60

1
2
3 443 A strength of this study is assessment of health impacts using the holistic WHO definition
4
5
6 444 of health, including not just disease status, but also mental, social, and physical well-being[83].
7
8 445 Outcomes along the causal chain include standard, but more subjective measures, such as
9
10 446 reported diarrheal diseases and respiratory infection, as well as more objective measures such as
11
12 447 fecal environmental contamination, soil transmitted helminth infection, and anthropometry.
13
14 448 Although there is risk of response bias for reported outcomes, it is unlikely to be differential by
15
16 449 intervention status since the study team is not directly linked to Gram Vikas. Even though field
17
18 450 workers may be aware of village intervention status, lab staff analyzing water, hand rinse, stool,
19
20 451 and blood samples will be blinded. In addition, this study includes the more novel use of
21
22 452 seroconversion for enteric pathogens, biomarkers of environmental enteric dysfunction, and
23
24 453 measures of collective efficacy in an evaluation assessment. While there are limitations inherent
25
26 454 to observational studies, the matched study design and multivariate modeling analysis plan
27
28 455 reduces the potential for confounding. However, there is still the potential for residual
29
30 456 unmeasured confounding.
31
32
33
34
35
36
37
38

39 458 **Ethics and Dissemination.** This study has been reviewed and approved by the Ethics Committee
40
41 459 of the London School of Hygiene and Tropical Medicine, U.K (No. 9071) and Institute Ethics
42
43 460 Committee of the Kalinga Institute of Medical Sciences of KIIT University, Bhubaneswar, India
44
45 461 (KIMS/KIIT/IEC/053/2015). Efforts will be made to communicate the central findings and
46
47 462 implications with study communities, the implementing organization and government officials in
48
49 463 India. The results of this study will be submitted for publication in peer reviewed journals and
50
51 464 presented at conferences. The data collected in the study will be publicly available, with personal
52
53
54
55
56
57
58
59
60

1
2
3 465 identifiable data redacted, following the publication of the primary results within 24 months of
4
5
6 466 the final data collection date.
7

8 467
9
10 468 **Funding.** This study is supported by a grant from the Bill & Melinda Gates Foundation to the
11
12 469 London School of Hygiene & Tropical Medicine (OPP1008048) and to Emory University.
13
14 470 (OOP1125067).
15
16
17

18 471
19
20 472 **Competing Interests:** None declared.
21

22 473
23
24 474 **Contributions from authors:** TC, HR, PR, BT, and HC contributed to study design. HR, LZ
25
26 475 and BT developed laboratory protocols. HR, BT, GS, MD, SS, LZ, and BC developed data
27
28 476 collection tools. All authors contributed to editing and revising the manuscript.
29
30
31

32 477
33
34 478
35
36 479
37
38 480 **Figure 1.** Description of the three phases of the Gram Vikas MANTRA water and sanitation
39
40 481 intervention.
41
42
43 482

44
45 483 **Figure 2.** Study sites in Ganjam and Gajapati districts, Odisha, India with intervention villages
46
47 484 in black and control villages in white. Inset shows location of districts in India.
48
49
50

51 485
52 486 **Figure 3.** Restriction, matching and exclusion process for selection of intervention and control
53
54 487 villages (1), and timeline for study data collection (2).
55
56
57 488

489 **REFERENCES**

- 490 1 UNICEF, WHO. Progress and Sanitation and Drinking Water: 2015 Update and MDG
491 Assessment. 2015.
- 492 2 Water and Sanitation Program. A Decade of the Total Sanitation Campaign: Rapid
493 Assessment of Processes and Outcomes. 2010.
- 494 3 Coffey D, Gupta A, Hathi P, *et al*. Culture and the health transition: Understanding
495 sanitation behavior in rural north India. 2015.
- 496 4 Clasen T, Pruss-Ustun A, Mathers CD, *et al*. Estimating the impact of unsafe water,
497 sanitation and hygiene on the global burden of disease: evolving and alternative methods.
498 *Trop Med Int Heal* 2014;**19**:884–93. doi:10.1111/tmi.12330
- 499 5 Arnold BF, Khush RS, Ramaswamy P, *et al*. Causal inference methods to study
500 nonrandomized, preexisting development interventions. *Proc Natl Acad Sci U S A*
501 2010;**107**:22605–10. doi:10.1073/pnas.1008944107
- 502 6 Ercumen A, Arnold BF, Kumpel E, *et al*. Upgrading a Piped Water Supply from
503 Intermittent to Continuous Delivery and Association with Waterborne Illness: A Matched
504 Cohort Study in Urban India. *PLOS Med* 2015;**12**:e1001892.
505 doi:10.1371/journal.pmed.1001892
- 506 7 Fewtrell L, Kaufmann RB, Kay D, *et al*. Water, sanitation, and hygiene interventions to
507 reduce diarrhoea in less developed countries: a systematic review and meta-analysis.
508 *Lancet Infect Dis* 2005;**5**:42–52. doi:10.1016/S1473-3099(04)01253-8
- 509 8 Engell RE, Lim SS. Does clean water matter? An updated meta-analysis of water supply
510 and sanitation interventions and diarrhoeal diseases. *Lancet* 2013;**381**:S44.
511 doi:10.1016/S0140-6736(13)61298-2

- 1
2
3 512 9 Clasen T, Bostoen K, Schmidt W, *et al.* Interventions to improve disposal of human
4
5
6 513 excreta for preventing diarrhoea. *Cochrane database Syst Rev* Published Online First:
7
8 514 2010.<http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD007180.pub2/pdf/standard>
9
10 515 (accessed 4 Jan2015).
- 11
12 516 10 Clasen T, Schmidt W-P, Rabie T, *et al.* Interventions to improve water quality for
13
14 517 preventing diarrhoea: systematic review and meta-analysis. *BMJ* 2007;**334**:782.
15
16 518 doi:10.1136/bmj.39118.489931.BE
- 17
18 519 11 Cairncross S, Hunt C, Boisson S, *et al.* Water, sanitation and hygiene for the prevention of
19
20 520 diarrhoea. *Int J Epidemiol* 2010;**39 Suppl 1**:i193–205. doi:10.1093/ije/dyq035
- 21
22 521 12 Clasen TF, Alexander KT, Sinclair D, *et al.* Interventions to improve water quality for
23
24 522 preventing diarrhoea. *Cochrane database Syst Rev* 2015;**10**:CD004794.
25
26 523 doi:10.1002/14651858.CD004794.pub3
- 27
28 524 13 Routray P, Schmidt W-P, Boisson S, *et al.* Socio-cultural and behavioural factors
29
30 525 constraining latrine adoption in rural coastal Odisha: an exploratory qualitative study.
31
32 526 *BMC Public Health* 2015;**15**:880. doi:10.1186/s12889-015-2206-3
- 33
34 527 14 Barnard S, Routray P, Majorin F, *et al.* Impact of Indian Total Sanitation Campaign on
35
36 528 Latrine Coverage and Use: A Cross-Sectional Study in Orissa Three Years following
37
38 529 Programme Implementation. *PLoS One* 2013;**8**. doi:10.1371/journal.pone.0071438
- 39
40 530 15 Patil SR, Arnold BF, Salvatore AL, *et al.* The effect of India's total sanitation campaign
41
42 531 on defecation behaviors and child health in rural Madhya Pradesh: a cluster randomized
43
44 532 controlled trial. *PLoS Med* 2014;**11**:e1001709. doi:10.1371/journal.pmed.1001709
- 45
46 533 16 Clasen T, Boisson S, Routray P, *et al.* Effectiveness of a rural sanitation programme on
47
48 534 diarrhoea, soil-transmitted helminth infection, and child malnutrition in Odisha, India: a
49
50
51
52
53
54
55
56
57
58
59
60

- 1
2
3 535 cluster-randomised trial. *Lancet Glob Heal* 2014;**2**:e645–53. doi:10.1016/S2214-
4
5 536 109X(14)70307-9
6
7
8 537 17 Eisenberg JNS, Trostle J, Sorensen RJD, *et al*. Towards a Systems Approach to Enteric
9
10 538 Pathogen Transmission: From Individual Independence to Community Interdependence.
11
12 539 *Annu Rev Public Health* 2013;**33**:239–57. doi:10.1016/j.micinf.2011.07.011.Innate
13
14 540 18 Root GPM. Sanitation, Community Environments, and Childhood Diarrhea in Rural
15
16 541 Zimbabwe. *J Heal Popul Nutr* 2001;**19**:73–82.
17
18 542 19 Bateman OM, Smith S. A Comparison of the Health Effects of Water Supply and
19
20 543 Sanitation in Urban and Rural Guatemala. 1999.
21
22 544 20 Huq A, Yunus M, Sohel SS, *et al*. Simple Sari Cloth Filtration of Water Is Sustainable and
23
24 545 Continues to Protect Villagers from Cholera in Matlab, Bangladesh. *MBio* 2010;**1**:1–5.
25
26 546 doi:10.1128/mBio.00034-10.Invited
27
28 547 21 UN Inter-agency Group for Child Mortality Estimation. Levels and trends in child
29
30 548 mortality: Report 2015. 2015.
31
32 549 22 World Health Organization. Estimated total deaths by cause, sex and WHO Member State,
33
34 550 2008. Geneva, Switzerland: 2011.
35
36 551 23 Lin A, Arnold BF, Afreen S, *et al*. Household environmental conditions are associated
37
38 552 with enteropathy and impaired growth in rural Bangladesh. *Am J Trop Med Hyg*
39
40 553 2013;**89**:130–7. doi:10.4269/ajtmh.12-0629
41
42 554 24 Korpe PS, Petri WA. Environmental enteropathy: critical implications of a poorly
43
44 555 understood condition. *Trends Mol Med* 2012;**18**:328–36.
45
46 556 doi:10.1016/j.molmed.2012.04.007
47
48 557 25 Guerrant RL, Oriá RB, Moore SR, *et al*. Malnutrition as an enteric infectious disease with
49
50
51
52
53
54
55
56
57
58
59
60

- 1
2
3 558 long-term effects on child development. 2008;**66**:487–505. doi:10.1111/j.1753-
4
5 559 4887.2008.00082.x.Malnutrition
6
7
8 560 26 Crane RJ, Jones KDJ, Berkley JA. Environmental enteric dysfunction : An overview.
9
10 561 2015;**36**:76–87.
11
12 562 27 Humphrey JH. Child undernutrition, tropical enteropathy, toilets, and handwashing.
13
14 563 *Lancet* 2009;**374**:1032–5. doi:10.1016/S0140-6736(09)60950-8
15
16
17 564 28 Mbuya MNN, Humphrey JH. Preventing environmental enteric dysfunction through
18
19 565 improved water, sanitation and hygiene: An opportunity for stunting reduction in
20
21 566 developing countries. *Matern Child Nutr* 2015;**12**:1–15. doi:10.1111/mcn.12220
22
23
24 567 29 Budge PJ, Griffin MR, Edwards KM, *et al.* Impact of home environment interventions on
25
26 568 the risk of influenza-associated ARI in Andean children: observations from a prospective
27
28 569 household-based cohort study. *PLoS One* 2014;**9**:e91247.
29
30 570 doi:10.1371/journal.pone.0091247
31
32
33 571 30 Aiello AE, Coulborn RM, Perez V, *et al.* Effect of hand hygiene on infectious disease risk
34
35 572 in the community setting: a meta-analysis. *Am J Public Health* 2008;**98**:1372–81.
36
37 573 doi:10.2105/AJPH.2007.124610
38
39
40 574 31 Caruso BA, Sevilimedu V, Fung IC-H, *et al.* Gender disparities in water, sanitation, and
41
42 575 global health. *Lancet (London, England)* 2015;**386**:650–1. doi:10.1016/S0140-
43
44 576 6736(15)61497-0
45
46
47 577 32 Sahoo KC, Hulland KR, Caruso B a., *et al.* Sanitation-related psychosocial stress: a
48
49 578 grounded theory study of women across the life-course in Odisha, India. *Soc Sci Med*
50
51 579 2015;**139**:80–9. doi:10.1016/j.socscimed.2015.06.031
52
53
54 580 33 Hulland KRS, Chase RP, Caruso BA, *et al.* Sanitation, stress, and life stage: A systematic
55
56
57
58
59
60

- 1
2
3 581 data collection study among women in Odisha, India. *PLoS One* 2015;**10**.
4
5
6 582 doi:10.1371/journal.pone.0141883
7
8 583 34 World Health Organization/The United Nations Children's Fund (UNICEF). Ending
9
10 584 Preventable Child Deaths from Pneumonia and Diarrhoea by 2025: The integrated Global
11
12 585 Action Plan for Pneumonia and Diarrhoea (GAPPD). 2013.
13
14
15 586 35 Gram Vikas. Annual Report 2013-2014. 2014.
16
17 587 36 Duflo E, Greenstone M, Guiteras R, *et al*. Toilets Can Work : Short and Medium Run
18
19 588 Health Impacts of Addressing Complementarities and Externalities in Water and
20
21 589 Sanitation.
22
23
24 590 37 Rosa G, Clasen T. Estimating the scope of household water treatment in low- and
25
26 591 medium-income countries. *Am J Trop Med Hyg* 2010;**82**:289–300.
27
28 592 doi:10.4269/ajtmh.2010.09-0382
29
30
31 593 38 International Institute for Population Sciences (IIPS). District Level Household and
32
33 594 Facility Survey (DLHS-3), 2007-08: India, Orissa. Mumbai: 2010.
34
35
36 595 39 Brady H, Caughey D, Dehejia R, *et al*. Genetic Matching for Estimating Causal Effects :
37
38 596 2012.
39
40
41 597 40 Sekhon JS. "Multivariate and Propensity Score Matching Software. *J Stat Softw*
42
43 598 2011;**42**.[http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Multivariate+a](http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Multivariate+and+Propensity+Score+Matching#9)
44
45 599 [nd+Propensity+Score+Matching#9](http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Multivariate+and+Propensity+Score+Matching#9) (accessed 30 May2014).
46
47
48 600 41 Austin PC. Balance diagnostics for comparing the distribution of baseline covariates
49
50 601 between treatment groups in propensity-score matched samples. *Stat Med* 2009;**28**:3083–
51
52 602 107. doi:10.1002/sim
53
54
55 603 42 Rubin DB. Using propensity scores to help design observational studies: Application to

- 1
2
3 604 the tobacco litigation. *Health Serv Outcomes Res Methodol* 2002;**2**:169–88.
- 4
5 605 43 Harder VS, Stuart EA, Anthony JC. Propensity score techniques and the assessment of
6
7
8 606 measured covariate balance to test causal associations in psychological research. *Psychol*
9
10 607 *Methods* 2010;**15**:997–1003. doi:10.1016/j.biotechadv.2011.08.021.Secreted
- 11
12 608 44 Faul F, Erdfelder E, Buchner A, *et al*. Statistical power analyses using G*Power 3.1: tests
13
14 609 for correlation and regression analyses. *Behav Res Methods* 2009;**41**:1149–60.
15
16 610 doi:10.3758/BRM.41.4.1149
- 17
18 611 45 Wolf J, Prüss-Ustün A, Cumming O, *et al*. Systematic review: Assessing the impact of
19
20 612 drinking water and sanitation on diarrhoeal disease in low- and middle-income settings:
21
22 613 Systematic review and meta-regression. *Trop Med Int Heal* 2014;**19**:928–42.
23
24 614 doi:10.1111/tmi.12331
- 25
26 615 46 Carl Hartung YAWBALCTGB. Open Data Kit: Tools to Build Information Services for
27
28 616 Developing Regions. <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.176.8017>
- 29
30 617 47 Feikin DR, Audi a., Olack B, *et al*. Evaluation of the optimal recall period for disease
31
32 618 symptoms in home-based morbidity surveillance in rural and urban Kenya. *Int J*
33
34 619 *Epidemiol* 2010;**39**:450–8. doi:10.1093/ije/dyp374
- 35
36 620 48 Arnold BF, Galiani S, Ram PK, *et al*. Optimal Recall Period for Caregiver-reported Illness
37
38 621 in Risk Factor and Intervention Studies: A Multicountry Study. *Am J Epidemiol*
39
40 622 2013;**177**:361–70. doi:10.1093/aje/kws281
- 41
42 623 49 World Health Organization. Integrated Management of Childhood Illness: Chart Booklet.
43
44 624 50 Cogill B. Anthropometric indicators measurement guide. Revised edition. *Washington,*
45
46 625 *DC, Acad Educ Dev [AED], Food Nutr Tech Assist Proj*
47
48 626 2003;:92.http://www.developmentgateway.org/download/202582/anthro_2003.pdf
49
50
51
52
53
54
55
56
57
58
59
60

- 1
2
3 627 51 de Onis M, Onyango AW, Van den Broeck J, *et al.* Measurement and standardization
4
5
6 628 protocols for anthropometry used in the construction of a new international growth
7
8 629 reference. *Food Nutr Bull* 2004;**25**:S27–36.
9
10
11 630 52 Truant a. L, Elliott SH, Kelly MT, *et al.* Comparison of formalin-ethyl ether
12
13 631 sedimentation, formalin-ethyl acetate sedimentation, and zinc sulfate flotation techniques
14
15 632 for detection of intestinal parasites. *J Clin Microbiol* 1981;**13**:882–4.
16
17
18 633 53 Kosek M, Haque R, Lima A, *et al.* Fecal markers of intestinal inflammation and
19
20 634 permeability associated with the subsequent acquisition of linear growth deficits in
21
22 635 infants. *Am J Trop Med Hyg* 2013;**88**:390–6. doi:10.4269/ajtmh.2012.12-0549
23
24
25 636 54 Crump J a, Mendoza CE, Priest JW, *et al.* Comparing serologic response against enteric
26
27 637 pathogens with reported diarrhea to assess the impact of improved household drinking
28
29 638 water quality. *Am J Trop Med Hyg* 2007;**77**:136–41.
30
31
32 639 55 Steinberg EB, Mendoza CE, Glass R, *et al.* Prevalence of Infection with Waterborne
33
34 640 Pathogens: a Seroepidemiologic Study in Children 6-36 Months Old in San Juan
35
36 641 Sacatepquez, Guatemala. 2004;**70**:83–8.
37
38
39 642 56 Brussow H, Sidoti J, Link H, *et al.* Age-specific prevalence of antibody to enterotoxigenic
40
41 643 *Escherichia coli* in Ecuadorian and German children. *J Infect Dis* 1990;**162**:974–7.
42
43
44 644 57 Khanna B, Cutler A, Israel N, *et al.* Use caution with serologic testing for *Helicobacter*
45
46 645 *pylori* infection in children. *J Infect Dis* 1998;**178**:460–5.
47
48
49 646 58 Lindkvist P, Asrat D, Nilsson I. Age at acquisition of *Helicobacter pylori* infection:
50
51 647 comparison of a high and a low prevalence country. *Scand J Infect Dis* 1996;**28**:181–4.
52
53
54 648 59 Ungar B, Gilman R, Lanata C, *et al.* Seroepidemiology of *Cryptosporidium* infection in
55
56 649 two Latin American populations. *J Infect Dis* 1988;**157**:551–6.
57
58
59
60

- 1
2
3 650 60 Vitral C, Yoshida C, Lemos E, *et al.* Age-specific prevalence of antibodies to hepatitis A
4
5
6 651 in children and adolescents from Rio de Janeiro, Brazil, 1978 and 1995: relationship of
7
8 652 prevalence to environmental factors. *Mem Inst Oswaldo Cruz* 1998;**93**:1–5.
9
10 653 61 Lammie PJ, Moss DM, Brook Goodhew E, *et al.* Development of a new platform for
11
12 654 neglected tropical disease surveillance. *Int J Parasitol* 2012;**42**:797–800.
13
14 655 doi:10.1016/j.ijpara.2012.07.002
15
16
17 656 62 Centers for Disease Control and Prevention. Laboratory Methods for the Diagnosis of
18
19 657 *Vibrio cholerae*. Atlanta, Georgia: 1999.
20
21
22 658 63 Centers for Disease Control and Prevention. Isolation and Identification of *Shigella*.
23
24 659 Atlanta, Georgia:
25
26
27 660 64 United States Environmental Protection Agency. Coliforms—Total and *E. coli*, Membrane
28
29 661 Filtration Method 10029. 1999.
30
31
32 662 65 Gronewold AD, Borsuk ME, Wolpert RL, *et al.* An Assessment of Fecal Indicator
33
34 663 Bacteria-Based Water Quality Standards. *Environ Sci Technol* 2008;**42**:4676–82.
35
36 664 doi:10.1021/es703144k
37
38
39 665 66 Gruber JS, Ercumen A, Colford JM. Coliform bacteria as indicators of diarrheal risk in
40
41 666 household drinking water: systematic review and meta-analysis. *PLoS One*
42
43 667 2014;**9**:e107429. doi:10.1371/journal.pone.0107429
44
45
46 668 67 Levy K, Nelson KL, Hubbard A, *et al.* Rethinking indicators of microbial drinking water
47
48 669 quality for health studies in tropical developing countries: case study in northern coastal
49
50 670 Ecuador. *Am J Trop Med Hyg* 2012;**86**:499–507. doi:10.4269/ajtmh.2012.11-0263
51
52
53 671 68 Livio S, Strockbine N a, Panchalingam S, *et al.* *Shigella* Isolates From the Global Enteric
54
55 672 Multicenter Study Inform Vaccine Development. *Clin Infect Dis* 2014;**59**.

- 1
2
3 673 doi:10.1093/cid/ciu468
4
5
6 674 69 Kotloff KL, Nataro JP, Blackwelder WC, *et al.* Burden and aetiology of diarrhoeal disease
7
8 675 in infants and young children in developing countries (the Global Enteric Multicenter
9
10 676 Study, GEMS): a prospective, case-control study. *Lancet* 2013;**382**:209–22.
11
12 677 doi:10.1016/S0140-6736(13)60844-2
13
14
15 678 70 Kanungo S, Sah BK, Lopez a. L, *et al.* Cholera in India: An analysis of reports, 1997-
16
17 679 2006. *Bull World Health Organ* 2010;**88**:185–91. doi:10.2471/BLT.09.073460
18
19
20 680 71 Cellini SR, Kee JE. Cost-Effectiveness and Cost-Benefit Analysis. In: Wholey JS, Hatry
21
22 681 HP, Newcomer KE, eds. *Handbook of Practical Program Evaluation*. San Francisco: :
23
24 682 John Wiley & Sons, Inc. 2010.
25
26
27 683 72 Bandura A. *Self-Efficacy: The Exercise of Control*. Worth Publishers 1997.
28
29 684 73 Malapit HJ, Sproule K, Kovarik C, *et al.* Measuring progress toward empowerment
30
31 685 Women’s Empowerment in Agriculture Index: Baseline Report. *IfpriOrg* 2014;:1–60.
32
33
34 686 74 Black RE, Victora CG, Walker SP, *et al.* Maternal and child undernutrition and
35
36 687 overweight in low-income and middle-income countries. *Lancet* 2013;**382**:427–51.
37
38 688 doi:10.1016/S0140-6736(13)60937-X
39
40
41 689 75 Smith LC, Ramakrishnan U, Ndiaye A, *et al.* The Importance of Women’s Status for
42
43 690 Child Nutrition in Developing Countries. 2003.
44
45 691 <http://www.ifpri.org/publication/importance-womens-status-child-nutrition-developing->
46
47 692 [countries](http://www.ifpri.org/publication/importance-womens-status-child-nutrition-developing-)
48
49
50
51 693 76 Sumpter C, Torondel B. A Systematic Review of the Health and Social Effects of
52
53 694 Menstrual Hygiene Management. *PLoS One* 2013;**8**:e62004.
54
55 695 doi:10.1371/journal.pone.0062004
56
57
58
59
60

- 1
2
3 696 77 Dasgupta A, Sarkar M. Menstrual Hygiene: How Hygienic is the Adolescent Girl? *Indian*
4
5 697 *J Community Med* 2008;**33**:77–80. doi:10.4103/0970-0218.40872
6
7
8 698 78 Das P, Baker KK, Dutta A, *et al.* Menstrual Hygiene Practices, WASH Access and the
9
10 699 Risk of Urogenital Infection in Women from Odisha, India. *PLoS One* 2015;**10**:e0130777.
11
12 700 doi:10.1371/journal.pone.0130777
13
14
15 701 79 van Eijk AM, Sivakami M, Thakkar MB, *et al.* Menstrual hygiene management among
16
17 702 adolescent girls in India: a systematic review and meta-analysis. *BMJ Open*
18
19 703 2016;**6**:e010290. doi:10.1136/bmjopen-2015-010290
20
21
22 704 80 Caruso B. *Sanitation Insecurity: Definition, Measurement, and Associations with*
23
24 705 *Women's Mental Health in Rural Orissa, India.*
25
26 706 2015.<http://pid.emory.edu/ark:/25593/rfhnt>
27
28
29 707 81 Peal A, Evans B, Blackett I, *et al.* Fecal sludge management (FSM): analytical tools for
30
31 708 assessing FSM in cities. *J Water, Sanit Hyg Dev* 2014;**4**:371.
32
33 709 doi:10.2166/washdev.2014.139
34
35
36 710 82 Peal A, Evans B, Blackett I, *et al.* Fecal Sludge Management: a comparative analysis of
37
38 711 12 cities. *J Water, Sanit Hyg Dev* 2014;**4**:563–75.
39
40
41 712 83 World Health Organization. WHO definition of health.
42
43 713 <http://www.who.int/about/definition/en/print.html> (accessed 26 Apr2016).
44
45
46 714

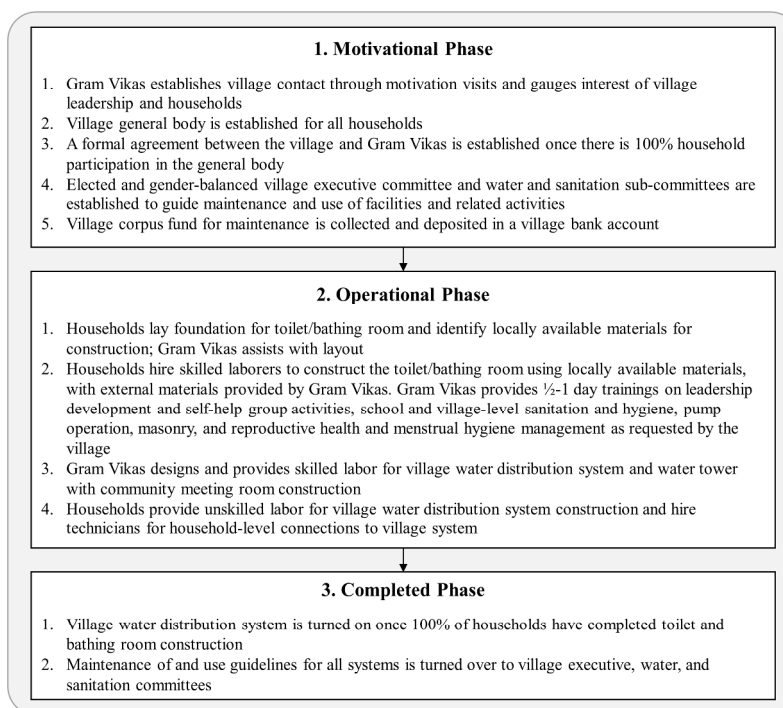


Figure 1. Description of the three phases of the Gram Vikas MANTRA water and sanitation intervention.

Figure 1

254x190mm (300 x 300 DPI)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

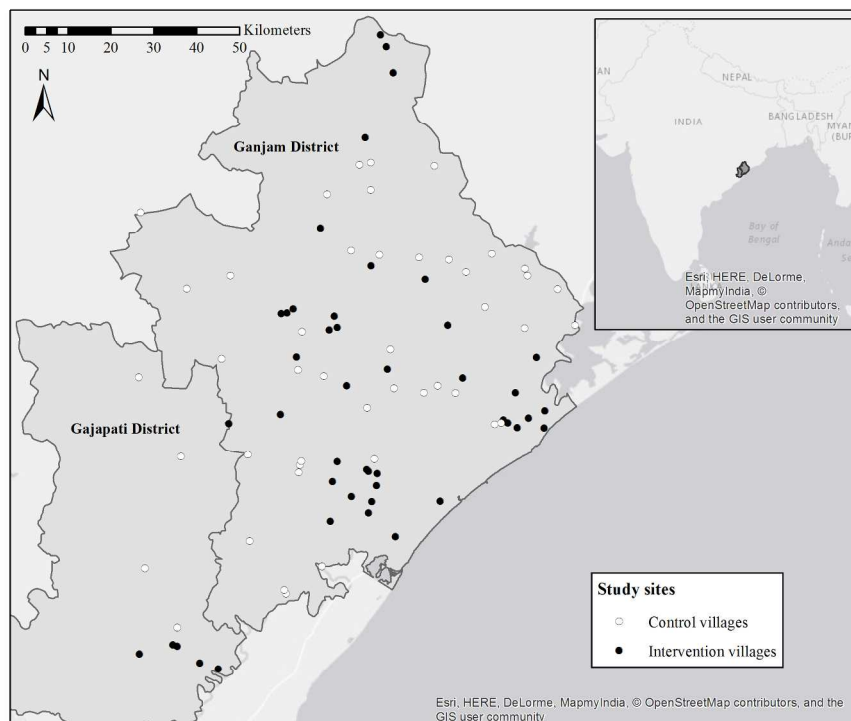


Figure 1. Study sites in Ganjam and Gajapati districts, Odisha, India with intervention villages in black and control villages in white. Inset shows location of districts in India.

Figure 1
215x279mm (300 x 300 DPI)

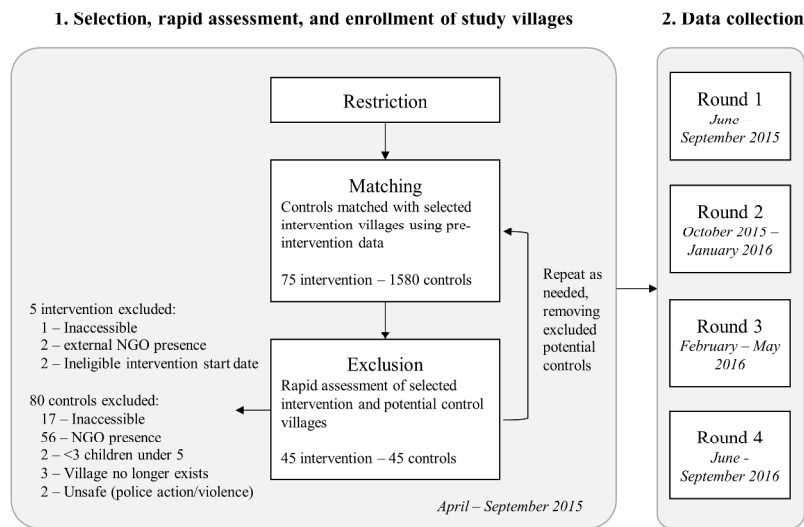


Figure 3. Restriction, matching and exclusion process for selection of intervention and control villages (1), and timeline for study data collection (2).

Figure 3
254x190mm (300 x 300 DPI)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

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) <i>Cohort study</i> —Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up <i>Case-control study</i> —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 <i>Cross-sectional study</i> —Give the eligibility criteria, and the sources and methods of selection of participants	8-11	
		(b) <i>Cohort study</i> —For matched studies, give matching criteria and number of exposed and unexposed <i>Case-control study</i> —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

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49

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) <i>Cohort study</i> —If applicable, explain how loss to follow-up was addressed <i>Case-control study</i> —If applicable, explain how matching of cases and controls was addressed <i>Cross-sectional study</i> —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) <i>Cohort study</i> —Summarise follow-up time (eg, average and total amount)	NA
Outcome data	15*	<i>Cohort study</i> —Report numbers of outcome events or summary measures over time	NA
		<i>Case-control study</i> —Report numbers in each exposure category, or summary measures of exposure	NA
		<i>Cross-sectional study</i> —Report numbers of outcome events or summary measures	NA
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:	<i>BMJ Open</i>
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

SCHOLARONE™
Manuscripts

1
2
3 1 **Design and rationale of a matched cohort study to assess the effectiveness of a combined**
4
5 2 **household-level piped water and sanitation intervention in rural Odisha, India**
6
7
8
9

10 4 Heather Reese,¹ Parimita Routray,² Belen Torondel,² Gloria Sclar,¹ Maryann G. Delea,^{1,2} Sheela
11
12 5 S. Sinharoy,³ Laura Zambrano,¹ Bethany Caruso,¹ Samir R. Mishra,⁴ Howard H. Chang,⁵
13
14 6 Thomas Clasen^{1,2}
15
16
17
18
19

20 8 ¹ Department of Environmental Health, Rollins School of Public Health, Emory University,
21
22 9 Atlanta, Georgia, United States of America
23

24 10 ² London School of Hygiene and Tropical Medicine, London, United Kingdom
25
26

27 11 ³ Nutrition and Health Sciences Program, Laney Graduate School, Emory University, Atlanta,
28
29 12 Georgia, United States of America
30

31 13 ⁴ School of Biotechnology, KIIT University, Bhubaneswar, Odisha, India
32
33

34 14 ⁵ Department of Biostatistics and Bioinformatics, Rollins School of Public Health, Emory
35
36 15 University, Atlanta, Georgia, United States of America
37
38
39
40

41 17 Contact information for corresponding author, Heather Reese:

42
43 18 Department of Environmental Health, Rollins School of Public Health, Emory University
44
45 19 1518 Clifton Rd, Atlanta, GA 30322
46
47

48 20 Email: heather.e.reese@emory.edu, Phone: (619) 379-8963
49

50
51 21 Keywords: rural, piped water, sanitation

52
53 22 Word count: 5442
54
55
56
57
58
59
60

1
2
3 **23 ABSTRACT**
4

5 **24 Introduction:** Government efforts to address massive shortfalls in rural water and sanitation in
6
7
8 India have centered on construction of community water sources and toilets for selected
9
10
11 households. However, deficiencies with water quality and quantity at the household level, and
12
13
14 community coverage and actual use of toilets has led Gram Vikas, a local NGO in Odisha, India,
15
16
17 to develop an approach that provides household-level piped water connections contingent on full
18
19
20 community-level toilet coverage.

21 **30 Methods:** This matched cohort study was designed to assess the effectiveness of a combined
22
23
24 piped water and sanitation intervention. Households with children under five years in 45
25
26
27 randomly selected intervention villages and 45 matched control villages will be followed over 17
28
29
30 months. The primary outcome is prevalence of diarrheal diseases; secondary health outcomes
31
32
33 include soil-transmitted helminth infection, nutritional status, seroconversion to enteric
34
35
36 pathogens, urogenital infections, and environmental enteric dysfunction. In addition, intervention
37
38
39 effects on sanitation and water coverage, access and use, environmental fecal contamination,
40
41
42 women's empowerment, as well as collective efficacy, and intervention cost and cost-
43
44
45 effectiveness will be assessed.

46 **39 Ethics and dissemination:** The study protocol has been reviewed and approved by the ethics
47
48
49 boards of the London School of Hygiene and Tropical Medicine, U.K. and KIIT University,
50
51
52 Bhubaneswar, India. Findings will be disseminated via peer-reviewed literature and presentation
53
54
55 to stakeholders, government officials, implementers and researchers.
56
57
58

59 **43 Trial registration identifier:** NCT02441699
60

46 STRENGTHS AND LIMITATIONS OF THIS STUDY

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

62 INTRODUCTION

63 Of the one billion people practicing open defecation worldwide, over half live in India[1]. While
64 international and national pressure on improving sanitation conditions in India has led to over
65 350 thousand people gaining access to improved toilets since 1990, it has barely kept up with
66 population growth[1,2]. Recent studies show that even in areas with access to household-level
67 improved sanitation, use of these toilets is low[3–5]. This may be due in part to a mismatch
68 between the culturally acceptable pour-flush toilets and the level of water access. Coverage of

1
2
3 69 improved water sources, usually community-level pumps or taps, is relatively high even in rural
4
5 70 areas in India, but it may not be sufficient for flushing purposes on top of other daily water
6
7
8 71 needs[1,6].
9

10 Although the effectiveness of water, sanitation and hygiene (WASH) interventions vary,
11
12 72 meta-analyses have found that individual or combined WASH interventions decrease diarrheal
13
14 73 disease prevalence by up to 48%[7–11]. While combined interventions would be expected to
15
16 74 have a greater influence on multiple exposure pathways and thus a greater combined impact on
17
18 75 health, there is limited evidence of additive benefits[12]. This may be due to poor uptake,
19
20 76 inconsistent use, or an incomplete understanding of relevant pathways[8–10]. In India,
21
22 77 combining water and sanitation interventions may be more critical than just interrupting multiple
23
24 78 transmission pathways for enteric infection; evidence suggests that household-level water access
25
26 79 is integral to the use of improved sanitation in this context[13].
27
28
29
30
31

32 While the intent of improved sanitation facilities is to separate human feces from human
33
34 82 contact, most of the focus is on constructing household toilets to increase improved sanitation
35
36 83 coverage—the primary metric used in monitoring progress toward international targets.
37
38
39 84 However, studies in India have further shown that toilet construction does not translate into toilet
40
41 85 use in this context[5,14–16]. Moreover, with the interdependence between members of
42
43 86 households and households within communities, safe water and sanitation is a community-level
44
45
46 87 issue. There is growing emphasis on assessing health risk from poor water and sanitation
47
48 88 conditions not simply due to individual or even household-level risk factors, but also from
49
50 89 conditions in the community environment[17]. There is evidence that even households without
51
52 90 toilets, and households which do not filter drinking water, show decreased health risk if they live
53
54
55 91 in communities with high levels of coverage and use[18–20].
56
57
58
59
60

1
2
3 92 Moreover, the effectiveness of community interventions may be higher in communities with
4
5
6 93 positive perceptions of their collective ability to come together to improve their conditions.
7
8 94 Collective efficacy, a latent construct comprised of the structural and cognitive components that
9
10 95 facilitate a community's shared belief in its ability to come together and execute actions related
11
12 96 to a common goal, may explain some variance in intervention effectiveness across communities
13
14 97 receiving WASH interventions[21].
15
16

17 98 A main risk of poor WASH conditions is enteric infection, caused by a diverse array of
18
19 99 bacteria, viruses, protozoa, and parasites, including soil-transmitted helminths. These infections
20
21 100 may cause diarrhea, the second leading cause of mortality for children under five years
22
23 101 worldwide and in India, a leading cause of mortality regardless of age[22,23]. There is also
24
25 102 growing evidence that asymptomatic enteric infections may pose a similar risk, with repeat
26
27 103 enteric infections contributing to chronic malnutrition, environmental enteric dysfunction, poor
28
29 104 cognitive outcomes, and poor vaccine uptake[24–29]. Poor WASH conditions are also linked to
30
31 105 increased risk of respiratory infection, the leading cause of mortality for children under five
32
33 106 years worldwide[22,30,31]. Poor water and sanitation access can also affect the social, physical
34
35 107 and mental well-being of women, acting through pathways ranging from unsafe menstrual
36
37 108 hygiene management practices and increased risk of violence[32–34].
38
39
40
41
42
43
44
45

46 110 **Description of the intervention**

47
48 111 Over the past decades there has been a global commitment to determine water and sanitation
49
50 112 interventions with demonstrated effectiveness, not just efficacy[35]. Gram Vikas, a non-
51
52 113 governmental organization (NGO) based in Odisha, India (<http://www.gramvikas.org/>), has
53
54 114 responded by implementing its MANTRA (Movement and Action Network for Transformation
55
56
57
58
59
60

1
2
3 115 of Rural Areas) water and sanitation program in more than 1000 villages since 2002[36]. This
4
5 116 approach includes both household-level piped water connections, and community-level
6
7
8 117 mobilization for culturally appropriate household toilets. A previous interrupted time series
9
10
11 118 analysis of the MANTRA intervention reported it to be protective against diarrheal diseases[37].
12
13 119 However, in addition to limitations of design, this study relied on outcome data collected and
14
15 120 reported by Gram Vikas, the intervention implementer, and did not assess intervention coverage
16
17
18 121 or impacts on environmental fecal contamination.

19
20 122 The MANTRA water and sanitation intervention is rolled out in a three-phase process over
21
22 123 an average three to five years (Figure 1). During the first, or Motivational, phase, representatives
23
24 124 of Gram Vikas visit the identified village several times to assess village interest and progress
25
26
27 125 towards a set of Gram Vikas requirements, including: 1) the commitment of every household to
28
29 126 participate, 2) creation of a village corpus fund from contributions from every household, and 3)
30
31 127 development of village guidelines for maintenance and use of facilities.

32
33
34 128 Once this set of requirements is achieved, the village progresses into the second, or
35
36 129 Operational, phase of the intervention. Each household constructs a pour-flush toilet with two
37
38
39 130 soak-pits and a separate bathing room. The households hire a local, skilled mason and provide
40
41 131 their own unskilled labor and locally available materials to complete the superstructure. Gram
42
43 132 Vikas provides external materials such as PVC pipes and porcelain pans. At the same time, a
44
45
46 133 water tank, community meeting space, and piped water distribution system connected to every
47
48 134 household, with taps in the toilet and bathing rooms and a separate tap in the kitchen, is
49
50
51 135 constructed through a similar collaborative process.

52
53 136 All households must construct a toilet and bathing room for the village to progress into
54
55 137 the final, or Completed, phase of the intervention, in which the water system is turned on.
56
57
58
59
60

1
2
3 138 Notably, this three-phase process only allows each household access to piped water once every
4
5 139 household in the village has a toilet and bathing room. This model contrasts with most previous
6
7
8 140 water and sanitation interventions, including those implemented under India's Total Sanitation
9
10 141 Campaign and other government programs, which do not require community-level sanitation
11
12 142 compliance and do not provide a piped water supply at the household level[38].
13
14

15 143

17 144 **Study aims**

18
19
20 145 The primary objective of this study is to evaluate the effectiveness of the combined household-
21
22 146 level water supply and sanitation intervention, as implemented by Gram Vikas in Odisha, India.
23

24 147 Toward that objective, this study aims to:

- 25
26
27 148 1) Assess the effectiveness of the intervention in improving water and sanitation
28
29 149 infrastructure coverage, access, and use, and to assess fecal sludge management practices
30
31 150 in intervention communities.
32
33
34 151 2) Assess the effectiveness of the intervention in reducing environmental fecal
35
36 152 contamination.
37
38
39 153 3) Assess the effectiveness of the intervention in improving health. This includes reported
40
41 154 diarrheal disease in children under 5 years (primary outcome), acute respiratory infection,
42
43 155 infection with soil-transmitted helminthes, nutritional status, environmental enteric
44
45 156 dysfunction, seroconversion for selected enteric pathogens, and urogenital diseases
46
47
48 157 associated with menstrual hygiene management practices. Mental and social well-being
49
50 158 will be explored through assessment of sanitation insecurity and women's empowerment.
51
52
53 159 4) Assess the cost and cost-effectiveness of the intervention.
54
55
56
57
58
59
60

1
2
3 160 5) Develop and assess a theoretically-grounded, empirically informed collective efficacy
4
5 161 scale; and determine the effect of collective efficacy on intervention effectiveness.
6
7
8 162

10 163 **METHODS**

12 164 **Setting**

13 165 The study is located in Ganjam and Gajapati districts located in eastern Odisha, India (Figure 2).
14
15 166 These two contiguous districts were a single district until 1992. Over 44% of the population in
16
17 167 these districts is recognized by the Government of India as being below the poverty line
18
19
20 168 (BPL)[39]. As of 2008, a majority of households in both districts had access to an improved,
21
22 169 likely community-level, drinking water source, with over 23% of households in Ganjam having
23
24 170 access to any sanitation facility, compared to only 8% of households in Gajapati[39]. The area is
25
26 171 primarily rural and agrarian, and the climate is characterized by a monsoon season from June to
27
28 172 September, with an average rainfall of ~1400 mms/year.
29
30
31
32
33
34 173

35 174 **Study design**

36 175 This study uses a matched cohort design to assess the effectiveness of a completed intervention
37
38 176 with data collected across four study rounds from June 2015 to October 2016 (Figure 3). Data
39
40 177 are collected continuously across all study rounds for diarrhea, acute respiratory infection,
41
42 178 nutritional status, and stored and source water outcomes to assess seasonality. Data are collected
43
44 179 in rounds 2 and 4 for environmental enteric dysfunction, seroconversion, and hand-rinses, and
45
46 180 cross-sectionally in one or more rounds for the remaining outcomes. As described below, control
47
48 181 villages were matched to randomly selected intervention villages through a multi-step restriction,
49
50
51 182 genetic matching, and exclusion process using the following eligibility criteria.
52
53
54
55
56
57
58
59
60

183

184 Eligibility criteria for villages

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

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

203 In addition, to be eligible for inclusion both intervention and control villages must: 1)
204 appear in the Government of India Census 2001 and the BPL Survey 2002, 2) have a population

1
2
3 205 of at least 20 households, and 3) be within approximately three hours travel from the study office
4
5 206 in Brahmapur, Ganjam District. This last criterion is due to logistical constraints.
6
7

8 207 *2. Matching.* After restriction, genetic matching was used to match potential control villages to
9
10 208 the randomly selected intervention villages without replacement[5,40,41]. Villages were exact
11
12 209 matched on district to limit any political or large scale geographic variation between district
13
14 210 populations, and were also matched on pre-intervention demographic, socioeconomic, sanitation,
15
16 211 and water access characteristics listed in Table 1[5]. These village-level matching variables were
17
18 212 selected due to their theorized association with the primary outcome, diarrheal diseases, as well
19
20 213 as data availability.
21
22
23

24 214 *3. Exclusion.* The field team visited matched potential control villages and intervention villages
25
26 215 to assess suitability for the study through a rapid assessment interview with village leadership
27
28 216 and to ensure accessibility. Villages were excluded if they are not within three hours travel of the
29
30 217 field office in Brahmapur, had sustained major infrastructure damage due to a natural disaster, or
31
32 218 if there was a current or planned sanitation or water intervention by an organization external to
33
34 219 the village in the next 12 months as determined through the rapid assessment interview with
35
36 220 village leadership. In addition, villages were excluded if there were fewer than three households
37
38 221 with children under five years old. As villages were removed from the pool of prospective
39
40 222 control villages, the matching process was repeated for all intervention villages and remaining
41
42 223 eligible control villages, and balance measures were assessed. The matching and exclusion
43
44 224 processes were repeated as necessary.
45
46
47
48
49

50 225 After the iterative matching and exclusion process was complete, covariate balance was
51
52 226 assessed for all matching variables for the final set of intervention and control villages through
53
54 227 examination of balance measures[42–44]. Matching resulted in an improvement in balance as
55
56
57
58
59
60

228 assessed through comparison of several measures including q-q plots, Kolmogorov-Smirnov
 229 bootstrap p-values, and standardized differences. After matching, there were no significant
 230 differences between intervention and control groups (Table 1).

231

232 Eligibility criteria for households

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

239

240 **Table 1.** Pre-intervention characteristics used in matching, and balance diagnostics before and
 241 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 [‡] (%)	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

242 All eligible: all villages that are eligible for the matching process after restriction

243 Std Diff (absolute standardized difference): a value greater than 0.1 is considered meaningful imbalance [42]

244 ‡ Ganjam villages only; no data available for Gajapati villages

245 Kolmogorov-Smirnov bootstrap p-values: * <0.05 ** <0.01

246

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

249

250 Sample Size

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

1
2
3 263
4
5
6 264 **Outcome Measurement**
7
8 265 Outcomes, and individual-, household-, and community-level risk factors, will be measured
9
10 266 through surveys, interviews, or through the collection and analysis of environmental, stool or
11
12 267 dried blood spot samples. All survey questions will be translated into the primary local language,
13
14
15 268 Odia, and back-translated to confirm wording. Household surveys will be verbally administered
16
17 269 by trained field workers to the mother or primary caregiver of the youngest child under five in
18
19
20 270 each household, unless otherwise specified below. Community surveys will be verbally
21
22 271 administered to the *sarpanch* (village head) or other member of village leadership. Survey data
23
24 272 will be collected on mobile phones using Open Data Kit[47]. GPS coordinates for households,
25
26 273 water sources and other relevant sites will be collected using Garmin eTrex 10 or 20 devices
27
28
29 274 (Garmin Ltd., Olathe, KS, USA).

30
31
32 275
33
34 276 Coverage, access and use of sanitation, water and hygiene infrastructure
35
36 277 Coverage, access and use of WASH infrastructure will be assessed in all four rounds. Presence
37
38 278 of and access to toilets, water sources and hand-washing stations will be assessed through
39
40 279 standard questions from the Demographic and Health Surveys and confirmed through spot
41
42
43 280 observations. Spot observations of household toilets and hand-washing stations will be further
44
45
46 281 used to assess indicators of functionality, maintenance, recent use. Reported water and sanitation
47
48 282 practices, including child feces disposal practices, will be captured through household survey
49
50 283 questions.

51
52
53 284
54
55 285 Diarrheal Diseases
56
57
58
59
60

1
2
3 286 The primary outcome for this study is prevalence of diarrheal diseases, recorded as both daily
4
5
6 287 point prevalence over the previous three days and seven-day period prevalence, for all household
7
8 288 members in all four rounds. Although self-reported diarrhea is a subjective outcome with a well-
9
10 289 established risk of bias, three-day recall reduces recall bias[48,49]. Diarrheal disease will be
11
12 290 measured using the World Health Organization (WHO) definition of three or more loose stools
13
14 291 in a 24-hour period, with or without the presence of blood. Field workers will use a simple
15
16 292 calendar as a visual aid to help respondents with recall. Each household member will be asked to
17
18 293 recall his or her own disease status and the mother or primary caregiver will be asked to report
19
20 294 disease for children.
21
22
23
24
25
26

27 296 Respiratory infection

28
29 297 Prevalence of respiratory infections will be recorded as both daily point prevalence over the
30
31 298 previous three days and seven-day period prevalence for all household members in all four
32
33 299 rounds. Respiratory infection is defined as the presence of cough and/or shortness of
34
35 300 breath/difficulty breathing according to WHO's Integrated Management of Childhood Illness
36
37 301 (IMCI)[50]. The full IMCI case definition for acute lower respiratory infection also includes
38
39 302 measurement of respiratory rate and observation of chest indrawing, stridor and other danger
40
41 303 signs; these criteria were excluded from our definition as there is concern about the technical
42
43 304 support required to produce consistent and accurate data within this context[50]. Our definition
44
45 305 will provide a broad assessment of respiratory illness burden. Each household member will be
46
47 306 asked to recall his or her own disease status and the mother or primary caregiver will be asked to
48
49 307 report disease for children.
50
51
52
53
54
55
56
57
58
59
60

309 Nutritional Status

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

321

322 Soil-transmitted helminth infection

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

1
2
3 331 STH infection prevalence between intervention and control villages, while the stool samples
4
5
6 332 collected approximately 8 months later in round 4 will provide a measure of re-infection rate.
7
8 333
9
10 334 Environmental enteric dysfunction
11
12 335 Stools from a randomly selected subset of 200 children under two years old, collected in rounds
13
14 336 2 and 4, will be used to assess environmental enteric dysfunction (EED) through quantification
15
16 337 of biomarkers of intestinal inflammation and permeability. Fecal myeloperoxidase (MPO),
17
18 338 alpha-1-antitrypsin (AAT), and neopterin (NEO), markers for neutrophil activity, intestinal
19
20 339 permeability and TH1 immune activation, respectively, were selected for this study based on
21
22 340 evidence of association with EED, subsequent linear growth deficits, and household
23
24 341 environmental fecal contamination[24,25,54].
25
26 342
27
28 343 Seroconversion for enteric pathogens
29
30 344 Serological assays that assess antibody production against various enteric pathogens can provide
31
32 345 an objective measure of exposure to enteric infections[55]. Enrolling children aged 6 to 18
33
34 346 months will reduce the potential for interference from maternally acquired antibodies and permit
35
36 347 analysis of seroconversion data in a critical window for young children who experience higher
37
38 348 diarrheal disease morbidity and mortality before two years of age[56–61]. Children who are 6 to
39
40 349 12 months during round 2 will have capillary blood drawn by fingerstick or heelstick, as
41
42 350 appropriate, and will be visited again during round 4 for a second capillary blood sample. All
43
44 351 blood samples will be preserved on TropBio (Sydney, Australia) filter discs and stored within 7
45
46 352 days of collection at -20°C. Seroconversion against markers for norovirus, *Giardia intestinalis*,
47
48 353 *Cryptosporidium parvum*, *Entamoeba histolytica*, enterotoxigenic *E. coli* heat-labile enterotoxin
49
50
51
52
53
54
55
56
57
58
59
60

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

356

357 Environmental fecal contamination

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

372

373 Cost and cost-effectiveness

374 Costs and potential cost savings (i.e., averted costs) associated with the intervention will be
375 assessed through an economic costing approach that recognizes and quantifies costs and benefits
376 from a societal perspective[72]. Data on program and point-of-delivery inputs will be collected at

1
2
3 377 household, community, and implementer levels in round 3. Field workers will administer
4
5 378 community surveys to a village leader, and household surveys to the household decision-maker
6
7
8 379 for toilet installation, in 20 randomly selected households in half of the intervention and control
9
10 380 villages. Given cost-effectiveness analyses require the effect of the intervention to be measured
11
12 381 against a counterfactual, and the intervention of interest is a community-based intervention, cost
13
14 382 and effectiveness measures will be summarized at the village level [73]. Surveys will collect data
15
16 383 on household- and community-level inputs related to materials and labor required to construct
17
18 384 household toilets and wash rooms, the community water tank and distribution system, and
19
20 385 household water connections; longer-term water supply and toilet maintenance costs; and
21
22 386 financing required for this infrastructure as well as perceived benefits, including averted social
23
24 387 opportunity costs. Implementer inputs from Gram Vikas will be collected through an
25
26 388 enumeration exercise, interviews, and examination of the implementer's financial records.
27
28
29
30
31
32
33

34 390 Collective efficacy

35
36 391 Collective efficacy (CE) is a latent construct comprised of the structural and cognitive
37
38 392 components that facilitate a community's shared belief in its ability to come together and execute
39
40 393 actions related to a common goal[21]. A review of the literature and established conceptual
41
42 394 frameworks will be performed to define the CE construct. A sequential exploratory mixed
43
44 395 qualitative and quantitative design will be used to develop and refine a scale to measure CE and
45
46 396 test hypotheses. Field workers will administer the refined, multi-item, Likert-type CE scale to
47
48 397 one randomly selected household member aged 18 years or older in each household in round 3.
49
50
51
52

53 398

54 399 Women's empowerment

1
2
3 400 Four dimensions of women's empowerment will be measured in rounds 3 and 4: group
4
5
6 401 participation, leadership, decision-making and freedom of movement. Group participation and
7
8 402 leadership will be measured using modules from the Women's Empowerment in Agriculture
9
10 403 Index (WEAI), which has been tested in South Asia[74]. Decision-making will be measured
11
12 404 using questions from the women's status module of the Demographic and Health Surveys.
13
14 405 Freedom of movement will be measured using questions from the project-level Women's
15
16 406 Empowerment in Agriculture Index (pro-WEAI). These measures will be collected for the
17
18 407 primary female caregiver of the youngest child under 5, and were selected based on the
19
20 408 importance of women's empowerment for child nutrition[75,76]. Women's empowerment is
21
22 409 conceptualized as both an outcome and a potential mediator along the pathway between the
23
24 410 Gram Vikas intervention and child health outcomes.
25
26
27 411
28
29
30
31 412 Menstrual hygiene management
32
33 413 Menstrual hygiene management practices vary worldwide and depend on personal preference,
34
35 414 socioeconomic status, local traditions and beliefs, and access to water and sanitation
36
37 415 resources[77]. Unhygienic washing practices are common in rural India and among women and
38
39 416 girls in lower socioeconomic groups, and may increase risk of urogenital infection[78–80].
40
41 417 However, the link between access to water and sanitation, menstrual hygiene management and
42
43 418 urogenital infections has been poorly studied. Household surveys will be administered in round 4
44
45 419 to a randomly selected woman aged 18 or older, in a subset of 800 households, and will capture
46
47 420 self-reported urogenital infection, defined as at least one of the following symptoms: 1) abnormal
48
49 421 vaginal discharge (unusual texture and color/more abundant than normal), 2) burning or itching
50
51 422 in the genitalia, 3) burning or itching when urinating, or 4) genital sores[79].
52
53
54
55
56
57
58
59
60

1
2
3 423
4
5
6 424 Sanitation Insecurity
7
8 425 This study will assess the associations between sanitation access and sanitation insecurity with
9
10 426 mental health among women. In previous research in Odisha, a contextually specific definition
11
12 427 and measure for sanitation insecurity was developed, with associations between facets of
13
14 428 sanitation insecurity and mental health independent of sanitation facility access[81]. This
15
16 429 previously developed measure will be used to determine if levels of sanitation insecurity differ
17
18 430 between intervention and control villages and how it may be associated with mental health
19
20 431 outcomes, specifically well-being, anxiety, depression, and distress. Household surveys will be
21
22 432 administered in round 4 to a randomly selected woman aged 18 or older, in a random subset of
23
24 433 800 households.
25
26
27
28

29 434
30
31 435 Fecal sludge management.
32
33
34 436 In sanitation systems where sewerage is not feasible, such as the household toilets constructed as
35
36 437 part of the MANTRA intervention, safe management of fecal waste is necessary. Although there
37
38 438 is growing emphasis on safe fecal sludge management (FSM), research has mainly focused on
39
40 439 urban settings[82,83]. Preliminary research in Odisha suggests that fecal sludge management in
41
42 440 this rural setting is a substantial challenge, and may impact household use of toilets. In round 3,
43
44 441 household surveys and spot checks of toilets in intervention villages will be used to assess toilet
45
46 442 use and fecal sludge management practices.
47
48
49

50 443

51 444 **STATISTICAL ANALYSES**

52
53
54
55
56
57
58
59
60

1
2
3 445 The effect of the intervention on infrastructure coverage, access, and use (aim 1), and the effect
4
5 446 of the intervention on improving health (aim 3), will be analyzed using logistic, linear, log
6
7
8 447 binomial, or negative binomial multilevel regression depending on the outcome, to compare
9
10 448 intervention versus control villages. Prevalence of fecal sludge management practices in
11
12 449 intervention communities will be assessed using multilevel regression (aim 1). For all models,
13
14 450 the hierarchical structure of the data will be accounted for using random effects. Estimation of
15
16 451 relative risks through Poisson regression or binary regression methods for binary outcomes will
17
18 452 be considered to ensure robustness of results. Mediation of the potential association between
19
20 453 intervention and nutritional status outcomes by women's empowerment will be assessed using
21
22 454 multilevel structural equation modeling, and statistical approaches to reduce bias will be
23
24 455 explored as needed[84].
25
26
27
28

29 456 The impact of intervention on reducing environmental fecal contamination (aim 2), will
30
31 457 be assessed through two methods. First, logistic and negative binomial multilevel regression to
32
33 458 estimate intervention effects on the relative scale will be used to compare intervention versus
34
35 459 control villages. Estimation of relative risks through Poisson regression or binary regression
36
37 460 methods for binary outcomes will be considered to ensure robustness of results. Second, a
38
39 461 stochastic microbial risk framework will be used to assess differential fecal environmental
40
41 462 contamination between intervention and control villages.
42
43
44

45 463 The cost and cost-effectiveness of the intervention (aim 4) will be assessed in two steps.
46
47 464 Incremental intervention benefits will be ascertained by combining health benefit data, from
48
49 465 analysis of health outcome data and established averted cost data, with other averted social
50
51 466 opportunity costs. An incremental cost-effectiveness ratio, expressed in cost per disease-specific
52
53
54
55
56
57
58
59
60

1
2
3 467 DALY, will be calculated by dividing the incremental intervention costs by the incremental
4
5 468 intervention benefits.

6
7
8 469 The collective efficacy scale will be analyzed using a psychometric approach in which
9
10 470 factor analytics are employed to identify an appropriate factor solution and test the reliability and
11
12 471 validity of the CE scores. Once a CE factor solution and an empirically derived multilevel data
13
14 472 structure have been identified, the association between CE and intervention effectiveness will be
15
16 473 analyzed using multilevel generalized linear mixed models to estimate relative risks (aim 5)
17
18 474 [85,86].

19
20 475 For all outcomes, variables used in the matching process may be considered as
21
22 476 covariates, as needed, in addition to individual-, household-, and community-level risk factors.
23
24 477 Covariates that are statistically associated with outcomes of interest in bivariate analyses will be
25
26 478 considered for inclusion in final multivariate models, following standard stepwise model-
27
28 479 building approaches. Secondary analyses may also evaluate models for effect modification as
29
30 480 relevant, including exposure-mediator interaction for mediation models and cross-level
31
32 481 interaction, by assessing changes in parameter values based on potential effect modifiers.
33
34 482 Potential effect modifiers may include breastfeeding for seroconversion outcomes, and climate
35
36 483 factors and population density for environmental fecal contamination and health outcomes.
37
38 484 However, this study was not designed to assess effect modification and therefore is not
39
40 485 specifically powered for these analyses. For all outcomes, unadjusted models will be presented
41
42 486 along with models adjusting for covariates.
43
44
45
46
47
48
49
50

51 487

52
53 488 **DISCUSSION**
54
55
56
57
58
59
60

1
2
3 489 This matched cohort study is one of the first to evaluate the effect of a rural combined
4
5 490 household-level piped water and sanitation intervention, implemented at the community level, on
6
7
8 491 a large scale. The matched design provides a rigorous means for estimating causal effects given
9
10 492 that randomization to intervention group was not feasible due to the several year implementation
11
12 493 process[5]. By focusing on an intervention where the implementation process is complete, it also
13
14 494 limits the risk presented by randomized controlled trials, where the intervention has little uptake,
15
16 495 an especially important study challenge given interdependence of exposure and outcomes within
17
18 496 communities, and a problem that has characterized previous trials of sanitation interventions in
19
20 497 India[15,16].

24 498 A strength of this study is the assessment of health impacts using the holistic WHO
25
26 499 definition of health, including not just disease status, but also mental, social, and physical well-
27
28 500 being[87]. Outcomes along the causal chain include standard, but more subjective measures,
29
30 501 such as reported diarrheal diseases and respiratory infection, as well as more objective measures
31
32 502 such as fecal environmental contamination, soil transmitted helminth infection, and
33
34 503 anthropometry. Although there is risk of response bias for reported outcomes, it is unlikely to be
35
36 504 differential by intervention status since the study team is not directly linked to Gram Vikas. Even
37
38 505 though field workers may be aware of village intervention status, lab staff analyzing water, hand
39
40 506 rinse, stool, and blood samples will be blinded. In addition, this study includes the more novel
41
42 507 use of seroconversion for enteric pathogens, biomarkers of environmental enteric dysfunction,
43
44 508 and measures of collective efficacy in an evaluation assessment. While there are limitations
45
46 509 inherent to observational studies, the matched study design and multivariate modeling analysis
47
48 510 plan reduce the potential for confounding. However, there is still the potential for residual
49
50 511 unmeasured confounding.
51
52
53
54
55
56
57
58
59
60

1
2
3 512
4
5
6 513 **Ethics and Dissemination.** This study has been reviewed and approved by the Ethics Committee
7
8 514 of the London School of Hygiene and Tropical Medicine, U.K (No. 9071) and Institute Ethics
9
10 515 Committee of the Kalinga Institute of Medical Sciences of KIIT University, Bhubaneswar, India
11
12 516 (KIMS/KIIT/IEC/053/2015). Efforts will be made to communicate the central findings and
13
14 517 implications with study communities, the implementing organization and government officials in
15
16 518 India. The results of this study will be submitted for publication in peer reviewed journals and
17
18 519 presented at conferences. The data collected in the study will be publicly available, with personal
19
20 520 identifiable data redacted, following the publication of the primary results within 24 months of
21
22 521 the final data collection date.
23
24
25
26

27 522
28
29 523 **Funding.** This study is supported by a grant from the Bill & Melinda Gates Foundation to the
30
31 524 London School of Hygiene & Tropical Medicine (OPP1008048) and to Emory University.
32
33 525 (OOP1125067).
34
35

36 526

37
38
39 527 **Competing Interests:** None declared.
40

41 528

42
43 529 **Contributions from authors:** TC, HR, PR, BT, and HC contributed to study design. HR, LZ
44
45 530 and BT developed laboratory protocols. HR, BT, GS, MD, SS, LZ, and BC developed data
46
47 531 collection tools. All authors contributed to editing and revising the manuscript.
48
49

50 532

51 533

52
53
54
55 534
56
57
58
59
60

1
2
3 535 **Figure 1.** Description of the three phases of the Gram Vikas MANTRA water and sanitation
4
5 536 intervention.
6
7

8 537
9
10 538 **Figure 2.** Study sites in Ganjam and Gajapati districts, Odisha, India with intervention villages
11
12 539 in black and control villages in white. Inset shows location of districts in India.
13
14

15 540
16
17 541 **Figure 3.** Restriction, matching and exclusion process for selection of intervention and control
18
19 542 villages (1), and timeline for study rounds and outcome data collection (2).
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

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

- 1
2
3 566 9 Clasen T, Bostoen K, Schmidt W, *et al.* Interventions to improve disposal of human
4
5
6 567 excreta for preventing diarrhoea. *Cochrane database Syst Rev* Published Online First:
7
8 568 2010.<http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD007180.pub2/pdf/standard>
9
10 569 (accessed 4 Jan2015).
- 11
12 570 10 Clasen T, Schmidt W-P, Rabie T, *et al.* Interventions to improve water quality for
13
14 571 preventing diarrhoea: systematic review and meta-analysis. *BMJ* 2007;**334**:782.
15
16 572 doi:10.1136/bmj.39118.489931.BE
- 17
18 573 11 Cairncross S, Hunt C, Boisson S, *et al.* Water, sanitation and hygiene for the prevention of
19
20 574 diarrhoea. *Int J Epidemiol* 2010;**39 Suppl 1**:i193–205. doi:10.1093/ije/dyq035
- 21
22 575 12 Clasen TF, Alexander KT, Sinclair D, *et al.* Interventions to improve water quality for
23
24 576 preventing diarrhoea. *Cochrane database Syst Rev* 2015;**10**:CD004794.
25
26 577 doi:10.1002/14651858.CD004794.pub3
- 27
28 578 13 Routray P, Schmidt W-P, Boisson S, *et al.* Socio-cultural and behavioural factors
29
30 579 constraining latrine adoption in rural coastal Odisha: an exploratory qualitative study.
31
32 580 *BMC Public Health* 2015;**15**:880. doi:10.1186/s12889-015-2206-3
- 33
34 581 14 Barnard S, Routray P, Majorin F, *et al.* Impact of Indian Total Sanitation Campaign on
35
36 582 Latrine Coverage and Use: A Cross-Sectional Study in Orissa Three Years following
37
38 583 Programme Implementation. *PLoS One* 2013;**8**. doi:10.1371/journal.pone.0071438
- 39
40 584 15 Patil SR, Arnold BF, Salvatore AL, *et al.* The effect of India's total sanitation campaign
41
42 585 on defecation behaviors and child health in rural Madhya Pradesh: a cluster randomized
43
44 586 controlled trial. *PLoS Med* 2014;**11**:e1001709. doi:10.1371/journal.pmed.1001709
- 45
46 587 16 Clasen T, Boisson S, Routray P, *et al.* Effectiveness of a rural sanitation programme on
47
48 588 diarrhoea, soil-transmitted helminth infection, and child malnutrition in Odisha, India: a
49
50
51
52
53
54
55
56
57
58
59
60

- 1
2
3 589 cluster-randomised trial. *Lancet Glob Heal* 2014;**2**:e645–53. doi:10.1016/S2214-
4
5 590 109X(14)70307-9
6
7
8 591 17 Eisenberg JNS, Trostle J, Sorensen RJD, *et al.* Towards a Systems Approach to Enteric
9
10 592 Pathogen Transmission: From Individual Independence to Community Interdependence.
11
12 593 *Annu Rev Public Health* 2013;**33**:239–57. doi:10.1016/j.micinf.2011.07.011.Innate
13
14 594 18 Root GPM. Sanitation, Community Environments, and Childhood Diarrhea in Rural
15
16 595 Zimbabwe. *J Heal Popul Nutr* 2001;**19**:73–82.
17
18 596 19 Bateman OM, Smith S. A Comparison of the Health Effects of Water Supply and
19
20 597 Sanitation in Urban and Rural Guatemala. 1999.
21
22 598 20 Huq A, Yunus M, Sohel SS, *et al.* Simple Sari Cloth Filtration of Water Is Sustainable and
23
24 599 Continues to Protect Villagers from Cholera in Matlab, Bangladesh. *MBio* 2010;**1**:1–5.
25
26 600 doi:10.1128/mBio.00034-10.Invited
27
28 601 21 Bandura A. *Self-Efficacy: The Exercise of Control*. Worth Publishers 1997.
29
30 602 22 UN Inter-agency Group for Child Mortality Estimation. Levels and trends in child
31
32 603 mortality: Report 2015. 2015.
33
34 604 23 World Health Organization. Estimated total deaths by cause, sex and WHO Member State,
35
36 605 2008. Geneva, Switzerland: 2011.
37
38 606 24 Lin A, Arnold BF, Afreen S, *et al.* Household environmental conditions are associated
39
40 607 with enteropathy and impaired growth in rural Bangladesh. *Am J Trop Med Hyg*
41
42 608 2013;**89**:130–7. doi:10.4269/ajtmh.12-0629
43
44 609 25 Korpe PS, Petri WA. Environmental enteropathy: critical implications of a poorly
45
46 610 understood condition. *Trends Mol Med* 2012;**18**:328–36.
47
48 611 doi:10.1016/j.molmed.2012.04.007
49
50
51
52
53
54
55
56
57
58
59
60

- 1
2
3 612 26 Guerrant RL, Oriá RB, Moore SR, *et al.* Malnutrition as an enteric infectious disease with
4
5 long-term effects on child development. 2008;**66**:487–505. doi:10.1111/j.1753-
6 613
7 4887.2008.00082.x.Malnutrition
8 614
9
10 615 27 Crane RJ, Jones KDJ, Berkley JA. Environmental enteric dysfunction : An overview.
11
12 616 2015;**36**:76–87.
13
14
15 617 28 Humphrey JH. Child undernutrition, tropical enteropathy, toilets, and handwashing.
16
17 618 *Lancet* 2009;**374**:1032–5. doi:10.1016/S0140-6736(09)60950-8
18
19
20 619 29 Mbuya MNN, Humphrey JH. Preventing environmental enteric dysfunction through
21
22 620 improved water, sanitation and hygiene: An opportunity for stunting reduction in
23
24 621 developing countries. *Matern Child Nutr* 2015;**12**:1–15. doi:10.1111/mcn.12220
25
26
27 622 30 Budge PJ, Griffin MR, Edwards KM, *et al.* Impact of home environment interventions on
28
29 623 the risk of influenza-associated ARI in Andean children: observations from a prospective
30
31 624 household-based cohort study. *PLoS One* 2014;**9**:e91247.
32
33 doi:10.1371/journal.pone.0091247
34 625
35
36 626 31 Aiello AE, Coulborn RM, Perez V, *et al.* Effect of hand hygiene on infectious disease risk
37
38 627 in the community setting: a meta-analysis. *Am J Public Health* 2008;**98**:1372–81.
39
40 628 doi:10.2105/AJPH.2007.124610
41
42
43 629 32 Caruso BA, Sevilimedu V, Fung IC-H, *et al.* Gender disparities in water, sanitation, and
44
45 630 global health. *Lancet (London, England)* 2015;**386**:650–1. doi:10.1016/S0140-
46
47 631 6736(15)61497-0
48
49
50 632 33 Sahoo KC, Hulland KR, Caruso B a., *et al.* Sanitation-related psychosocial stress: a
51
52 633 grounded theory study of women across the life-course in Odisha, India. *Soc Sci Med*
53
54 634 2015;**139**:80–9. doi:10.1016/j.socscimed.2015.06.031
55
56
57
58
59
60

- 1
2
3 635 34 Hulland KRS, Chase RP, Caruso BA, *et al.* Sanitation, stress, and life stage: A systematic
4
5 636 data collection study among women in Odisha, India. *PLoS One* 2015;**10**.
6
7
8 637 doi:10.1371/journal.pone.0141883
9
10 638 35 World Health Organization/The United Nations Children's Fund (UNICEF). Ending
11
12 639 Preventable Child Deaths from Pneumonia and Diarrhoea by 2025: The integrated Global
13
14 640 Action Plan for Pneumonia and Diarrhoea (GAPPD). 2013.
15
16
17 641 36 Gram Vikas. Annual Report 2013-2014. 2014.
18
19 642 37 Duflo E, Greenstone M, Guiteras R, *et al.* Toilets Can Work : Short and Medium Run
20
21 643 Health Impacts of Addressing Complementarities and Externalities in Water and
22
23 644 Sanitation.
24
25
26 645 38 Rosa G, Clasen T. Estimating the scope of household water treatment in low- and
27
28 646 medium-income countries. *Am J Trop Med Hyg* 2010;**82**:289–300.
29
30 647 doi:10.4269/ajtmh.2010.09-0382
31
32
33 648 39 International Institute for Population Sciences (IIPS). District Level Household and
34
35 649 Facility Survey (DLHS-3), 2007-08: India, Orissa. Mumbai: 2010.
36
37
38 650 40 Brady H, Caughey D, Dehejia R, *et al.* Genetic Matching for Estimating Causal Effects :
39
40 651 2012.
41
42
43 652 41 Sekhon JS. "Multivariate and Propensity Score Matching Software. *J Stat Softw*
44
45 653 2011;**42**.[http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Multivariate+a](http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Multivariate+and+Propensity+Score+Matching#9)
46
47 654 [nd+Propensity+Score+Matching#9](http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Multivariate+and+Propensity+Score+Matching#9) (accessed 30 May2014).
48
49
50 655 42 Austin PC. Balance diagnostics for comparing the distribution of baseline covariates
51
52 656 between treatment groups in propensity-score matched samples. *Stat Med* 2009;**28**:3083–
53
54 657 107. doi:10.1002/sim

- 1
2
3 658 43 Rubin DB. Using propensity scores to help design observational studies: Application to
4
5 659 the tobacco litigation. *Health Serv Outcomes Res Methodol* 2002;**2**:169–88.
6
7
8 660 44 Harder VS, Stuart EA, Anthony JC. Propensity score techniques and the assessment of
9
10 661 measured covariate balance to test causal associations in psychological research. *Psychol*
11
12 662 *Methods* 2010;**15**:997–1003. doi:10.1016/j.biotechadv.2011.08.021.Secreted
13
14
15 663 45 Faul F, Erdfelder E, Buchner A, *et al*. Statistical power analyses using G*Power 3.1: tests
16
17 664 for correlation and regression analyses. *Behav Res Methods* 2009;**41**:1149–60.
18
19 665 doi:10.3758/BRM.41.4.1149
20
21
22 666 46 Wolf J, Prüss-Ustün A, Cumming O, *et al*. Systematic review: Assessing the impact of
23
24 667 drinking water and sanitation on diarrhoeal disease in low- and middle-income settings:
25
26 668 Systematic review and meta-regression. *Trop Med Int Heal* 2014;**19**:928–42.
27
28 669 doi:10.1111/tmi.12331
29
30
31 670 47 Carl Hartung YAWBALCTGB. Open Data Kit: Tools to Build Information Services for
32
33 671 Developing Regions. <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.176.8017>
34
35
36 672 48 Feikin DR, Audi a., Olack B, *et al*. Evaluation of the optimal recall period for disease
37
38 673 symptoms in home-based morbidity surveillance in rural and urban Kenya. *Int J*
39
40 674 *Epidemiol* 2010;**39**:450–8. doi:10.1093/ije/dyp374
41
42
43 675 49 Arnold BF, Galiani S, Ram PK, *et al*. Optimal Recall Period for Caregiver-reported Illness
44
45 676 in Risk Factor and Intervention Studies: A Multicountry Study. *Am J Epidemiol*
46
47 677 2013;**177**:361–70. doi:10.1093/aje/kws281
48
49
50 678 50 World Health Organization. Integrated Management of Childhood Illness: Chart Booklet.
51
52 679 51 Cogill B. Anthropometric indicators measurement guide. Revised edition. *Washington,*
53
54 680 *DC, Acad Educ Dev [AED], Food Nutr Tech Assist Proj*

- 1
2
3 681 2003;:92.http://www.developmentgateway.org/download/202582/anthro_2003.pdf
4
5
6 682 52 de Onis M, Onyango AW, Van den Broeck J, *et al.* Measurement and standardization
7
8 683 protocols for anthropometry used in the construction of a new international growth
9
10 684 reference. *Food Nutr Bull* 2004;**25**:S27–36.
11
12
13 685 53 Truant a. L, Elliott SH, Kelly MT, *et al.* Comparison of formalin-ethyl ether
14
15 686 sedimentation, formalin-ethyl acetate sedimentation, and zinc sulfate flotation techniques
16
17 687 for detection of intestinal parasites. *J Clin Microbiol* 1981;**13**:882–4.
18
19
20 688 54 Kosek M, Haque R, Lima A, *et al.* Fecal markers of intestinal inflammation and
21
22 689 permeability associated with the subsequent acquisition of linear growth deficits in
23
24 690 infants. *Am J Trop Med Hyg* 2013;**88**:390–6. doi:10.4269/ajtmh.2012.12-0549
25
26
27 691 55 Crump J a, Mendoza CE, Priest JW, *et al.* Comparing serologic response against enteric
28
29 692 pathogens with reported diarrhea to assess the impact of improved household drinking
30
31 693 water quality. *Am J Trop Med Hyg* 2007;**77**:136–41.
32
33
34 694 56 Steinberg EB, Mendoza CE, Glass R, *et al.* Prevalence of Infection with Waterborne
35
36 695 Pathogens: a Seroepidemiologic Study in Children 6-36 Months Old in San Juan
37
38 696 Sacatepquez, Guatemala. 2004;**70**:83–8.
39
40
41 697 57 Brussow H, Sidoti J, Link H, *et al.* Age-specific prevalence of antibody to enterotoxigenic
42
43 698 *Escherichia coli* in Ecuadorian and German children. *J Infect Dis* 1990;**162**:974–7.
44
45
46 699 58 Khanna B, Cutler A, Israel N, *et al.* Use caution with serologic testing for *Helicobacter*
47
48 700 *pylori* infection in children. *J Infect Dis* 1998;**178**:460–5.
49
50
51 701 59 Lindkvist P, Asrat D, Nilsson I. Age at acquisition of *Helicobacter pylori* infection:
52
53 702 comparison of a high and a low prevalence country. *Scand J Infect Dis* 1996;**28**:181–4.
54
55
56 703 60 Ungar B, Gilman R, Lanata C, *et al.* Seroepidemiology of *Cryptosporidium* infection in
57
58
59
60

- 1
2
3 704 two Latin American populations. *J Infect Dis* 1988;**157**:551–6.
4
5
6 705 61 Vitral C, Yoshida C, Lemos E, *et al.* Age-specific prevalence of antibodies to hepatitis A
7
8 706 in children and adolescents from Rio de Janeiro, Brazil, 1978 and 1995: relationship of
9
10 707 prevalence to environmental factors. *Mem Inst Oswaldo Cruz* 1998;**93**:1–5.
11
12
13 708 62 Lammie PJ, Moss DM, Brook Goodhew E, *et al.* Development of a new platform for
14
15 709 neglected tropical disease surveillance. *Int J Parasitol* 2012;**42**:797–800.
16
17 710 doi:10.1016/j.ijpara.2012.07.002
18
19
20 711 63 Centers for Disease Control and Prevention. Laboratory Methods for the Diagnosis of
21
22 712 *Vibio cholerae*. Atlanta, Georgia: 1999.
23
24 713 64 Centers for Disease Control and Prevention. Isolation and Identification of *Shigella*.
25
26 714 Atlanta, Georgia:
27
28
29 715 65 United States Environmental Protection Agency. Coliforms—Total and *E. coli*, Membrane
30
31 716 Filtration Method 10029. 1999.
32
33
34 717 66 Gronewold AD, Borsuk ME, Wolpert RL, *et al.* An Assessment of Fecal Indicator
35
36 718 Bacteria-Based Water Quality Standards. *Environ Sci Technol* 2008;**42**:4676–82.
37
38 719 doi:10.1021/es703144k
39
40
41 720 67 Gruber JS, Ercumen A, Colford JM. Coliform bacteria as indicators of diarrheal risk in
42
43 721 household drinking water: systematic review and meta-analysis. *PLoS One*
44
45 722 2014;**9**:e107429. doi:10.1371/journal.pone.0107429
46
47
48 723 68 Levy K, Nelson KL, Hubbard A, *et al.* Rethinking indicators of microbial drinking water
49
50 724 quality for health studies in tropical developing countries: case study in northern coastal
51
52 725 Ecuador. *Am J Trop Med Hyg* 2012;**86**:499–507. doi:10.4269/ajtmh.2012.11-0263
53
54
55 726 69 Livio S, Strockbine N a, Panchalingam S, *et al.* *Shigella* Isolates From the Global Enteric

- 1
2
3 727 Multicenter Study Inform Vaccine Development. *Clin Infect Dis* 2014;**59**.
4
5 728 doi:10.1093/cid/ciu468
6
7
8 729 70 Kotloff KL, Nataro JP, Blackwelder WC, *et al*. Burden and aetiology of diarrhoeal disease
9
10 730 in infants and young children in developing countries (the Global Enteric Multicenter
11
12 731 Study, GEMS): a prospective, case-control study. *Lancet* 2013;**382**:209–22.
13
14 732 doi:10.1016/S0140-6736(13)60844-2
15
16
17 733 71 Kanungo S, Sah BK, Lopez a. L, *et al*. Cholera in India: An analysis of reports, 1997-
18
19 734 2006. *Bull World Health Organ* 2010;**88**:185–91. doi:10.2471/BLT.09.073460
20
21
22 735 72 Cellini SR, Kee JE. Cost-Effectiveness and Cost-Benefit Analysis. In: Wholey JS, Hatry
23
24 736 HP, Newcomer KE, eds. *Handbook of Practical Program Evaluation*. San Francisco: :
25
26 737 John Wiley & Sons, Inc. 2010.
27
28
29 738 73 Edejer TT, Baltussen R, Adam T, *et al*. Making Choices in Health: WHO Guide to Cost-
30
31 739 Effectiveness Analysis. 2003.
32
33
34 740 74 Malapit HJ, Sproule K, Kovarik C, *et al*. Measuring progress toward empowerment
35
36 741 Women’s Empowerment in Agriculture Index: Baseline Report. *IfpriOrg* 2014;:1–60.
37
38
39 742 75 Black RE, Victora CG, Walker SP, *et al*. Maternal and child undernutrition and
40
41 743 overweight in low-income and middle-income countries. *Lancet* 2013;**382**:427–51.
42
43 744 doi:10.1016/S0140-6736(13)60937-X
44
45
46 745 76 Smith LC, Ramakrishnan U, Ndiaye A, *et al*. The Importance of Women’s Status for
47
48 746 Child Nutrition in Developing Countries. 2003.
49
50 747 <http://www.ifpri.org/publication/importance-womens-status-child-nutrition-developing->
51
52 748 [countries](http://www.ifpri.org/publication/importance-womens-status-child-nutrition-developing-)
53
54
55 749 77 Sumpter C, Torondel B. A Systematic Review of the Health and Social Effects of
56
57
58
59
60

- 1
2
3 750 Menstrual Hygiene Management. *PLoS One* 2013;**8**:e62004.
4
5
6 751 doi:10.1371/journal.pone.0062004
7
8 752 78 Dasgupta A, Sarkar M. Menstrual Hygiene: How Hygienic is the Adolescent Girl? *Indian*
9
10 753 *J Community Med* 2008;**33**:77–80. doi:10.4103/0970-0218.40872
11
12 754 79 Das P, Baker KK, Dutta A, *et al.* Menstrual Hygiene Practices, WASH Access and the
13
14
15 755 Risk of Urogenital Infection in Women from Odisha, India. *PLoS One* 2015;**10**:e0130777.
16
17 756 doi:10.1371/journal.pone.0130777
18
19 757 80 van Eijk AM, Sivakami M, Thakkar MB, *et al.* Menstrual hygiene management among
20
21
22 758 adolescent girls in India: a systematic review and meta-analysis. *BMJ Open*
23
24 759 2016;**6**:e010290. doi:10.1136/bmjopen-2015-010290
25
26 760 81 Caruso B. *Sanitation Insecurity: Definition, Measurement, and Associations with*
27
28 761 *Women's Mental Health in Rural Orissa, India.*
29
30 762 2015.<http://pid.emory.edu/ark:/25593/rfhnt>
31
32 763 82 Peal A, Evans B, Blackett I, *et al.* Fecal sludge management (FSM): analytical tools for
33
34 764 assessing FSM in cities. *J Water, Sanit Hyg Dev* 2014;**4**:371.
35
36 765 doi:10.2166/washdev.2014.139
37
38 766 83 Peal A, Evans B, Blackett I, *et al.* Fecal Sludge Management: a comparative analysis of
39
40 767 12 cities. *J Water, Sanit Hyg Dev* 2014;**4**:563–75.
41
42 768 84 Richiardi L, Bellocco R, Zugna D. Mediation analysis in epidemiology: Methods,
43
44 769 interpretation and bias. *Int J Epidemiol* 2013;**42**:1511–9. doi:10.1093/ije/dyt127
45
46 770 85 McNutt LA, Wu C, Xue X, *et al.* Estimating the relative risk in cohort studies and clinical
47
48 771 trials of common outcomes. *Am J Epidemiol* 2003;**157**:940–3. doi:10.1093/aje/kwg074
49
50 772 86 Thompson ML, Myers JE, Kriebel D. Prevalence odds ratio or prevalence ratio in the
51
52
53
54
55
56
57
58
59
60

1
2
3 773 analysis of cross sectional data: what is to be done? *Occup Environ Med* 1998;**55**:272–7.

4
5 774 doi:10.1136/oem.55.4.272

6
7
8 775 87 World Health Organization. WHO definition of health.

9
10 776 <http://www.who.int/about/definition/en/print.html> (accessed 26 Apr2016).

11
12
13 777
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For peer review only

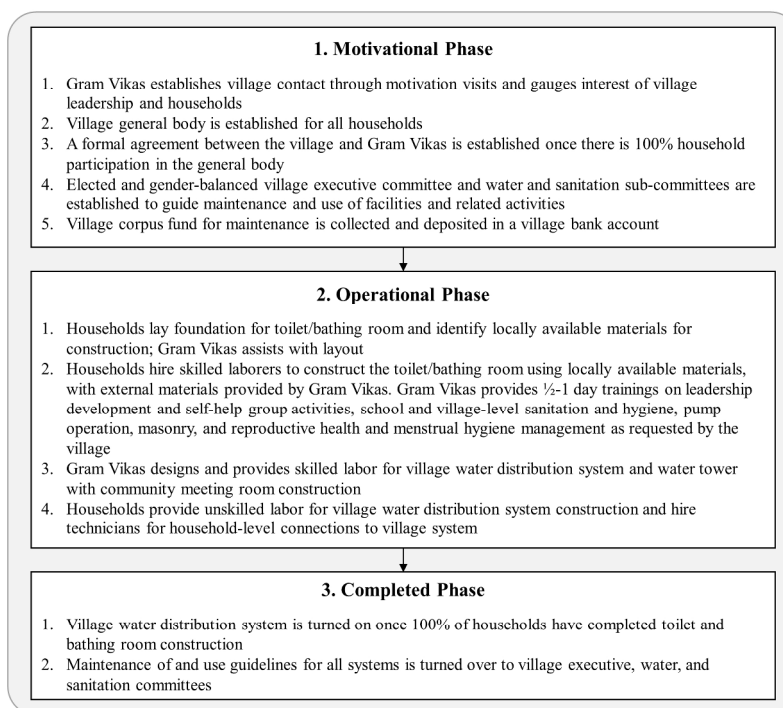


Figure 1. Description of the three phases of the Gram Vikas MANTRA water and sanitation intervention.

Figure 1
254x190mm (300 x 300 DPI)

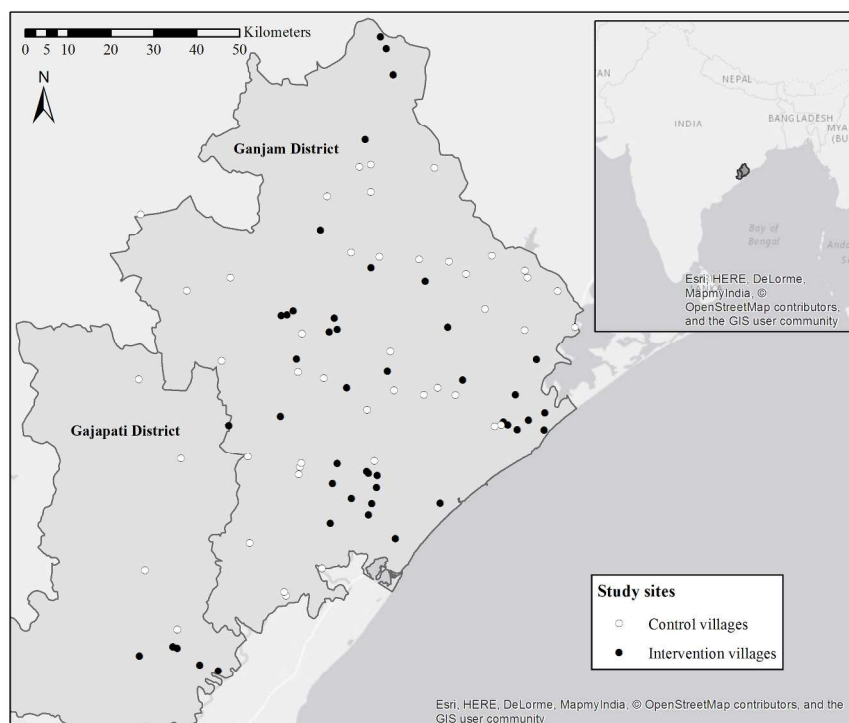


Figure 2. Study sites in Ganjam and Gajapati districts, Odisha, India with intervention villages in black and control villages in white. Inset shows location of districts in India.

Figure 2
215x279mm (300 x 300 DPI)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

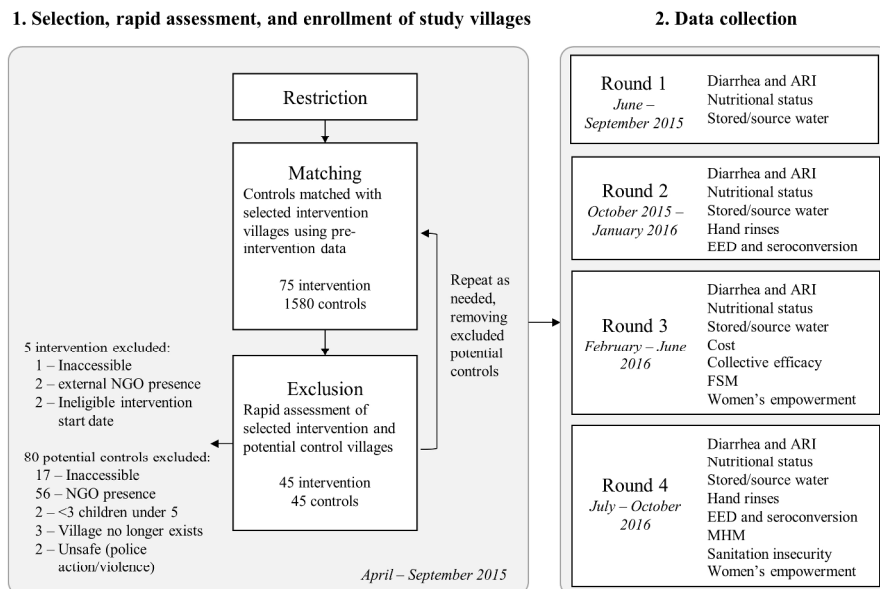


Figure 3. Restriction, matching and exclusion process for selection of intervention and control villages (1), and timeline for study rounds and outcome data collection (2).

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) <i>Cohort study</i> —Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up <i>Case-control study</i> —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 <i>Cross-sectional study</i> —Give the eligibility criteria, and the sources and methods of selection of participants	8-11	
		(b) <i>Cohort study</i> —For matched studies, give matching criteria and number of exposed and unexposed <i>Case-control study</i> —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

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49

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) <i>Cohort study</i> —If applicable, explain how loss to follow-up was addressed <i>Case-control study</i> —If applicable, explain how matching of cases and controls was addressed <i>Cross-sectional study</i> —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) <i>Cohort study</i> —Summarise follow-up time (eg, average and total amount)	NA
Outcome data	15*	<i>Cohort study</i> —Report numbers of outcome events or summary measures over time	NA
		<i>Case-control study</i> —Report numbers in each exposure category, or summary measures of exposure	NA
		<i>Cross-sectional study</i> —Report numbers of outcome events or summary measures	NA
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:	<i>BMJ Open</i>
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

1
2
3 1 **Design and rationale of a matched cohort study to assess the effectiveness of a combined**
4
5 2 **household-level piped water and sanitation intervention in rural Odisha, India**
6
7
8
9

10 4 Heather Reese,¹ Parimita Routray,² Belen Torondel,² Gloria Sclar,¹ Maryann G. Delea,^{1,2} Sheela
11
12 5 S. Sinharoy,³ Laura Zambrano,¹ Bethany Caruso,¹ Samir R. Mishra,⁴ Howard H. Chang,⁵
13
14 6 Thomas Clasen^{1,2}
15
16
17
18
19

20 8 ¹ Department of Environmental Health, Rollins School of Public Health, Emory University,
21
22 9 Atlanta, Georgia, United States of America
23

24 10 ² London School of Hygiene and Tropical Medicine, London, United Kingdom
25
26

27 11 ³ Nutrition and Health Sciences Program, Laney Graduate School, Emory University, Atlanta,
28
29 12 Georgia, United States of America
30

31 13 ⁴ School of Biotechnology, KIIT University, Bhubaneswar, Odisha, India
32
33

34 14 ⁵ Department of Biostatistics and Bioinformatics, Rollins School of Public Health, Emory
35
36 15 University, Atlanta, Georgia, United States of America
37
38
39
40

41 17 Contact information for corresponding author, Heather Reese:

42
43 18 Department of Environmental Health, Rollins School of Public Health, Emory University
44
45 19 1518 Clifton Rd, Atlanta, GA 30322
46
47

48 20 Email: heather.e.reese@emory.edu, Phone: (619) 379-8963
49

50
51 21 Keywords: rural, piped water, sanitation

52
53 22 Word count: 5442
54
55
56
57
58
59
60

1
2
3 **ABSTRACT**
4

5 **Introduction:** Government efforts to address massive shortfalls in rural water and sanitation in
6
7
8 India have centered on construction of community water sources and toilets for selected
9
10 households. However, deficiencies with water quality and quantity at the household level, and
11
12 community coverage and actual use of toilets has led Gram Vikas, a local NGO in Odisha, India,
13
14 to develop an approach that provides household-level piped water connections contingent on full
15
16 community-level toilet coverage.
17
18

19 **Methods:** This matched cohort study was designed to assess the effectiveness of a combined
20
21 piped water and sanitation intervention. Households with children under five years in 45
22
23 randomly selected intervention villages and 45 matched control villages will be followed over 17
24
25 months. The primary outcome is prevalence of diarrheal diseases; secondary health outcomes
26
27 include soil-transmitted helminth infection, nutritional status, seroconversion to enteric
28
29 pathogens, urogenital infections, and environmental enteric dysfunction. In addition, intervention
30
31 effects on sanitation and water coverage, access and use, environmental fecal contamination,
32
33 women's empowerment, as well as collective efficacy, and intervention cost and cost-
34
35 effectiveness will be assessed.
36
37

38 **Ethics and dissemination:** The study protocol has been reviewed and approved by the ethics
39
40 boards of the London School of Hygiene and Tropical Medicine, U.K. and KIIT University,
41
42 Bhubaneswar, India. Findings will be disseminated via peer-reviewed literature and presentation
43
44 to stakeholders, government officials, implementers and researchers.
45

46 **Trial registration identifier:** NCT02441699
47
48
49
50
51
52
53
54
55
56
57
58
59
60

46 STRENGTHS AND LIMITATIONS OF THIS STUDY

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

62 INTRODUCTION

63 Of the one billion people practicing open defecation worldwide, over half live in India[1]. While
64 international and national pressure on improving sanitation conditions in India has led to over
65 350 thousand people gaining access to improved toilets since 1990, it has barely kept up with
66 population growth[1,2]. Recent studies show that even in areas with access to household-level
67 improved sanitation, use of these toilets is low[3–5]. This may be due in part to a mismatch
68 between the culturally acceptable pour-flush toilets and the level of water access. Coverage of

1
2
3 69 improved water sources, usually community-level pumps or taps, is relatively high even in rural
4
5 70 areas in India, but it may not be sufficient for flushing purposes on top of other daily water
6
7
8 71 needs[1,6].
9

10
11 72 Although the effectiveness of water, sanitation and hygiene (WASH) interventions vary,
12
13 73 meta-analyses have found that individual or combined WASH interventions decrease diarrheal
14
15 74 disease prevalence by up to 48%[7–11]. While combined interventions would be expected to
16
17 75 have a greater influence on multiple exposure pathways and thus a greater combined impact on
18
19 76 health, there is limited evidence of additive benefits[12]. This may be due to poor uptake,
20
21 77 inconsistent use, or an incomplete understanding of relevant pathways[8–10]. In India,
22
23 78 combining water and sanitation interventions may be more critical than just interrupting multiple
24
25 79 transmission pathways for enteric infection; evidence suggests that household-level water access
26
27 80 is integral to the use of improved sanitation in this context[13].
28
29
30
31

32 81 While the intent of improved sanitation facilities is to separate human feces from human
33
34 82 contact, most of the focus is on constructing household toilets to increase improved sanitation
35
36 83 coverage—the primary metric used in monitoring progress toward international targets.
37
38 84 However, studies in India have further shown that toilet construction does not translate into toilet
39
40 85 use in this context[5,14–16]. Moreover, with the interdependence between members of
41
42 86 households and households within communities, safe water and sanitation is a community-level
43
44 87 issue. There is growing emphasis on assessing health risk from poor water and sanitation
45
46 88 conditions not simply due to individual or even household-level risk factors, but also from
47
48 89 conditions in the community environment[17]. There is evidence that even households without
49
50 90 toilets, and households which do not filter drinking water, showed decreased health risk if they
51
52 91 live in communities with high levels of coverage and use[18–20].
53
54
55
56
57
58
59
60

1
2
3 92 Moreover, the effectiveness of community interventions may be higher in communities with
4
5
6 93 positive perceptions of their collective ability to come together to improve their conditions.
7
8 94 Collective efficacy, a latent construct comprised of the structural and cognitive components that
9
10 95 facilitate a community's shared belief in its ability to come together and execute actions related
11
12 96 to a common goal, may explain some variance in intervention effectiveness across communities
13
14 97 receiving WASH interventions[21].
15
16

17 98 A main risk of poor WASH conditions is enteric infection, caused by a diverse array of
18
19 99 bacteria, viruses, protozoa, and parasites, including soil-transmitted helminths. These infections
20
21 100 may cause diarrhea, the second leading cause of mortality for children under five years
22
23 101 worldwide and in India, a leading cause of mortality regardless of age[22,23]. There is also
24
25 102 growing evidence that asymptomatic enteric infections may pose a similar risk, with repeat
26
27 103 enteric infections contributing to chronic malnutrition, environmental enteric dysfunction, poor
28
29 104 cognitive outcomes, and poor vaccine uptake[24–29]. Poor WASH conditions are also linked to
30
31 105 increased risk of respiratory infection, the leading cause of mortality for children under five
32
33 106 years worldwide[22,30,31]. Poor water and sanitation access can also affect the social, physical
34
35 107 and mental well-being of women, acting through pathways ranging from unsafe menstrual
36
37 108 hygiene management practices and increased risk of violence[32–34].
38
39
40
41
42
43
44
45

46 110 **Description of the intervention**

47
48 111 Over the past decades there has been a global commitment to determine water and sanitation
49
50 112 interventions with demonstrated effectiveness, not just efficacy[35]. Gram Vikas, a non-
51
52 113 governmental organization based in Odisha, India (<http://www.gramvikas.org/>), has responded
53
54 114 by implementing its MANTRA (Movement and Action Network for Transformation of Rural
55
56
57
58
59
60

1
2
3 115 Areas) water and sanitation program in more than 1000 villages since 2002[36]. This approach
4
5 116 includes both household-level piped water connections, and community-level mobilization for
6
7
8 117 culturally appropriate household toilets. A previous interrupted time series analysis of the
9
10 118 MANTRA intervention reported it to be protective against diarrheal diseases[37]. However, in
11
12 119 addition to limitations of design, this study relied on outcome data collected and reported by
13
14 120 Gram Vikas, the intervention implementer, and did not assess intervention coverage or impacts
15
16 121 on environmental fecal contamination.
17

18
19
20 122 The MANTRA water and sanitation intervention is rolled out in a three-phase process over
21
22 123 an average of three years. During the first, or Motivational, phase (approximately 8-12 mo),
23
24 124 representatives of Gram Vikas visit the identified village several times to assess village interest
25
26 125 and progress towards a set of Gram Vikas requirements, including: 1) the commitment of every
27
28 126 household to participate, 2) creation of a village corpus fund from contributions from every
29
30 127 household, and 3) development of village guidelines for maintenance and use of facilities.
31
32

33
34 128 Once this set of requirements is achieved, the village progresses into the second, or
35
36 129 Operational, phase of the intervention (approximately 17-35 mo). Each household constructs a
37
38 130 pour-flush toilet with two soak-pits and a separate bathing room. The households hire a local,
39
40 131 skilled mason and provide their own unskilled labor and locally available materials to complete
41
42 132 the superstructure. Gram Vikas provides external materials such as PVC pipes and porcelain
43
44 133 pans. At the same time, a water tank, community meeting space, and piped water distribution
45
46 134 system connected to every household, with taps in the toilet and bathing rooms and a separate tap
47
48 135 in the kitchen, is constructed through a similar collaborative process.
49
50

51
52
53 136 All households must construct a toilet and bathing room for the village to progress into
54
55 137 the final, or Completed, phase of the intervention, in which the water system is turned on.
56
57
58
59
60

1
2
3 138 Notably, this three-phase process only allows each household access to piped water once every
4
5 139 household in the village has a toilet and bathing room. This model contrasts with most previous
6
7
8 140 water and sanitation interventions, including those implemented under India's Total Sanitation
9
10 141 Campaign and other government programs, which do not require community-level sanitation
11
12 142 compliance and do not provide a piped water supply at the household level[38].
13
14

15 143

17 144 **Study aims**

18
19
20 145 The primary objective of this study is to evaluate the effectiveness of the combined household-
21
22 146 level water supply and sanitation intervention, as implemented by Gram Vikas in Odisha, India.
23

24 147 Toward that objective, this study aims to:

- 25
26
27 148 1) Assess the effectiveness of the intervention in improving water and sanitation
28
29 149 infrastructure coverage, access, and use, and to assess fecal sludge management practices
30
31 150 in intervention communities.
32
33
34 151 2) Assess the effectiveness of the intervention in reducing environmental fecal
35
36 152 contamination.
37
38
39 153 3) Assess the effectiveness of the intervention in improving health. This includes reported
40
41 154 diarrheal disease in children under 5 years (primary outcome), acute respiratory infection,
42
43 155 infection with soil-transmitted helminthes, nutritional status, environmental enteric
44
45 156 dysfunction, seroconversion for selected enteric pathogens, and urogenital diseases
46
47
48 157 associated with menstrual hygiene management practices. Mental and social well-being
49
50 158 will be explored through assessment of sanitation insecurity and women's empowerment.
51
52
53 159 4) Assess the cost and cost-effectiveness of the intervention.
54
55
56
57
58
59
60

1
2
3 160 5) Develop and assess a theoretically-grounded, empirically informed collective efficacy
4
5 161 scale; and determine the effect of collective efficacy on intervention effectiveness.
6
7
8 162

10 163 **METHODS**

12 164 **Setting**

13 165 The study is located in Ganjam and Gajapati districts in eastern Odisha, India (Figure 1). These
14
15 166 two contiguous districts were a single district until 1992. Over 44% of the population in these
16
17 167 districts is recognized by the Government of India as being below the poverty line (BPL)[39]. As
18
19 168 of 2008, a majority of households in both districts had access to an improved, likely community-
20
21 169 level, drinking water source, with over 23% of households in Ganjam having access to any
22
23 170 sanitation facility, compared to only 8% of households in Gajapati[39]. The area is primarily
24
25 171 rural and agrarian, and the climate is characterized by a monsoon season from June to
26
27 172 September, with an average rainfall of ~1400 mms/year.
28
29
30
31
32
33

34 173

36 174 **Study design**

37 175 This study uses a matched cohort design to assess the effectiveness of a completed intervention
38
39 176 with data collected across four study rounds from June 2015 to October 2016 (Figure 2). Data
40
41 177 was collected in all study rounds for diarrhea, acute respiratory infection, nutritional status, and
42
43 178 stored and source water outcomes to assess seasonality. Data was collected in rounds 2 and 4 for
44
45 179 environmental enteric dysfunction, seroconversion, and hand-rinses, and cross-sectionally in one
46
47 180 or more rounds for the remaining outcomes. As described below, control villages were matched
48
49 181 to randomly selected intervention villages through a multi-step restriction, genetic matching, and
50
51 182 exclusion process using the following eligibility criteria.
52
53
54
55
56
57
58
59
60

183

184 Eligibility criteria for villages

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

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

203 In addition, to be eligible for inclusion both intervention and control villages must: 1)
204 appear in the Government of India Census 2001 and the BPL Survey 2002, 2) have a population

1
2
3 205 of at least 20 households, and 3) be within approximately three hours travel from the study office
4
5
6 206 in Brahmapur, Ganjam District. This last criterion is due to logistical constraints.

7
8 207 *2. Matching.* After restriction, genetic matching was used to match potential control villages to
9
10 208 the randomly selected intervention villages without replacement[5,40,41]. Villages were exact
11
12 209 matched on district to limit any political or large scale geographic variation between district
13
14 210 populations, and were also matched on pre-intervention demographic, socioeconomic, sanitation,
15
16 211 and water access characteristics listed in Table 1[5]. These village-level matching variables were
17
18 212 selected due to their theorized association with the primary outcome, diarrheal diseases, as well
19
20 213 as data availability.

21
22 214 *3. Exclusion.* The field team visited matched potential control villages and intervention villages
23
24 215 to assess suitability for the study through a rapid assessment interview with village leadership
25
26 216 and to ensure accessibility. Villages were excluded if they are not within three hours travel of the
27
28 217 field office in Brahmapur, had sustained major infrastructure damage due to a natural disaster, or
29
30 218 if there was a current or planned sanitation or water intervention by an organization external to
31
32 219 the village in the next 12 months as determined through the rapid assessment interview with
33
34 220 village leadership. In addition, villages were excluded if there were fewer than three households
35
36 221 with children under five years old. As villages were removed from the pool of prospective
37
38 222 control villages, the matching process was repeated for all intervention villages and remaining
39
40 223 eligible control villages, and balance measures were assessed. The matching and exclusion
41
42 224 processes were repeated as necessary.

43
44 225 After the iterative matching and exclusion process was complete, covariate balance was
45
46 226 assessed for all matching variables for the final set of intervention and control villages through
47
48 227 examination of balance measures[42–44]. Matching resulted in an improvement in balance as
49
50
51
52
53
54
55
56
57
58
59
60

228 assessed through comparison of several measures including q-q plots, Kolmogorov-Smirnov
 229 bootstrap p-values, and standardized differences. After matching, there were no significant
 230 differences between intervention and control groups (Table 1).

231

232 Eligibility criteria for households

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

239

240 **Table 1.** Pre-intervention characteristics used in matching, and balance diagnostics before and
 241 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 [‡] (%)	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

242 All eligible: all villages that are eligible for the matching process after restriction

243 Std Diff (absolute standardized difference): a value greater than 0.1 is considered meaningful imbalance [42]

244 ‡ Ganjam villages only; no data available for Gajapati villages

245 Kolmogorov-Smirnov bootstrap p-values: * <0.05 ** <0.01

246

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

249

250 Sample Size

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

1
2
3 263
4
5
6 264 **Outcome Measurement**
7
8 265 Outcomes, and individual, household, and community-level risk factors, will be measured
9
10 266 through surveys, interviews, or through the collection and analysis of environmental, stool or
11
12 267 dried blood spot samples. All survey questions will be translated into the primary local language,
13
14
15 268 Odia, and back-translated to confirm wording. Household surveys include household and
16
17 269 individual factors and will be verbally administered by trained field workers to the mother or
18
19
20 270 primary caregiver of the youngest child under five in each household, unless otherwise specified
21
22 271 below. Community surveys will be verbally administered to the *sarpanch* (village head) or other
23
24 272 member of village leadership. Survey data will be collected on mobile phones using Open Data
25
26
27 273 Kit[47]. GPS coordinates for households, water sources and other relevant sites will be collected
28
29 274 using Garmin eTrex 10 or 20 devices (Garmin Ltd., Olathe, KS, USA).
30
31

32 275
33
34 276 Coverage, access and use of sanitation, water and hygiene infrastructure
35
36 277 Coverage, access and use of WASH infrastructure will be assessed in all four rounds. Presence
37
38 278 of and access to toilets, water sources and hand-washing stations will be assessed through
39
40 279 standard questions from the Demographic and Health Surveys (DHS) and confirmed through
41
42 280 spot observations. Spot observations of household toilets and hand-washing stations will be
43
44 281 further used to assess indicators of functionality, maintenance, recent use. Reported water and
45
46 282 sanitation practices, including child feces disposal practices, will be captured through household
47
48 283 survey questions.
49
50

51 284
52
53
54
55 285 Diarrheal Diseases
56
57
58
59
60

1
2
3 286 The primary outcome for this study is prevalence of diarrheal diseases, recorded as both daily
4
5
6 287 point prevalence over the previous three days and seven-day period prevalence, for all household
7
8 288 members in all four rounds. Although self-reported diarrhea is a subjective outcome with a well-
9
10 289 established risk of bias, three-day recall reduces recall bias[48,49]. Diarrheal disease will be
11
12 290 measured using the World Health Organization (WHO) definition of three or more loose stools
13
14 291 in a 24-hour period, with or without the presence of blood. Field workers will use a simple
15
16 292 calendar as a visual aid to help respondents with recall. Each household member will be asked to
17
18 293 recall his or her own disease status and the mother or primary caregiver will be asked to report
19
20 294 disease for children.
21
22
23
24
25
26

27 296 Respiratory infection

28
29 297 Prevalence of respiratory infections will be recorded as both daily point prevalence over the
30
31 298 previous three days and seven-day period prevalence for all household members in all four
32
33 299 rounds. Respiratory infection is defined as the presence of cough and/or shortness of
34
35 300 breath/difficulty breathing according to WHO's Integrated Management of Childhood Illness
36
37 301 (IMCI)[50]. The full IMCI case definition for acute lower respiratory infection also includes
38
39 302 measurement of respiratory rate and observation of chest indrawing, stridor and other danger
40
41 303 signs; these criteria were excluded from our definition as there was concern about the technical
42
43 304 support required to produce consistent and accurate data within this context[50]. Our definition
44
45 305 provides a broad assessment of respiratory illness burden. Each household member will be asked
46
47 306 to recall his or her own disease status and the mother or primary caregiver will be asked to report
48
49 307 disease for children.
50
51
52
53
54
55
56
57
58
59
60

309 Nutritional Status

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

321

322 Soil-transmitted helminth infection

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

1
2
3 331 infection prevalence between intervention and control villages, while the stool samples collected
4
5 332 approximately 8 months later in round 4 will provide a measure of re-infection rate.
6
7

8 333

9
10 334 Environmental enteric dysfunction

11
12 335 Stools from a randomly selected subset of 200 children under two years old, collected in rounds

13 336 2 and 4, will be used to assess environmental enteric dysfunction (EED) through quantification

14
15 337 of biomarkers of intestinal inflammation and permeability. Fecal myeloperoxidase (MPO),

16
17 338 alpha-1-antitrypsin (AAT), and neopterin (NEO), markers for neutrophil activity, intestinal

18
19 339 permeability and TH1 immune activation, respectively, were selected for this study based on

20
21 340 evidence of association with EED, subsequent linear growth deficits, and household

22
23 341 environmental fecal contamination[24,25,54].
24
25
26
27
28

29 342

30
31 343 Seroconversion for enteric pathogens

32
33 344 Serological assays that assess antibody production against various enteric pathogens can provide

34
35 345 an objective measure of exposure to enteric infections[55]. Enrolling children aged 6 to 18

36
37 346 months will reduce the potential for interference from maternally acquired antibodies and permit

38
39 347 analysis of seroconversion data in a critical window for young children who experience higher

40
41 348 diarrheal disease morbidity and mortality before two years of age[56–61]. Children who are 6 to

42
43 349 12 months during round 2 will have capillary blood drawn by fingerstick or heelstick, as

44
45 350 appropriate, and will be visited again during round 4 for a second capillary blood sample. All

46
47 351 blood samples will be preserved on TropBio (Sydney, Australia) filter discs and stored within 7

48
49 352 days of collection at -20°C. Seroconversion against markers for norovirus, *Giardia intestinalis*,

50
51 353 *Cryptosporidium parvum*, *Entamoeba histolytica*, enterotoxigenic *E. coli* heat-labile enterotoxin
52
53
54
55
56
57
58
59
60

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

356

357 Environmental fecal contamination

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

372

373 Cost and cost-effectiveness

374 Costs and potential cost savings (i.e., averted costs) associated with the intervention will be
375 assessed through an economic costing approach that recognizes and quantifies costs and benefits
376 from a societal perspective[72]. Data on program and point-of-delivery inputs will be collected at

1
2
3 377 household, community, and implementer levels in round 3. Field workers will administer
4
5 378 community surveys to a village leader, and household surveys to the household decision-maker
6
7
8 379 for toilet installation, in 20 randomly selected households in twenty matched intervention and
9
10 380 control villages. Given cost-effectiveness analyses require the effect of the intervention to be
11
12 381 measured against a counterfactual, and the intervention of interest is a community-based
13
14 382 intervention, cost and effectiveness measures will be summarized at the village level [73].
15
16 383 Surveys will collect data on household- and community-level inputs related to materials and
17
18 384 labor required to construct household toilets and wash rooms, the community water tank and
19
20 385 distribution system, and household water connections; longer-term water supply and toilet
21
22 386 maintenance costs; and financing required for this infrastructure as well as perceived benefits,
23
24 387 including averted social opportunity costs. Implementer inputs from Gram Vikas will be
25
26 388 collected through an enumeration exercise, interviews, and examination of the implementer's
27
28 389 financial records.
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

391 Collective efficacy

392 Collective efficacy (CE) is a latent construct comprised of the structural and cognitive
393 components that facilitate a community's shared belief in its ability to come together and execute
394 actions related to a common goal[21]. A review of the literature and established conceptual
395 frameworks will be performed to define the CE construct. A sequential exploratory mixed
396 qualitative and quantitative design will be used to develop and refine a scale to measure CE and
397 test hypotheses. Field workers will administer the refined, multi-item, Likert-type CE scale to
398 one randomly selected household member aged 18 years or older in each household in round 3.
399

1
2
3 400 Women's empowerment

4
5 401 Four dimensions of women's empowerment will be measured in rounds 3 and 4: group
6
7
8 402 participation, leadership, decision-making and freedom of movement. Group participation and
9
10 403 leadership will be measured using modules from the Women's Empowerment in Agriculture
11
12 404 Index (WEAI), which has been tested in South Asia[74]. Decision-making will be measured
13
14 405 using questions from the women's status module of Demographic and Health Surveys. Freedom
15
16 406 of movement will be measured using questions from the project-level Women's Empowerment
17
18 407 in Agriculture Index (pro-WEAI). These measures will be collected for the primary female
19
20 408 caregiver of the youngest child under 5, and were selected based on the importance of women's
21
22 409 empowerment for child nutrition[75,76]. Women's empowerment is conceptualized as both an
23
24 410 outcome and a potential mediator along the pathway between the Gram Vikas intervention and
25
26 411 child health outcomes.
27
28
29
30
31

32 412
33
34 413 Menstrual hygiene management

35
36 414 Menstrual hygiene management practices vary worldwide and depend on personal preference,
37
38 415 socioeconomic status, local traditions and beliefs, and access to water and sanitation
39
40 416 resources[77]. Unhygienic washing practices are common in rural India and among women and
41
42 417 girls in lower socioeconomic groups, and may increase risk of urogenital infection[78–80].
43
44 418 However, the link between access to water and sanitation, menstrual hygiene management and
45
46 419 urogenital infections has been poorly studied. Household surveys will be administered in round 4
47
48 420 to a randomly selected woman aged 18 or older, in a subset of 800 households, and will capture
49
50 421 self-reported urogenital infection, defined as at least one of the following symptoms: 1) abnormal
51
52
53
54
55
56
57
58
59
60

1
2
3 422 vaginal discharge (unusual texture and color/more abundant than normal), 2) burning or itching
4
5 423 in the genitalia, 3) burning or itching when urinating, or 4) genital sores[79].
6
7

8 424

9
10 425 Sanitation Insecurity

11
12 426 This study will assess the associations between sanitation access and sanitation insecurity with
13
14 427 mental health among women. In previous research in Odisha, a contextually specific definition
15
16 428 and measure for sanitation insecurity was developed, with associations between facets of
17
18 429 sanitation insecurity and mental health independent of sanitation facility access[81]. This
19
20 430 previously developed measure will be used to determine if levels of sanitation insecurity differ
21
22 431 between intervention and control villages and how it may be associated with mental health
23
24 432 outcomes, specifically well-being, anxiety, depression, and distress. Household surveys will be
25
26 433 administered in round 4 to a randomly selected woman aged 18 or older, in a random subset of
27
28 434 800 households.
29
30
31
32

33 435

34
35
36 436 Fecal sludge management.

37
38 437 In sanitation systems where sewerage is not feasible, such as the household toilets constructed as
39
40 438 part of the MANTRA intervention, safe management of fecal waste is necessary. Although there
41
42 439 is growing emphasis on safe fecal sludge management (FSM), research has mainly focused on
43
44 440 urban settings[82,83]. Preliminary research in Odisha suggests that fecal sludge management in
45
46 441 this rural setting is a substantial challenge, and may impact household use of toilets. In round 3,
47
48 442 household surveys and spot checks of toilets in intervention villages will be used to assess toilet
49
50 443 use and fecal sludge management practices.
51
52

53 444
54
55
56
57
58
59
60

445 STATISTICAL ANALYSES

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

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

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

1
2
3 468 DALY, will be calculated by dividing the incremental intervention costs by the incremental
4
5
6 469 intervention benefits.

7
8 470 The collective efficacy scale will be analyzed using a psychometric approach in which
9
10 471 factor analytics are employed to identify an appropriate factor solution and test the reliability and
11
12 472 validity of the CE scores. Once a CE factor solution and an empirically derived multilevel data
13
14 473 structure have been identified, the association between CE and intervention effectiveness will be
15
16 474 analyzed using multilevel generalized linear mixed models to estimate relative risks[85,86]. (aim
17
18 475 5). For all outcomes, variables used in the matching process may be considered as covariates, as
19
20 476 needed, in addition to individual, household, and community-level risk factors. Covariates that
21
22 477 are statistically associated with outcomes of interest in bivariate analyses will be considered for
23
24 478 inclusion in final multivariable models, following standard stepwise model-building approaches.
25
26 479 Secondary analyses may also evaluate models for effect modification as relevant, including
27
28 480 exposure-mediator interaction for mediation models and cross-level interaction, by assessing
29
30 481 changes in parameter values based on potential effect modifiers. Potential effect modifiers may
31
32 482 include breastfeeding for seroconversion outcomes, and climate factors and population density
33
34 483 for environmental fecal contamination and health outcomes. However, this study was not
35
36 484 designed to assess effect modification and therefore is not specifically powered for these
37
38 485 analyses. For all outcomes, unadjusted models will be presented along with models adjusting for
39
40 486 covariates.
41
42
43
44
45
46
47
48
49

50 488 **DISCUSSION**

51
52 489 This matched cohort study is one of the first to evaluate the effect of a rural combined
53
54 490 household-level piped water and sanitation intervention, implemented at the community level, on
55
56
57
58
59
60

1
2
3 491 a large scale. The matched design provides a rigorous means for estimating causal effects given
4
5 492 that randomization to intervention group was not feasible due to the several year implementation
6
7
8 493 process[5]. By focusing on an intervention where the implementation process is complete, it also
9
10 494 limits the risk presented by randomized controlled trials, where the intervention has little uptake,
11
12 495 an especially important study challenge given interdependence of exposure and outcomes within
13
14 496 communities, and a problem that has characterized previous trials of sanitation interventions in
15
16
17 497 India[15,16].

18
19
20 498 A strength of this study is the assessment of health impacts using the holistic WHO
21
22 499 definition of health, including not just disease status, but also mental, social, and physical well-
23
24 500 being[87]. Outcomes along the causal chain include standard, but more subjective measures,
25
26 501 such as reported diarrheal diseases and respiratory infection, as well as more objective measures
27
28 502 such as fecal environmental contamination, soil transmitted helminth infection, and
29
30 503 anthropometry. Although there is risk of response bias for reported outcomes, it is unlikely to be
31
32 504 differential by intervention status since the study team is not directly linked to Gram Vikas. Even
33
34 505 though field workers may be aware of village intervention status, lab staff analyzing water, hand
35
36 506 rinse, stool, and blood samples will be blinded. In addition, this study includes the more novel
37
38 507 use of seroconversion for enteric pathogens, biomarkers of environmental enteric dysfunction,
39
40 508 and measures of collective efficacy in an evaluation assessment. While there are limitations
41
42 509 inherent to observational studies, the matched study design and multivariable modeling analysis
43
44 510 plan reduce the potential for confounding. However, there is still the potential for residual
45
46 511 unmeasured confounding.
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 513 **Ethics and Dissemination.** This study has been reviewed and approved by the Ethics Committee
4
5
6 514 of the London School of Hygiene and Tropical Medicine, U.K (No. 9071) and Institute Ethics
7
8 515 Committee of the Kalinga Institute of Medical Sciences of KIIT University, Bhubaneswar, India
9
10 516 (KIMS/KIIT/IEC/053/2015). Efforts will be made to communicate the central findings and
11
12 517 implications with study communities, the implementing organization and government officials in
13
14 518 India. The results of this study will be submitted for publication in peer reviewed journals and
15
16 519 presented at conferences. The data collected in the study will be publicly available, with personal
17
18 520 identifiable data redacted, following the publication of the primary results within 24 months of
19
20 521 the final data collection date.
21
22
23
24
25
26

27 523 **Funding.** This study is supported by a grant from the Bill & Melinda Gates Foundation to the
28
29 524 London School of Hygiene & Tropical Medicine (OPP1008048) and to Emory University.
30
31 525 (OOP1125067).
32
33
34
35

36 527 **Competing Interests:** None declared.
37
38
39
40

41 529 **Contributions from authors:** TC, HR, PR, BT, and HC contributed to study design. HR, LZ
42
43 530 and BT developed laboratory protocols. HR, BT, GS, MD, SS, LZ, and BC developed data
44
45 531 collection tools. All authors contributed to editing and revising the manuscript.
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 536 **Figure 1.** Study sites in Ganjam and Gajapati districts, Odisha, India with intervention villages
4
5
6 537 in black and control villages in white. Inset shows location of districts in India.
7

8 538

9
10 539 **Figure 2.** Restriction, matching and exclusion process for selection of intervention and control
11
12 540 villages (1), and timeline for study rounds and outcome data collection (2).
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For peer review only

541 **REFERENCES**

- 542 1 UNICEF, WHO. Progress and Sanitation and Drinking Water: 2015 Update and MDG
543 Assessment. 2015.
- 544 2 Water and Sanitation Program. A Decade of the Total Sanitation Campaign: Rapid
545 Assessment of Processes and Outcomes. 2010.
- 546 3 Coffey D, Gupta A, Hathi P, *et al*. Culture and the health transition: Understanding
547 sanitation behavior in rural north India. 2015.
- 548 4 Clasen T, Pruss-Ustun A, Mathers CD, *et al*. Estimating the impact of unsafe water,
549 sanitation and hygiene on the global burden of disease: evolving and alternative methods.
550 *Trop Med Int Heal* 2014;**19**:884–93. doi:10.1111/tmi.12330
- 551 5 Arnold BF, Khush RS, Ramaswamy P, *et al*. Causal inference methods to study
552 nonrandomized, preexisting development interventions. *Proc Natl Acad Sci U S A*
553 2010;**107**:22605–10. doi:10.1073/pnas.1008944107
- 554 6 Ercumen A, Arnold BF, Kumpel E, *et al*. Upgrading a Piped Water Supply from
555 Intermittent to Continuous Delivery and Association with Waterborne Illness: A Matched
556 Cohort Study in Urban India. *PLOS Med* 2015;**12**:e1001892.
557 doi:10.1371/journal.pmed.1001892
- 558 7 Fewtrell L, Kaufmann RB, Kay D, *et al*. Water, sanitation, and hygiene interventions to
559 reduce diarrhoea in less developed countries: a systematic review and meta-analysis.
560 *Lancet Infect Dis* 2005;**5**:42–52. doi:10.1016/S1473-3099(04)01253-8
- 561 8 Engell RE, Lim SS. Does clean water matter? An updated meta-analysis of water supply
562 and sanitation interventions and diarrhoeal diseases. *Lancet* 2013;**381**:S44.
563 doi:10.1016/S0140-6736(13)61298-2

- 1
2
3 564 9 Clasen T, Bostoen K, Schmidt W, *et al.* Interventions to improve disposal of human
4
5
6 565 excreta for preventing diarrhoea. *Cochrane database Syst Rev* Published Online First:
7
8 566 2010.<http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD007180.pub2/pdf/standard>
9
10 567 (accessed 4 Jan2015).
- 11
12 568 10 Clasen T, Schmidt W-P, Rabie T, *et al.* Interventions to improve water quality for
13
14
15 569 preventing diarrhoea: systematic review and meta-analysis. *BMJ* 2007;**334**:782.
16
17 570 doi:10.1136/bmj.39118.489931.BE
- 18
19
20 571 11 Cairncross S, Hunt C, Boisson S, *et al.* Water, sanitation and hygiene for the prevention of
21
22 572 diarrhoea. *Int J Epidemiol* 2010;**39 Suppl 1**:i193–205. doi:10.1093/ije/dyq035
- 23
24
25 573 12 Clasen TF, Alexander KT, Sinclair D, *et al.* Interventions to improve water quality for
26
27 574 preventing diarrhoea. *Cochrane database Syst Rev* 2015;**10**:CD004794.
28
29 575 doi:10.1002/14651858.CD004794.pub3
- 30
31
32 576 13 Routray P, Schmidt W-P, Boisson S, *et al.* Socio-cultural and behavioural factors
33
34 577 constraining latrine adoption in rural coastal Odisha: an exploratory qualitative study.
35
36 578 *BMC Public Health* 2015;**15**:880. doi:10.1186/s12889-015-2206-3
- 37
38
39 579 14 Barnard S, Routray P, Majorin F, *et al.* Impact of Indian Total Sanitation Campaign on
40
41 580 Latrine Coverage and Use: A Cross-Sectional Study in Orissa Three Years following
42
43 581 Programme Implementation. *PLoS One* 2013;**8**. doi:10.1371/journal.pone.0071438
- 44
45
46 582 15 Patil SR, Arnold BF, Salvatore AL, *et al.* The effect of India's total sanitation campaign
47
48 583 on defecation behaviors and child health in rural Madhya Pradesh: a cluster randomized
49
50 584 controlled trial. *PLoS Med* 2014;**11**:e1001709. doi:10.1371/journal.pmed.1001709
- 51
52
53 585 16 Clasen T, Boisson S, Routray P, *et al.* Effectiveness of a rural sanitation programme on
54
55 586 diarrhoea, soil-transmitted helminth infection, and child malnutrition in Odisha, India: a

- 1
2
3 587 cluster-randomised trial. *Lancet Glob Heal* 2014;**2**:e645–53. doi:10.1016/S2214-
4
5 588 109X(14)70307-9
6
7
8 589 17 Eisenberg JNS, Trostle J, Sorensen RJD, *et al.* Towards a Systems Approach to Enteric
9
10 590 Pathogen Transmission: From Individual Independence to Community Interdependence.
11
12 591 *Annu Rev Public Health* 2013;**33**:239–57. doi:10.1016/j.micinf.2011.07.011.Innate
13
14
15 592 18 Root GPM. Sanitation, Community Environments, and Childhood Diarrhea in Rural
16
17 593 Zimbabwe. *J Heal Popul Nutr* 2001;**19**:73–82.
18
19
20 594 19 Bateman OM, Smith S. A Comparison of the Health Effects of Water Supply and
21
22 595 Sanitation in Urban and Rural Guatemala. 1999.
23
24 596 20 Huq A, Yunus M, Sohel SS, *et al.* Simple Sari Cloth Filtration of Water Is Sustainable and
25
26 597 Continues to Protect Villagers from Cholera in Matlab, Bangladesh. *MBio* 2010;**1**:1–5.
27
28 598 doi:10.1128/mBio.00034-10.Invited
29
30
31 599 21 Bandura A. *Self-Efficacy: The Exercise of Control*. Worth Publishers 1997.
32
33
34 600 22 UN Inter-agency Group for Child Mortality Estimation. Levels and trends in child
35
36 601 mortality: Report 2015. 2015.
37
38
39 602 23 World Health Organization. Estimated total deaths by cause, sex and WHO Member State,
40
41 603 2008. Geneva, Switzerland: 2011.
42
43
44 604 24 Lin A, Arnold BF, Afreen S, *et al.* Household environmental conditions are associated
45
46 605 with enteropathy and impaired growth in rural Bangladesh. *Am J Trop Med Hyg*
47
48 606 2013;**89**:130–7. doi:10.4269/ajtmh.12-0629
49
50
51 607 25 Korpe PS, Petri WA. Environmental enteropathy: critical implications of a poorly
52
53 608 understood condition. *Trends Mol Med* 2012;**18**:328–36.
54
55 609 doi:10.1016/j.molmed.2012.04.007
56
57
58
59
60

- 1
2
3 610 26 Guerrant RL, Oriá RB, Moore SR, *et al.* Malnutrition as an enteric infectious disease with
4
5 long-term effects on child development. 2008;**66**:487–505. doi:10.1111/j.1753-
6 611
7 4887.2008.00082.x.Malnutrition
8 612
9
10 613 27 Crane RJ, Jones KDJ, Berkley JA. Environmental enteric dysfunction : An overview.
11
12 614 2015;**36**:76–87.
13
14 615 28 Humphrey JH. Child undernutrition, tropical enteropathy, toilets, and handwashing.
15
16 616 *Lancet* 2009;**374**:1032–5. doi:10.1016/S0140-6736(09)60950-8
17
18 617 29 Mbuya MNN, Humphrey JH. Preventing environmental enteric dysfunction through
19
20 618 improved water, sanitation and hygiene: An opportunity for stunting reduction in
21
22 619 developing countries. *Matern Child Nutr* 2015;**12**:1–15. doi:10.1111/mcn.12220
23
24 620 30 Budge PJ, Griffin MR, Edwards KM, *et al.* Impact of home environment interventions on
25
26 621 the risk of influenza-associated ARI in Andean children: observations from a prospective
27
28 622 household-based cohort study. *PLoS One* 2014;**9**:e91247.
29
30 623 doi:10.1371/journal.pone.0091247
31
32 624 31 Aiello AE, Coulborn RM, Perez V, *et al.* Effect of hand hygiene on infectious disease risk
33
34 625 in the community setting: a meta-analysis. *Am J Public Health* 2008;**98**:1372–81.
35
36 626 doi:10.2105/AJPH.2007.124610
37
38 627 32 Caruso BA, Sevilimedu V, Fung IC-H, *et al.* Gender disparities in water, sanitation, and
39
40 628 global health. *Lancet (London, England)* 2015;**386**:650–1. doi:10.1016/S0140-
41
42 629 6736(15)61497-0
43
44 630 33 Sahoo KC, Hulland KR, Caruso B a., *et al.* Sanitation-related psychosocial stress: a
45
46 631 grounded theory study of women across the life-course in Odisha, India. *Soc Sci Med*
47
48 632 2015;**139**:80–9. doi:10.1016/j.socscimed.2015.06.031
49
50
51
52
53
54
55
56
57
58
59
60

- 1
2
3 633 34 Hulland KRS, Chase RP, Caruso BA, *et al.* Sanitation, stress, and life stage: A systematic
4
5 634 data collection study among women in Odisha, India. *PLoS One* 2015;**10**.
6
7
8 635 doi:10.1371/journal.pone.0141883
9
10 636 35 World Health Organization/The United Nations Children’s Fund (UNICEF). Ending
11
12 637 Preventable Child Deaths from Pneumonia and Diarrhoea by 2025: The integrated Global
13
14 638 Action Plan for Pneumonia and Diarrhoea (GAPPD). 2013.
15
16
17 639 36 Gram Vikas. Annual Report 2013-2014. 2014.
18
19 640 37 Duflo E, Greenstone M, Guiteras R, *et al.* Toilets Can Work : Short and Medium Run
20
21 641 Health Impacts of Addressing Complementarities and Externalities in Water and
22
23 642 Sanitation.
24
25
26 643 38 Rosa G, Clasen T. Estimating the scope of household water treatment in low- and
27
28 644 medium-income countries. *Am J Trop Med Hyg* 2010;**82**:289–300.
29
30 645 doi:10.4269/ajtmh.2010.09-0382
31
32
33 646 39 International Institute for Population Sciences (IIPS). District Level Household and
34
35 647 Facility Survey (DLHS-3), 2007-08: India, Orissa. Mumbai: 2010.
36
37
38 648 40 Brady H, Caughey D, Dehejia R, *et al.* Genetic Matching for Estimating Causal Effects :
39
40 649 2012.
41
42
43 650 41 Sekhon JS. “Multivariate and Propensity Score Matching Software. *J Stat Softw*
44
45 651 2011;**42**.[http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Multivariate+a](http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Multivariate+and+Propensity+Score+Matching#9)
46
47 652 [nd+Propensity+Score+Matching#9](http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Multivariate+and+Propensity+Score+Matching#9) (accessed 30 May2014).
48
49
50 653 42 Austin PC. Balance diagnostics for comparing the distribution of baseline covariates
51
52 654 between treatment groups in propensity-score matched samples. *Stat Med* 2009;**28**:3083–
53
54 655 107. doi:10.1002/sim

- 1
2
3 656 43 Rubin DB. Using propensity scores to help design observational studies: Application to
4
5 657 the tobacco litigation. *Health Serv Outcomes Res Methodol* 2002;**2**:169–88.
6
7
8 658 44 Harder VS, Stuart EA, Anthony JC. Propensity score techniques and the assessment of
9
10 659 measured covariate balance to test causal associations in psychological research. *Psychol*
11
12 660 *Methods* 2010;**15**:997–1003. doi:10.1016/j.biotechadv.2011.08.021.Secreted
13
14
15 661 45 Faul F, Erdfelder E, Buchner A, *et al*. Statistical power analyses using G*Power 3.1: tests
16
17 662 for correlation and regression analyses. *Behav Res Methods* 2009;**41**:1149–60.
18
19 663 doi:10.3758/BRM.41.4.1149
20
21
22 664 46 Wolf J, Prüss-Ustün A, Cumming O, *et al*. Systematic review: Assessing the impact of
23
24 665 drinking water and sanitation on diarrhoeal disease in low- and middle-income settings:
25
26 666 Systematic review and meta-regression. *Trop Med Int Heal* 2014;**19**:928–42.
27
28 667 doi:10.1111/tmi.12331
29
30
31 668 47 Carl Hartung YAWBALCTGB. Open Data Kit: Tools to Build Information Services for
32
33 669 Developing Regions. <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.176.8017>
34
35
36 670 48 Feikin DR, Audi a., Olack B, *et al*. Evaluation of the optimal recall period for disease
37
38 671 symptoms in home-based morbidity surveillance in rural and urban Kenya. *Int J*
39
40 672 *Epidemiol* 2010;**39**:450–8. doi:10.1093/ije/dyp374
41
42
43 673 49 Arnold BF, Galiani S, Ram PK, *et al*. Optimal Recall Period for Caregiver-reported Illness
44
45 674 in Risk Factor and Intervention Studies: A Multicountry Study. *Am J Epidemiol*
46
47 675 2013;**177**:361–70. doi:10.1093/aje/kws281
48
49
50 676 50 World Health Organization. Integrated Management of Childhood Illness: Chart Booklet.
51
52 677 51 Cogill B. Anthropometric indicators measurement guide. Revised edition. *Washington,*
53
54 678 *DC, Acad Educ Dev [AED], Food Nutr Tech Assist Proj*

- 1
2
3 679 2003;:92.http://www.developmentgateway.org/download/202582/anthro_2003.pdf
4
5
6 680 52 de Onis M, Onyango AW, Van den Broeck J, *et al.* Measurement and standardization
7
8 681 protocols for anthropometry used in the construction of a new international growth
9
10 682 reference. *Food Nutr Bull* 2004;**25**:S27–36.
11
12
13 683 53 Truant a. L, Elliott SH, Kelly MT, *et al.* Comparison of formalin-ethyl ether
14
15 684 sedimentation, formalin-ethyl acetate sedimentation, and zinc sulfate flotation techniques
16
17 685 for detection of intestinal parasites. *J Clin Microbiol* 1981;**13**:882–4.
18
19
20 686 54 Kosek M, Haque R, Lima A, *et al.* Fecal markers of intestinal inflammation and
21
22 687 permeability associated with the subsequent acquisition of linear growth deficits in
23
24 688 infants. *Am J Trop Med Hyg* 2013;**88**:390–6. doi:10.4269/ajtmh.2012.12-0549
25
26
27 689 55 Crump J a, Mendoza CE, Priest JW, *et al.* Comparing serologic response against enteric
28
29 690 pathogens with reported diarrhea to assess the impact of improved household drinking
30
31 691 water quality. *Am J Trop Med Hyg* 2007;**77**:136–41.
32
33
34 692 56 Steinberg EB, Mendoza CE, Glass R, *et al.* Prevalence of Infection with Waterborne
35
36 693 Pathogens: a Seroepidemiologic Study in Children 6-36 Months Old in San Juan
37
38 694 Sacatepquez, Guatemala. 2004;**70**:83–8.
39
40
41 695 57 Brussow H, Sidoti J, Link H, *et al.* Age-specific prevalence of antibody to enterotoxigenic
42
43 696 *Escherichia coli* in Ecuadorian and German children. *J Infect Dis* 1990;**162**:974–7.
44
45
46 697 58 Khanna B, Cutler A, Israel N, *et al.* Use caution with serologic testing for *Helicobacter*
47
48 698 *pylori* infection in children. *J Infect Dis* 1998;**178**:460–5.
49
50
51 699 59 Lindkvist P, Asrat D, Nilsson I. Age at acquisition of *Helicobacter pylori* infection:
52
53 700 comparison of a high and a low prevalence country. *Scand J Infect Dis* 1996;**28**:181–4.
54
55
56 701 60 Ungar B, Gilman R, Lanata C, *et al.* Seroepidemiology of *Cryptosporidium* infection in
57
58
59
60

- 1
2
3 702 two Latin American populations. *J Infect Dis* 1988;**157**:551–6.
4
5
6 703 61 Vitral C, Yoshida C, Lemos E, *et al.* Age-specific prevalence of antibodies to hepatitis A
7
8 704 in children and adolescents from Rio de Janeiro, Brazil, 1978 and 1995: relationship of
9
10 705 prevalence to environmental factors. *Mem Inst Oswaldo Cruz* 1998;**93**:1–5.
11
12
13 706 62 Lammie PJ, Moss DM, Brook Goodhew E, *et al.* Development of a new platform for
14
15 707 neglected tropical disease surveillance. *Int J Parasitol* 2012;**42**:797–800.
16
17 708 doi:10.1016/j.ijpara.2012.07.002
18
19
20 709 63 Centers for Disease Control and Prevention. Laboratory Methods for the Diagnosis of
21
22 710 *Vibio cholerae*. Atlanta, Georgia: 1999.
23
24
25 711 64 Centers for Disease Control and Prevention. Isolation and Identification of *Shigella*.
26
27 712 Atlanta, Georgia:
28
29 713 65 United States Environmental Protection Agency. Coliforms—Total and *E. coli*, Membrane
30
31 714 Filtration Method 10029. 1999.
32
33
34 715 66 Gronewold AD, Borsuk ME, Wolpert RL, *et al.* An Assessment of Fecal Indicator
35
36 716 Bacteria-Based Water Quality Standards. *Environ Sci Technol* 2008;**42**:4676–82.
37
38 717 doi:10.1021/es703144k
39
40
41 718 67 Gruber JS, Ercumen A, Colford JM. Coliform bacteria as indicators of diarrheal risk in
42
43 719 household drinking water: systematic review and meta-analysis. *PLoS One*
44
45 720 2014;**9**:e107429. doi:10.1371/journal.pone.0107429
46
47
48 721 68 Levy K, Nelson KL, Hubbard A, *et al.* Rethinking indicators of microbial drinking water
49
50 722 quality for health studies in tropical developing countries: case study in northern coastal
51
52 723 Ecuador. *Am J Trop Med Hyg* 2012;**86**:499–507. doi:10.4269/ajtmh.2012.11-0263
53
54
55 724 69 Livio S, Strockbine N a, Panchalingam S, *et al.* *Shigella* Isolates From the Global Enteric

- 1
2
3 725 Multicenter Study Inform Vaccine Development. *Clin Infect Dis* 2014;**59**.
4
5 726 doi:10.1093/cid/ciu468
6
7
8 727 70 Kotloff KL, Nataro JP, Blackwelder WC, *et al*. Burden and aetiology of diarrhoeal disease
9
10 728 in infants and young children in developing countries (the Global Enteric Multicenter
11
12 729 Study, GEMS): a prospective, case-control study. *Lancet* 2013;**382**:209–22.
13
14 730 doi:10.1016/S0140-6736(13)60844-2
15
16
17 731 71 Kanungo S, Sah BK, Lopez a. L, *et al*. Cholera in India: An analysis of reports, 1997-
18
19 732 2006. *Bull World Health Organ* 2010;**88**:185–91. doi:10.2471/BLT.09.073460
20
21
22 733 72 Cellini SR, Kee JE. Cost-Effectiveness and Cost-Benefit Analysis. In: Wholey JS, Hatry
23
24 734 HP, Newcomer KE, eds. *Handbook of Practical Program Evaluation*. San Francisco: :
25
26 735 John Wiley & Sons, Inc. 2010.
27
28
29 736 73 Edejer TT, Baltussen R, Adam T, *et al*. Making Choices in Health: WHO Guide to Cost-
30
31 737 Effectiveness Analysis. 2003.
32
33
34 738 74 Malapit HJ, Sproule K, Kovarik C, *et al*. Measuring progress toward empowerment
35
36 739 Women’s Empowerment in Agriculture Index: Baseline Report. *IfpriOrg* 2014;:1–60.
37
38
39 740 75 Black RE, Victora CG, Walker SP, *et al*. Maternal and child undernutrition and
40
41 741 overweight in low-income and middle-income countries. *Lancet* 2013;**382**:427–51.
42
43 742 doi:10.1016/S0140-6736(13)60937-X
44
45
46 743 76 Smith LC, Ramakrishnan U, Ndiaye A, *et al*. The Importance of Women’s Status for
47
48 744 Child Nutrition in Developing Countries. 2003.
49
50 745 <http://www.ifpri.org/publication/importance-womens-status-child-nutrition-developing->
51
52 746 [countries](http://www.ifpri.org/publication/importance-womens-status-child-nutrition-developing-)
53
54
55 747 77 Sumpter C, Torondel B. A Systematic Review of the Health and Social Effects of
56
57
58
59
60

- 1
2
3 748 Menstrual Hygiene Management. *PLoS One* 2013;**8**:e62004.
4
5
6 749 doi:10.1371/journal.pone.0062004
7
8 750 78 Dasgupta A, Sarkar M. Menstrual Hygiene: How Hygienic is the Adolescent Girl? *Indian*
9
10 751 *J Community Med* 2008;**33**:77–80. doi:10.4103/0970-0218.40872
11
12 752 79 Das P, Baker KK, Dutta A, *et al.* Menstrual Hygiene Practices, WASH Access and the
13
14 Risk of Urogenital Infection in Women from Odisha, India. *PLoS One* 2015;**10**:e0130777.
15 753
16 doi:10.1371/journal.pone.0130777
17 754
18
19 755 80 van Eijk AM, Sivakami M, Thakkar MB, *et al.* Menstrual hygiene management among
20
21 adolescent girls in India: a systematic review and meta-analysis. *BMJ Open*
22 756
23 2016;**6**:e010290. doi:10.1136/bmjopen-2015-010290
24 757
25
26 758 81 Caruso B. *Sanitation Insecurity: Definition, Measurement, and Associations with*
27
28 *Women's Mental Health in Rural Orissa, India.*
29 759
30 2015.<http://pid.emory.edu/ark:/25593/rfhnt>
31 760
32
33 761 82 Peal A, Evans B, Blackett I, *et al.* Fecal sludge management (FSM): analytical tools for
34
35 assessing FSM in cities. *J Water, Sanit Hyg Dev* 2014;**4**:371.
36 762
37 doi:10.2166/washdev.2014.139
38 763
39
40 764 83 Peal A, Evans B, Blackett I, *et al.* Fecal Sludge Management: a comparative analysis of
41
42 12 cities. *J Water, Sanit Hyg Dev* 2014;**4**:563–75.
43 765
44
45 766 84 Richiardi L, Bellocco R, Zugna D. Mediation analysis in epidemiology: Methods,
46
47 interpretation and bias. *Int J Epidemiol* 2013;**42**:1511–9. doi:10.1093/ije/dyt127
48 767
49
50 768 85 McNutt LA, Wu C, Xue X, *et al.* Estimating the relative risk in cohort studies and clinical
51
52 trials of common outcomes. *Am J Epidemiol* 2003;**157**:940–3. doi:10.1093/aje/kwg074
53 769
54
55 770 86 Thompson ML, Myers JE, Kriebel D. Prevalence odds ratio or prevalence ratio in the
56
57
58
59
60

1
2
3 771 analysis of cross sectional data: what is to be done? *Occup Environ Med* 1998;**55**:272–7.

4
5 772 doi:10.1136/oem.55.4.272

6
7
8 773 87 World Health Organization. WHO definition of health.

9
10 774 <http://www.who.int/about/definition/en/print.html> (accessed 26 Apr2016).

11
12
13 775
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For peer review only

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

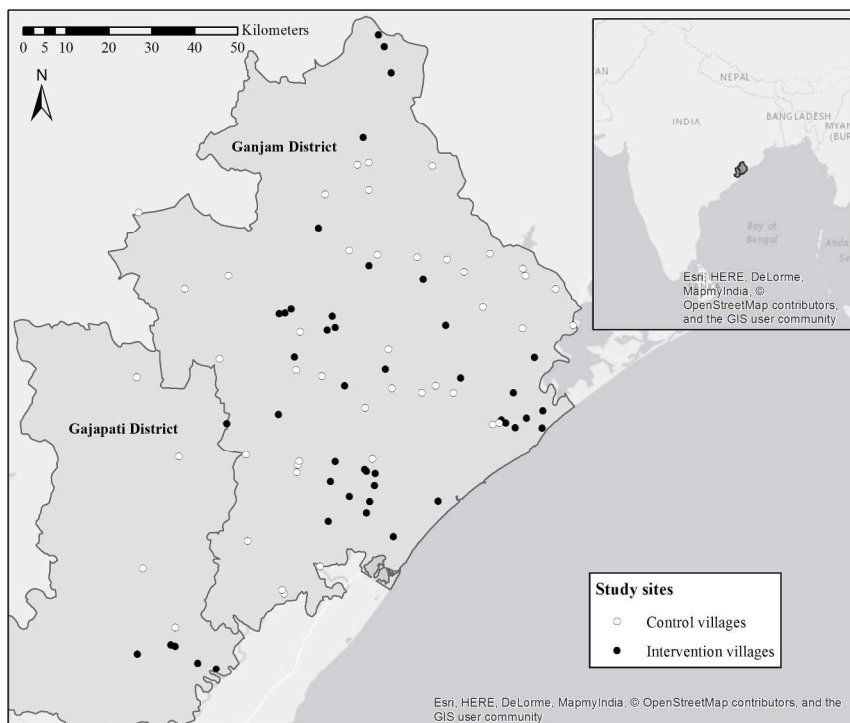


Figure 1. Study sites in Ganjam and Gajapati districts, Odisha, India with intervention villages in black and control villages in white. Inset shows location of districts in India.

Figure 1
215x279mm (300 x 300 DPI)

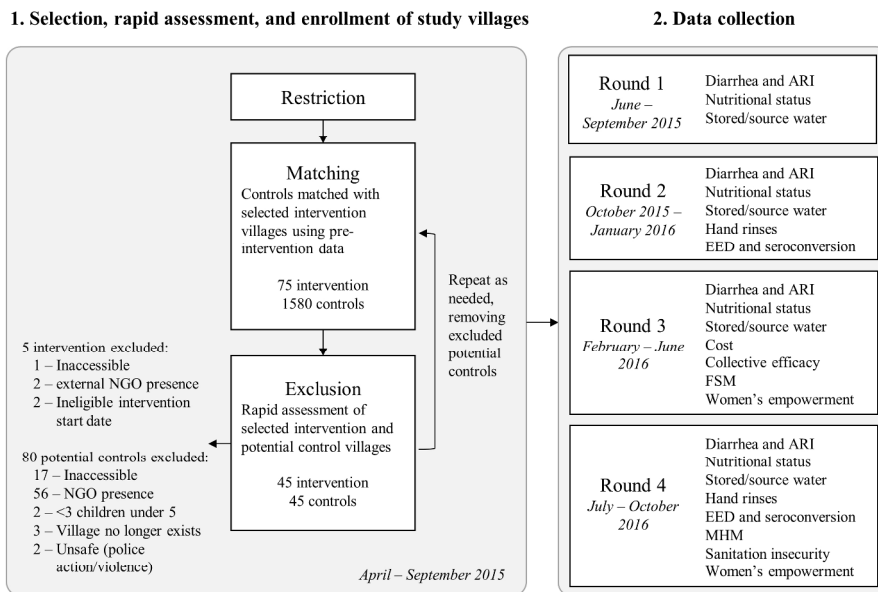


Figure 2. Restriction, matching and exclusion process for selection of intervention and control villages (1), and timeline for study rounds and outcome data collection (2).

Figure 2
254x190mm (300 x 300 DPI)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

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) <i>Cohort study</i> —Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up <i>Case-control study</i> —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 <i>Cross-sectional study</i> —Give the eligibility criteria, and the sources and methods of selection of participants	8-11	
		(b) <i>Cohort study</i> —For matched studies, give matching criteria and number of exposed and unexposed <i>Case-control study</i> —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) <i>Cohort study</i> —If applicable, explain how loss to follow-up was addressed <i>Case-control study</i> —If applicable, explain how matching of cases and controls was addressed <i>Cross-sectional study</i> —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) <i>Cohort study</i> —Summarise follow-up time (eg, average and total amount)	NA
Outcome data	15*	<i>Cohort study</i> —Report numbers of outcome events or summary measures over time	NA
		<i>Case-control study</i> —Report numbers in each exposure category, or summary measures of exposure	
		<i>Cross-sectional study</i> —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.