medRxiv preprint doi: https://doi.org/10.1101/2022.08.19.22278991; this version posted August 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Title

Long-term impacts of an urban sanitation intervention on enteric pathogens in children in Maputo city, Mozambique: study protocol for a cross-sectional follow-up to the Maputo Sanitation (MapSan) trial five years post-intervention

Authors

David Holcomb[1,2], Vanessa Monteiro[3], Drew Capone[1,4], Virgílio António[5], Márcia Chiluvane,[3] Victória Cumbane[3], Nália Ismael[5], Jackie Knee[6], Erin Kowalsky[1], Amanda Lai[1], Yarrow Linden[1], Elly Mataveia[3], Rassul Nalá[7], Gouthami Rao[1], Jorge Ribeiro[3], Oliver Cumming[6], Edna Viegas[3], Joe Brown[1,*]

Affiliations

1. Department of Environmental Sciences and Engineering, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA

2. Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA

3. Centro de Investigação e Treino em Saúde da Polana Caniço, Instituto Nacional de Saúde, Maputo, Mozambique

4. Department of Environmental and Occupational Health, School of Public Health, Indiana University, Bloomington, Indiana, USA

5. Division of Biotechnology and Genetics, Instituto Nacional de Saúde, Marracuene, Mozambique

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2022.08.19.22278991; this version posted August 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

6. Department of Disease Control, Faculty of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, London, UK

7. Division of Parasitology, Instituto Nacional de Saúde, Marracuene, Mozambique

Corresponding Author

* Joe Brown; 135 Dauer Drive, Chapel Hill, NC, USA, 27599-7435; T: (919) 966-7892; E: joebrown@unc.edu

medRxiv preprint doi: https://doi.org/10.1101/2022.08.19.22278991; this version posted August 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Abstract

Background

We previously assessed the effect of an onsite sanitation intervention in informal neighborhoods of urban Maputo, Mozambique on enteric pathogen detection in children after two years of follow-up (Maputo Sanitation [MapSan] trial, clinicaltrials.gov: NCT02362932). We found significant reductions in *Shigella* and *Trichuris* prevalence but only among children born after the intervention was delivered. In this study, we assess the health impacts of the sanitation intervention after five years among children born into study households post-intervention.

Methods

We are conducting a cross-sectional household study of enteric pathogen detection in child stool and the environment in the 16 MapSan study neighborhoods at compounds (household clusters sharing sanitation and outdoor living space) that received the pour-flush toilet and septic tank intervention at least five years prior or meet the original criteria for trial control sites. We are enrolling at least 400 children (ages 29 days – 60 months) in each treatment arm. Our primary outcome is the prevalence of 22 bacterial, protozoan, and soil transmitted helminth (STH) enteric pathogens in child stool using the pooled odds ratio across the outcome set to assess the intervention effect. Secondary outcomes include the individual pathogen detection prevalence and gene copy density of 27 enteric pathogens (including viruses); mean height-for-age (HAZ), weight-for-age (WAZ), and weight-for-height (WHZ) z-scores; prevalence of stunting, underweight, and wasting; and the 7-day period-prevalence of caregiver-reported diarrhea. All analyses are adjusted for pre-specified covariates and examined for effect measure modification by age. Environmental samples from study households and the public domain are assessed for

medRxiv preprint doi: https://doi.org/10.1101/2022.08.19.22278991; this version posted August 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

pathogens and fecal indicators to explore environmental exposures and monitor disease transmission.

Discussion

We are using stool-based detection of multiple enteric pathogens as an objective measure of exposure to evaluate an existing sanitation intervention. We propose a novel pooled estimate of the treatment effect across a pre-specified outcome set to summarize the overall impact of the intervention on enteric pathogen exposure as the primary trial outcome.

Trial Registration

This trial was prospectively registered on 16 March 2022 (ISRCTN86084138).

Keywords

Sanitation; urban slum; child health; enteric infection; diarrheal disease; stunting; pooled effect estimates; molecular pathogen detection; environmental microbiology; wastewater-based epidemiology

medRxiv preprint doi: https://doi.org/10.1101/2022.08.19.22278991; this version posted August 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Background

There is a high burden of childhood enteric infections associated with poor environmental conditions, with multiple enteric pathogens frequently detected in stool within the first year of life [1–3]. This is also the period with the highest incidence of diarrhea, which remains a leading cause of child mortality in low- and middle-income countries (LMIC) and has long been associated with stunted growth [4–6]. However, many diarrheal episodes are not attributable to infectious etiologies [3, 7], while the etiology of attributable diarrhea varies widely by setting and only certain diarrheal pathogens are consistently implicated in reduced linear growth [2, 8, 9]. Far more prevalent is asymptomatic enteric pathogen shedding in stool [3, 7, 10–14], which may be more strongly associated with poor growth than diarrheal illness [2, 15–17], potentially by contributing to intestinal inflammation, gut permeability, and nutrient malabsorption in a condition known as environmental enteric dysfunction (EED) [18–21]. In addition to the diverse negative impacts of stunting [22], specific adverse health outcomes associated with enteric infection and EED include delayed cognitive development [23–25] and reduced oral vaccine efficacy [26, 27].

Water, sanitation, and hygiene (WASH) interventions aim to prevent diarrhea and improve child growth by interrupting fecal-oral pathogen transmission [28]. However, recent rigorous WASH intervention trials have demonstrated inconsistent and often limited impacts on child health [29–34, 13]. Methodological limitations of such distal outcome measures notwithstanding [35, 36], these findings are consistent with the interventions insufficiently interrupting environmental transmission and exposure to enteric pathogens [37]. Combined WASH interventions with high fidelity and adherence reduced stool-based detection of *Giardia*, hookworm, and possibly

medRxiv preprint doi: https://doi.org/10.1101/2022.08.19.22278991; this version posted August 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Ascaris among 30 month-old children (but not younger children) and enteric viruses in rural Bangladesh [10, 38, 39], *Ascaris* in rural Kenya [40], and the number of co-detected parasites in rural Zimbabwe [11], but not bacterial pathogens or those most associated with stunting in any setting [2]. The sanitation-only intervention arm in rural Bangladesh reduced *Trichuris* prevalence in child stool [38] and may have reduced pathogenic *Escherichia coli* on child hands in more crowded households [41], while a sanitation intervention in rural Cambodia did not impact any pathogens measured in child stool [13]. Viewed collectively, a clear picture emerges of pervasive childhood polymicrobial exposures that were not meaningfully prevented by low-cost WASH interventions [42–44].

The preceding trials were all conducted in rural settings, but rapid urbanization has led to exceptional growth of densely populated informal settlements that lack basic services and present unique health challenges [45, 46]. We have previously investigated whether an onsite sanitation intervention delivered in low-income neighborhoods of urban Maputo, Mozambique reduced enteric pathogen prevalence in child stool [47]. Although we similarly found no evidence of an effect on combined prevalence of pre-specified enteric pathogens in our primary analyses [34], additional evidence suggests that the sanitation intervention may have reduced exposures to enteric pathogens in the environment. The intervention was delivered with high fidelity and was widely used by intervention households after two years [48]. Exploratory sub-group analyses indicated that 24 months after the intervention, *Shigella* prevalence was halved (-51%; 95% confidence interval [CI]: -15%, -72%) and *Trichuris* prevalence reduced by three-quarters (-76%; 95% CI: -40%, -90%) among children born into the study compounds after the intervention was implemented, relative to children born into the control compounds after baseline [34].

medRxiv preprint doi: https://doi.org/10.1101/2022.08.19.22278991; this version posted August 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Furthermore, overall pathogen prevalence, pathogen counts, *E. coli* gene copy density, and the individual prevalence of *Ascaris* and pathogenic *E. coli* were all significantly reduced in soil at the intervention latrine entrance [49, 50], suggesting the intervention effectively contained human excreta. While animals have been implicated as major sources of pathogen exposure in other settings [51], only companion animals were frequently present at study households and a locally validated indicator of poultry fecal contamination (the most commonly observed non-companion animal type) was rarely detected in household environments [52]. Conversely, indicators of human fecal contamination were widespread [50] and the two pathogens most impacted by the intervention—*Shigella* and *Trichuris*—are considered anthroponotic [53], suggesting that human excreta was a primary source of enteric pathogens in our dense, urban setting and that human-associated pathogens were impacted by sustained exposure to the intervention.

Enteric pathogen carriage is highly age-dependent [8, 12], but we were previously limited by the timing of our follow-up survey (24-months post-intervention) to children under two years of age when assessing impacts on those born into the intervention. Furthermore, it has been suggested that the typically short (1 – 2 year) follow-up periods often used in WASH evaluations may be inadequate for any longer-term benefits to manifest, potentially contributing to the limited observed effects [42]. Accordingly, we are conducting a cross-sectional follow-up study to better understand the long-term impacts of the sanitation intervention five years after it was implemented on child enteric pathogen exposures and shedding, diarrheal disease, and growth among children born into study households post-intervention.

medRxiv preprint doi: https://doi.org/10.1101/2022.08.19.22278991; this version posted August 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Study Objectives

Primary objective

To measure the long-term effect of a shared, onsite urban sanitation intervention on the pooled prevalence of pre-specified enteric pathogen targets detected in children's stool at least 60 months post intervention.

Hypotheses

H1. The risk of stool-based enteric pathogen detection among children 29 days – 60 months old is reduced for children born into households that previously received the sanitation intervention.

H2. Children born into households that previously received the sanitation intervention experience delayed exposure to enteric pathogens relative to comparably aged children from non-intervention households, reflected in a greater reduction in the risk of enteric pathogen detection among younger age groups and attenuated reduction in risk among older children.

Secondary objectives

1. To measure the effect of a shared, onsite urban sanitation intervention program on the individual prevalence and density of pre-specified enteric pathogen targets detected in children's stool and environmental samples at least 60 months post intervention.
2. To measure the effect of a shared, onsite urban sanitation intervention program on child growth and the prevalence of caregiver-reported diarrhea in children at least 60 months post intervention.

medRxiv preprint doi: https://doi.org/10.1101/2022.08.19.22278991; this version posted August 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Methods

Study setting

This household-based study is being conducted in 11 *bairros* (neighborhoods) in the *Nhlamankulu* District and 5 *bairros* in the *KaMaxaquene* District of Maputo city, Mozambique. Households in these densely populated, low-income neighborhoods typically cluster into self-defined compounds delineated by a wall or boundary that share outdoor living space and sanitation facilities. Study activities with participating households are conducted primarily in the compound shared outdoor space by staff from the *Centro de Investigação e Treino em Saúde da Polana Caniço* (Polana Caniço Health Research and Training Center; CISPOC) in Maputo, Mozambique, from which the study is implemented and managed.

Study design

Urban sanitation intervention

The nongovernmental organization (NGO) Water and Sanitation for the Urban Poor (WSUP) implemented a sanitation intervention in 2015-2016 at approximately 300 compounds (groups of households typically delineated by a wall or barrier, that shared sanitation and outdoor living space) in densely populated, low-income, informal urban neighborhoods of Maputo under a larger program led by the Water and Sanitation Programme (WSP) of the World Bank [47]. In these compounds, ranging between three and twenty-five households each, existing unhygienic shared latrines were replaced with pour-flush toilets and a septic tank with a soak-away pit for the liquid effluent. Two intervention designs employing the same sanitation technology were implemented, with approximately 50 communal sanitation blocks (CSBs) and 250 shared latrines

medRxiv preprint doi: https://doi.org/10.1101/2022.08.19.22278991; this version posted August 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

(SLs) constructed across 11 neighborhoods that exhibited diversity across density and other key characteristics (susceptibility to flooding, relative poverty, access to water and sanitation infrastructure) [12, 52, 54]. SLs served compounds with fewer than 21 residents and provided a single cabin with the intervention toilet, while CSBs provided an additional cabin for every 20 compound residents and other amenities including rainwater harvesting, municipal water connections and storage, and washing, bathing and laundry facilities [34, 55].

Original controlled before-and-after study

To evaluate the impact of the WSUP sanitation intervention on child health, we conducted the Maputo Sanitation (MapSan) trial, a controlled before-and-after study of enteric pathogen detection in child stool (clinicaltrials.gov: NCT02362932) [47]. WSUP selected intervention compounds using the following criteria: (1) residents shared sanitation in poor condition as determined by an engineer; (2) the compound was located in the pre-defined implementation neighborhoods; (3) there were no fewer than 12 residents; (4) residents were willing to contribute financially to construction costs; (5) sufficient space was available for construction of the new facility; (6) the compound was accessible for transportation of construction materials and tank-emptying activities; (7) the compound had access to a legal piped water supply; and (8) the groundwater level was deep enough for construction of a septic tank. Control compounds were selected according to criteria 1, 3, 4, and 7 from the 11 neighborhoods where the intervention was implemented and 5 additional neighborhoods with comparable characteristics [12].

The MapSan trial recruited an open cohort at three time-points: baseline (pre-intervention), 12 months post-intervention, and 24 months post-intervention; children were eligible to participate

medRxiv preprint doi: https://doi.org/10.1101/2022.08.19.22278991; this version posted August 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

if older than 1 month at the time of enrollment and under 48 months at baseline. We enrolled intervention and control compounds concurrently to limit any differential effects of seasonality or other secular trends on the outcomes. We found no evidence that the sanitation intervention reduced the combined prevalence of 12 bacterial and protozoan enteric pathogens (the pre-specified primary outcome), the individual prevalence of any single bacterial, protozoan, STH, or viral pathogen, or the period prevalence of caregiver-reported diarrhea 12 and 24 months post-intervention [34]. However, exploratory analyses indicated the intervention may have been protective against bacterial and soil-transmitted helminth infections among the cohort who were born into the intervention and may have reduced the spread of some pathogens into latrine entrance soils.

Long-term cross-sectional follow-up study

We are revisiting MapSan trial compounds at least five years after the sanitation intervention to conduct a cross-sectional survey of children who were born after the intervention was implemented. Due to substantial population turnover [34], both intervention and control compounds are being identified using previously collected geolocation data and their locations confirmed by study staff during site identification visits. Any compounds identified in the study neighborhoods during the course of these mapping efforts that have received the sanitation intervention (built by WSUP to the same specifications) or match the original enrollment criteria for MapSan control compounds are also invited to participate, irrespective of previous participation in the MapSan trial. We anticipate enrollment to continue for approximately one year and are again enrolling intervention and control compounds concurrently to limit any differential effects of seasonality and other secular trends across the enrollment period.

medRxiv preprint doi: https://doi.org/10.1101/2022.08.19.22278991; this version posted August 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Eligibility criteria

We attempt to enroll all eligible children in each compound that either previously participated in the MapSan trial, has previously received an intervention latrine, or meets the original enrollment criteria for MapSan control compounds [34]. Participant inclusion criteria include:

1. Child aged 29 days – 60 months old
2. Born into and residing in a compound enrolled in the five-year follow-up study; in intervention compounds, the child must have been born following the delivery of the sanitation intervention
3. Has continuously resided in the study compound for the preceding 6 months, or since birth if under 6 months of age
4. Has a parent or guardian who is able to understand and complete the written informed consent process and allow their child to participate

Children are excluded if they have any caregiver-indicated medical condition or disability that precludes participation in the study.

Participant enrollment

Enrollment is conducted by trained study staff in either Portuguese or Changana, according to the respondent's preference, with written materials provided in Portuguese. We first obtain verbal consent from the compound leader to approach households in the compound for enrollment. We then seek written, informed consent to participate from the parent or guardian of each eligible child. Participation is entirely voluntary; guardians may decline their child's

medRxiv preprint doi: https://doi.org/10.1101/2022.08.19.22278991; this version posted August 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

participation for any reason (and are under no obligation to provide a reason) and can withdraw their children at any point during the study. We began enrolling participants on 28 March 2022.

Household visits and procedures

After initially locating study compounds, the household of a participating child is visited twice, typically on two consecutive days. On the first day, trained study staff conduct written consent procedures; administer compound, household, and child questionnaires; record child anthropometry measures; collect environmental samples (in a sub-set of households); and request the child's caregiver to retain a sample of the child's stool. The household is revisited the following day to collect a stool sample from the child and complete environmental sample collection, if necessary. If the stool sample is unavailable on the second day, the study staff coordinate an additional visit child's caregiver to retrieve the stool. In the event that 7 or more days pass since the initial visit without collection of a stool sample, a qualified nurse visits the child to obtain a rectal swab, which is explained during the initial informed consent process along with all other household visit procedures.

After collection of the stool sample, deworming is offered to all household members >1 year old who have not been dewormed in the past year, with the exception of pregnant and breastfeeding women. Deworming consultation and medication provision is conducted by Ministry of Health staff following the national guidelines for deworming procedures. Deworming is offered in-kind to all household members and leverages the household interaction to provide an important public health service. Besides deworming, this study offers no direct benefit to children participating in this study. No incentives are provided to study participants, but we provide 50 meticais

medRxiv preprint doi: https://doi.org/10.1101/2022.08.19.22278991; this version posted August 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

(approximately US$1) of mobile phone credit on the caregiver's preferred network for each child to compensate for the costs incurred in communicating with the study team to arrange household visits and retrieve the child's stool. Households from which we collect food and/or large-volume water samples are reimbursed an additional 50 meticais to offset these expenses.

Environmental sample collection

We are sampling environmental compartments at a randomly selected subset of 100 intervention and 100 control compounds to represent compound- and household-level exposures [37]. At the entrance to the compound latrine we collect soil, flies, and a large volume air sample, as well as fecal sludge from the latrine or septic tank and any animal feces observed in the shared outdoor space [49, 50, 56, 57]. One household is selected at random among the households in the compound with children enrolled in the child health study, from which we collect swabs of flooring at the household entrance, flies in cooking area, prepared child's food, stored drinking water, and water from the household's primary source [52, 58, 59].

We are also collecting environmental samples from the public domain, including wastewater, open surface water, and soils in the vicinity of solid and fecal waste disposal locations, as well as wastewater effluent from hospitals treating COVID-19 patients. These samples are anonymous, are not linked to specific individuals in any way, and are used to conduct environmental surveillance of pathogens circulating in the community [56, 60].

medRxiv preprint doi: https://doi.org/10.1101/2022.08.19.22278991; this version posted August 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Study outcomes

Enteric pathogen detection in stool

Stool-based molecular detection is performed for 27 enteric pathogens commonly implicated in both symptomatic and asymptomatic childhood infections globally, including those identified at the Global Enteric Multicenter Study (GEMS) study site in Manhiça, Mozambique [12, 36, 56, 61]. Reverse-transcription quantitative polymerase chain reaction (RT-qPCR) is conducted by custom TaqMan Array Card (TAC; Thermo Fisher, Carlsbad, CA, USA) to simultaneously quantify genetic targets corresponding to 13 bacterial pathogens (*Aeromonas* spp.; *Campylobacter jejuni*/*coli*; *Clostridioides difficile*; *E. coli* O157; enteroaggregative *E. coli* (EAEC); enteropathogenic *E. coli* (EPEC); enterotoxigenic *E. coli* (ETEC); Shiga toxin-producing *E. coli* (STEC); enteroinvasive *E. coli* (EIEC)/*Shigella* spp.; *Helicobacter pylori*; *Plesiomonas shigelloides*; *Salmonella enterica*; *Vibrio cholerae*), 4 protozoan parasites (*Cryptosporidium* spp.; *Cyclospora cayetanensis*; *Entamoeba histolytica; Giardia* spp.), 5 soil transmitted helminths (*Ascaris lumbricoides*; *Ancylostoma duodenale*; *Necator americanus*; *Strongyloides stercolaris*; *Trichuris trichiura*), and 5 enteric viruses (adenovirus 40/41; astrovirus; norovirus GI/GII; rotavirus; sapovirus) [62]. We also include the respiratory virus SARS-CoV-2 on the custom TAC, which is shed in the stool of many infected individuals, enabling surveillance through fecal waste streams [63, 64].

Our primary outcome is the prevalence of a pre-specified subset of 22 enteric pathogens; as in the original MapSan study, we exclude enteric viruses and SARS-CoV-2 from the primary outcome due to the potential for direct contact transmission, which is unlikely to be impacted by the intervention [34, 47]. Because we anticipate a combined prevalence (detection of at least one

medRxiv preprint doi: https://doi.org/10.1101/2022.08.19.22278991; this version posted August 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

of the 22 primary outcome pathogens in a given stool sample) near 100%, and the intervention could plausibly increase the prevalence of some pathogens while reducing others, we do not define a composite prevalence outcome [34, 65]. Rather, we use mixed effects models to analyze all pathogens concurrently, thereby estimating the pooled effect of the intervention on enteric pathogen prevalence across the outcome set [66–68].

Secondary outcomes include the individual prevalence of all 27 pre-specified enteric pathogens, including *Shigella* spp. and *T. trichiura*, the two pathogens most impacted at 24 months among children born after the intervention was delivered [34]. A similar mixed effects model is used to analyze all pathogens concurrently to estimate separate treatment effects on the prevalence of each pathogen while accounting for correlations between outcomes, which alleviates concerns about *post-hoc* adjustments for multiple comparisons [10, 66]. Additional secondary outcomes include the continuous gene copy density in stool of the 22 primary outcome pathogens, for which a pooled treatment effect on the mean gene copy density is estimated, as well as the individual gene copy densities for all 27 enteric pathogens assessed on the TAC.

Anthropometry

Child weight and recumbent length (child age < 24 months) or standing height (24 – 60 months) are assessed according to standard World Health Organization (WHO) protocols and transformed to age-adjusted z-scores using WHO reference populations to obtain height-for-age (HAZ), weight-for-age (WAZ), and weight-for-height (WHZ) z-scores [69, 70]. Secondary outcomes include continuous HAZ, WAZ, and WHZ, as well as prevalence of binary growth outcomes stunting (HAZ < -2), underweight (WAZ < -2), and wasting (WHZ < -2) [13].

medRxiv preprint doi: https://doi.org/10.1101/2022.08.19.22278991; this version posted August 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Caregiver reported illness

Caregiver surveys are administered to ascertain child diarrheal disease, defined as the passage of three or more loose or watery stools in a 24-hour period, or any bloody stool, in the past 7 days [71]. We also assess two caregiver-reported negative control outcomes for each child: the 7-day period-prevalence of bruises, scrapes, or abrasions and the 7-day period-prevalence of toothache [72]. We do not expect the intervention to impact either child bruising or toothache prevalence, so significant differences in these outcomes by treatment arm would suggest possible bias in our caregiver-reported outcomes.

Pathogen detection in environmental matrices

Molecular detection of selected enteric and non-enteric pathogens and fecal source tracking (FST) markers is performed for environmental samples from both the private (compound & household) and public domains using a second custom TAC [49, 56, 73, 74]. A subset of the enteric pathogens assessed in stool are included on the environmental TAC (adenovirus 40/41; astrovirus; norovirus GI/GII; rotavirus; sapovirus; *Aeromonas* spp.; *C. jejuni*/*coli*; *C. difficile*; *E. coli* O157; EAEC; EPEC; ETEC; STEC; EIEC/*Shigella* spp.; *H. pylori*; *S. enterica*; *V. cholerae; Cryptosporidium* spp.; *E. histolytica; Giardia* spp.; *A. lumbricoides*; *A. duodenale*; *N. americanus*; *T. trichiura*), as well as select environmental and zoonotic pathogens (*Leptospira* spp; *Toxocara* spp.) [53, 75, 76], other human pathogens detectable in feces (SARS-CoV-2, Zika virus, HIV proviral DNA; *Plasmodium* spp.; *Mycobacterium tuberculosis*) [77–80], FST markers (human, poultry, and canine mitochondrial DNA; avian 16S rRNA) [81, 82], and general bacterial and anthropogenic pollution/antimicrobial resistance markers (bacterial 16S rRNA;

medRxiv preprint doi: https://doi.org/10.1101/2022.08.19.22278991; this version posted August 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

class 1 integron-integrase gene *intl1*) [83, 84]. We also culture fecal indicator bacteria (total coliforms and *E. coli*) using the IDEXX Colilert-18 and Quanti-Tray 2000 system [82, 85]. We conduct ongoing genomic surveillance of SARS-CoV-2 viral lineages in public domain wastewater samples using amplicon-based Illumina next generation sequencing [86].

Statistical analysis

We pre-specified a statistical analysis plan that was deposited in a permanent online repository (https://osf.io/e7pvk/) and linked with the prospective trial registration (ISRCTN86084138) prior to commencing enrollment [68].

Enteric pathogen outcomes

We will use mixed effects models with varying slopes and intercepts to estimate the weighted-average intervention effect across all pathogens included in the model, in essence treating each pathogen as a separate study of the intervention effect on enteric pathogen detection. Additional random effects will be included to account for clustering by compound and child. The prevalence odds ratio (POR) for children living in intervention compounds, relative to children in control compounds, will be estimated as the measure of effect for binary enteric pathogen detection outcomes, including the primary outcome and secondary individual pathogen detection outcomes. Prevalence differences (PD) between children in intervention and control compounds will also be estimated from the posterior predictive distribution at representative values of other model covariates [87]. Mean differences in standard deviation-scaled gene copy density will be estimated as the measure of effect for continuous enteric pathogen outcomes, with non-detects considered true zeros and censoring used to create a zero class (as in Tobit regression) [88].

medRxiv preprint doi: https://doi.org/10.1101/2022.08.19.22278991; this version posted August 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Effect estimates will be summarized using the mean to represent the expected effect size and the central 95% probability interval to capture the range of effect sizes compatible with the data (the 95% compatibility interval [CI]). Parameters with 95% CIs that exclude the null will be considered significant, although the magnitude and uncertainty of parameter estimates will also be considered holistically in evaluating the evidence for clinically meaningful effects [89].

Growth and caregiver reported outcomes

The effects of the intervention on mean HAZ, WAZ, and WHZ; the prevalence of stunting, underweight, and wasting; and the period-prevalence of caregiver-reported diarrhea and negative control outcomes (bruising, scrapes, and abrasions; toothache) will be analyzed separately as secondary outcomes using generalized estimating equations (GEE) and robust standard errors with exchangeable correlation structure and clustering by compound (the level at which the sanitation intervention was delivered) [13, 34, 90]. The estimated difference in age-adjusted z-scores by treatment assignment will be used as the measure of effect for continuous anthropometry outcomes. The prevalence ratio (PR) will be estimated by Poisson regression for binary growth status and caregiver-reported outcomes. We will not adjust for multiple comparisons [31, 91].

Covariates and effect measure modification

As a cross-sectional study of an existing intervention, all analyses will be adjusted for a set of covariates selected *a priori* as potential confounders of the sanitation-enteric pathogen carriage relationship [92]. The adjustment set will include child age and sex, caregiver's education, and household wealth index [34, 93]. Additional covariates will be considered in exploratory

medRxiv preprint doi: https://doi.org/10.1101/2022.08.19.22278991; this version posted August 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

adjusted analyses [10, 11]. The specific enteric pathogens detected are expected to be strongly related to child age [8, 34]. We will examine effect measure modification of the primary and secondary outcomes stratifying by age group (1-11 months, 12-23 months, and 24-60 months) [3, 94].

Independently upgraded controls

We anticipate some of the control compounds may have independently upgraded their sanitation facilities to conditions comparable to the intervention. Control compounds with sanitation facilities observed to possess cleanable, intact hardscape slabs; pour-flush or water-sealed toilets; a functional ventilation pipe; and a fixed superstructure with sturdy walls and a secure door that ensure privacy during use are considered to have independently upgraded to conditions comparable to the intervention [55]. Children living in control compounds with independently upgraded latrines are enrolled but will be excluded from the main analyses of the intervention effects. Two sets of subgroup analyses will instead be conducted that include all participants: one in which children in independently upgraded controls are considered as part of the control (non-intervention) arm, and again considered as part of the intervention arm. We will compare parameter estimates from the three sets of analyses (independent upgrades excluded, independent upgrades as controls, and independent upgrades as interventions) to investigate whether the sanitation improvements independently available in the study communities, which may represent more accessible options for achieving greater coverage of high-quality sanitation infrastructure, are comparable to the full sanitation intervention package assessed in the MapSan trial in terms of child health impacts.

medRxiv preprint doi: https://doi.org/10.1101/2022.08.19.22278991; this version posted August 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Eligible children residing in any compound that has received a WSUP intervention latrine will be considered part of the intervention arm in primary analyses. We are assessing the current conditions of intervention facilities but will not exclude or otherwise adjust for either upgraded or degraded sanitation facilities in intervention compounds in order to evaluate the long-term impacts of the intervention following extended use.

Minimum detectable effect size

The number of participants will be constrained by the number of compounds in the study neighborhoods that have received the sanitation intervention or meet the eligibility requirements for MapSan control compounds, most of which were previously enrolled in the MapSan trial. At the 24 month follow-up, an average of 2.5 children per compound were enrolled from 408 compounds [34]. Compound-level intra-class correlation coefficients (ICCs) were generally less than 0.1 for individual pathogens, corresponding to cluster variances of ~0.05. We calculate the minimum detectable effect size (MDES) on individual pathogen prevalence with 80% power, 5% significance level, and 0.05 compound cluster variance for a conservative scenario with 200 compounds per treatment arm and 2 children enrolled per compound (for 800 children total, 400 per arm), a moderate scenario of 220 compounds per treatment arm and 2.5 children enrolled per compound (550 children per arm, 1100 total), and a maximal scenario of 300 compounds per arm, 2.5 children per compound (750 children per arm, 1500 total) [95]. Across all scenarios, a minimum baseline (untreated) prevalence of 6-8% is required to reach 80% power for even the largest theoretical effect (nearly 100% reduction). The minimum detectable relative reduction in prevalence decreases (that is, smaller effect sizes are more readily detected) as pathogen prevalence increases towards 100% (Figure 1). The difference between the scenarios on the

medRxiv preprint doi: https://doi.org/10.1101/2022.08.19.22278991; this version posted August 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

multiplicative scale is relatively minor, with the relative reduction MDES largely driven by pathogen baseline prevalence. A pathogen with baseline prevalence below 15% must have its prevalence halved (PR < 0.5) in order to attain 80% power, while a 25% reduction is detectable with 80% power for baseline prevalence of 34-46% under the maximal and conservative scenarios, respectively. We expect the simultaneous consideration of multiple pathogens to reduce the MDES for the primary outcome pooled intervention effect by effectively increasing the sample size. Because this pooled effect is dependent on the prevalence of each pathogen considered and the correlations between them, we will conduct simulation analyses to characterize plausible MDES ranges for the pooled primary outcome treatment effect [65, 95].

medRxiv preprint doi: https://doi.org/10.1101/2022.08.19.22278991; this version posted August 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

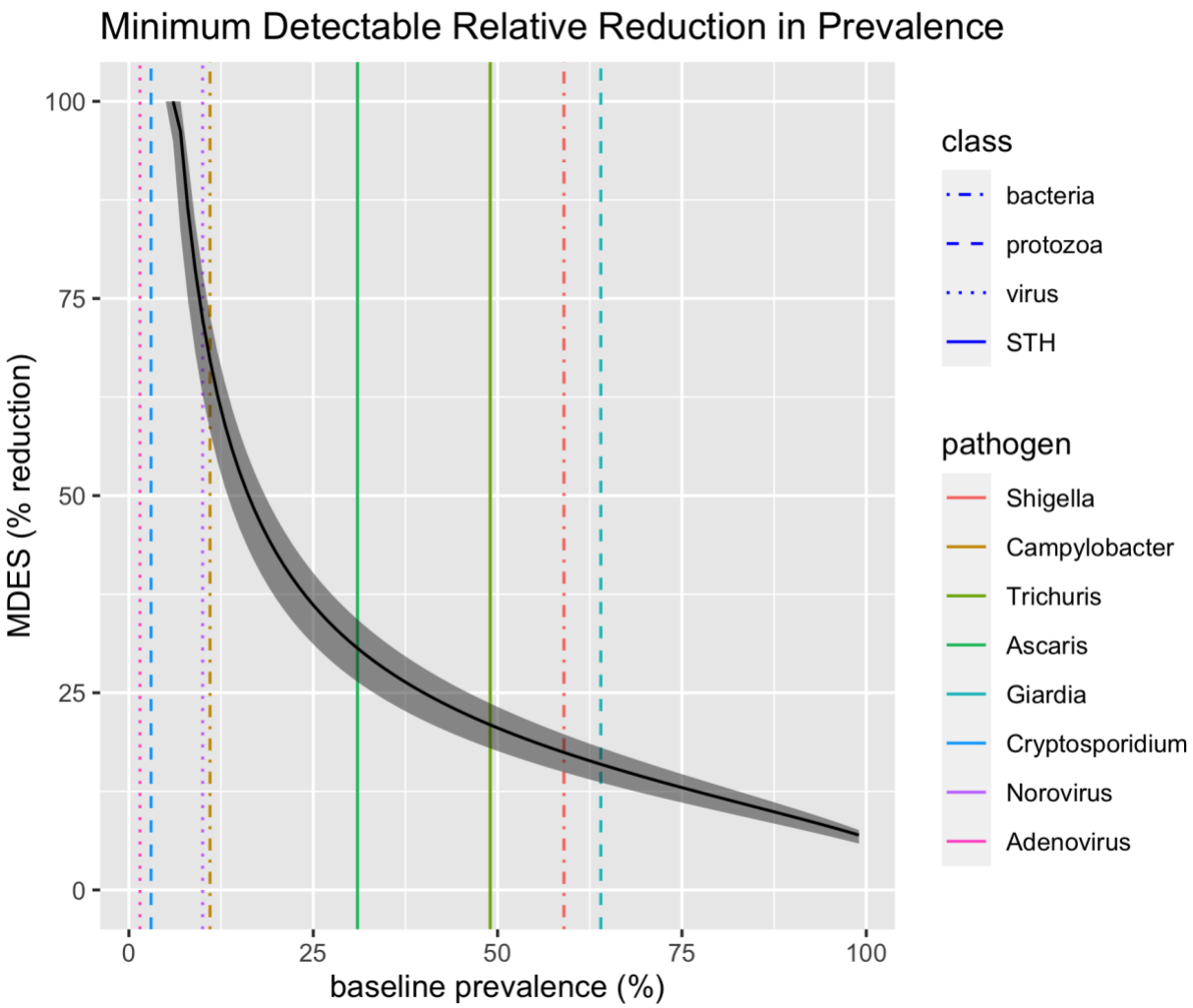

Figure 1. Range of minimum detectable effect sizes (shaded region) for the percent reduction in pathogen prevalence with 80% power, 5% significance level, and 0.05 cluster variance across three sample size scenarios. The upper edge of the shaded area represents a conservative scenario with 800 total participants (2 per compound, 200 compounds per arm) while the lower edge corresponds to a maximal scenario with 1500 total participants (2.5 per compound, 300 compounds per arm). A moderate scenario with 1100 total participants (2.5 per compound, 220 compounds per arm) is represented by the black curve within the shaded area. The vertical lines show the prevalence of a subset of pathogens assessed in control compound children during the 24-month follow-up in the original MapSan trial. Line color indicates the specific pathogen and line pattern reflects the pathogen class.

medRxiv preprint doi: https://doi.org/10.1101/2022.08.19.22278991; this version posted August 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Discussion

Defining a primary trial outcome

Caregiver-reported diarrhea has persisted as a primary outcome for health impact evaluations of WASH interventions no doubt in part because it provides single, non-invasive clinically relevant endpoint that can be relatively easily assessed by simple questionnaire. Similar to the fecal indicator paradigm for monitoring fecal pollution, the sheer number of potential pathogens that must be considered to comprehensively assess efforts to interrupt fecal-oral disease transmission demands a simple indicator of intervention impact [82]. Recent developments in molecular diagnostic platforms enabling simultaneous detection of multiple enteric pathogens in stool and environmental matrices have been rapidly adopted [10, 11, 13, 34, 36, 37, 49] and the inherent limitations of symptomatic diarrhea as a trial endpoint increasingly recognized [14, 35]. In addition to providing objective measures of prior pathogen exposure, these efforts have already offered more granular insight into the behaviors and responses of particular organisms in different settings. For example, the reduction in enteric viruses attributable to a combined WASH intervention in rural Bangladesh aligns with the reduction in diarrhea but lack of impact on linear growth observed in the parent trial [10, 31], as enteric viruses drive diarrhea in this setting but are much less associated with stunting than other pathogens that were not affected by the intervention [2].

However, such granularity complicates the interpretation of the intervention's overall effect on enteric pathogen exposure, yielding as many as several dozen separate—but not entirely independent—estimates of the intervention effect that can be highly heterogenous. Although procedures to control the false discovery rate can alleviate concerns over multiple hypothesis

medRxiv preprint doi: https://doi.org/10.1101/2022.08.19.22278991; this version posted August 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

testing [10, 11], the power to detect meaningful effects for each target depends on the background prevalence of each pathogen assessed, which can vary widely by setting and over time [3, 7, 8]. Because multiple comparison adjustments trade power to control false positives and studies are likely to be underpowered to detect individual effects on lower-prevalence pathogens, the potential for magnitude ("Type M") and sign ("Type S") errors may be uncomfortably high [96]. That is, the estimated effects on individual pathogens that are large enough to clear a statistical significance threshold are likely to be inflated in magnitude relative to the true value (Type M error) and may also appear to act in the opposite direction than is actually the case (Type S error) [97, 98]. It is quite possible that seemingly strong protective effects against individual pathogens observed in one realization of a study could dissipate or even reverse direction in a later follow-up or replication in a comparable setting and study population.

As an alternative to multiple individual effect estimates, many studies have defined composite outcomes such as the detection of one or more pathogens (possibly of a particular class) [34], the count of co-detected pathogens [11, 49], and the aggregate intensity of a select pathogens of interest [10]. These metrics are accessible and relatively easily interpreted, but also assume homogeneity of the effect across the (equally weighted) component outcomes, with implications for statistical power and sensitivity to the selection of pathogens included in the composite metric [65].

The challenge of synthesizing evidence across heterogenous effect estimates is common throughout the health, social, and environmental sciences, and is typically addressed using meta-

medRxiv preprint doi: https://doi.org/10.1101/2022.08.19.22278991; this version posted August 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

analysis, for which a robust and active methodological literature exists [99]. We propose adapting the concept of estimating a pooled effect across multiple separate effect estimates to obtain a summary of an intervention's impact that incorporates uncertainty and variability in the set of outcomes analyzed [68]. Rather than estimating a pooled effect across multiple studies of similar interventions on the same outcome, we estimate the pooled effect of a single intervention across a set of related outcomes, which may be thought of as the treatment effect on a "generic" enteric pathogen. Just as studies obtaining more precise effect estimates receive greater weight in meta-analyses, the pathogens for which individual effects can be more precisely estimated (such as those with higher background prevalence) contribute more to the pooled estimate. The pooled effect is recovered alongside the individual effects for each pathogen and importantly shares the same scale, allowing this summary metric to be directly interpreted in the context of its components, augmenting rather than replacing the individual estimates. This stands in contrast to composite outcomes, which construct new metrics that are related to, but fundamentally differ from, their component outcomes—the prevalence of *any* pathogen being conceptually distinct from the prevalence of a *particular* pathogen, for instance.

Although pooled effect estimates offer several advantages over composite outcomes as a summary of multiple related outcomes, both approaches are sensitive to the selection of the set of outcomes to be analyzed. Including many pathogens with little chance of being affected by the intervention in our analysis will bias the summary effect towards the null, in an inversion of the familiar meta-analysis challenge of publication bias against reporting null results leading to inflated meta-analytic effect estimates [100]. Conversely, we wish to avoid undue influence—and avoid limiting opportunities for discovery and learning—from too narrowly defining the

medRxiv preprint doi: https://doi.org/10.1101/2022.08.19.22278991; this version posted August 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

outcome set to include only those targets expected to provide a desired result [101]. Striking the appropriate balance between these considerations will likely remain an exercise in scientific judgement, predicated on the careful consideration of subject matter knowledge and validated to the extent possible through sensitivity and bias analyses [102–104].

For our present study, we choose to exclude enteric viruses from the primary outcome set because we believe the possibility of direct contact transmission limits the potential for the sanitation intervention to prevent exposure [105]. A combined WASH intervention reduced several enteric viruses (and no other pathogens) in younger children in rural Bangladesh [10]; because that intervention included hand hygiene and other components that could feasibly prevent direct transmission in addition to sanitation infrastructure, we argue it remains defensible to exclude enteric viruses from the primary analysis of our onsite sanitation intervention. However, we will assess the individual and pooled effects on enteric viruses and all measured enteric pathogens collectively in secondary analyses. Otherwise, we have chosen to be relatively expansive with the outcome set for the primary pooled intervention effect analysis, including all bacteria, protozoa, and STHs that are being measured in child stool. Nevertheless, we have reservations about including a particular parasite, *Giardia* spp., in the outcome set, as discussed below.

Should *Giardia* have been excluded as a trial outcome?

Giardia is one of the most commonly detected pathogens in child stool in low- and middle-income countries, including in the original MapSan study cohort and elsewhere in southern Mozambique and it was included as one of a panel of enteric pathogens in the original MapSan

medRxiv preprint doi: https://doi.org/10.1101/2022.08.19.22278991; this version posted August 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

trial [12, 34, 61, 106]. The high prevalence of *Giardia*, which has been found to increase rapidly with age, demonstrates a failure to prevent exposure [107]. However, persistent shedding of *Giardia* has also been observed in endemic areas, to an extent that it has been suggested to function as something of a gut commensal in such settings, even potentially protecting against diarrheal illness [108]. Whether arising from persistent infection or rapid re-infection [106, 109], this extended shedding suggests that detection of *Giardia* may not serve as a meaningful indicator of recent exposure in the context of household sanitation infrastructure. Although household finished flooring was associated with reduced *Giardia* prevalence in both Bangladesh and Kenya and a combined WASH intervention also reduced *Giardia* prevalence among older children in Bangladesh, onsite sanitation interventions were not associated with *Giardia* in rural Bangladesh or Zimbabwe, nor in urban Mozambique [10, 11, 34, 110]. Recognizing that its unique and insufficiently understood epidemiology may limit interpretation, we will repeat the primary outcome analysis with *Giardia* excluded from the outcome set, using the remaining 21 non-viral pathogens to estimate the pooled intervention effect on pathogen prevalence.

Limitations

As an observational, cross-sectional evaluation of an existing intervention, this study faces a number of limitations that may impact the validity and generalizability of our findings. By design, all of the study participants will have been born after the intervention was implemented and it will not be possible to account for pre-intervention characteristics. However, the opportunity to evaluate intervention impacts among individuals across a range of ages who have been exposed to intervention for their entire lives (and potentially also *in utero* for many, although not an eligibility criterion nor assessed) is sufficiently valuable to warrant the approach.

medRxiv preprint doi: https://doi.org/10.1101/2022.08.19.22278991; this version posted August 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Relatedly, there is potential for confounding bias in our estimates of the intervention impacts on child health, particularly concerning socioeconomic factors that may be associated with increased pathogen exposures. While non-randomized, the criteria for intervention and control compounds differed only by engineering considerations that we believe to be independent of the outcome. However, in the years since the intervention was implemented, the presence of the intervention itself may have influenced the desirability of the intervention sites and thus the socioeconomic status (SES) of their residents, particularly in light of the high population turnover previously observed after only 1 – 2 years. All analyses will be adjusted for a location-specific wealth index to account for potential differential SES between treatment arms [93], in addition to other pre-specified covariates associated with enteric pathogen exposure that may plausibly be related to treatment status, but the possibility of unmeasured and insufficiently controlled confounding remains.

Because childhood diarrhea is a leading cause of child mortality, selection bias arising from survivor effects may be present in our sample, particularly among older age groups [111]. However, loss to follow-up due to mortality was exceptionally rare in the previous assessment—far less common than emigration, which was not previously differential by treatment arm [34]. Our use of stool-based enteric pathogen detection as an objective primary outcome mitigates the potential for measurement bias [36] and we are assessing multiple negative control outcomes to account for potential response bias in our secondary caregiver-reported outcomes [72]. As a long-term evaluation of an existing intervention, the potential for exposure measurement error is low—interventions that have degraded substantially may shift results towards the null, but in so doing would represent a valid assessment of intervention sustainability [112]. The potential for

medRxiv preprint doi: https://doi.org/10.1101/2022.08.19.22278991; this version posted August 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

exposure misclassification among controls is higher, in that they may have independently upgraded to sanitation infrastructure comparable to the intervention. We are actively monitoring this possibility at all study sites, have pre-specified criteria for identifying sites achieving such conditions, and will conduct sensitivity analyses with these sites excluded, considered as controls, and considered as equivalent to the intervention to characterize the impact of any such potential exposure misclassification among controls.

As discussed previously, a key challenge is identifying an appropriate set of outcomes against which to evaluate the intervention; in this regard, we preferred inclusivity given the pathogens likely to be observed in our study setting, enteric viruses notwithstanding, at the risk of unduly shifting our eventual results towards the null. This challenge persists, regardless of study design, so long as the study aims include assessing the effect of some condition or intervention on exposure to, or infection by, multiple pathogens that share a common route of transmission [103].

Contribution

Although observational, this study is unique in evaluating sanitation intervention effects at least five years after the intervention was implemented and up to five years after participants were borne into the intervention conditions. Previous studies have focused on WASH intervention impacts up to 2 – 3 years after delivery [13, 31–34], which may be insufficient time to realize impacts [42]. Furthermore, most previous studies have been conducted in rural settings, while we are investigating the long-term effects of an urban sanitation intervention that is broadly representative of the types of infrastructure improvements likely to be available in the rapidly

medRxiv preprint doi: https://doi.org/10.1101/2022.08.19.22278991; this version posted August 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

growing urban informal settlements in coming years. Finally, we use stool-based detection of multiple enteric pathogens as an objective outcome and propose a novel pooled estimate of the treatment effect across a pre-specified outcome set to summarize the overall impact of the intervention on enteric pathogen exposure as the primary trial outcome.

medRxiv preprint doi: https://doi.org/10.1101/2022.08.19.22278991; this version posted August 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

List of Abbreviations

Abbreviation	Description
CISPOC	*Centro de Investigação e Treino em Saúde da Polana Caniço* (Polana Caniço Health Research and Training Center)
CSB	communal sanitation block
EAEC	enteroaggregative *E. coli*
EED	environmental enteric dysfunction
EIEC	enteroinvasive *E. coli*
EMM	effect measure modification
EPEC	enteropathogenic *E. coli*
ETEC	enterotoxigenic *E. coli*
FST	fecal source tracking
GEE	generalized estimating equations
GEMS	Global Enteric Multicenter Study
HAZ	height-for-age z-score
ICC	intra-class correlation coefficient
INS	*Instituto Nacional de Saúde* (National Institute of Health, Republic of Mozambique)
LMIC	low- and middle-income countries
LSHTM	London School of Hygiene and Tropical Medicine
MapSan	Maputo Sanitation trial
MDES	minimum detectable effect size
NGO	nongovernmental organization
PD	prevalence difference
POR	prevalence odds ratio
PR	prevalence ratio
RT-qPCR	Reverse-transcription quantitative [real-time] polymerase chain reaction
SES	socioeconomic status
SL	shared latrine
STEC	Shiga toxin-producing *E. coli*
STH	soil transmitted helminth
TAC	TaqMan Array Card
UNC	University of North Carolina at Chapel Hill
WASH	water, sanitation, and hygiene
WAZ	weight-for-age z-score
WHO	World Health Organization
WHZ	weight-for-height z-score
WSP	Water and Sanitation Programme

medRxiv preprint doi: https://doi.org/10.1101/2022.08.19.22278991; this version posted August 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

WSUP	Water and Sanitation for the Urban Poor

medRxiv preprint doi: https://doi.org/10.1101/2022.08.19.22278991; this version posted August 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Declarations

Ethics approval and consent to participate

This study was approved by Comité Nacional de Bioética para a Saúde, Ministério da Saúde de Moçambique (FWA#: 00003139, IRB00002657, 326/CNBS/21; approved: 15 June 2021) and University of North Carolina at Chapel Hill Ethics Committee (IRB#: 21-1119; approved 19 August 2021).

Consent for publication

Not applicable.

Availability of data and materials

Upon publication of study results, the underlying individual participant data will be fully de-identified according to the Safe Harbor method and made freely available in a permanent online repository (https://osf.io/e7pvk/) in accordance with the funder's open data policies.

Competing interests

The authors declare that they have no competing interests.

Funding

This study is funded by the Bill & Melinda Gates Foundation (OPP1137224) with additional support from a National Institute of Environmental Health Sciences training grant

medRxiv preprint doi: https://doi.org/10.1101/2022.08.19.22278991; this version posted August 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

(T32ES007018). The funders had no role in the study design; data collection, analysis, and interpretation; or decision to publish.

Authors' contributions

JB, EV, OC, RN, and JK designed the study, secured funding, and provided ongoing supervision. VM and DC drafted the initial study protocol and obtained ethical approvals. DH developed the statistical analysis plan and prepared the prospective trial registration and protocol manuscript. EK, GR, EM, JR, VA, AL, and YL developed and implemented the study procedures, which were overseen by VC, VM, MC, and NI. All authors contributed to and approved the final manuscript.

Acknowledgements

We thank all the participants, their families, and neighbors for graciously welcoming us into their communities, and our implementing partner, Water and Sanitation for the Urban Poor, for their continued support. We also gratefully acknowledge the hard work of the CISPOC survey team, including Jorge Binguane, Noémia Come, Anelsa Dunhe, Alice Fumo, Antônio Johane, Evelin Matos, Eloisa Mula, Mariza Rachid, and Líria Sambo; and of Filipe Fazenda, Claúdia Machume, and Alfredo Muchanga in the CISPOC laboratory.

medRxiv preprint doi: https://doi.org/10.1101/2022.08.19.22278991; this version posted August 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

References

1. Taniuchi M, Sobuz SU, Begum S, Platts-Mills J a., Liu J, Yang Z, et al. Etiology of Diarrhea in Bangladeshi Infants in the First Year of Life Analyzed Using Molecular Methods. J Infect Dis. 2013;208:1794–802. https://doi.org/10.1093/infdis/jit507.

2. Rogawski ET, Liu J, Platts-Mills JA, Kabir F, Lertsethtakarn P, Siguas M, et al. Use of quantitative molecular diagnostic methods to investigate the effect of enteropathogen infections on linear growth in children in low-resource settings: longitudinal analysis of results from the MAL-ED cohort study. Lancet Glob Health. 2018;6:e1319–28. https://doi.org/10.1016/S2214-109X(18)30351-6.

3. Liu J, Platts-Mills JA, Juma J, Kabir F, Nkeze J, Okoi C, et al. Use of quantitative molecular diagnostic methods to identify causes of diarrhoea in children: a reanalysis of the GEMS case-control study. The Lancet. 2016;388:1291–301. https://doi.org/10.1016/S0140-6736(16)31529-X.

4. Troeger CE, Khalil IA, Blacker BF, Biehl MH, Albertson SB, Zimsen SRM, et al. Quantifying risks and interventions that have affected the burden of diarrhoea among children younger than 5 years: an analysis of the Global Burden of Disease Study 2017. Lancet Infect Dis. 2020;20:37–59. https://doi.org/10.1016/s1473-3099(19)30401-3.

5. Black RE, Brown KH, Becker S. Effects of Diarrhea Associated with Specific Enteropathogens on the Growth of Children in Rural Bangladesh. Pediatrics. 1984;73:799–805. https://doi.org/10.1542/peds.73.6.799.

6. Benjamin-Chung J, Mertens A, Colford JM, Hubbard AE, van der Laan MJ, Coyle J, et al. Early childhood linear growth faltering in low- and middle-income countries. medRxiv. 2021. https://doi.org/10.1101/2020.06.09.20127001.

7. Platts-Mills JA, Liu J, Rogawski ET, Kabir F, Lertsethtakarn P, Siguas M, et al. Use of quantitative molecular diagnostic methods to assess the aetiology, burden, and clinical characteristics of diarrhoea in children in low-resource settings: a reanalysis of the MAL-ED cohort study. Lancet Glob Health. 2018;6:e1309–18. https://doi.org/10.1016/S2214-109X(18)30349-8.

8. Baker JM, Hasso-Agopsowicz M, Pitzer VE, Platts-Mills JA, Peralta-Santos A, Troja C, et al. Association of enteropathogen detection with diarrhoea by age and high versus low child mortality settings: a systematic review and meta-analysis. Lancet Glob Health. 2021;9:e1402–10. https://doi.org/10.1016/S2214-109X(21)00316-8.

9. Schnee AE, Haque R, Taniuchi M, Uddin MJ, Alam MM, Liu J, et al. Identification of Etiology-Specific Diarrhea Associated With Linear Growth Faltering in Bangladeshi Infants. Am J Epidemiol. 2018;187:2210–8. https://doi.org/10.1093/aje/kwy106.

10. Grembi JA, Lin A, Karim MA, Islam MO, Miah R, Arnold BF, et al. Effect of Water, Sanitation, Handwashing, and Nutrition Interventions on Enteropathogens in Children 14 Months

medRxiv preprint doi: https://doi.org/10.1101/2022.08.19.22278991; this version posted August 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Old: A Cluster-Randomized Controlled Trial in Rural Bangladesh. J Infect Dis. 2020. https://doi.org/10.1093/infdis/jiaa549.

11. Rogawski McQuade ET, Platts-Mills JA, Gratz J, Zhang J, Moulton LH, Mutasa K, et al. Impact of Water Quality, Sanitation, Handwashing, and Nutritional Interventions on Enteric Infections in Rural Zimbabwe: The Sanitation Hygiene Infant Nutrition Efficacy (SHINE) Trial. J Infect Dis. 2019;221:1379–86. https://doi.org/10.1093/infdis/jiz179.

12. Knee J, Sumner T, Adriano Z, Berendes D, de Bruijn E, Schmidt W, et al. Risk factors for childhood enteric infection in urban Maputo, Mozambique: A cross-sectional study. PLoS Negl Trop Dis. 2018;12:e0006956. https://doi.org/10.1371/journal.pntd.0006956.

13. Lai A, Velez I, Ambikapathi R, Seng K, Levy K, Kowalsky E, et al. Independent and combined effects of nutrition and sanitation interventions on child growth and enteric pathogen infection in rural Cambodia: a factorial cluster-randomised controlled trial. Prep. 2022.

14. Rego R, Watson S, Alam MAU, Abdullah SA, Yunus M, Alam IT, et al. A comparison of traditional diarrhoea measurement methods with microbiological and biochemical indicators: A cross-sectional observational study in the Cox's Bazar displaced persons camp. eClinicalMedicine. 2021;42:101205. https://doi.org/10.1016/j.eclinm.2021.101205.

15. Berendes D, Capone D, Knee J, Holcomb D, Sultana S, Pickering AJ, et al. Associations between enteric pathogen carriage and height-for-age, weight-for-age and weight-for-height in children under 5 years old in urban Dhaka, Bangladesh. Epidemiol Infect. 2020;148:e39. https://doi.org/10.1017/s0950268820000369.

16. Collard J-M, Andrianonimiadana L, Habib A, Rakotondrainipiana M, Andriantsalama P, Randriamparany R, et al. High prevalence of small intestine bacteria overgrowth and asymptomatic carriage of enteric pathogens in stunted children in Antananarivo, Madagascar. PLoS Negl Trop Dis. 2022;16:1–22. https://doi.org/10.1371/journal.pntd.0009849.

17. Garzón M, Pereira-da-Silva L, Seixas J, Papoila AL, Alves M. Subclinical Enteric Parasitic Infections and Growth Faltering in Infants in São Tomé, Africa: A Birth Cohort Study. Int J Environ Res Public Health. 2018;15. https://doi.org/10.3390/ijerph15040688.

18. Prendergast AJ, Kelly P. Interactions between intestinal pathogens, enteropathy and malnutrition in developing countries. Curr Opin Infect Dis. 2016;:1. https://doi.org/10.1097/QCO.0000000000000261.

19. Prendergast AJ, Rukobo S, Chasekwa B, Mutasa K, Ntozini R, Mbuya MNN, et al. Stunting Is Characterized by Chronic Inflammation in Zimbabwean Infants. Plos One. 2014;9:11. https://doi.org/10.1371/journal.pone.0086928.

20. George CM, Burrowes V, Perin J, Oldja L, Biswas S, Sack D, et al. Enteric Infections in Young Children are Associated with Environmental Enteropathy and Impaired Growth. Trop Med Int Health. 2017;12:3218–21. https://doi.org/10.1111/tmi.13002.

medRxiv preprint doi: https://doi.org/10.1101/2022.08.19.22278991; this version posted August 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

21. Kosek MN, Ahmed T, Bhutta Z, Caulfield L, Guerrant R, Houpt E, et al. Causal Pathways from Enteropathogens to Environmental Enteropathy: Findings from the MAL-ED Birth Cohort Study. EBioMedicine. 2017;18:109–17. https://doi.org/10.1016/j.ebiom.2017.02.024.

22. Hoddinott J, Behrman JR, Maluccio JA, Melgar P, Quisumbing AR, Ramirez-Zea M, et al. Adult consequences of growth failure in early childhood. Am J Clin Nutr. 2013;98:1170–8. https://doi.org/10.3945/ajcn.113.064584.

23. MAL-ED Network Investigators. Early childhood cognitive development is affected by interactions among illness, diet, enteropathogens and the home environment: findings from the MAL-ED birth cohort study. BMJ Glob Health. 2018;3:e000752. https://doi.org/10.1136/bmjgh-2018-000752.

24. Oriá RB, Murray-Kolb LE, Scharf RJ, Pendergast LL, Lang DR, Kolling GL, et al. Early-life enteric infections: relation between chronic systemic inflammation and poor cognition in children. Nutr Rev. 2016;74:374–86. https://doi.org/10.1093/nutrit/nuw008.

25. Donowitz JR, Drew J, Taniuchi M, Platts-Mills JA, Alam M, Ferdous T, et al. Diarrheal Pathogens Associated With Growth and Neurodevelopment. Clin Infect Dis. 2021;73:e683–91. https://doi.org/10.1093/cid/ciaa1938.

26. Parker EP, Ramani S, Lopman BA, Church JA, Iturriza-Gómara M, Prendergast AJ, et al. Causes of impaired oral vaccine efficacy in developing countries. Future Microbiol. 2018;13:97–118. https://doi.org/10.2217/fmb-2017-0128.

27. Taniuchi M, Platts-Mills JA, Begum S, Uddin MJ, Sobuz SU, Liu J, et al. Impact of enterovirus and other enteric pathogens on oral polio and rotavirus vaccine performance in Bangladeshi infants. Vaccine. 2016;34:3068–75. https://doi.org/10.1016/j.vaccine.2016.04.080.

28. Cumming O, Cairncross S. Can water, sanitation and hygiene help eliminate stunting? Current evidence and policy implications. Matern Child Nutr. 2016;12:91–105. https://doi.org/10.1111/mcn.12258.

29. Wolf J, Hubbard S, Brauer M, Ambelu A, Arnold BF, Bain R, et al. Effectiveness of interventions to improve drinking water, sanitation, and handwashing with soap on risk of diarrhoeal disease in children in low-income and middle-income settings: a systematic review and meta-analysis. The Lancet. 2022;400:48–59. https://doi.org/10.1016/s0140-6736(22)00937-0.

30. Contreras JD, Eisenberg JNS. Does Basic Sanitation Prevent Diarrhea? Contextualizing Recent Intervention Trials through a Historical Lens. Int J Environ Res Public Health. 2020;17. https://doi.org/10.3390/ijerph17010230.

31. Luby SP, Rahman M, Arnold BF, Unicomb L, Ashraf S, Winch PJ, et al. Effects of water quality, sanitation, handwashing, and nutritional interventions on diarrhoea and child growth in rural Bangladesh: a cluster randomised controlled trial. Lancet Glob Health. 2018;6:e302–15. https://doi.org/10.1016/s2214-109x(17)30490-4.

medRxiv preprint doi: https://doi.org/10.1101/2022.08.19.22278991; this version posted August 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

32. Null C, Stewart CP, Pickering AJ, Dentz HN, Arnold BF, Arnold CD, et al. Effects of water quality, sanitation, handwashing, and nutritional interventions on diarrhoea and child growth in rural Kenya: a cluster-randomised controlled trial. Lancet Glob Health. 2018;6:e316–29. https://doi.org/10.1016/s2214-109x(18)30005-6.

33. Humphrey JH, Mbuya MNN, Ntozini R, Moulton LH, Stoltzfus RJ, Tavengwa NV, et al. Independent and combined effects of improved water, sanitation, and hygiene, and improved complementary feeding, on child stunting and anaemia in rural Zimbabwe: a cluster-randomised trial. Lancet Glob Health. 2019;7:e132–47. https://doi.org/10.1016/s2214-109x(18)30374-7.

34. Knee J, Sumner T, Adriano Z, Anderson C, Bush F, Capone D, et al. Effects of an urban sanitation intervention on childhood enteric infection and diarrhea in Maputo, Mozambique: a controlled before-and-after trial. eLife. 2021;10:e62278. https://doi.org/10.7554/elife.62278.

35. Watson SI, Rego RTT, Hofer T, Lilford RJ. Evaluations of water, sanitation and hygiene interventions should not use diarrhoea as (primary) outcome. BMJ Glob Health. 2022;7. https://doi.org/10.1136/bmjgh-2022-008521.

36. Brown J, Cumming O. Stool-Based Pathogen Detection Offers Advantages as an Outcome Measure for Water, Sanitation, and Hygiene Trials. Am J Trop Med Hyg. 2020;102:260–1. https://doi.org/10.4269/ajtmh.19-0639.

37. Goddard F, Ban R, Barr DB, Brown J, Cannon J, Colford JM, et al. Measuring environmental exposure to enteric pathogens in low-income settings: review and recommendations of an interdisciplinary working group. Environ Sci Technol. 2020;:acs.est.0c02421. https://doi.org/10.1021/acs.est.0c02421.

38. Ercumen A, Benjamin-Chung J, Arnold BF, Lin A, Hubbard AE, Stewart C, et al. Effects of water, sanitation, handwashing and nutritional interventions on soil-transmitted helminth infections in young children: A cluster-randomized controlled trial in rural Bangladesh. PLoS Negl Trop Dis. 2019;13:1–24. https://doi.org/10.1371/journal.pntd.0007323.

39. Lin A, Ercumen A, Benjamin-Chung J, Arnold BF, Das S, Haque R, et al. Effects of water, sanitation, handwashing, and nutritional interventions on child enteric protozoan infections in rural Bangladesh: A cluster-randomized controlled trial. Clin Infect Dis. 2018;67:1515–22. https://doi.org/10.1093/cid/ciy320.

40. Pickering AJ, Njenga SM, Steinbaum L, Swarthout J, Lin A, Arnold BF, et al. Effects of single and integrated water, sanitation, handwashing, and nutrition interventions on child soil-transmitted helminth and Giardia infections: A cluster-randomized controlled trial in rural Kenya. PLOS Med. 2019;16:e1002841. https://doi.org/10.1371/journal.pmed.1002841.

41. Fuhrmeister ER, Ercumen A, Pickering AJ, Jeanis KM, Crider Y, Ahmed M, et al. Effect of sanitation improvements on pathogens and microbial source tracking markers in the rural Bangladeshi household environment. Environ Sci Technol. 2020;:acs.est.9b04835. https://doi.org/10.1021/acs.est.9b04835.

medRxiv preprint doi: https://doi.org/10.1101/2022.08.19.22278991; this version posted August 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

42. Pickering AJ, Null C, Winch PJ, Mangwadu G, Arnold BF, Prendergast AJ, et al. The WASH Benefits and SHINE trials: interpretation of WASH intervention effects on linear growth and diarrhoea. Lancet Glob Health. 2019;7:e1139–46. https://doi.org/10.1016/s2214-109x(19)30268-2.

43. Amebelu A, Ban R, Bhagwan J, Brown J, Chilengi R, Chandler C, et al. The Lancet Commission on water, sanitation and hygiene, and health. The Lancet. 2021;398:1469–70. https://doi.org/10.1016/S0140-6736(21)02005-5.

44. Budge S, Ambelu A, Bartram J, Brown J, Hutchings P. Environmental sanitation and the evolution of water, sanitation and hygiene. Bull World Health Organ. 2022;100:286–8. https://doi.org/10.2471/BLT.21.287137.

45. Ezeh A, Oyebode O, Satterthwaite D, Chen Y-FF, Ndugwa R, Sartori J, et al. The history, geography, and sociology of slums and the health problems of people who live in slums. The Lancet. 2017;389:547–58. https://doi.org/10.1016/S0140-6736(16)31650-6.

46. Sinharoy SS, Pittluck R, Clasen T. Review of drivers and barriers of water and sanitation policies for urban informal settlements in low-income and middle-income countries. Util Policy. 2019;60 July:100957. https://doi.org/10.1016/j.jup.2019.100957.

47. Brown J, Cumming O, Bartram J, Cairncross S, Ensink J, Holcomb D, et al. A controlled, before-and-after trial of an urban sanitation intervention to reduce enteric infections in children: research protocol for the Maputo Sanitation (MapSan) study, Mozambique. BMJ Open. 2015;5:e008215–e008215. https://doi.org/10.1136/bmjopen-2015-008215.

48. Bick S, Buxton H, Chase RP, Ross I, Adriano Z, Capone D, et al. Using path analysis to test theory of change: a quantitative process evaluation of the MapSan trial. BMC Public Health. 2021;21:1411. https://doi.org/10.1186/s12889-021-11364-w.

49. Capone D, Berendes D, Cumming O, Holcomb D, Knee J, Konstantinidis KT, et al. Impact of an Urban Sanitation Intervention on Enteric Pathogen Detection in Soils. Environ Sci Technol. 2021;55:9989–10000. https://doi.org/10.1021/acs.est.1c02168.

50. Holcomb DA, Knee J, Capone D, Sumner T, Adriano Z, Nalá R, et al. Impacts of an Urban Sanitation Intervention on Fecal Indicators and the Prevalence of Human Fecal Contamination in Mozambique. Environ Sci Technol. 2021;55:11667–79. https://doi.org/10.1021/acs.est.1c01538.

51. Fuhrmeister ER, Ercumen A, Pickering AJ, Jeanis KM, Ahmed M, Brown S, et al. Predictors of Enteric Pathogens in the Domestic Environment from Human and Animal Sources in Rural Bangladesh. Environ Sci Technol. 2019;53:10023–33. https://doi.org/10.1021/acs.est.8b07192.

52. Holcomb DA, Knee J, Sumner T, Adriano Z, de Bruijn E, Nalá R, et al. Human fecal contamination of water, soil, and surfaces in households sharing poor-quality sanitation facilities in Maputo, Mozambique. Int J Hyg Environ Health. 2020;226:113496. https://doi.org/10.1016/j.ijheh.2020.113496.

medRxiv preprint doi: https://doi.org/10.1101/2022.08.19.22278991; this version posted August 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

53. Delahoy MJ, Wodnik B, McAliley L, Penakalapati G, Swarthout J, Freeman MC, et al. Pathogens Transmitted in Animal Feces in Low- and Middle-Income Countries. Int J Hyg Environ Health. 2018. https://doi.org/10.1016/j.ijheh.2018.03.005.

54. Capone D, Adriano Z, Berendes D, Cumming O, Dreibelbis R, Holcomb DA, et al. A localized sanitation status index as a proxy for fecal contamination in urban Maputo, Mozambique. PLOS ONE. 2019;14:e0224333. https://doi.org/10.1371/journal.pone.0224333.

55. Shiras T, Cumming O, Brown J, Muneme B, Nala R, Dreibelbis R. Shared latrines in Maputo, Mozambique: exploring emotional well-being and psychosocial stress. BMC Int Health Hum Rights. 2018;18:30. https://doi.org/10.1186/s12914-018-0169-z.

56. Capone D, Berendes D, Cumming O, Knee J, Nalá R, Risk BB, et al. Analysis of Fecal Sludges Reveals Common Enteric Pathogens in Urban Maputo, Mozambique. Environ Sci Technol Lett. 2020;:acs.estlett.0c00610. https://doi.org/10.1021/acs.estlett.0c00610.

57. Ginn O, Rocha-Melogno L, Bivins A, Lowry S, Cardelino M, Nichols D, et al. Detection and Quantification of Enteric Pathogens in Aerosols Near Open Wastewater Canals in Cities with Poor Sanitation. Environ Sci Technol. 2021;55:14758–71. https://doi.org/10.1021/acs.est.1c05060.

58. Bick S, Perieres L, D'Mello-Guyett L, Baker KK, Brown J, Muneme B, et al. Risk factors for child food contamination in low-income neighbourhoods of Maputo, Mozambique: An exploratory, cross-sectional study. Matern Child Nutr. 2020; February:e12991. https://doi.org/10.1111/mcn.12991.

59. Exum NG, Olórtegui MP, Yori PP, Davis MF, Heaney CD, Kosek M, et al. Floors and Toilets: Association of Floors and Sanitation Practices with Fecal Contamination in Peruvian Amazon Peri-Urban Households. Environ Sci Technol. 2016;50:7373–81. https://doi.org/10.1021/acs.est.6b01283.

60. Schang C, Crosbie ND, Nolan M, Poon R, Wang M, Jex A, et al. Passive Sampling of SARS-CoV-2 for Wastewater Surveillance. Environ Sci Technol. 2021;55:10432–41. https://doi.org/10.1021/acs.est.1c01530.

61. Kotloff KL, Nataro JP, Blackwelder WC, Nasrin D, Farag TH, Panchalingam S, et al. Burden and aetiology of diarrhoeal disease in infants and young children in developing countries (the Global Enteric Multicenter Study, GEMS): A prospective, case-control study. The Lancet. 2013;382:209–22. https://doi.org/10.1016/s0140-6736(13)60844-2.

62. Liu J, Gratz J, Amour C, Nshama R, Walongo T, Maro A, et al. Optimization of Quantitative PCR Methods for Enteropathogen Detection. PLOS ONE. 2016;11:e0158199. https://doi.org/10.1371/journal.pone.0158199.

63. Natarajan A, Zlitni S, Brooks EF, Vance SE, Dahlen A, Hedlin H, et al. Gastrointestinal symptoms and fecal shedding of SARS-CoV-2 RNA suggest prolonged gastrointestinal infection. Med. 2022;3:371-387.e9. https://doi.org/10.1016/j.medj.2022.04.001.

medRxiv preprint doi: https://doi.org/10.1101/2022.08.19.22278991; this version posted August 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

64. Bivins A, North D, Ahmad A, Ahmed W, Alm E, Been F, et al. Wastewater-Based Epidemiology: Global Collaborative to Maximize Contributions in the Fight Against COVID-19. Environ Sci Technol. 2020;:acs.est.0c02388. https://doi.org/10.1021/acs.est.0c02388.

65. Vickerstaff V, Ambler G, Omar RZ. A comparison of methods for analysing multiple outcome measures in randomised controlled trials using a simulation study. Biom J. 2021;63:599–615. https://doi.org/10.1002/bimj.201900040.

66. Gelman A, Hill J, Yajima M. Why We (Usually) Don't Have to Worry About Multiple Comparisons. J Res Educ Eff. 2012;5:189–211. https://doi.org/10.1080/19345747.2011.618213.

67. Mitchell S, Gelman A, Ross R, Chen J, Bari S, Huynh UK, et al. The Millennium Villages Project: a retrospective, observational, endline evaluation. Lancet Glob Health. 2018;6:e500–13. https://doi.org/10.1016/s2214-109x(18)30065-2.

68. Holcomb D, Capone C, Cumbane V, Knee J, Kowaslky E, Lai A, et al. Data Analysis Plan: Long-term impact of an urban sanitation intervention on child health in low-income neighborhoods of Maputo city, Mozambique: a cross-sectional follow-up five years post-intervention in the Maputo Sanitation (MapSan) trial. Version 3.1 (revised 2022-08-18). Open Science Framework. 2022. https://osf.io/e7pvk/.

69. de Onis M, Garza C, Victora CG, Onyango AW, Frongillo EA, Martines J. The WHO Multicentre Growth Reference Study: Planning, Study Design, and Methodology. Food Nutr Bull. 2004;25 1_suppl_1:S15–26. https://doi.org/10.1177/15648265040251S104.

70. WHO. Length/height-for-age, weight-for-age, weight-for-length, weight-for-height and body mass index-for-age: Methods and development. Geneva: World Health Organization; 2006.

71. Arnold BF, Galiani S, Ram PK, Hubbard AE, Briceño B, Gertler PJ, et al. Optimal Recall Period for Caregiver-reported Illness in Risk Factor and Intervention Studies: A Multicountry Study. Am J Epidemiol. 2013;177:361–70. https://doi.org/10.1093/aje/kws281.

72. Arnold BF, Ercumen A, Benjamin-Chung J, Colford JM. Brief Report: Negative Controls to Detect Selection Bias and Measurement Bias in Epidemiologic Studies. Epidemiology. 2016;27:637–41. https://doi.org/10.1097/EDE.0000000000000504.

73. Lappan R, Henry R, Chown SL, Luby SP, Higginson EE, Bata L, et al. Monitoring of diverse enteric pathogens across environmental and host reservoirs with TaqMan array cards and standard qPCR: a methodological comparison study. Lancet Planet Health. 2021;5:e297–308. https://doi.org/10.1016/S2542-5196(21)00051-6.

74. Baker KK, Senesac R, Sewell D, Sen Gupta A, Cumming O, Mumma J. Fecal Fingerprints of Enteric Pathogen Contamination in Public Environments of Kisumu, Kenya, Associated with Human Sanitation Conditions and Domestic Animals. Environ Sci Technol. 2018;52:10263–74. https://doi.org/10.1021/acs.est.8b01528.

medRxiv preprint doi: https://doi.org/10.1101/2022.08.19.22278991; this version posted August 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

75. Rostami A, Ma G, Wang T, Koehler AV, Hofmann A, Chang BCH, et al. Human toxocariasis – A look at a neglected disease through an epidemiological 'prism.' Infect Genet Evol. 2019;74:104002. https://doi.org/10.1016/j.meegid.2019.104002.

76. Schneider AG, Casanovas-Massana A, Hacker KP, Wunder EA Jr, Begon M, Reis MG, et al. Quantification of pathogenic Leptospira in the soils of a Brazilian urban slum. PLoS Negl Trop Dis. 2018;12:1–15. https://doi.org/10.1371/journal.pntd.0006415.

77. Muirhead A, Zhu K, Brown J, Basu M, Brinton MA, Costa F, et al. Zika Virus RNA Persistence in Sewage. Environ Sci Technol Lett. 2020;7:659–64. https://doi.org/10.1021/acs.estlett.0c00535.

78. Chakrabarti AK, Caruso L, Ding M, Shen C, Buchanan W, Gupta P, et al. Detection of HIV-1 RNA/DNA and CD4 mRNA in feces and urine from chronic HIV-1 infected subjects with and without anti-retroviral therapy. AIDS Res Ther. 2009;6:20. https://doi.org/10.1186/1742-6405-6-20.

79. Jirků M, Pomajbíková K, Petrželková KJ, Hůzová Z, Modrý D, Lukeš J. Detection of *Plasmodium* spp. in Human Feces. Emerg Infect Dis. 2012;18. https://doi.org/10.3201/eid1804.110984.

80. Sultana S, Ansar A, Saif-Ur-Rahman KM. Stool specimen for diagnosis of pulmonary tuberculosis in adults: protocol for a systematic review and meta-analysis. BMJ Open. 2021;11:e052212. https://doi.org/10.1136/bmjopen-2021-052212.

81. Zhu K, Suttner B, Pickering A, Konstantinidis KT, Brown J. A novel droplet digital PCR human mtDNA assay for fecal source tracking. Water Res. 2020;:116085. https://doi.org/10.1016/j.watres.2020.116085.

82. Holcomb DA, Stewart JR. Microbial Indicators of Fecal Pollution: Recent Progress and Challenges in Assessing Water Quality. Curr Environ Health Rep. 2020;7:311–24. https://doi.org/10.1007/s40572-020-00278-1.

83. Gillings MR, Gaze WH, Pruden A, Smalla K, Tiedje JM, Zhu Y-G. Using the class 1 integron-integrase gene as a proxy for anthropogenic pollution. ISME J. 2015;9:1269–79. https://doi.org/10.1038/ismej.2014.226.

84. Keenum I, Liguori K, Calarco J, Davis BC, Milligan E, Harwood VJ, et al. A framework for standardized qPCR-targets and protocols for quantifying antibiotic resistance in surface water, recycled water and wastewater. Crit Rev Environ Sci Technol. 2022;:1–25. https://doi.org/10.1080/10643389.2021.2024739.

85. Julian TR, Islam MA, Pickering AJ, Roy S, Fuhrmeister ER, Ercumen A, et al. Genotypic and Phenotypic Characterization of Escherichia coli Isolates from Feces, Hands, and Soils in Rural Bangladesh via the Colilert Quanti-Tray System. Appl Environ Microbiol. 2015;81:1735–43. https://doi.org/10.1128/AEM.03214-14.

medRxiv preprint doi: https://doi.org/10.1101/2022.08.19.22278991; this version posted August 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

86. Karthikeyan S, Levy JI, De Hoff P, Humphrey G, Birmingham A, Jepsen K, et al. Wastewater sequencing reveals early cryptic SARS-CoV-2 variant transmission. Nature. 2022. https://doi.org/10.1038/s41586-022-05049-6.

87. Gelman A, Pardoe I. 2. Average Predictive Comparisons for Models with Nonlinearity, Interactions, and Variance Components. Sociol Methodol. 2007;37:23–51. https://doi.org/10.1111/j.1467-9531.2007.00181.x.

88. Clark JS, Nemergut D, Seyednasrollah B, Turner PJ, Zhang S. Generalized joint attribute modeling for biodiversity analysis: median-zero, multivariate, multifarious data. Ecol Monogr. 2017;87:34–56. https://doi.org/10.1002/ecm.1241.

89. Rafi Z, Greenland S. Semantic and cognitive tools to aid statistical science: replace confidence and significance by compatibility and surprise. BMC Med Res Methodol. 2020;20:244. https://doi.org/10.1186/s12874-020-01105-9.

90. Bottomley C, Kirby MJ, Lindsay SW, Alexander N. Can the buck always be passed to the highest level of clustering? BMC Med Res Methodol. 2016;16:29. https://doi.org/10.1186/s12874-016-0127-1.

91. Rothman KJ. No Adjustments Are Needed for Multiple Comparisons. Epidemiology. 1990;1. https://doi.org/10.1097/00001648-199001000-00010.

92. VanderWeele TJ. Principles of confounder selection. Eur J Epidemiol. 2019;34:211–9. https://doi.org/10.1007/s10654-019-00494-6.

93. Schreiner M. Simple Poverty Scorecard Poverty-Assessment Tool Mozambique. Microfinance Risk Management, LLC; 2017. https://www.simplepovertyscorecard.com/MOZ_2014_ENG.pdf.

94. VanderWeele TJ, Knol MJ. A Tutorial on Interaction. Epidemiol Methods. 2014;3:33–72. https://doi.org/10.1515/em-2013-0005.

95. Arnold BF, Hogan DR, Colford JMJ, Hubbard AE. Simulation methods to estimate design power: an overview for applied research. BMC Med Res Methodol. 2011;11:94. https://doi.org/10.1186/1471-2288-11-94.

96. Gelman A, Carlin J. Beyond Power Calculations: Assessing Type S (Sign) and Type M (Magnitude) Errors. Perspect Psychol Sci. 2014;9:641–51. https://doi.org/10.1177/1745691614551642.

97. Ioannidis JPA. Why Most Discovered True Associations Are Inflated. Epidemiology. 2008;19:640–8. https://doi.org/10.1097/EDE.0b013e31818131e7.

98. Lash TL. The Harm Done to Reproducibility by the Culture of Null Hypothesis Significance Testing. Am J Epidemiol. 2017;186:627–35. https://doi.org/10.1093/aje/kwx261.

medRxiv preprint doi: https://doi.org/10.1101/2022.08.19.22278991; this version posted August 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

99. Gurevitch J, Koricheva J, Nakagawa S, Stewart G. Meta-analysis and the science of research synthesis. Nature. 2018;555:175–82. https://doi.org/10.1038/nature25753.

100. Mathur MB, VanderWeele TJ. Sensitivity analysis for publication bias in meta-analyses. J R Stat Soc Ser C Appl Stat. 2020;69:1091–119. https://doi.org/10.1111/rssc.12440.

101. Gelman A, Loken E. The Statistical Crisis in Science. Am Sci. 2014;102:460. https://doi.org/10.1511/2014.111.460.

102. Greenland S. For and Against Methodologies: Some Perspectives on Recent Causal and Statistical Inference Debates. Eur J Epidemiol. 2017;32:3–20. https://doi.org/10.1007/s10654-017-0230-6.

103. Lash TL, Fox MP, Maclehose RF, Maldonado G, Mccandless LC, Greenland S. Good practices for quantitative bias analysis. Int J Epidemiol. 2014;43:1969–85. https://doi.org/10.1093/ije/dyu149.

104. Freedman DA. Statistical Models and Shoe Leather. Sociol Methodol. 1991;21:291. https://doi.org/10.2307/270939.

105. Wikswo ME, Kambhampati A, Shioda K, Walsh KA, Bowen A, Hall AJ. Outbreaks of Acute Gastroenteritis Transmitted by Person-to-Person Contact, Environmental Contamination, and Unknown Modes of Transmission — United States, 2009–2013. MMWR Surveill Summ. 2015;64:1–16. https://doi.org/10.15585/mmwr.ss6412a1.

106. Messa A, Köster PC, Garrine M, Gilchrist C, Bartelt LA, Nhampossa T, et al. Molecular diversity of Giardia duodenalis in children under 5 years from the Manhiça district, Southern Mozambique enrolled in a matched case-control study on the aetiology of diarrhoea. PLoS Negl Trop Dis. 2021;15:e0008987. https://doi.org/10.1371/journal.pntd.0008987.

107. Donowitz JR, Alam M, Kabir M, Ma JZ, Nazib F, Platts-Mills JA, et al. A Prospective Longitudinal Cohort to Investigate the Effects of Early Life Giardiasis on Growth and All Cause Diarrhea. Clin Infect Dis. 2016;63:792–7. https://doi.org/10.1093/cid/ciw391.

108. Bartelt LA, Platts-Mills JA. Giardia: a pathogen or commensal for children in high-prevalence settings? Curr Opin Infect Dis. 2016;29:502–7. https://doi.org/10.1097/qco.0000000000000293.

109. Rogawski ET, Bartelt LA, Platts-Mills JA, Seidman JC, Samie A, Havt A, et al. Determinants and Impact of Giardia Infection in the First 2 Years of Life in the MAL-ED Birth Cohort. J Pediatr Infect Dis Soc. 2017;6:153–60. https://doi.org/10.1093/jpids/piw082.

110. Benjamin-Chung J, Crider YS, Mertens A, Ercumen A, Pickering AJ, Lin A, et al. Household finished flooring and soil-transmitted helminth and Giardia infections among children in rural Bangladesh and Kenya: a prospective cohort study. Lancet Glob Health. 2021;9:e301–8. https://doi.org/10.1016/s2214-109x(20)30523-4.

medRxiv preprint doi: https://doi.org/10.1101/2022.08.19.22278991; this version posted August 22, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

111. Hernán MA, Hernández-Díaz S, Robins JM. A Structural Approach to Selection Bias. Epidemiology. 2004;15:615–25. https://doi.org/10.1097/01.ede.0000135174.63482.43.

112. Goddard FGB, Chang HH, Clasen TF, Sarnat JA. Exposure measurement error and the characterization of child exposure to fecal contamination in drinking water. Npj Clean Water. 2020;3:19. https://doi.org/10.1038/s41545-020-0063-9.